

Complex Self-organized Multi-pulse Dynamics in a Fiber Laser: The Rain of Solitons

S. Chouli and Ph. Grelu

Laboratoire Interdisciplinaire Carnot de Bourgogne, UMR 5209 CNRS, Université de Bourgogne
9 Avenue A. Savary, 21000 Dijon, France

Abstract— We present experimental studies of new dynamics observed in a fiber laser operated in a highly-pumped but weakly mode-locked regime. In this situation, both soliton pulses and cw background coexist and interact. In a specific dynamics which we have called “soliton rain”, additional solitons arise spontaneously in the cavity from the fluctuations of a background and drift until they reach a condensed phase of aggregated solitons. This process can go on forever in a quasi-stationary fashion. We have characterized the “soliton rain” dynamics, explored neighboring dynamics, and demonstrated a convenient way to control the soliton rain threshold by injecting an external cw laser inside the fiber laser cavity.

1. INTRODUCTION

Passively mode-locked fiber lasers are ideal tools for the exploration of new areas of solitons nonlinear dynamics in an open, dissipative environment [1, 2]. A common understanding is that mode locking appears as an abrupt transition from a noisy cw operation to a clean — background-free — short-pulsed laser operation. However, there can be significant deviations to such scenario. The existence of dispersive waves that are radiated by the pulse as it travels through the discontinuities of the laser cavity medium can significantly alter the spectral and temporal pulsed features [3, 4]. For instance, the observation of noiselike pulses is an example of puzzling dynamics that does not fall in the conventional picture of mode-locked operation [5]. Such intriguing is the recent finding of the soliton rain dynamics, which corresponds to an intermediate regime where soliton pulses and cw components of comparable strengths not only coexist, but also interact in a dramatic way [6, 7].

2. OBSERVATION AND CHARACTERIZATION OF SOLITONS RAINS

Experiments are performed with a passively mode-locked fiber ring laser detailed in [6]. Passive mode locking (ML) is obtained through the ultrafast saturable absorber effect that results from nonlinear polarization evolution. The laser dynamics can be controlled through paddle orientations of polarization controllers (PC).

We have observed at large pumping power a gradual transition from the cw regime to the ML regime. This transition is described as follows. For a given initial orientation of the set of PCs, the laser operates in a noisy-highly multimode-cw regime. Change of the orientation of a PC produces first the appearance of weak ML components next to the cw components. As the PC is turned further the ML components become stronger, until the cw components disappear. In the intermediate regime, namely where the cw components and soliton pulses coexist in the cavity, *rains of solitons* are observed: new solitons arise spontaneously from fluctuations of the background and drift at constant velocity until they merge with a condensed phase of jittering bound solitons.

Soliton rain is due to the interaction between three elements — see Fig. 1: the noisy background, drifting solitons and the condensed phase. The noisy background is produced by a large number of quasi-cw components. In principle, these quasi-cw components can be produced out of: the amplified spontaneous emission (ASE), cavity cw modes and dispersive waves produced by solitons. On average, drifting solitons are created far apart each others (nanosecond separation) with respect to their temporal widths (around 1 ps). The probability of soliton creation depends on the background level which is not uniform. Once formed, one soliton drifts at a rather low constant speed of the order of 10 m/s. Eventually, it collides with the condensed soliton phase. The condensed phase comprises several tens of solitons in a bunch wherein there is a large timing jitter. Soliton rains detailed features -size of the condensed phase, number of drifting solitons and their drifting velocity can be adjusted along with cavity parameters [7].

In terms of global dynamics and energy flows we get a picture of the interactions between the three elements that make a soliton rain. The background, through the amplification of its fluctuations, seeds the creation of drifting solitons like a droplet would be formed from a vapor cloud. Then, drifting solitons feed the condensed soliton phase. Since the size of the condensed

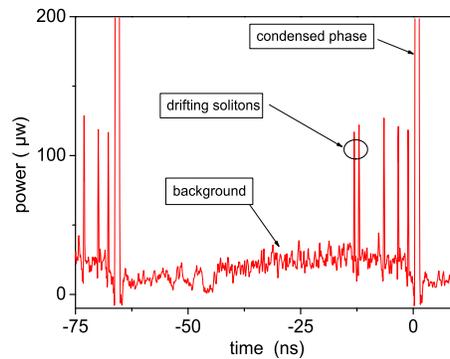


Figure 1: Output laser intensity — one cavity roundtrip time is 67 ns — showing the three field components of the soliton rain: the noisy background, drifting solitons and the condensed soliton phase.

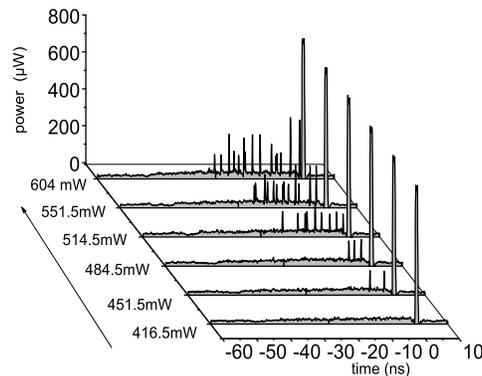


Figure 2: Influence of pumping power to soliton rain dynamics.

phase remains constant, this means that the arrival of new solitons causes the dissipation of a similar number of solitons. This dissipation releases additional radiation and dispersive waves that contribute to the background, such as the evaporation of liquid would contribute to the clouds. Because of these analogies with the cycle of water we have dubbed this fiber laser dynamics “rain of solitons”.

3. CONTROL OF SOLITONS RAINS

Since the soliton rain starts above a certain level of background fluctuations, it is possible to control its appearance. One way of control is through the pumping power. The increase of the pumping power increases the level of fluctuations in the background. For a sufficient pumping power, the solitons rain starts with 1 or 2 drifting solitons at a time per cavity roundtrip, see Fig. 2. Increasing the pumping power further (450 mW–551 mW), the number of drifting solitons increases up to making up a heavy soliton rain (604 mW). At this point, the probability that two solitons are created closer together becomes significant, so that interactions between neighboring drifting solitons can manifest.

The second way in which we can increase the level of cw fluctuations, is via the injection of light into the cavity. We have demonstrated the control of the appearance and disappearance of the soliton rain by injecting an external cw laser, whose wavelength and intensity can be varied.

Here follows the experimental procedure: firstly, the pumping power is adjusted to be close below the soliton rain threshold, with the injected laser “off” see Figs. 3(a) and 3(b). After switching “on” the injected laser, the soliton rain starts and lasts as long as injection continues Fig. 3(c). When injection is turned off, the soliton rain stops almost immediately, after the last soliton created finishes its drift to the condensed phase. We note that the injected laser can trigger soliton rain only when its wavelength is on the short-wavelength side of the soliton spectrum. A small injected power level as low as 10 microwatts is sufficient to triggered the soliton rain Fig. 3(d). In addition to controlling the appearance and disappearance of soliton rain, it is possible to tune the soliton drifting speed, from 12 to 5 m/s when the injected power level increases from 100 to 220 microwatts.

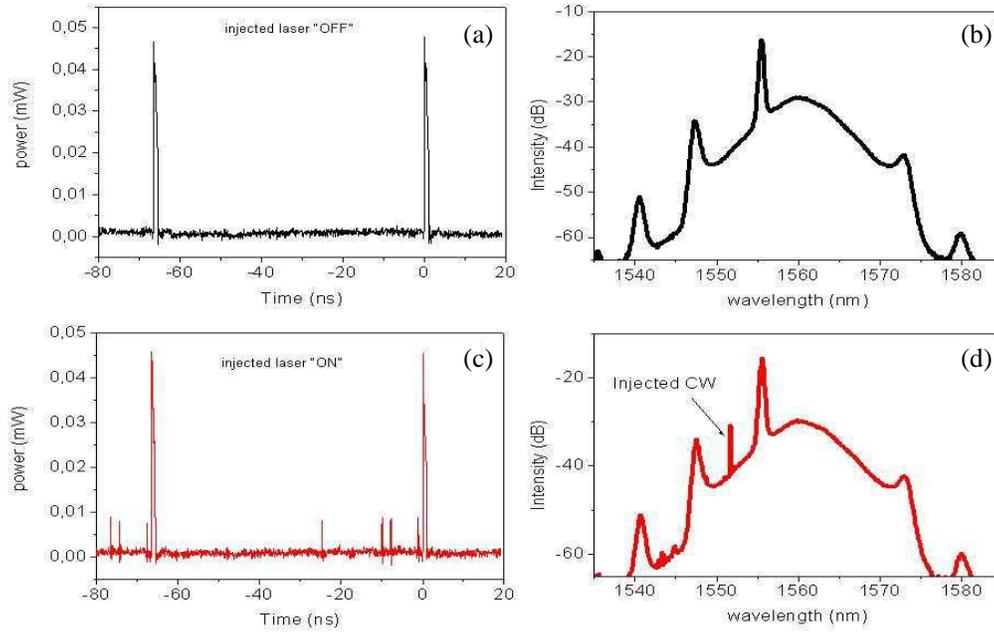


Figure 3: (a) Temporal optical intensity, injected laser off. (b) Optical spectrum, injected laser off. (c): Temporal optical intensity, injected laser on. (d) Optical spectrum, injected laser on.

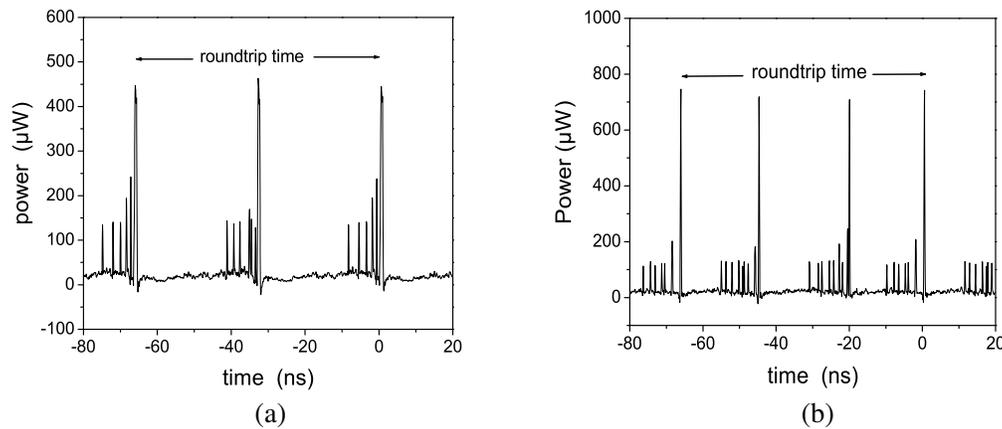


Figure 4: (a) Second cavity soliton rain harmonic at 30 MHz repetition rate. (b) Third cavity soliton rain harmonic at 45 MHz.

4. SOLITON RAIN HARMONICS

Close to the soliton rain, at high pumping power, we have observed soliton rain harmonics such as second (Fig. 4(a)) and third (Fig. 4(b)) soliton rain harmonics. The existence of soliton rain harmonics is due to the instability of the condensed phase when its size exceeds a certain level typically 40 solitons. The condensed phase then breaks into parts, and each new condensed part becomes related to its own background part that produces drifting solitons.

5. COMPARISON WITH OTHER SOLITON SELF-ORGANIZATIONS

5.1. Release of Solitons from the Condensed Phase

A typical soliton rain spectrum is presented in Fig. 5(a): it always features a cw peak on the short-wavelength side close to the center of the soliton spectrum, as well as stronger soliton sidebands on the short-wavelength side. However, we found specific settings of the PCs that produced an inversion with respect to the dominant spectral asymmetry. Both the cw and soliton sideband components now dominate on the long-wavelength side of the soliton spectrum — see Fig. 5(c). In the temporal domain, we observe the following: instead of drifting towards the condensed phase

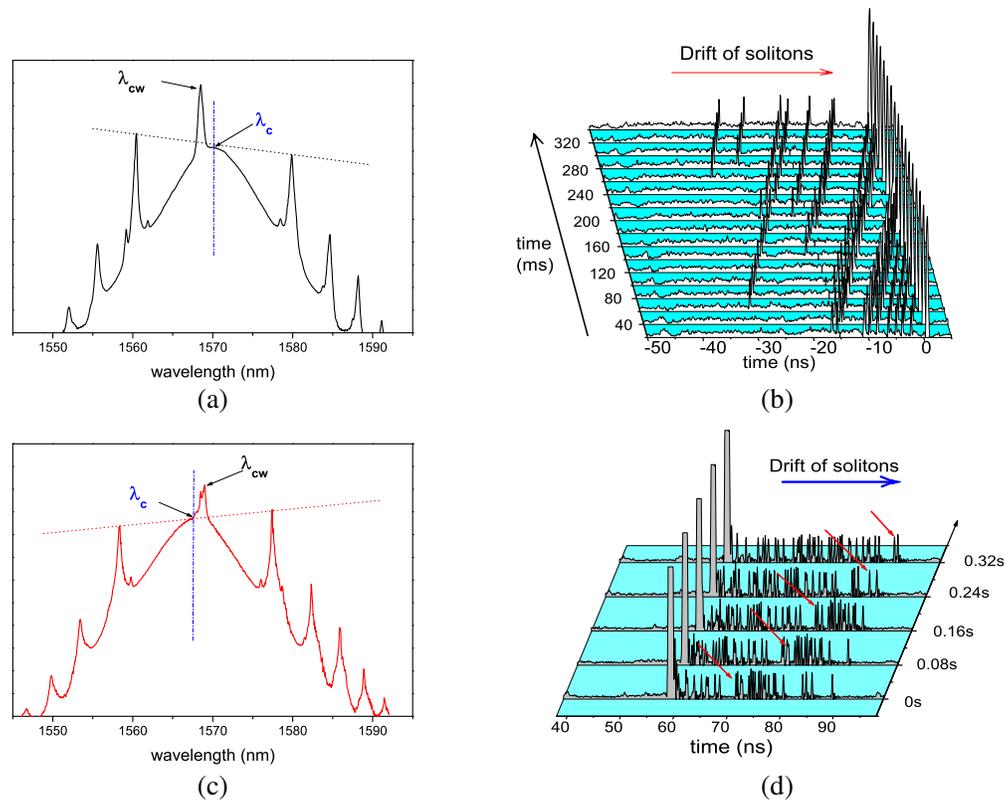


Figure 5: comparison between soliton rain dynamics where individual solitons flow towards the condensed phase and quasi-stationary release of solitons from the condensed phase: (a) optical spectrum of soliton rain, (b) corresponding stroboscopic recording of the temporal dynamics; and (c) optical spectrum of release of solitons, (d) corresponding stroboscopic recording of the temporal dynamics.

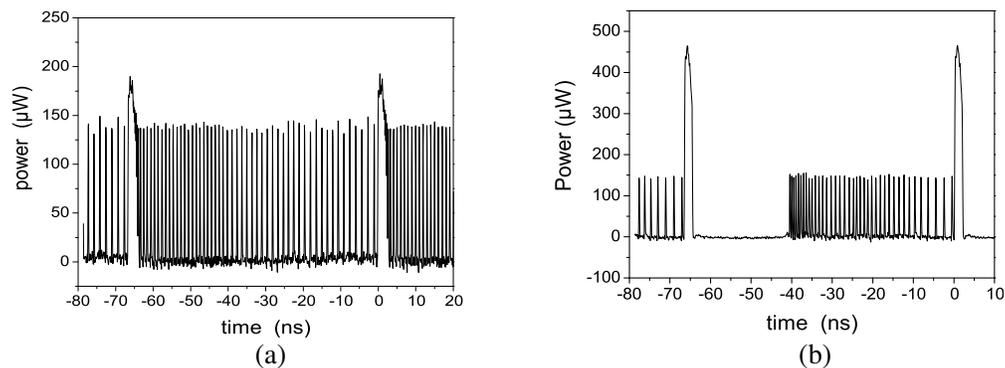


Figure 6: Chirped soliton trains and condensed phase: (a) chirped train filling the entire cavity; (b) chirped train partly filling the cavity and attached to a single condensed phase.

(Fig. 5(b)), the solitons are coming out of the condensed phase, flowing outward until they vanish inside the background (Fig. 5(d)). It becomes clear that the precise composition of the spectrum and the impact of chromatic dispersion dictate the temporal dynamics.

5.2. Coexistence of Stationary Chirped Trains and the Condensed Phase

Close to the domain of existence of soliton rains, we observed other examples of self-organization of dissipative solitons, where the drifting motion is absent and the observed patterns appear stationary.

One example is the *chirped trains and condensed phase* pattern: in this dynamics the individual solitons are stationary, they get organized with their neighbors in a train of pulses whose separation increases with time. In Fig. 6(a), the chirped train fills the entire cavity and is attached to the condensed phase at both ends. In Fig. 6(b), the chirped train occupies two-thirds of one cavity round trip and is attached at one end of the condensed phase. The extension of the chirped train

can be almost continually selected by a fine tuning of the orientation of PC. We assume that the non-uniformity of soliton-soliton spacing is caused by the non-uniformity of the background [6]. Tuning the PC is indeed able to translate or alter the effective non-linear transfer function so as to increase or decrease the level of background and radiation that live in the cavity along with solitons.

6. CONCLUSION

Soliton rain is a complex but beautiful illustration of self-organization among large number of solitons and radiation components. It is observed in a fiber laser operated at anomalous chromatic dispersion, high pumping power, and weakly mode-locked regime. The dynamics is due to interactions between three field components: from **background fluctuations**, additional **drifting solitons** are created and drift at constant relative velocity until they merge with a bunch of several tens of bound jittering solitons which we called the **condensed phase**. Since the size of condensed phase remains constant, the condensed phase emits a large amount of radiation that contribute to the background. The whole scenario repeats in a quasi-stationary fashion. Building a comprehensive theoretical model able to retrieve such complex interactions among large number of cw and solitons components is a challenging task that needs to be undertaken.

REFERENCES

1. Soto-Crespo, J. M., P. Grelu, N. Akhmediev, and N. Devine, “Soliton complexes in dissipative systems: Vibrating, shaking and mixed soliton pairs,” *Phys. Rev. E*, Vol. 75, 016613, 2007.
2. Grelu, P. and N. Akhmediev, “Group interactions of dissipative solitons in a laser cavity: the case of $2 + 1$,” *Opt. Express*, Vol. 12, 3184–3189, 2004.
3. Gordon, J. P., “Dispersive perturbations of solitons of the nonlinear Schrodinger equation,” *J. Opt. Soc. Am. B*, Vol. 9, No. 1, 91–97, 1992.
4. Soto-Crespo, J. M., N. Akhmediev, P. Grelu, and F. Belhache, “Quantized separations of phase-locked soliton pairs in fiber lasers,” *Opt. Lett.*, Vol. 28, No. 19, 1757–1759, 2003.
5. Horowitz, M., Y. Barad, and Y. Silberberg, “Noiselike pulses with a broadband spectrum generated from an erbium-doped fiber laser,” *Opt. Lett.*, Vol. 22, No. 10, 799–801, 1997.
6. Chouli, S. and P. Grelu, “Rains of solitons in a fiber laser,” *Opt. Express*, Vol. 17, 11776–11781, 2009.
7. Chouli, S. and P. Grelu, “Soliton rains in a fiber laser: An experimental study,” *Phys. Rev. A*, Vol. 81, 063829–063829, 2010.

Principal Component Analysis for Low-dimensional Modeling of Mode-locked Lasers

J. N. Kutz, M. Williams, E. Shlizerman, and E. Ding

Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA

Abstract— The ubiquitous onset of multi-pulsing in a laser cavity is considered in two representative laser models. Using a low-dimensional reduction constructed by the method of proper orthogonal decomposition (principal components), a complete characterization of the multi-pulsing transition is given, including the onset of periodic (Hopf) solutions and a period doubling and Neimark-Sacker (torus) bifurcation route to chaos. This is the first low-dimensional construction of the entire multi-pulsing transition from N to $N + 1$ pulses per round trip. The reduced model qualitatively reproduces the dynamics observed in the multi-pulse transition of the mode-locked laser and confirms recent experimental observations of periodic and chaotic behavior preceding the multi-pulsing transition.

1. INTRODUCTION

The onset of multi-pulsing in mode-locked lasers is a commonly observed phenomenon in pulsed lasers. The multi-pulsing instability effectively determines the maximum power and energy that can be delivered by the mode-locked pulses, thus limiting the performance of a given laser. Only a limited theoretical description currently exists to describe the multi-pulsing transition observed in mode-locking [1, 2]. However, these energy based theoretical arguments are sufficient to highlight the complicated and chaotic nature of the transition dynamics which have been recently observed experimentally [3, 4]. It is the fundamental nature of this transition that this manuscript attempts to address. Within the context of a low-dimensional model, generated by the so-called proper orthogonal decomposition, a complete description of the laser cavity dynamics can be generated, giving valuable insight into the engineering and performance features of mode-locked lasers.

Alternative methods exist for producing dimensionality reduction in optics. As an example, the variational method can be modified [5] to account for large dissipative effects associated with mode-locking [6]. In this reduction technique, the validity of the resulting model is determined by the accuracy of the chosen solution (pulse) ansatz. For single-pulse solutions, a chirped hyperbolic secant ansatz has been used to reproduce the dynamics observed in full simulations of several mode-locking models [5, 6]. However, no viable ansatz has been developed for characterizing the multi-pulsing transition dynamics. Motivated by dimensionality reduction techniques in other fields, such as turbulent flow and pattern formation, the proper orthogonal decomposition (POD), which is also known as principal component analysis (PCA), has been used in combination with Galerkin projection to obtain a low dimensional model of systems that exhibit complex dynamics. In the context of mode-locking, this method has only recently been used to model the multi-pulse transition in ring cavity lasers using waveguide arrays [7] or within context of the cubic quintic Ginzburg-Landau equation [8]. In this case, the low-dimensional reduction captures the transition in the laser cavity that is characterized by a fold bifurcation of a limit cycle that did not exhibit chaos.

2. GOVERNING EQUATIONS: WAVEGUIDE ARRAY LASER

The waveguide array mode-locked laser (WGAML) describes the mode-locking process in a fiber-waveguide system that combines the nonlinear mode coupling of waveguides (saturable absorption) with the saturating and bandwidth limited gain of erbium-doped fiber. Although this is the specific model considered, the multi-pulsing transition phenomenon observed for the WGAML appears to be ubiquitous in mode-locked lasers [3, 4]. The governing equations for the WGAML are given by

$$i \frac{\partial A_0}{\partial Z} + \frac{D}{2} \frac{\partial^2 A_0}{\partial T^2} + \beta |A_0|^2 A_0 - ig(Z) \left(1 + \tau \frac{\partial^2}{\partial T^2} \right) A_0 + i\gamma_0 A_0 + CA_1 = 0 \quad (1a)$$

$$i \frac{\partial A_1}{\partial Z} + C(A_0 + A_2) + i\gamma_1 A_1 = 0 \quad (1b)$$

$$i \frac{\partial A_2}{\partial Z} + CA_1 + i\gamma_2 A_2 = 0 \quad (1c)$$

with

$$g(Z) = \frac{2g_0}{1 + ||A_0||^2/e_0}. \quad (2)$$

where A_0 , A_1 and A_2 are the envelopes of the electric field in waveguides 0, 1 and 2 respectively. In waveguides 1 and 2, which are modeled by Eqs. (1b)–(1c), there is no injected gain and the cubic and the dispersive terms are assumed to be negligible to leading order. Due to the injection of gain into waveguide 0, its amplitude is not small and the cubic Kerr nonlinearity ($|A_0|^2 A_0$) and chromatic dispersion ($\partial_T^2 A_0$) terms appear in the governing equation along with the gain term ($-ig(Z)(1 + \tau\partial_T^2)A_0$). Common to all the waveguides are the effects of linear loss (attenuation) within each of the individual waveguides and linear (evanescent) coupling between nearest neighbors.

3. LOW-DIMENSIONAL PROJECTION

To construct such a reduced model, the numerical solutions of Eq. (1) are projected onto a set of orthogonal functions (a Galerkin projection). Due to the complicated nature of the solutions, the minimal orthogonal basis to represent the solutions cannot be trivially determined. For example, a large number of Fourier modes are required to represent the chaotic solution during the multipulsing transition process. Therefore, we use the proper orthogonal decomposition to obtain the minimal orthogonal basis and project the WGAML dynamics onto this basis. This method will be called the POD-Galerkin projection.

The POD method is related to the singular value decomposition (SVD). Computationally, the SVD is implemented as a built-in routine in many scientific software packages, such as MATLAB or NumPy. To generate a complete set of POD modes, a data set is compiled and represented as a matrix \mathbf{X} . Each row of the matrix consists of a sample solution taken at a specific value of Z , and the number of rows in the matrix is the number of samples taken at different but evenly spaced values of Z . Therefore, if the data consists of m samples with n points per sample, then $\mathbf{X} \in \mathbb{C}^{m \times n}$. The SVD factorizes the matrix \mathbf{X} into three matrices.

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^* \quad (3)$$

where $\mathbf{U} \in \mathbb{C}^{m \times m}$, $\mathbf{V} \in \mathbb{C}^{n \times n}$ and $\mathbf{\Sigma} \in \mathbb{R}^{m \times n}$ and the asterisk denotes the conjugate transpose.

One possible way to reduce the dimensionality of the matrix \mathbf{X} is to use a subset of the POD basis. The relative importance of the j -th POD mode $\vec{\phi}_j$ in the approximation of the matrix \mathbf{X} is determined by the relative energy E_j of that mode, defined as

$$E_j = \frac{\sigma_j^2}{\sum_{i=1}^n \sigma_i^2} \quad (4)$$

where the total energy is normalized such that $\sum_{j=1}^n E_j = 1$. If the sum of the energies of the retained modes is unity, then these modes can be used to completely reconstruct \mathbf{X} . Typically, the number of modes required to capture all of the energy is very large and does not result in a significant dimensionality reduction. In practice, however, it has been found that the matrix can be accurately approximated by using POD modes whose corresponding energies sum to almost all of the total energy. Then the POD basis that is used in the approximation is a truncated basis, where the POD modes with negligible energy are neglected. In practice, the truncated basis whose energies sum to 99% of the total energy is considered as a plausible truncation. The advantage of using a truncated set of POD modes rather than any other set of modes is that the representation of the data generated by the POD modes is guaranteed to have a smaller least squares error than the representation of the data generated by any other orthonormal set of the same size.

For the WGAML, the matrix \mathbf{X} is composed of numerical solutions of the WGAML taken from a single waveguide at evenly spaced values of Z and T . The SVD is then performed on each data set individually and the largest N modes, i.e., those that contain the most energy as characterized

by their singular values, are retained. The solution of the WGAML is then approximated by

$$A_0(T, Z) = \sum_{j=1}^N a_j(Z) \phi_{0,j}(T) \quad (5a)$$

$$A_1(T, Z) = \sum_{j=1}^N b_j(Z) \phi_{1,j}(T) \quad (5b)$$

$$A_2(T, Z) = \sum_{j=1}^N c_j(Z) \phi_{2,j}(T) \quad (5c)$$

where $\phi_{i,j}$ is the j -th POD mode from the i -th waveguide and N are the number of modes retained. Eq. (5) is substituted into the WGAML in Eq. (1), and the inner product of the resulting equations with $\phi_{i,j}$ is taken. The resulting set of differential equations are

$$\frac{\partial a_n}{\partial Z} = \left(\frac{D}{2} + g\tau \right) \sum_{j=1}^N \langle \phi_{0,n} \partial_T^2 \phi_{0,j} \rangle a_j + (g - \gamma_0) a_n \quad (6a)$$

$$+ iC \sum_{j=1}^N \langle \phi_{0,n} \phi_{0,j} \rangle b_j + i\beta \sum_{j=1}^N \sum_{k=1}^N \sum_{m=1}^N \langle \phi_{0,n} \phi_{0,k} \phi_{0,j} \phi_{0,m}^* \rangle a_j a_k a_m^*$$

$$\frac{\partial b_n}{\partial Z} = iC \sum_{j=1}^N (\langle \phi_{1,n} \phi_{0,j} \rangle a_j + \langle \phi_{1,n} \phi_{2,j} \rangle c_j) - \gamma_1 b_n \quad (6b)$$

$$\frac{\partial c_n}{\partial Z} = iC \sum_{j=1}^N \langle \phi_{2,n} \phi_{1,j} \rangle b_j - \gamma_2 c_n \quad (6c)$$

where $\langle u, v \rangle = \int_{-\infty}^{\infty} u^* v dT$ and

$$g = \frac{2g_0}{1 + \sum_{j=1}^N |a_j|^2 / e_0}. \quad (7)$$

Note that POD modes are obtained for each of the waveguides separately because this allows for changes in the relative phase between waveguides to be accurately and explicitly captured as g_0 is increased.

4. LOW-DIMENSIONAL DYNAMICS

For POD modes computed for solutions of the WGAML with a fixed value of gain g_0 , the reduced model obtained by a Galerkin projection onto a small number of POD modes accurately reproduces the observed behavior. However, the validity and accuracy of the resulting models diminishes when g_0 is varied. In addition, the accuracy of the model can be reduced if a bifurcation in the WGAML model occurs and the truncated set of the POD modes does not include the POD modes of the bifurcating dynamics. To study the multi-pulse transition, the reduced model must be valid for a range of g_0 and capture the bifurcations in the transition. Therefore, multiple individual runs, with various values of g_0 , need to be combined into a single amalgamated data set. Each individual run consists of a short-time, high-resolution solution, $Z \in [0, 10]$ and $\Delta Z = 5 \times 10^{-3}$. The initial conditions chosen are a small perturbation away from the attractor at the fixed value of g_0 considered. By combining these individual runs into a single data set, the SVD provides the set of POD modes that best captures the energy of the global dynamics through the transition from one- to two-pulses.

Figure 1 shows the first six POD modes generated using this methodology for solutions of the even-WGAML model. Only the modes of the 0th waveguide are shown. However, there are two additional sets of modes, one for each of the other two waveguides. The data set includes information from the single-pulse, the breather, the chaotic, and the double-pulse solutions. These six modes capture over 99% of the modal energy, and 99.9% may be captured by including an additional four modes. The POD modes appear to be a nontrivial combination of the modes from

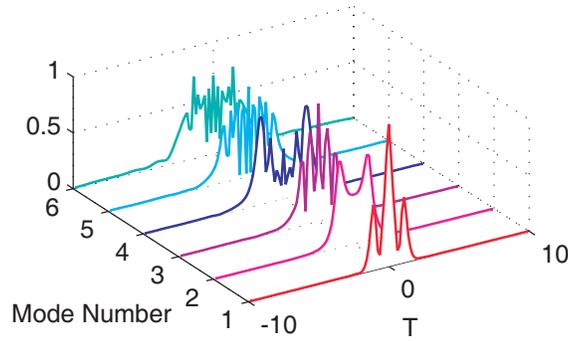


Figure 1: POD modes taken from the combined data set for the 0th waveguide that are capable of qualitatively reproducing the dynamics observed in the full WGAML model.

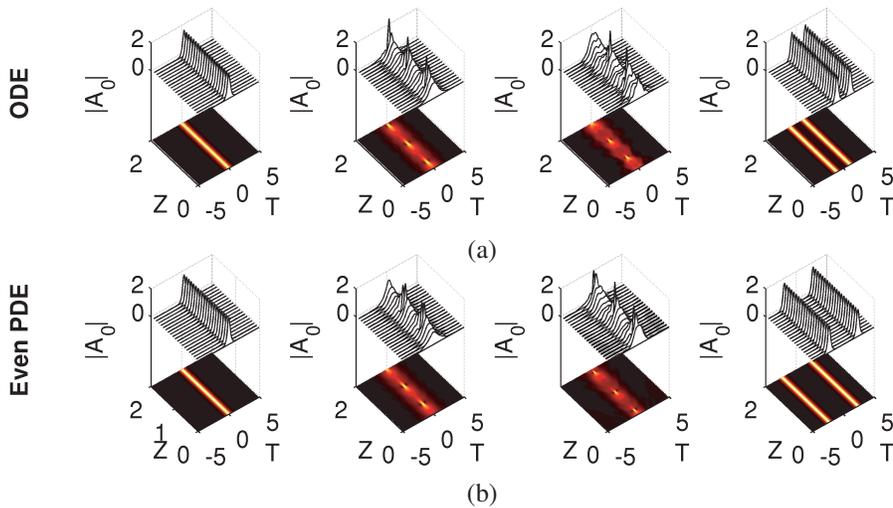


Figure 2: (a) Surface and pseudo-color plots of the single-pulse, breather, chaotic, and double pulse solutions computed for the finite-dimensional model at $g_0 = 1.5, 2.5, 3.495,$ and 3.5 respectively. (b) The same plots for the even WGAML model taken at $g_0 = 2.3, 2.5, 3.1,$ and 3.2 respectively. The reduced model accurately reproduces the four behaviors observed.

all the runs at different gain values. These modes would be impossible to predict *a priori* even with knowledge about the solutions of the system.

The POD-reduced model is used to study the multi-pulse transition of the WGAML. The reduced model is obtained by the Galerkin projection of WGAML model in Eq. (1) onto the global POD modes in Figure 1. This produces a system of differential equations of the form shown in Eq. (6). To determine the validity of the reduced model, we compare the full WGAML dynamics and the POD models for fixed values of g_0 in the relevant ranges for single-pulse, breather, chaotic, and double-pulse solutions using low amplitude noise as the initial condition in both the full and reduced dynamics. The evolution considered is for a long period of time so that the attractor is reached for each value g_0 considered. In this manner, the full and reduced dynamics of the WGAML starting from a cold-cavity configuration can be compared. In addition to comparing the solutions at a single value of g_0 , we also compare the branches of single-pulse and breather solutions in both models over a range of g_0 values to reveal the route to chaos during the multi-pulse transition.

The first row of Figure 2 shows the single-pulse, breather, chaotic, and double-pulse solutions reconstructed using a six-mode POD model, ordered from left to right. To obtain these solutions, the reduced model was evolved forward for 1000 units in Z starting with a low-amplitude white-noise initial condition. These results compare favorably with the same four regimes of the even WGAML dynamics shown in the second row of Figure 2, whose solutions were obtained by evolving for 200 units in Z starting from low amplitude even white-noise. The reduced model qualitatively captures the dynamics and profile of the solution. The primary difference is the value of g_0 at which these solutions are obtained. The stationary solutions of the POD model lose stability more quickly than

the stationary solutions of the WGAML. Further, the breather solutions in the POD model remain stable for larger values of g_0 than the WGAML. Due to the vastly smaller dimension of the reduced model and the range of g_0 for which the POD generated differential equations are valid, some disparity between the models is inevitable. Although the reduced model does not quantitatively match the full dynamics, it still qualitatively matches all the observed dynamics. Furthermore, these results also compare well with the unrestricted WGAML solution except for the translations in T that were disallowed by imposing evenness during the chaotic regime.

5. CONCLUSIONS

The POD-Galerkin can be a useful tool for the studies of the dynamics of nonlinear evolution equations describing various optical systems. Indeed, our study here and reproduce the dynamics observed, including a faithful reproduction of the multi-pulsing instability and transition. In many cases where the observed dynamics are coherent but not trivial and exhibit various phenomena, the POD-Galerkin is the natural choice to construct a reduced and tractable model. If the model correctly reproduces the observed dynamics, it can be then studied using regular computational methods and in special cases analytically, to identify the bifurcations responsible for these phenomena. The POD-Galerkin methodology can be easily modified to changes in the model. Such robustness of the method, allows one to study the operating regimes as a function of the parameters in the system. Here the WGAML was modeled as a three-waveguide array. However, using the same methodology a system of N -waveguide arrays could be studied. Only minor modifications to Eq. (6) are required to model a system with a different number of waveguides. Then parameters for optimal mode-locking in the such a system or the impact of the parameters on the multi-pulse transition can be studied. In addition, models based on the POD-Galerkin are sometimes able to predict dynamics in regions outside of the data set where the POD modes were computed, the low-dimensional model can be used to find interesting regions within the system without having to develop an appropriate ansatz. Once these regions have been identified, by altering the used data set, a more accurate representation of the region can be developed. In summary, the POD reduction gives a completely new and insightful way to explore the dynamics and bifurcations of a given system.

ACKNOWLEDGMENT

J. N. Kutz acknowledges support from the National Science Foundation (DMS-1007621) and the US Air Force Office of Scientific Research (AFOSR) (FA9550-09-0174).

REFERENCES

1. Namiki, S., E. P. Ippen, H. A. Haus, and C. X. Yu, "Energy rate equations for mode-locked lasers," *J. Opt. Soc. Am. B*, Vol. 14, 2099–2111, 1997.
2. Feng, L., P. K. A. Wai, and J. N. Kutz, "A geometrical description of the onset of multi-pulsing in mode-locked laser cavities," *J. Opt. Soc. Am. B*, Vol. 27, 2068–2077, 2010.
3. Bale, B. G., K. Kieu, J. N. Kutz, and F. Wise, "Transition dynamics for multi-pulsing in mode-locked lasers," *Opt. Express*, Vol. 17, 23137–23146, 2009.
4. Soto-Crespo, J. M., M. Grapinet, P. Grelu, and N. Akhmediev, "Bifurcations and multiple period soliton pulsations in a passively mode-locked ber laser," *Phys. Rev. E*, Vol. 70, 66612, 2010.
5. Anderson, D., M. Lisak, and A. Berntson, "A variational approach to nonlinear evolution equations in optics," *Pramana*, Vol. 57, 917–936, 2001.
6. Bale, B. G. and J. N. Kutz, "Variational method for mode-locked lasers," *J. Opt. Soc. Am. B*, Vol. 25, 1193–1202, 2008.
7. Williams, M., E. Shlizerman, and J. N. Kutz, "The multi-pulsing transition in mode-locked lasers: A low-dimensional approach using waveguide arrays," *J. Op. Soc. Am. B*, Vol. 27, 2471–2481, 2010.
8. Ding, E., E. Shlizerman, and J. N. Kutz, "Modeling multi-pulsing transition in ring cavity lasers with proper orthogonal decomposition," *Phys. Rev. A*, Vol. 82, 023823, 2010.

Extraordinary Transmission and Light Confinement in Subwavelength Metallic Films Apertures

R. Ortuño, C. García-Meca, F. J. Rodríguez-Fortuño, and A. Martínez
Nanophotonics Technology Center, Universidad Politécnica de Valencia, Spain

Abstract— We have studied the relationship between internal-surface plasmon polaritons and negative permeability in cascaded patterned metallic layers. By properly selecting the dielectric thicknesses of the multilayer structure, the negative effective index can broaden which could be of potential use in the design of metamaterials. The excitations of cladding modes are also shown to cause extraordinary light transmission. Several subwavelength hole arrays are fabricated and characterized showing extraordinary light transmission peaks. With the presented results we hope to extend the potential of these structures for use in modern developing applications.

1. INTRODUCTION

Light-matter interactions in a metal layer patterned at the subwavelength scale give rise to a wide variety of optical resonances. Since the discovery of extraordinary light transmission (EOT) [1], numerous experiments on such subwavelength apertures of opaque metallic films have demonstrated the ability to control light confinement and light propagation at the subwavelength scale.

The pattern in periodic structures allows coupling mechanisms between impinging light and surface plasmon waves. The excitation of horizontal coupled surface plasmon polaritons (SPP) on the surfaces of the grating is pointed out to explain the extraordinary transmission of light through periodic arrays of subwavelength apertures in metallic films [2]. However, other mechanisms are also responsible for the high transmission and huge confinement of light in metallic gratings. For instance, waveguide modes, which appear in one-dimensional gratings made of narrow slits, also play an important role in the onset of extraordinary transmission [3]. In addition, when the patterned metallic film is surrounded by a dielectric material, the excitation of cladding modes by the incoming light also cause extraordinary transmission resonances [4].

Moreover, cascading extraordinary transmission structures leads to a magnetic resonance response which yields a left-handed behavior and a negative refraction over a specific frequency range [5–7]. By properly adjusting the separation distance among the cascaded metallic layers the negative refractive index can be extended over a larger frequency range [8].

These phenomena allow us to design planar optical elements having different refractive indices and spectral dispersion properties. This is very attractive for many optical applications in fundamental optics and in optoelectronics, like lenses and filters, and open up new possibilities in controlling and manipulating light by a wide range of plasmonic materials. In this sense, we have designed subwavelength structures for filtering purposes in the mid-infrared frequency range with high transmission levels. In addition, one advantage of such filters is their simple and compact implementation by comparison with conventional multilayer filters due to their huge number of layers and their large thicknesses required in the mid-infrared domain.

2. EXPERIMENTAL AND SIMULATION RESULTS

Though extraordinary transmission can be observable through isolated holes [9] or surrounded by periodic corrugations [10], the most common structure employed for the onset of EOT is a one- or two-dimensional aperture array drilled in an opaque metal film due to the simplicity with which their spectral properties can be tuned and scaled. Traditionally, EOT studies have been focused on single-layer structures. It is until quite recently when studies of multilayer aperture arrays have gained more attention, including lateral incidence configuration achieving broadband all-angle negative refraction [11]. Other works have also analyzed analytically the propagation in stacked devices [12] and reinterpreted the model through an equivalent circuit [12, 13]. In this way, Figure 1 shows the transmission through a dielectric layer sandwiched between two metal films drilled with a two-dimensional hole array. As it is already known, periodic metallic structures can convert impinging light into SPP by providing the necessary momentum conservation for the coupling process [14]. It is therefore not surprising that periodic aperture arrays such as those shown in Figure 1 can give rise to EOT phenomenon where the transmission spectrum contains a set of peaks larger than one (when normalized to the lattice-hole area ratio) even when the

individual apertures are so small that they do not allow propagation of light. When several metal layers are placed close enough, SPP modes can be excited among the metallic films, giving rise to the so-called internal-SPP (I-SPP). In addition, external-SPP (E-SPP) i.e., the similar SPP mode that appears in single metallic films structures is also supported by the structure. Both the I-SPP and the E-SPP resonances can be seen in the spectrum transmission of Figure 1(b) (top panel). However, the most interesting result is that a negative permeability is obtained at those frequencies at which an I-SPP is excited as Figure 1(b) (middle panel) shows. In the case that the negative permeability region overlaps with a negative permittivity region a negative effective index is achieved (see Figure 1(b) (bottom panel)) [5] so a negative index EOT (NEOT) is obtained [15].

At those frequencies at which NEOT phenomenon occurs a left-handed propagation is obtained, merging the field of plasmonics with that of metamaterials. From a practical point of view, it could be desirable to obtain a negative index over a large frequency range. However, in order to accomplish that, we have to face with the problem that the negative index comes from resonating effects and so it spreads over a limited frequency range. One possibility we have exploit to dodge this problem is to take advantage of the hybridization of SPP. In this sense, if several metal layers separated with dielectrics films of different thickness are stacked, so that different -but close enough- I-SPP could be excited, a broader negative index bandwidth with negative permeability values can be achieved [8]. Figure 2(a) shows the structure under consideration where two different thicknesses (t_1 and t_2) are considered for the dielectric layers. On the other hand, the optical properties of the structure were studied as a function of the number of stacked unit cells N in order to check the convergence of the parameters when stacking unit cells. Figure 2(b) shows new resonances appearing due to plasmon hybridization, coming from the vertical electromagnetic coupling of neighboring layers [8]. Due to these hybridization effects which lead the negative permeability to spread as the number of stacked unit cells increase, the double-negative range (i.e., $\epsilon, \mu < 0$) widens. And, consequently, the negative index bandwidth and the FOM broaden (see Figure 2(c)).

Figure 3(a) depicts a patterned metallic layer surrounded by a dielectric material. Only in the case the structure is periodically patterned, multiple peaks emerge in the transmission spectrum showing EOT, as shown in Figure 3(b). Due to the periodicity, the incoming light can couple to cladding modes which couple evanescently through the drilled metallic film leading to multiple EOT. Figure 3(c) shows the calculated transmission coefficient $T(f, k_x)$ for the same dielectric thickness as in Figure 3(a). The multiple EOT peaks appeared in the transmission spectrum of Figure 3(b) can be explained if we just consider the values in k_x associated to the periodicity in the metallic hole array, for which light coupling can be possible. This is of great significance as it relates the important role of the periodicity not only to surface modes, like SPPs, but also to propagating modes in the dielectric claddings in the onset of EOT resonances [4]. It is worth mentioning that a very similar result has been analytically obtained for sandwiched periodic distributions of slits employing an equivalent circuit model to explain the transmission of electromagnetic waves through the perforated conducting screen sandwiched between dielectrics layers [16].

In our aim of exploiting the EOT phenomena in the design of planar optical elements we have fabricated several subwavelength hole arrays made on silver and on gold for filtering purposes in the

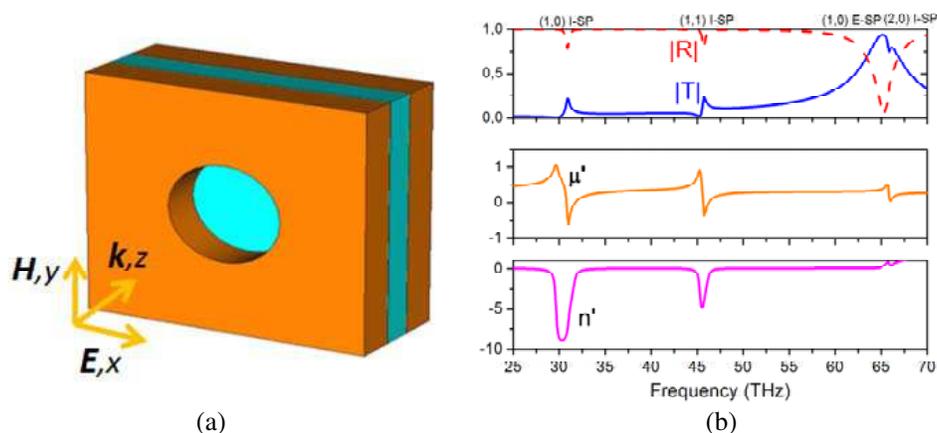


Figure 1: (a) Schematic diagram of the double-layer hole array unit cell. (b) Calculated transmission and reflection spectrum (top), and retrieved effective permeability (middle) and refractive index (down).

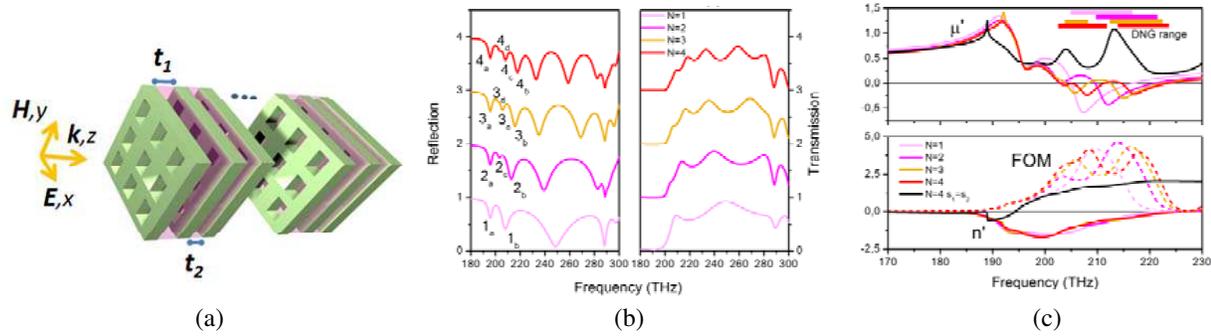


Figure 2: Scheme of the analyzed multilayer metamaterial. (b) Simulated reflection and transmission spectra. (c) Retrieved effective material parameters of permeability and refractive index along with figure of merit.

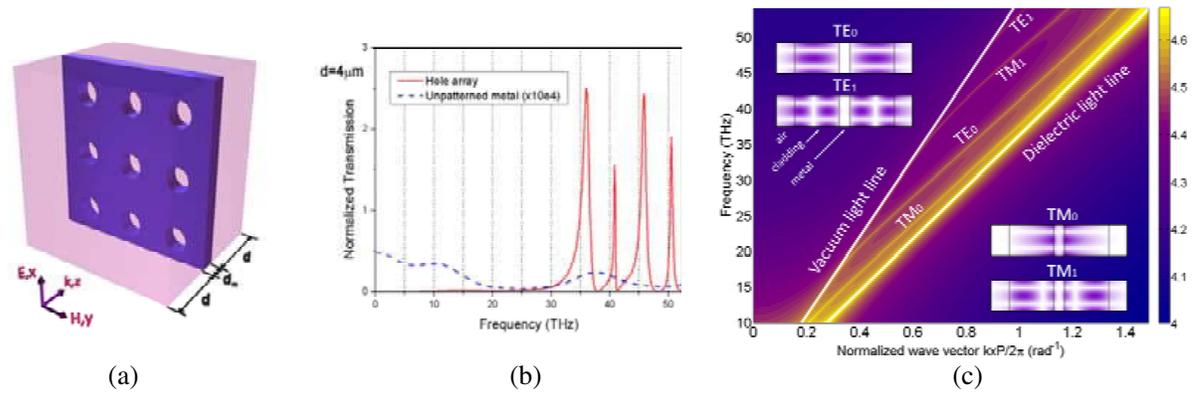


Figure 3: (a) Schematic view of the metallic patterned film surrounded by dielectric materials. (b) Transmission through the hole array with dielectric thickness $d = 4 \mu\text{m}$. (c) Theoretical transmission coefficient $T(f, k_x)$ for a dielectric thickness of $d = 4 \mu\text{m}$.

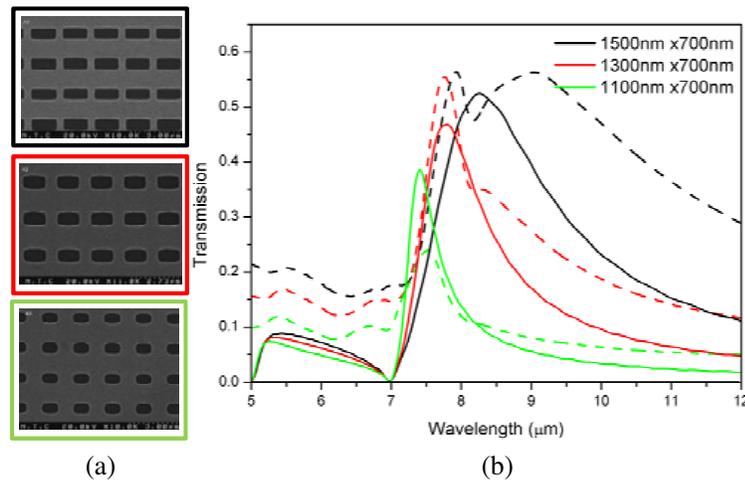


Figure 4: (a) SEM images of three subwavelength periodic arrays and (b) their measured (dashed) and simulated (solid) transmission spectrum.

mid-infrared range with high transmission levels. Figure 4 shows three scanning electron microscope images (SEM) of subwavelength periodic arrays fabricated with 100 nm of gold film evaporated on a silicon wafer. All samples have a square lattice array spacing of $2 \mu\text{m}$. Different hole openings are considered for each sample. The corresponding hole dimensions are listed in the legend of the figure of the transmission spectrum. The measurements were recorded using the BrukerTM FTIR spectrometer (dashed lines) and show a good agreement with the expected theoretical transmission (solid lines). EOT can be clearly seen on the transmission spectra of all samples [17]. These

devices present the additional advantages of being very compact in the propagation direction and their scalability and CMOS compatible fabrication.

3. CONCLUSIONS

We have studied the relation between the I-SPP excitation and the onset of negative permeability in cascaded structured metallic films. By exploiting this phenomenon we have demonstrated the possibility of enlarging the negative effective refractive index in fishnet metamaterials by selecting the convenient dielectric thicknesses comprising the structure. In addition, we have shown how the cladding modes, excited thanks to the lattice sketched on the metal film, are responsible of the multiple peaks which appear in the transmission spectra. On the other hand, several metallic subwavelength hole arrays have been fabricated and characterized showing EOT. This behaviour can be used to implement ultracompact optical filters for several applications such as gas detection. The presented results will allow for properly understanding and controlling the EOT phenomenon, and will extend the potential of these structures for use in modern developing applications.

ACKNOWLEDGMENT

Financial support by the Spanish MICINN under Contracts No. PET2007-0505, TEC2008-06871-C02-02 and CSD2008-00066 (Consolider EMET) is gratefully acknowledged. The authors also acknowledge grant program FPI of UPV, FPU of MICINN and FPI of GVA, respectively.

REFERENCES

1. Ebbesen, T. W., H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, "Extraordinary optical transmission through sub-wavelength hole arrays," *Nature*, Vol. 391, No. 6668, 667–669, 1998.
2. Genet, C. and T. W. Ebbesen, "Light in tiny holes," *Nature*, Vol. 445, No. 7123, 39–46, 2007.
3. Porto, C. A., F. J. García-Vidal, and J. B. Pendry, "Transmission resonances on metallic gratings with very narrow slits," *Phys. Rev. Lett.*, Vol. 83, No. 14, 2845–2848, 1999.
4. Ortuño, R., C. Gracia-Meca, F. J. Rodríguez-Fortuño, and A. Martínez, "Multiple extraordinary optical transmission peaks from evanescent coupling in perforated metal plates surrounded by dielectrics," *Opt. Express*, Vol. 18, No. 8, 7893–7898, 2010.
5. Ortuño, R., C. Gracia-Meca, F. J. Rodríguez-Fortuño, J. Martí, and A. Martínez, "Role of surface plasmon polaritons on optical transmission through double layer metallic hole arrays," *Phys. Rev. B*, Vol. 79, No. 7, 075425, 2009.
6. Beruete, M., M. Sorolla, and I. Campillo, "Left-handed extraordinary optical transmission through a photonic crystal of subwavelength hole arrays," *Opt. Express*, Vol. 14, No. 12, 5445–5455, 2006.
7. Beruete, M., M. Sorolla, M. Navarro-Cía, F. Falcone, I. Campillo, and V. Lomakin, "Extraordinary transmission and left-handed propagation in miniaturized stacks of doubly periodic subwavelength hole arrays," *Opt. Express*, Vol. 15, No. 3, 1107–1114, 2007.
8. Ortuño, R., C. Gracia-Meca, F. J. Rodríguez-Fortuño, and A. Martínez, "Enlarging the negative index bandwidth of optical metamaterials by hybridized plasmon resonances," *Opt. Lett.*, Vol. 35, No. 23, 2010 (to be published).
9. Degiron, A., H. J. Lezec, N. Yamamoto, and T. W. Ebbesen, "Optical transmission properties of a single subwavelength aperture in a real metal," *Opt. Commun*, Vol. 239, No. 1–3, 61–66, 2004.
10. Lezec, H. J., A. Degiron, E. Devaux, R. A. Linke, L. Martín-Moreno, F. J. García-Vidal, and T. W. Ebbesen, "Beaming light from a subwavelength aperture," *Science*, Vol. 297, No. 5582, 820–822, 2002.
11. Wei, Z., Y. Cao, J. Han, C. Wu, Y. Fan, and H. Li, "Broadband negative refraction in stacked fishnet metamaterial," *Appl. Phys. Lett.*, Vol. 15, No. 14, 141901, 2010.
12. Marqués, R., L. Jelinek, F. Mesa, and F. Medina, "Analytical theory of wave propagation through stacked fishnet metamaterials," *Opt. Express*, Vol. 17, No. 14, 11582–11593, 2009.
13. Carbonell, J., C. Croënne, F. Garet, E. Lheurette, J. L. Coutaz, and D. Lippens, "Lumped elements circuit of terahertz fishnet-like arrays with composite dispersion," *J. Appl. Phys.*, Vol. 108, No. 1, 014907, 2010.
14. Maier, S. A., *Plasmonics, Fundamentals and Applications*, Springer, New York, 2007.
15. García-Meca, C., R. Ortuño, F. J. Rodríguez-Fortuño, J. Martí, and A. Martínez, "Negative refractive index metamaterials aided by extraordinary optical transmission," *Opt. Express*, Vol. 17, No. 8, 6026–6031, 2009.

16. Rodríguez-Berral, R., F. Mesa, and F. Medina, “Circuit model for a periodic array of slits sandwiched between two dielectric slabs,” *Appl. Phys. Lett.*, Vol. 96, No. 16, 161104, 2010.
17. Ortuño, R., C. García-Meca, F. J. Rodríguez-Fortuño, A. Hakansson, A. Griol, J. Hurtado, J. A. Ayucar, L. Bellieres, P. J. Rodríguez, F. López-Royo, J. Martí, and A. Martínez, “Mid-infrared filters based on extraordinary optical transmission through subwavelength structured gold films,” *J. Appl. Phys.*, Vol. 106, No. 12, 124313, 2009.

Transmission Properties of Dual-period Arrays of Cylinders

D. C. Skigin¹ and M. Lester²

¹Applied Electromagnetics Group, Physics Dept., FCEN, University of Buenos Aires
and IFIBA (CONICET), Argentina

²Grupo Optica de Sólidos-Elfo, Instituto de Física Arroyo Seco, Facultad de Ciencias Exactas
Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina

Abstract— We investigate the potential of dual-period structures to control and manipulate the transmitted intensity. We consider supergratings (periodic arrays with a compound unit cell) in which each period comprises several cylinders. We show that this kind of structures permits one to control the diffracted response, regardless of the cylinder material and the incident polarization. A given diffraction order can be either enhanced or suppressed by appropriately choosing the geometrical parameters of the structure, and this property is basically a geometric effect. For subwavelength cylinders, if their axes are aligned in a plane tilted with respect to the periodicity direction, the structure behaves like a blazed grating in the sense of its capability to enhance the intensity in a pre-designed direction. This blazed-like behaviour is found for both incident polarization modes and for dielectric as well as for metallic cylinders. If the array is illuminated by an evanescent wave, the system exhibits resonances and the inhomogeneous wave is converted into propagating waves that radiate to the far field. This effect can be controlled by varying the geometrical parameters of the structure, such as the period and the inclination angle, what makes these structures good potential candidates for many applications.

1. INTRODUCTION

Dual-period arrays in which each period is formed by a subarray of cylinders forming a finite grating, have not received much attention until recently. These structures offer more versatility for the design of optical devices than usual periodic structures since they have more geometrical parameters that can be tailored to make them adequate for a particular purpose. The dual-period characteristic of gratings has been recently studied in connection with the coupling of surface plasmon polaritons to the incident wave for different applications [1–5]. The existence of phase resonances in dual-period arrays has been theoretically predicted for reflection gratings with rectangular grooves [6] as well as for transmitting structures formed by rectangular metallic wires [7–9]. Very recently, experimental evidence of this kind of resonances, which are only allowed for p -polarization (magnetic field parallel to the rulings), have been given [10, 11].

In this paper we investigate the electromagnetic response of dual-period structures comprising several cylinders in each period, which can be either metallic or dielectric, and analyze their properties to control the transmitted response. In Sec. 2 we investigate the ability of such structures to enhance the intensity of a given diffraction order [12] and show numerical results that evidence this property. In Sec. 3 we analyze the behaviour of these dual-period structures when they are illuminated by an evanescent wave. Our results show that these structures behave like optical antennas that convert evanescent into propagating waves, what suggests that this structure is a good candidate to be used in highly sensitive detection devices and other applications. Finally, concluding remarks are given in Sec. 4.

2. PROPAGATING WAVE INCIDENCE

The structure under study is a finite array of N cylinders of radius r immersed in vacuum. The cylinders are grouped in M subarrays, each of which comprises J cylinders ($N = M \times J$) whose axes are separated a distance d . In the scheme of Fig. 1, $N = 15$, $M = 3$ and $J = 5$. The cylinders' axes of each subarray are contained in the same plane and the planes corresponding to different subarrays are parallel to each other, forming an angle ϕ with respect to the periodicity direction (x), as shown in Fig. 1(a). The subarrays are equally spaced, forming a finite grating with period D , and the system is normally illuminated by a Gaussian beam of wavelength λ and half-width W . The scattering problem is rigorously solved by means of an integral method based on the application of Green's second identity to the 2D Helmholtz equation, followed by the imposition of the appropriate boundary conditions at the interfaces, and the application of the Extinction Theorem [13, 14]. In the following examples we show that dual-period arrays comprising circular cylinders exhibit particular characteristics in their transmitted response.

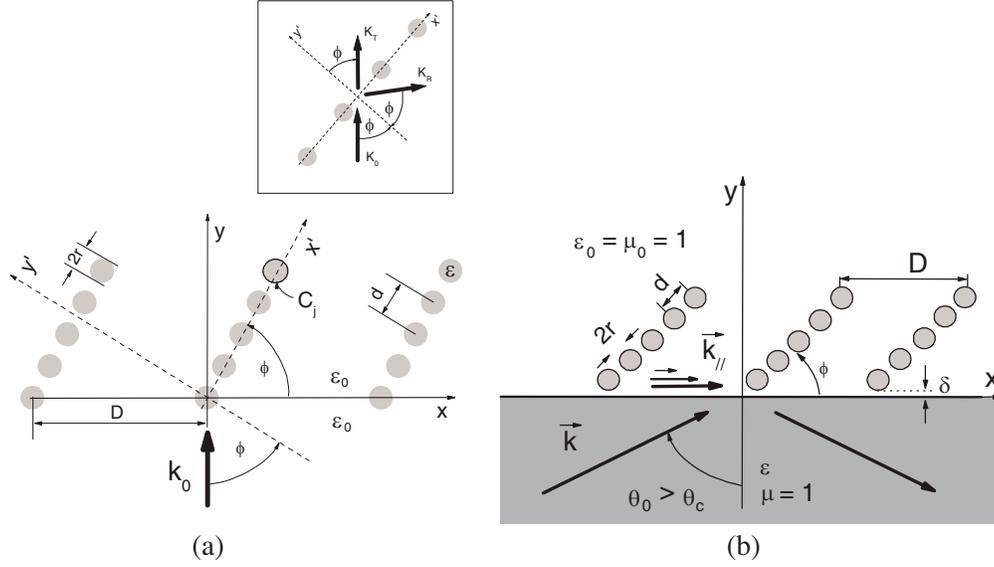


Figure 1: Configuration of the scattering problem. In this scheme $N = 15$, $M = 3$ and $J = 5$. (a) Propagating wave incidence. Inset: propagation directions of incident \mathbf{k}_0 , transmitted \mathbf{k}_T and reflected \mathbf{k}_R wavevectors in the x', y' reference frame; (b) evanescent wave incidence.

In Fig. 2 we show transmitted intensity maps as a function of the observation angle (θ) and the inclination angle (ϕ) (see caption for the parameters) for both polarization modes, for silver cylinders (Figs. 2(a) and 2(b)) and for silica cylinders (Figs. 2(c) and 2(d)). Three vertical fringes are observed for θ corresponding to the forward and to the ± 1 transmitted orders. The θ locations of these maxima are independent of ϕ , since these directions — diffraction orders in the infinite grating case — are determined by the grating equation:

$$\sin \theta_m = \sin \theta_0 + m \frac{\lambda}{D}. \quad m \in Z \quad (1)$$

For normal illumination and for the parameters considered in this example, these angles are $\theta_{\pm 1} = 36.8^\circ$. Although the inclination angle ϕ does not affect the maxima directions, it is evident from the plots that ϕ modifies the intensity distribution. For $\phi = 0$ the symmetry of the structure produces a symmetric power distribution, most of which propagates in the specular direction. As ϕ increases, more power propagates in the positive region of observation angles. It is evident from Fig. 2 that the inclination angles that maximize the intensity in the $+1$ transmitted order are $\phi \approx 62.5^\circ$ for s polarization and $\phi \approx 71.5^\circ$ for p polarization. For these large values of ϕ it is to be expected that resonant mechanisms such as cavity modes and coupling will be involved in the diffraction process. To get some physical insight into these results we apply a simple model. If each subarray is regarded as a plane for the incident wave — we recall that the distance between adjacent cylinders is much smaller than the wavelength of the incident wave —, an intensification of the transmitted beam should occur for the direction specularly reflected at this plane. For $\phi > 45^\circ$ the beam specularly reflected by each plane propagates towards the $y > 0$ half-space (see inset in Fig. 1), i.e., it becomes part of the transmitted field. According to the definition of the observation angle, the intensification direction is $\theta = \pi - 2\phi$. To enhance the m -th order, this angle should be the same as that of the transmitted order given by the grating Eq. (1). Then, the optimum ϕ value is given by:

$$\phi = \frac{\pi - \arcsin\left(m \frac{\lambda}{D}\right)}{2}. \quad (2)$$

For the parameters of Fig. 2 this estimation leads to $\phi = 71.5^\circ$, which is in very good agreement with the maximum intensity shown in Figs. 2(b) and 2(d) for p polarization. However, for the s case the maximum is slightly shifted from this value. This shift can be attributed to the excitation of a cavity mode within the structure: since the diameter of the cylinders and the distance between them are subwavelength, each subarray can be regarded as a slab, with an effective dielectric constant. Therefore, an open cavity is formed between subarrays, whose length and width are approximately

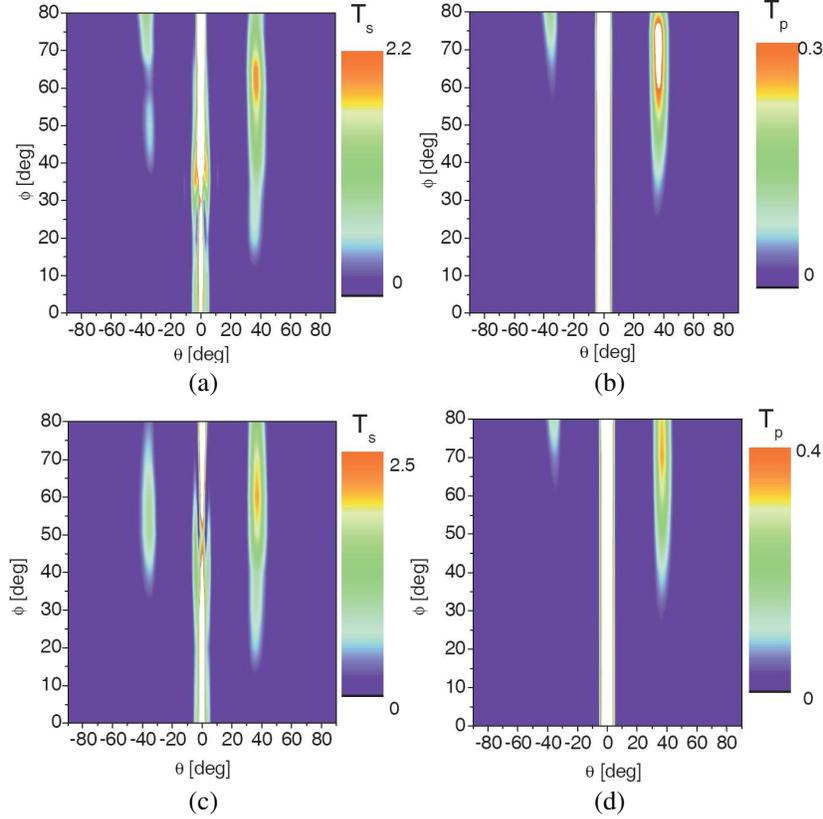


Figure 2: Contour plots of the transmission coefficient as a function of the observation (θ) and inclination (ϕ) angles, for $M = J = 5$ ($N = 25$), $r = 50$ nm, $\lambda = 800$ nm, $\lambda/D = 0.6$ ($D = 1333.33$ nm), $\lambda/d = 3$ ($d = 266.66$ nm), $W = 5000$ nm. (a) Silver cylinders ($\epsilon_{Ag} = -27.995 + i1.523$), s polarization; (b) silver cylinders, p polarization; (c) silica cylinders ($\epsilon_S = 3$), s polarization; (d) silica cylinders, p polarization. The white zones represent values that exceed those displayed in the right scale.

$l = D(1 - \cos \phi)$ and $w = D \sin \phi$, respectively. For $l/\lambda \approx 1$, a longitudinal mode is excited while the field within the cavity is intensified, producing an enhancement of the transmitted field in the allowed directions. For the parameters of Fig. 2 and for $\phi = 62.5^\circ$, $l \simeq 0.9\lambda$ and $w \simeq 1.5\lambda$, establishing an eigenmode between subarrays, as can be observed in the near field plot of Fig. 3(a). It is evident from Fig. 3(b) that the near field is still dominated by this mode for $\phi = 71.5^\circ$, which inhibits the intensification of the transmitted $+1$ order by the specular effect described before. For this value of ϕ the picture is different for p polarization, where the main intensification mechanism is a significant coupling between neighboring cylinders of the same subarray, as can be observed in Fig. 3(c).

3. EVANESCENT WAVE INCIDENCE

We consider that the system is illuminated by an evanescent wave generated by total internal reflection (TIR) at the plane $y = 0$. In Fig. 4 we show contour plots of the transmitted intensity as a function of the observation angle and of the incident wavelength for a system of silver cylinders with $M = J = 5$, $r = 20$ nm, $d = 50$ nm and $D = 800$ nm, for $\phi = \pm 45^\circ$. The distance from the closest wire to the interface is $\delta = 5$ nm, $\epsilon = 2.32$ and the incidence angle is $\theta_0 = 42^\circ$ (see Fig. 1(b)). The angular distribution of the transmitted intensity is dominated by the period D of the structure, and it follows the dispersion relation of a one-dimensional periodic system (diffraction grating), as observed for propagating incidence. The finitely periodic character of the system adds an extra momentum $n\kappa$, with n integer and $\kappa = 2\pi/D\hat{x}$, to the momentum of the incident wave \mathbf{k}_{\parallel} . The black lines in Fig. 4 correspond to the diffraction orders calculated as:

$$k \sin(\theta_n) = k_{\parallel} \pm n\kappa. \quad (3)$$

An excellent agreement is observed between the transmittance maxima distribution and the curves given by Eq. (3). Notice that if the inclination angle ϕ of the wires' chains is changed, the direction

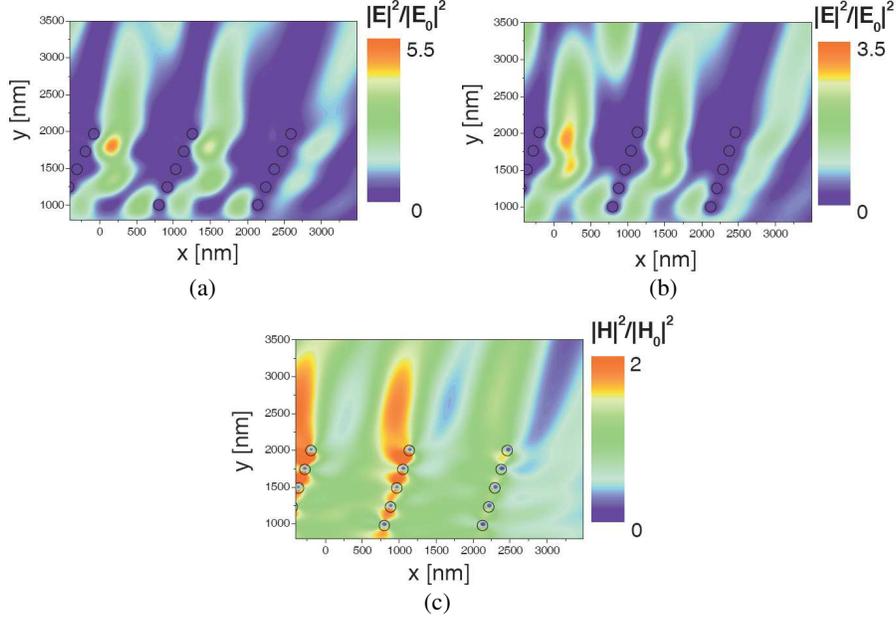


Figure 3: Near field plots for the silver structure considered in Fig. 2 for different ϕ -values. (a) s polarization and $\phi = 62.5^\circ$; (b) s polarization and $\phi = 71.5^\circ$; (c) p polarization and $\phi = 71.5^\circ$.

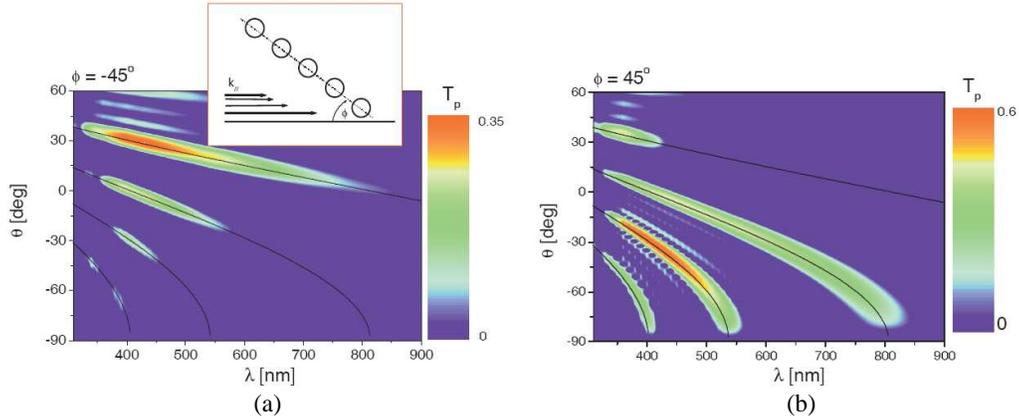


Figure 4: Transmitted intensity as a function of the observation angle and of the incident wavelength for a compound array of wires ($N = 25$, $M = 5$, $J = 5$) with $r = 20$ nm, $d = 50$ nm, $D = 800$ nm, $\epsilon = 2.32$ and $\theta_0 = 42^\circ$. (a) $\phi = -45^\circ$; (b) $\phi = 45^\circ$.

of the maximum transmitted intensity also changes, such that this system behaves somehow like a blazed grating [12]. By inspection of the near field at resonant conditions, it becomes evident that the mechanism responsible for the intensification is a surface plasmon excitation (not shown).

4. CONCLUSION

We have investigated the behaviour of dual-period arrays of cylinders under different illumination conditions. We found that this kind of structures have a great potential to control and predesign the intensity and direction of the transmitted response. In particular, a given diffraction order can be enhanced by appropriately choosing the geometrical parameters of the structure. In the case of an incident propagating wave, a blazed-like effect is obtained for subwavelength cylinders, and our results suggest that the same behaviour would also be present for other shapes of the individual scatterers. For evanescent illumination, the system converts the inhomogeneous wave into propagating waves that radiate to the far field, and the enhancement is produced by surface plasmon excitation. In conclusion, dual-period arrays constitute a realizable alternative not only for the microwave and millimeter wave regions but also for optical devices, and they could be used in highly sensitive detection devices, among other applications.

ACKNOWLEDGMENT

The authors gratefully acknowledge partial support from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad de Buenos Aires, Universidad Nacional del Centro de la Provincia de Buenos Aires and Instituto de Física Arroyo Seco, and Agencia Nacional de Promoción Científica y Tecnológica.

REFERENCES

1. Tan, W. -C., J. R. Sambles, and T. W. Preist, "Double-period zero-order metal gratings as effective selective absorbers," *Phys. Rev. B*, Vol. 61, No. 19, 13177–13182, 2000.
2. Hibbins, A. and J. R. Sambles, "Excitation of remarkably nondispersive surface plasmons on a nondiffracting, dual-pitch metal grating," *Appl. Phys. Lett.*, Vol. 80, No. 13, 2410–2412, 2002.
3. Lockyear, M. J., A. P. Hibbins, J. R. Sambles, and C. R. Lawrence, "Low angular-dispersion microwave absorption of a metal dual-period nondiffracting hexagonal grating," *Appl. Phys. Lett.*, Vol. 86, No. 18, 184103, 2005.
4. Lepage, J.-F. and N. McCarthy, "Analysis of the diffractive properties of dual-period apodizing gratings: Theoretical and experimental results," *Appl. Opt.*, Vol. 43, No. 17, 3504–3512, 2004.
5. Crouse, D. and P. Keshavareddy, "A method for designing electromagnetic resonance enhanced silicon-on-insulator metalsemiconductor-metal photodetectors," *J. Opt. A: Pure Appl. Opt.*, Vol. 8, No. 2, 175–181, 2006.
6. Skigin, D. C., V. V. Veremey, and R. Mittra, "Superdirective radiation from finite gratings of rectangular grooves," *IEEE Trans. Antennas Propag.*, Vol. 47, No. 2, 376–383, 1999.
7. Skigin, D. C. and R. A. Depine, "Transmission resonances in metallic compound gratings with subwavelength slits," *Phys. Rev. Lett.*, Vol. 95, No. 21, 217402, 2005.
8. Skigin, D. C. and R. A. Depine, "Narrow gaps for transmission through metallic structures gratings with subwavelength slits," *Phys. Rev. E*, Vol. 74, No. 4, 046606, 2006.
9. Skigin, D. C., H. Loui, Z. Popovic, and E. Kuester, "Bandwidth control of forbidden transmission gaps in compound structures with subwavelength slits," *Phys. Rev. E*, Vol. 76, No. 1, 016604, 2007.
10. Navarro-Cía, M., D. C. Skigin, M. Beruete, and M. Sorolla, "Experimental demonstration of phase resonances in metallic compound gratings with subwavelength slits in the millimeter wave regime," *Appl. Phys. Lett.*, Vol. 94, No. 9, 091107, 2009.
11. Beruete, M., M. Navarro-Cía, D. C. Skigin, and M. Sorolla, "Millimeter-wave phase resonances in compound reflection gratings with subwavelength grooves," *Opt. Express*, Vol. 18, No. 23, 23957–23964, 2010.
12. Lester M., D. C. Skigin, and R. A. Depine, "Blaze produced by a dual-period array of subwavelength cylinders," *J. Opt. A: Pure Appl. Opt.*, Vol. 11, No. 4, 045705, 2009.
13. Madrazo, A. and M. Nieto-Vesperinas, "Scattering of electromagnetic waves from a cylinder in front of a conducting plane," *J. Opt. Soc. Am. A*, Vol. 12, No. 6, 1298–1302, 1995.
14. Scaffardi, L. B., M. Lester, D. C. Skigin, and J. O. Tocho, "Optical extinction spectroscopy used to characterize metallic nanowires," *Nanotechnology*, Vol. 18, No. 31, 315402, 2007.

Scattering Properties of Elliptical Cylinder Coated by Lossy DNG Metamaterial

M. I. Hussein

Department of Electrical Engineering, United Arab Emirates University, Al-Ain, UAE

Abstract— This document gives formatting instructions for authors preparing papers for publication in the Proceedings of an IEEE conference. The authors must follow the instructions given in the document for the papers to be published. You can use this document as both an instruction set and as a template into which you can type your own text.

1. INTRODUCTION

Recently there has been a great interest among the electromagnetic society to explore and exploit the electromagnetic characteristics and properties of artificial materials, generally known as metamaterials. Metamaterials are artificial materials engineered to provide properties which “may not be readily available in nature”. The different classes of metamaterial include single-negative (SNG) materials, such as epsilon-negative (ENG) and mu-negative (MNG) materials, and double-negative (DNG) materials.

Among all the double negative (DNG) materials [1–4] in which the real part of both the permittivity and permeability are negative is the most interesting. In such materials the electric and magnetic field intensity vectors of a plane wave form a left-handed set of vectors with its wave number vector. This is in contrast to double positive (DPS) materials, in which these vectors form a right-handed set.

It is known that DNG materials are not found in nature and therefore must be made artificially. At microwave frequency one possible realization of DNG materials is based on periodic arrangements of conducting wires and split-ring resonators (SRRs). The relative effective permittivity and the relative effective permeability of such realization were found to possess a negative real part and an imaginary part as follow [5–7],

$$\varepsilon_r(\omega) = 1 - \frac{\omega_{pe}^2}{\omega^2 - j\omega\Gamma_e} \quad (1)$$

$$\mu_r(\omega) = 1 - \frac{F\omega^2}{\omega^2 - \omega_{0m}^2 - j\omega\Gamma_m} \quad (2)$$

The strong interest in DNG materials is due to their unfamiliar electromagnetic properties and the potential applications due to these properties. Researchers’ interest was directed in particular toward lossless DNG metamaterial [8–11]. In this work the scattering properties of a conducting elliptical cylinder coated by lossy DNG metamaterial will be investigated. The analysis is based on the well known separation of variable and exact boundary value technique, in the elliptic cylindrical coordinates system and Mathieu functions. Electric and Magnetic fields in all regions are expressed in terms of Fourier series with unknown expansion coefficients. The formulation of the problem of *TM* case is introduced. The system of equations is solved using standard numerical techniques to generate numerical results. The far scattered field pattern is calculated after using the large argument approximation of Mathieu Hankel function. The scattering by a conducting elliptical cylinder coated with lossy DNG metamaterials is compared with a conventional material and lossless DNG metamaterial.

2. PROBLEM FORMULATION

The electromagnetic wave interaction with elliptical cylinder is typically expressed in terms of the elliptic cylinder coordinates system (ξ, η, z) . Consider the case of a linearly polarized electromagnetic plane wave incident at an angle ϕ_i with respect to the positive x axis on a conducting elliptic cylinder coated by a metamaterials, where both permittivity and permeability are complex, and have negative real part, at certain frequency, as shown in Fig. 1, with $e^{j\omega t}$ time dependence. The electric field component of the *TM* polarized plane wave of amplitude E_0 is given by,

$$E_z^i = E_0 e^{jk_o\rho \cos(\phi - \phi_i)} \quad (3)$$

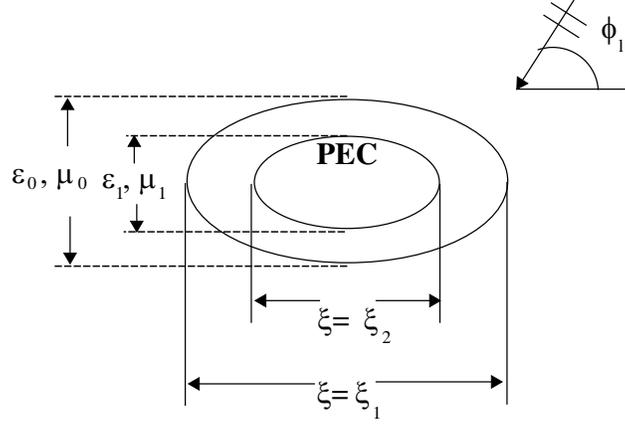


Figure 1: Geometry of the elliptic cylinder coated by metamaterial.

where k_0 is the wave number in free space. The incident electric field may be expressed as follows

$$E_z^i = \sum_{m=0}^{\infty} A_{em} R_{em}^{(1)}(c_0, \xi) S_{em}(c_0, \eta) + \sum_{m=1}^{\infty} A_{om} R_{om}^{(1)}(c_0, \xi) S_{om}(c_0, \eta) \quad (4)$$

where,

$$A_{om}^{em} = E_0 j^m \frac{\sqrt{8\pi}}{N_{om}^{em}(c_0)} S_{om}^{em}(c_0, \cos \phi_i) \quad (5)$$

$$N_{om}^{em}(c) = \int_0^{2\pi} [S_{om}^{em}(c, \eta)]^2 dv \quad (6)$$

$c_0 = k_0 F$, F is the semi-focal length of the elliptical cross section, S_{em} and S_{om} are the even and odd angular Mathieu functions of order m , respectively, $R_{em}^{(1)}$ and $R_{om}^{(1)}$ are the even and odd radial Mathieu functions of the first kind with complex argument, and N_{em} and N_{om} are the even and odd normalized functions.

The scattered electric field outside the coated elliptic cylinder for ($\xi > \xi_1$) can be expressed in terms of Mathieu functions as follows

$$E_z^s = \sum_{m=0}^{\infty} B_{em} R_{em}^{(4)}(c_0, \xi) S_{em}(c_0, \eta) + \sum_{m=1}^{\infty} B_{om} R_{om}^{(4)}(c_0, \xi) S_{om}(c_0, \eta) \quad (7)$$

where B_{em} and B_{om} are the unknown scattered field expansion coefficients, $R_{em}^{(4)}$ and $R_{om}^{(4)}$ are the even and odd Mathieu functions of the fourth kind. Similarly, the transmitted electric field inside the cylinder for ($\xi < \xi_1$) can be written as

$$E_z^t = \sum_{m=0}^{\infty} [C_{em} R_{em}^{(1)}(c_1, \xi) + D_{em} R_{em}^{(2)}(c_1, \xi)] S_{em}(c_1, \eta) + \sum_{m=1}^{\infty} [C_{om} R_{om}^{(1)}(c_1, \xi) + D_{om} R_{om}^{(2)}(c_1, \xi)] S_{om}(c_1, \eta) \quad (8)$$

where $c_1 = k_1 F$, $k_1 = \omega \sqrt{\mu_1 \epsilon_1}$, C_{em} , C_{em} , D_{em} , D_{om} the unknown transmitted field expansion coefficients, and $R_{em}^{(2)}$ and $R_{om}^{(2)}$ are the radial Mathieu functions of the second type. The magnetic field component inside and outside the cylinder can be obtained using Maxwell's equations.

The unknown expansion coefficients given in Equations (7) and (8) can be obtained by imposing the boundary condition first at $\xi = \xi_2$ and continuity of the tangential field components at $\xi = \xi_1$

along with the orthogonality property of the angular Mathieu function, one gets the unknown scattering coefficients B_{em} and B_{om} follows

$$\begin{aligned} & \sum_{m=0}^{\infty} M_{enm}(c_1, c_0) B_{em} \left[\mu_r R_{em}^{(4)'}(c_0, \xi_1) - R_{em}^{(4)}(c_0, \xi_1) \frac{\Delta'_{en}}{\Delta_{en}} \right] \\ = & \sum_{m=0}^{\infty} M_{enm}(c_1, c_0) A_{em} \left[R_{em}^{(1)}(c_0, \xi_1) \frac{\Delta'_{en}}{\Delta_{en}} - \mu_r R_{em}^{(1)'}(c_0, \xi_1) \right] \end{aligned} \quad (9)$$

$$\begin{aligned} & \sum_{m=1}^{\infty} M_{onm}(c_1, c_0) B_{om} \left[\mu_r R_{om}^{(4)'}(c_0, \xi_1) - R_{om}^{(4)}(c_0, \xi_1) \frac{\Delta'_{on}}{\Delta_{on}} \right] \\ = & \sum_{m=0}^{\infty} M_{onm}(c_1, c_0) A_{om} \left[R_{om}^{(1)}(c_0, \xi_1) \frac{R_{on}^{(1)'}(c_1, \xi_1)}{R_{on}^{(1)}(c_1, \xi_1)} - \mu_r R_{om}^{(1)'}(c_0, \xi_1) \right] \end{aligned} \quad (10)$$

where

$$\begin{aligned} \Delta_{en} &= R_{en}^{(1)}(c_1, \xi_1) - \frac{R_{en}^{(1)}(c_1, \xi_2)}{R_{en}^{(2)}(c_1, \xi_2)} R_{en}^{(2)}(c_1, \xi_1) \\ \Delta'_{en} &= R_{en}^{(1)'}(c_1, \xi_1) - \frac{R_{en}^{(1)}(c_1, \xi_2)}{R_{en}^{(2)}(c_1, \xi_2)} R_{en}^{(2)'}(c_1, \xi_1) \\ \Delta_{on} &= R_{on}^{(1)}(c_1, \xi_1) - \frac{R_{on}^{(1)}(c_1, \xi_2)}{R_{on}^{(2)}(c_1, \xi_2)} R_{on}^{(2)}(c_1, \xi_1) \\ \Delta'_{on} &= R_{on}^{(1)'}(c_1, \xi_1) - \frac{R_{on}^{(1)}(c_1, \xi_2)}{R_{on}^{(2)}(c_1, \xi_2)} R_{on}^{(2)'}(c_1, \xi_1) \end{aligned}$$

The scattered near and far fields for the TM can be calculated once the scattered fields expansion coefficients are known. The far scattered field expressions can be obtained as follows

$$E_z^s = \sqrt{\frac{j}{k_0 \rho}} e^{-jk_0 \rho} \sum_m j^m [B_{em}^{TM} S_{em}(c_0, \eta) + B_{om}^{TM} S_{om}(c_0, \eta)] \quad (11)$$

Far Field data are usually expressed in terms of the scattering cross section per unit length, i.e., the echo width is defined as

$$\sigma_{TM} = 2\pi \lim_{\rho \rightarrow \infty} \frac{|E_z^s|^2}{|E_z^i|^2} \quad (12)$$

Equation (12) can take simpler form as follows,

$$\sqrt{\frac{\sigma_{TM}}{\lambda}} = \sum_m j^m [B_{em} S_{em}(c_0, \eta) + B_{om} S_{om}(c_0, \eta)] \quad (13)$$

3. NUMERICAL RESULTS AND DISCUSSION

The analytic solution of the electromagnetic scattering by a conducting elliptic cylinder coated with metamaterials is given in terms of two, uncoupled, linear system of equations for the even and odd ordered expansion scattered field coefficients for both even and odd functions.

In order to verify the formulation and the accuracy of the developed program, the problem of electromagnetic scattering by a circular conducting cylinder coated with metamaterial was tested using the developed code. The back scattering cross section of a conducting circular cylinder coated with a one-layer dielectric versus the coating material permittivity and permeability is examined in Fig. 2 and Fig. 3, respectively, the results are in excellent agreement with [11]. Figs. [4–6] compare the echo width pattern against the scattering angle ϕ . The coated elliptical cylinder is assumed to have $k_o a = 2.5$, $k_o a_1 = 3.51$, $k_o b = 1.25$, and $k_o b_1 = 2.76$, the incident angle is $\phi_i = 0^\circ$. Fig. [4] compares the echo width of a conducting cylinder which has no coating with a cylinder

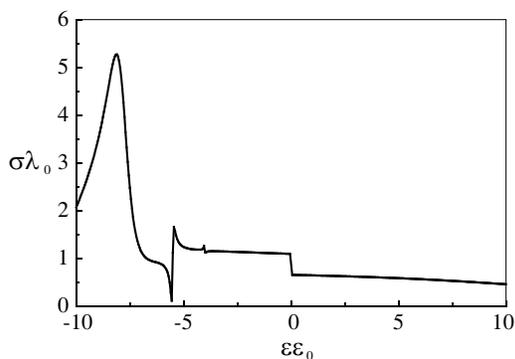


Figure 2: The back scattered cross section of a dielectric coated cylinder. ($a = 50$ mm, $a_1 = 70$ mm, $f = 1$ GHz, $\mu/\mu_0 = -1$, when $\varepsilon/\varepsilon_0 < 0$ and $\mu/\mu_0 = 1$, when $\varepsilon/\varepsilon_0 > 0$)

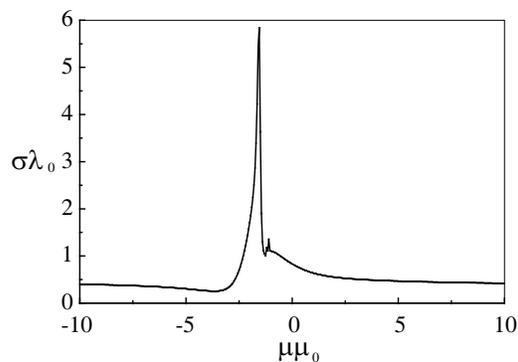


Figure 3: The back scattered cross section of a dielectric coated cylinder. ($a = 50$ mm, $a_1 = 70$ mm, $f = 1$ GHz, $\mu/\mu_0 = -1$, when $\varepsilon/\varepsilon_0 < 0$ and $\mu/\mu_0 = 1$, when $\varepsilon/\varepsilon_0 > 0$)

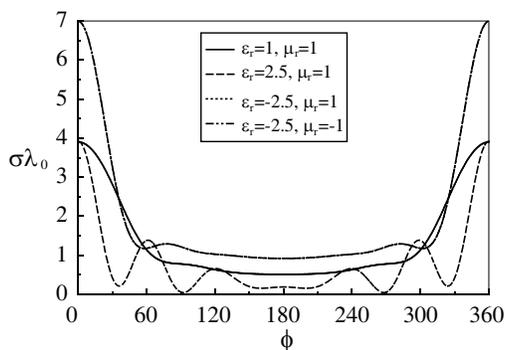


Figure 4: Echo width pattern against the scattering angle ϕ of a dielectric/DNG coated elliptic cylinder with $k_o a = 2.5$, $k_o a_1 = 3.51$, $k_o b = 1.25$, and $k_o b_1 = 2.76$, and $\phi_i = 0^\circ$.

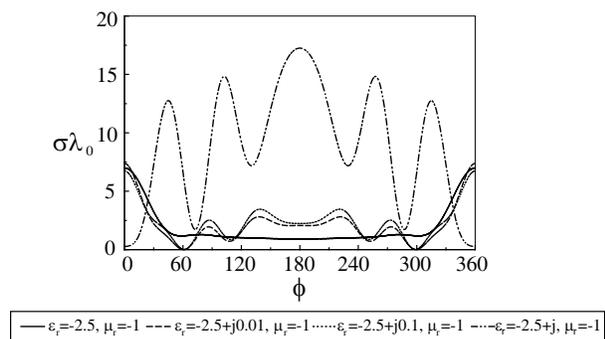


Figure 5: Echo width pattern against the scattering angle ϕ of DNG and lossy DNG metamaterial coated elliptic cylinder with $k_o a = 2.5$, $k_o a_1 = 3.51$, $k_o b = 1.25$, and $k_o b_1 = 2.76$, and $\phi_i = 0^\circ$.

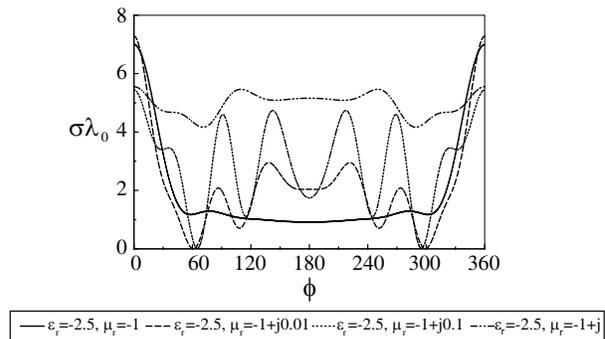


Figure 6: Echo width pattern against the scattering angle ϕ of DNG and lossy DNG metamaterial coated elliptic cylinder with $k_o a = 2.5$, $k_o a_1 = 3.51$, $k_o b = 1.25$, and $k_o b_1 = 2.76$, and $\phi_i = 0^\circ$.

that is coated with the dielectric material and a cylinder with DNG lossless metamaterial. Fig. [5] compares the echo width of lossless DNG metamaterial with lossy permittivity DNG metamaterial, while Fig. [6] compares the echo width of lossless DNG metamaterial with lossy permeability DNG metamaterial.

ACKNOWLEDGMENT

The author wish to acknowledge the support provided by the United Arab Emirates University.

REFERENCES

1. Veselago, V. G., “The electrodynamics of substances with simultaneously negative values of ε and μ ,” *Soviet Physics Usp.*, Vol. 10, No. 4, 509–514, January 1968.
2. Shelby, R. A., D. R. Smith, and S. Schultz, “Experimental verification of a negative index of refraction,” *Science*, Vol. 292, 77–79, April 6, 2001.
3. Lu, J., T. M. Grzegorzczak, Y. Zhang, J. J. Pacheco, B.-I. Wu, J. A. Kong, and M. Chen, “Cerenkov, adiation in materials with negative permittivity and permeability,” *Opt. Express*, Vol. 11, 723–734, April 2003.
4. Alù, A. and N. Engheta, “Radiation from a traveling-wave current sheet at the interface between a conventional material and a material with negative permittivity and permeability,” *Microwave Opt. Tech. Lett.*, Vol. 35, No. 6, 460–463, December 2002.
5. Pendry, J. B., A. J. Holden, W. J. Stewart, and I. Youngs, “Extremely low frequency plasmons in metallic meso structures,” *Phys. Rev. Lett.*, Vol. 76, No. 25, 4773–4776, June 1996.
6. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, “Low frequency plasmons in thin wire structures,” *J. Phys. Condens. Matter*, Vol. 10, 4785–4809, 1998.
7. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, “Magnetism from conductors and enhanced nonlinear phenomena,” *IEEE Trans. Microwave Theory Techniques*, Vol. 47, No. 11, 2075–2084, November 1999.
8. Alù, A., N. Engheta, A. Erentok, and R. W. Ziolkowski, “Single-negative, double-negative, and low-index metamaterials and their electromagnetic applications,” *IEEE Transactions on Antennas and Propagation Magazine*, Vol. 49, No. 1, 23–36, February 2007.
9. Pastorino, M., M. Raffetto, and A. Randazzo, “Interactions between electromagnetic waves and elliptically shaped metamaterials,” *IEEE Antennas and Wireless Propagation Letters*, Vol. 4, 165–168, 2005.
10. Hamid, A.-K. and M. I. Hussein , “Exact adiation from slotted circular or elliptical antenna coated by a concentric isorefractive metamaterials,” *International Journal of Applied Electromagnetics and Mechanics*, Vol. 26, 1–2, 101–111, 2007.
11. Li, C. and Z. Shen, “Electromagnetic scattering by a conducting cylinder coated with metamaterials,” *PIER*, Vol. 42, 91–105, 2003.

Enhanced Ferromagnetic and Ferroelectric Properties of La Doped Multiferroic $\text{Bi}_5\text{Fe}_{0.5}\text{Co}_{0.5}\text{Ti}_3\text{O}_{15}$ Ceramics

Xiangyu Mao¹, Wei Wang¹, Xiaobing Chen¹, and Yalin Lu²

¹College of Physics Science and Technology, Yangzhou University, Yangzhou 225002, China

²Department of Physics, Laser and Optics Research Center, U.S. Air Force Academy, Colorado 80840, USA

Abstract— A novel four-layer $\text{Bi}_{4.25}\text{La}_{0.75}\text{Fe}_{0.5}\text{Co}_{0.5}\text{Ti}_3\text{O}_{15}$ (BLFCT) ceramics was synthesized via sintering the mixture of Fe_2O_3 , Co_2O_3 , Bi_2O_3 oxides and already-prepared $\text{Bi}_{3.25}\text{La}_{0.75}\text{Ti}_3\text{O}_{12}$ (BLT) powders. The ferromagnetic transition is seen at $T_{CM} \approx 210^\circ\text{C}$ in M - T curves. At room temperature, the remanent magnetization ($2M_r$), and the coercive fields ($2H_c$) in BLFCT samples were about 49.7 memu/g and 475 Oe, respectively. The remnant polarization ($2P_r$) and the coercive field ($2E_c$) were about $15.4 \mu\text{C}/\text{cm}^2$ and $129 \text{ kV}/\text{cm}$ at $180 \text{ kV}/\text{cm}$, respectively.

1. INTRODUCTION

Multiferroics, sometimes called ferroelectromagnets (FM), magnetoelectrics (ME) and/or ferroelastics, with simultaneous electric and magnetic ordering, have attracted much attention for their potential applications and rich fundamental physics [1–3]. The coupling between the magnetic and electronic structures in multiferroics, namely, magnetoelectric effect, could provide an additional degree of freedom in device design, which is more attracting than the ferroelectricity or magnetism itself [4]. There are very few materials exhibiting both ferromagnetic (FM) and ferroelectric (FE) properties at room temperature (RT), such as (BiFeO_3) , abbreviated as BFO) and bismuth titanium iron ($\text{Bi}_5\text{FeTi}_3\text{O}_{15}$, abbreviated as BFTO) [5–10]. It is widely known that the $[\text{Bi}_2\text{O}_2]^{2+}$ layers have critical influence on the properties of bismuth layer structured ferroelectrics (BLSFs): the $[\text{Bi}_2\text{O}_2]^{2+}$ layers act as the insulating layers and are self regulated to compensate for space charge due to their net electric charge [11, 12]. The reported FE Curie point of BFTO is about 75°C with the transition from the $A2_{1am}$ group to the $I4mmm$ group. On the other hand, it was reported that BFTO favors the antiferromagnetic (AFM) ordering with its Neel point T_N of about 80 K [13]. In the near future, we have reported that the samples of $\text{Bi}_4\text{FeTi}_3\text{O}_{15}$ and $\text{Bi}_4\text{Fe}_{0.5}\text{Co}_{0.5}\text{Ti}_3\text{O}_{15}$ (BFCT) prepared by inserting BFO and BFO/ BiCoO_3 (BCO, BFO:BCO = 1:1) into $\text{Bi}_4\text{Ti}_3\text{O}_{12}$ showed stability in the structure of four-layer Aurivillius [14, 15]. At room temperature, remnant magnetization ($2M_r$) has been increased more than three times by substituting half Fe sites by Co ions, the $2M_r$ of the BFTO and BFCT samples were about 1.36×10^{-3} memu/g and 7.8 memu/g, respectively. The coercive field ($2H_C$) of BFCT is about 410 Oe, such a small value implies a soft FM property on BFCT samples. Additional, the remnant polarization ($2P_r$) of two samples were about $11.8 \mu\text{C}/\text{cm}^2$ (under $270 \text{ kV}/\text{cm}$) and $13 \mu\text{C}/\text{cm}^2$ (under $230 \text{ kV}/\text{cm}$), respectively.

Many investigations have manifested that doping is an effective approach to improve the properties of BTO. The remnant polarization of $\text{Bi}_{4-x}\text{La}_x\text{Ti}_3\text{O}_{12}$ (BLT- x) shows dependence on the La content: it increases first, maximizes at $x = 0.75$ ($\text{Bi}_{3.25}\text{La}_{0.75}\text{Ti}_3\text{O}_{12}$, BLT) and then decreases with further doping [16–18]. On the other hand, the magnetic and ferroelectric properties of multiferroic samples, such as BiMnO_3 , and BiFeO_3 et al., are known to be also dependent on their oxygen content [19–21]. From the crystallographic view of point, the Ladoped BFCT compound can be regarded as BLT- x + 0.5 BiFeO_3 (BFO) + 0.5 BiCoO_3 (BCO). When $x = 0.75$ the chemical formula turns out to be $\text{Bi}_{4.25}\text{La}_{0.75}\text{Fe}_{0.5}\text{Co}_{0.5}\text{Ti}_3\text{O}_{15}$ (BLFCT). In this letter, the BLFCT ceramics were prepared by the solid-state reaction method by inserting $\text{BiFeO}_3/\text{BiCoO}_3$ into BLT-0.75 (BLT-0.75: BFO: BCO = 1:0.5:0.5). The properties of BLFCT samples on the ferroelectric, magnetism, and dielectric measurement are reported.

The samples of BLCFT were processed in two steps [14, 15]: first making BLT-0.75 powders using multi-calcination-method and then synthesizing BLFCT samples in the traditional solid state reaction route. The detailed process is as follows: (1) The BLT-0.75 powders were prepared by a modified solid-state reaction process. The starting materials of Bi_2O_3 (analytic pure), La_2O_3 (analytic pure), and TiO_2 (spectral pure) powders in stoichiometry were finely ground (15 wt% excess of Bi_2O_3 were added to compensate for the volatilization of Bi_2O_3 in the sintering process). The oxide mixtures were ball-milled together for 24 h in ethanol and then calcined three times at $780 \sim 820^\circ\text{C}$ for $8 \sim 24$ h, respectively. (2) The as-obtained BLT powders were mixed with

Fe_2O_3 , Co_2O_3 , and Bi_2O_3 agents (analytic pure) in stoichiometry ($\text{BLT-0.75} : \text{Fe}_2\text{O}_3 : \text{Co}_2\text{O}_3 : \text{Bi}_2\text{O}_3 = 1 : 1/4 : 1/4 : 1/2$). The mixture were ball-milled together for 24 h in ethanol, presintered at 660°C for 8 h, and then ball-milled again for 24 h. Finally, the pretreated powders were pressed into pellets and sintered at 820°C for 6 h in air. The crystal structures of the samples were characterized by X-ray diffraction (XRD) using Bruker-D8 diffractometry, FESEM (field emission scanning electron microscope Hitachi, S-4800), HR-TEM (high-resolution transmission electron microscopy, T-f30). A vibrating samples magnetometer (VSM) (EV-7, ADE Co., USA) was used to measure magnetic properties. The ferroelectric properties were measured using a Precision LC ferroelectric analyzer (Radiant Technology product, USA). The dielectric properties of the BLFCT samples were obtained using a dielectric spectrometer (Novocontrol Technologies, Germany).

Figure 1 presents the XRD patterns of (a) BLT-0.75 powder and (b) BLFCT ceramics at room temperature (RT). The BLT-0.75 powder and BLFCT samples exhibit consistently the single-phase perovskite structure for three and four layer Aurivillius structure, respectively, no other impurity diffraction peaks are observed, and the detected XRD pattern is analogous to earlier reports for BLT-0.75 [17] and BFCT [15] samples. The SEM image (Figs. 1(a) and 1(b) inset) indicates that the as-obtained BLT-0.75 powder (before ball-milled) and BLFCT sample (on the cross-sectional surface) is a polycrystalline. According to the SEM images, the BLT-0.75 powder mainly consists of platelike grains as it was reported in the literature [22] and the microstructure of BLFCT is seen as densely compact plate-like grains.

Figure 2(a) shows the HRTEM image of BLFCT. The four-layer Aurivillius structure is further evidenced by the well-defined image in the three Ti-O layers and one Fe/Co-O layer are sandwiched by two Bi_2O_2 layers, the lattice parameter c is 4.02 nm which is consistent with about XRD result. The electron diffraction pattern also shows an undamaged Bi_2O_2 layer after inserting Fe^{3+} and Co^{3+} ions. Selected area electric (SAED) pattern for the $[010]$ zone axic of BLFCT [Fig. 2(b)] exhibits a regular and clear square diffraction. The BLFCT sample has a four-layered perovskite unit of nominal composition $(\text{Bi}_{2.25}\text{La}_{0.75}\text{FeTi}_3)^{2-}$ or $(\text{Bi}_{2.25}\text{La}_{0.75}\text{CoTi}_3)^{2-}$, also three TiO_6 and one FeO_6 or one CoO_6 octahedron sandwiched by two $(\text{Bi}_2\text{O}_2)^{2+}$ layers along the c axis.

At RT, the magnetic hysteresis loops of BLCFT and BCFT [15] samples are displayed in inset Fig. 3. The BLFCT sample exhibits a typical ferromagnetic M - H loop, with the $2M_r$ of ~ 49.7 memu/g and coercive field ($2H_c$) of ~ 475 Oe, respectively. Thus, the $2M_r$ of BLFCT sample is greatly increased compared to that of BFCT polycrystalline specimen ($2M_r \approx 7.8$ memu/g) [15]. To further investigate the magnetic properties of BLFCT, the zero field cooling (ZFC) and the field cooling (FC) magnetization (M) measurements are performed under a magnetic field of 500 Oe. As shown in Fig. 3, the M_{FC} and M_{ZFC} curve decreases when increasing temperature and good reversibility is maintained until the temperature about 143°C , below which an obvious difference between ZFC and FC magnetization appeared and the phenomenon might be suggest as a relaxation process [23, 24]. The magnetic Curie temperature (T_{CM}) is thus estimated to be $\sim 210^\circ\text{C}$. Since the ionic radius of La^{3+} (A -site) is slightly less than that of Bi^{3+} , the La doping with Bi is expected to lower the tolerance factor t . The presence of La^{3+} iron in the samples causes a de-

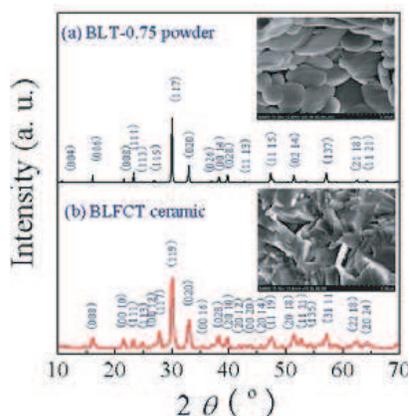


Figure 1: XRD spectrum of (a) BLT-0.75 powder and (b) BLFCT ceramics. Inset shows SEM image (a) BLT-0.75 powder and (b) BLFCT samples.

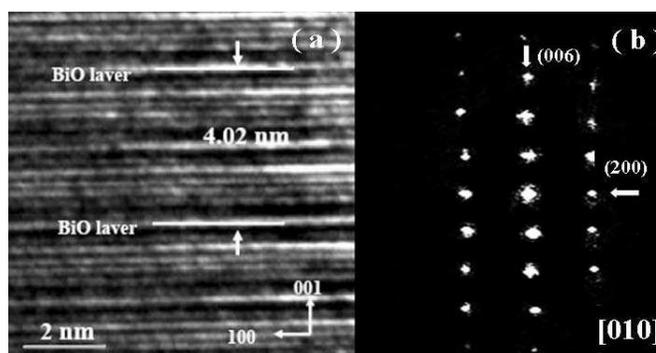


Figure 2: (a) HRTEM electron diffraction and (b) the SAED pattern of the BLFCT samples.

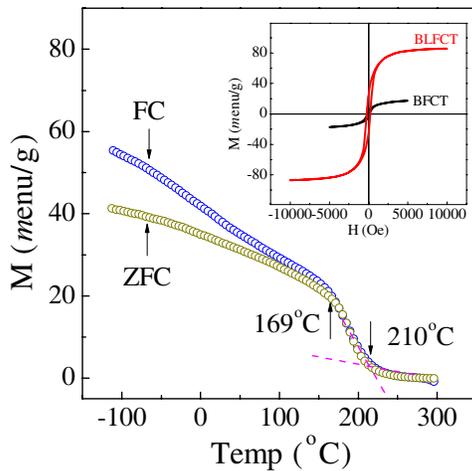


Figure 3: FC and ZFC magnetization curves of the BLFCT samples. Inset shows field dependence of magnetization at RT for BLFCT and BFCT [15] samples.

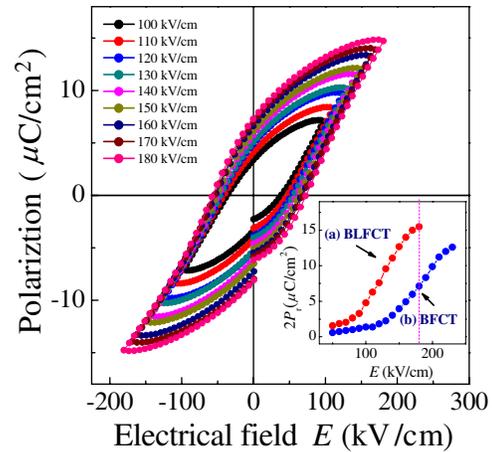


Figure 4: RT ferroelectric hysteresis loop of BLFCT measured. Inset: $2P_r$ - E curve of (a) BLFCT and (b) BFCT [15] samples.

crease of around 130°C in the magnetic ordering, showing a general weakening of the ferromagnetic interactions together with an increasing contribution of the AFM ones arising from coupling [25].

The results obtained from the current integration reveal the ferroelectric property at RT shown in Fig. 4. Fig. 4 shows P - E hysteresis loop of BLFCT measured at different electric field from 100 kV/cm to 180 kV/cm. Under the electric field of 180 kV/cm, the $2P_r$ and the coercive field ($2E_c$) were about $15.4 \mu\text{C}/\text{cm}^2$ and 129 kV/cm, respectively. The $2P_r$ of BLFCT samples increase by increasing the applied field (shown inset (a) of Fig. 4). The advantage of BLFCT samples comes mainly from the two factors: (1) Many investigations of bismuth layer-structured ferroelectrics (BLSFs), the doping of lanthanum was effective to the reduced oxygen vacancy by the stabilized oxygen octahedron. (2) It is known that the BFO and BCO crystals are rhombohedrally distorted. La doping causes phase transition from rhombohedral to orthorhombic and ferroelectric remanent polarization was therefore increased [26].

In conclusion, the $\text{Bi}_{4.25}\text{La}_{0.75}\text{Fe}_{0.5}\text{Co}_{0.5}\text{Ti}_3\text{O}_{15}$ ceramics was the conventional solid-state sintering method by BiFeO_3 and BiCoO_3 into the $\text{Bi}_{3.25}\text{La}_{0.75}\text{Ti}_3\text{O}_{12}$ were prepared. The BLFCT samples was four-layer structure, the material's exhibits fine multiferroic properties. At room temperature, the $2M_r$ and $2P_r$ were about 49.7 memu/g and $15.4 \mu\text{C}/\text{cm}^2$, respectively. The ferro-magnetic transition is seen at $T_{CM} \approx 210^\circ\text{C}$ in M - T curves.

ACKNOWLEDGMENT

This work was supported by the Chinese National Natural Science Foundation (Grant No. 51072177) and Jiangsu Province Department of Education (Grant No. 08KJB140011).

REFERENCES

- Smolenskii, G. A. and I. E. Chupis, *Sov. Phys. Usp.*, Vol. 25, 475, 1982.
- Fiebig, M., T. Lottermoser, D. Frohlich, A. V. Goltsev, and R. V. Pisarev, *Nature*, Vol. 419, 818, 2002.
- Wood, V. E. and A. E. Austin, *Int. J. Magn.*, Vol. 5, 303, 1974.
- Efremov, D. V., J. Van den Brink, and D. I. Khomskii, *Nat. Mater.*, Vol. 3, 853, 2004.
- Smolenskii, G. A., *Sov. Phys. Usp.*, Vol. 25, 475, 1982.
- Wang, J., J. B. Neaton, H. Zheng, V. Nagarajan, S. B. Ogale, B. Liu, D. Viehland, V. Vaithyanathan, D. G. Schlom, U. V. Waghmare, N. A. Spaldin, K. M. Rabe, M. Wuttig, and R. Ramesh, *Science*, Vol. 299, 1719, 2003.
- Subbarao, E. C., *J. Am. Ceram. Soc.*, Vol. 45, 166, 1962.
- Desu, S. B., P. C. Joshi, X. Zhang, and S. O. Ryu, *Appl. Phys. Lett.*, Vol. 71, 1041, 1997.
- Park, B. H., S. J. Hyun, S. D. Bu, T. W. Noh, J. Lee, H.-D. Kim, T. H. Kim, and W. Jo, *Appl. Phys. Lett.*, Vol. 74, 1907, 1999.
- Kubel, F. and H. Schmid, *Ferroelectrics*, Vol. 129, 101, 1992.

11. Kim, S. K., M. Miyayama, and H. Yanagida, *Mater. Res. Bull.*, Vol. 31, 121, 1996.
12. Porob, D. G. and P. A. Maggard, *Mater. Res. Bull.*, Vol. 41, 1513, 2006.
13. Singh, R. S., T. Bhimasankaram, G. S. Kumar, and S. V. Suryananrayana, *Solid State Commun.*, Vol. 91, 567, 1994.
14. Mao, X. Y., W. Wan, and X. B. Chen, *Solid State Commun.*, Vol. 147, 186, 2008.
15. Mao, X. Y., W. Wang, X. B. Chen, and Y. L. Lu, *Appl. Phys. Lett.*, Vol. 95, 82901, 2009.
16. Takenaka, T. and K. Sakata, *Ferroelectrics*, Vol. 38, 769, 1981.
17. Park, B. H., B. S. Kang, S. D. Bu, T. W. Noh, J. Lee, and W. Jo, *Nature*, Vol. 401, 682, 1999.
18. Shimakawa, Y., Y. Kubo, Y. Tauchi, H. Asano, T. Kamiyama, F. Izumi, and Z. Hiroi, *Appl. Phys. Lett.*, Vol. 79, 2791, 2001.
19. Sundaresan, A., R. V. K. Mangalam, A. Iyo, Y. Tanaka, and C. N. R. Rao, *J. Mater. Chem.*, Vol. 18, 2191, 2008.
20. Palkar, V. R., J. John, and R. Pinto, *Appl. Phys. Lett.*, Vol. 80, 1628, 2002.
21. Wang, Y. P., L. Zhou, M. F. Zhang, X. Y. Chen, J.-M. Liu, and Z. G. Liu, *Appl. Phys. Lett.*, Vol. 84, 1731, 2004.
22. Chen, X. B., R. Hui, J. Zhu, W. P. Lu, and X. Y. Mao, *J. Appl. Phys.*, Vol. 96, 5697, 2004.
23. Prasad, N. V. and G. S. Kumar, *J. Magn. Magn. Mater.* Vol. 213, 349, 2000.
24. Ani Kumar, P. S., P. A. Joy, and S. K. Date, *J. Phys.: Condens. Matter*, Vol. 10, L487, 1998.
25. Pefia, A., J. J. Blanco, M. Insausti, L. Lezama, J. Gutierrez, J. M. Barandiar, and T. Rojo, *J. Magn. Magn. Mater.*, Vol. 197, 543, 1999.
26. Wang, W., J. B. Sun, X. Y. Mao, and X. B. Chen, *J. Phys. D: Appl. Phys.*, Vol. 41, 155418, 2008.

Enhanced Absorption in Si Solar Cells via Adding Thin Surface Plasmonic Layers and Surface Microstructures

Yalin Lu¹, W. J. Mandeville², M. K. Shaffer¹, R. J. Knize¹, and Kitt Reinhardt³

¹Laser and Optics Research Center, Department of Physics
U.S. Air Force Academy, CO 80840, USA

²MITRE Corporation, 1155 Academy Park Loop, Colorado Springs, Colorado 80910, USA

³AFOSR/NE, 875 North Randolph Street, Ste 326, Arlington, VA 22203, USA

Abstract— In order further to understand the mechanisms and to reduce the fabrication cost, experimental demonstrations of light absorption affection by adding thin metallic plasmonic layers and laser-created surface microstructures to the thin film a-Si solar cells was performed. For adding the plasmonic layer, by adjusting those structural parameters such as thicknesses of the a-Si thin layer and the thin plasmonic layer's thickness, and the incident angle, light absorption enhancement by cavity resonances and by the planar plasmonic coupling was carefully analyzed. For creating the dense microstructures on silicon surface, frequency doubled, nanosecond (*nsec*) Nd:YAG laser pulses in the SF₆ ambient, were used. The infrared absorption is increased to near unity and extends well below the original bandgap far into the infrared. These results are interesting to making advanced solar cells on both efficiency and cost, as comparing to previous results reported using more complicated and less economical femtosecond (*fsec*) titanium sapphire and picosecond (*psec*) and nanosecond excimer lasers.

1. INTRODUCTION

Thin film solar cells using amorphous silicon (a-Si) still have the room to further reduce the cost, if their efficiency can be further improved. They have additional advantages including the overall high performance, good stability, abundance in nature and non-toxicity. Various approaches and new designs, which focus mainly on enhancing both light absorption and photon-generated carrier collection, have been extensively investigated. Among them, using the surface plasmonic effect is one of the major ways. They include using metallic nanoparticles [1], arrayed metallic/ nanostructures [2], and surface plasmonic resonance (SPR) enhanced luminescence [3], etc. Another approach is to create surface micro-features via pulsed laser or chemical etching, which is essentially able to enhance the absorption [4].

Enhancement of energy conversion efficiency in SPR-associated approaches actually utilizes many individual and cooperative effects such as much-strengthened localized electric field, enhanced scattering cross-sections, lateral coupling into the active layer, etc [5]. It is normally a combination of a few or all of them. Therefore, it will be complicated and a bit difficult to identify the precise mechanism. Angle acceptance of such nanostructured devices is an important criterion to the solar cell's performance. Moreover, using advanced surface nanostructures will be inevitably accompanied with an increase in cost.

Chemical [6] and laser [7] introduced microstructuring on silicon and the resulting enhancement of silicon's optical properties over the visible to the near-infrared band are attractive to solar cells. In particular, it has been shown that processing silicon by irradiating the surface with intense laser pulses, in an environment of sulfur hexafluoride (SF₆) gas, produces microstructured silicon, commonly referred to as 'black silicon'. This structured silicon, compared to unstructured silicon, exhibits a decrease in reflectivity as well as a dramatic increase in absorptivity. This increase in absorption extends below the 1.1 μm bandgap of the untextured silicon well into the infrared spectral region. However, the major issue inhibiting the widespread research and implementation of microstructured silicon technology is the efficiency and cost of the used fabrication, upkeep and technical expertise associated with lasers used to produce them. In this research, we are exploring the potential to integrate both surface plasmonic and surface micro-feature effects together on the a-Si thin film solar cells.

2. EXPERIMENTAL

In this research, a thin metallic layer was inserted and sandwiched by the active layer topped with a thin ITO surface electrode and the glass substrate, which forms the simplest surface plasmon

structure and provides the possibility to provide an in-depth understanding of possible absorption mechanisms. In addition, microstructured silicon is produced using an inexpensive, off the shelf, frequency doubled neodymium doped yttrium aluminum garnet (Nd:YAG) laser with the *nsec* pulse width.

The solar cell structure used in our analyses includes a glass substrate, a middle a-Si layer and a top ITO surface electrode layer. Between the a-Si and the substrate glass, a thin Ag layer was inserted. The ITO layer acts as the top electrode, and its thickness is fixed at 10 nm. The substrate is a 1/16" thick optical glass. The a-Si layer has a thickness of 200 nm. When making those samples, the Ag layer thickness will change from 25 nm to over 200 nm. Each layer in the sample was evaporated using a CT-1 Series evaporation system from VPT Inc. Optical reflection measurement was performed in the Shimadzu SolidSpec-3700 UV-VIS-IR spectrophotometer unit. The selected wavelength range is from 240 to 1200 nm, with an intention to cover the entire solar spectrum below the Si's bandgap.

Figure 2 shows the normal incidence optical reflection spectra of samples adding with silver (Ag). The a-Si layer was fixed at 200 nm, and the Ag layer changes from 25 to 150 nm at an interval of 25 nm. For comparison, the spectrum from the sample having the same ITO/a-Si/glass structure but no Ag layer was listed. This spectrum presents the typical spectral characteristics of optical reflections from multiple interfaces of air/ITO, ITO/a-Si, a-Si/glass, and glass/air, displays two

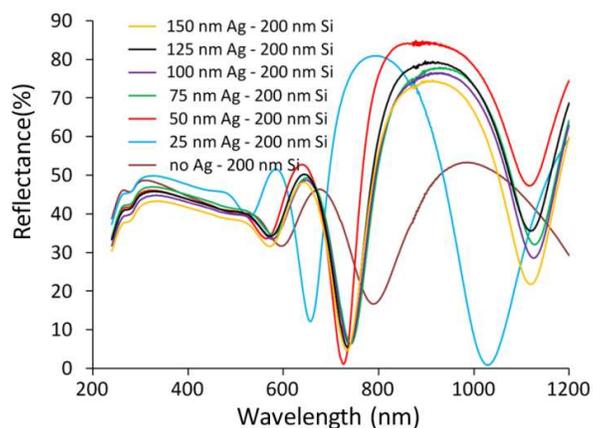


Figure 1: Reflection spectra from Ag-added samples.

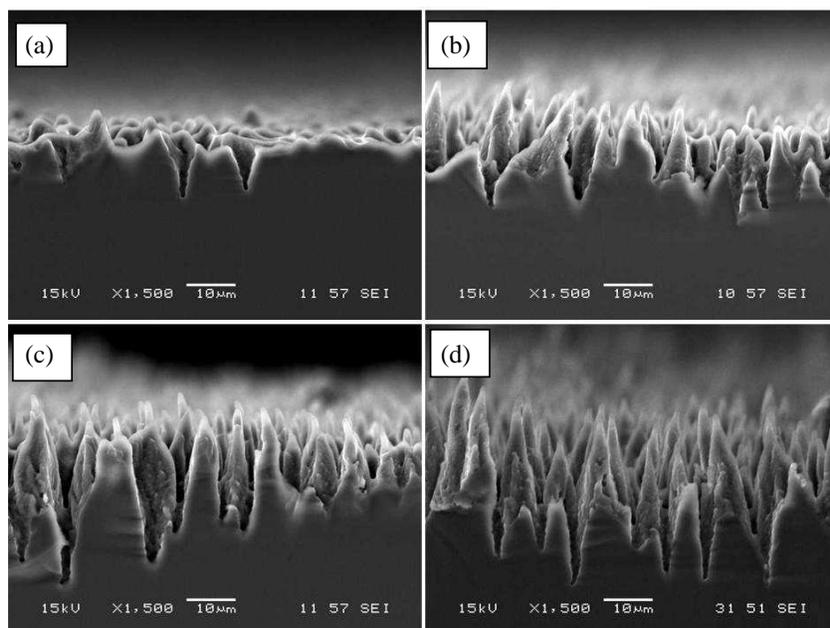


Figure 2: SEM images of a cleaved edge of a silicon wafer exposed to (a) 25, (b) 50, (c) 75, (d) 100 pulses per spot for the laser parameters described in the text.

typical absorptions at 598 nm and 797 nm. The other one is out of the 1200 nm range. Apparently adding the thin Ag layer enhances the 797 nm-peak absorption and shifts all three absorption peaks toward blue.

Adding a thin metallic layer to the solar cell structure can actually enhance the absorption of solar light. The apparent contribution to this enhancement is from the surface plasmon resonance (SPR) at the thin metallic layer's surfaces. At the resonance frequency, the metal dielectric constant is equal in magnitude and opposite in sign to that of the surrounding dielectric. Therefore, the SPR frequency is a function of metal, surrounding dielectric, constant and its dimension. In our designs, the metals are Ag and Al, the surrounding dielectrics are a-Si and glass, and the dimension is the metallic layer's thickness. It will change when the incident angle changes.

In order to introduce the microstructures on Si, we have irradiated n-type silicon wafers with 4 nanosecond laser pulses in an SF₆ environment. The wafers were (100) crystal orientation and 280 microns thick. The wafers were arsenic doped with a resistivity of 10–20 Ω-cm. Our samples were placed in a vacuum chamber which was evacuated to < 1 torr. The chamber was then backfilled with 600 torr of SF₆. The laser we used (Brilliant by Big Sky Laser) produces 200 mJ pulses at 532 nm with a pulse repetition rate of 20 Hz and a pulse length of 4 ns. The beam was focused to a 1 mm spot on the surface of the wafer producing an average fluence of 255 kJ/m² per pulse. We controlled the dose, or total energy deposited to a unit area, during processing by adjusting the translation speed of the wafer. The morphology of the samples was examined with a scanning electron microscope (SEM). Absolute reflectivity, R , and absolute transmission, T , of each sample were measured with a Carry model 6000i spectrophotometer equipped with an integrating sphere.

Figure 2 shows the surface of silicon wafers exposed to an increasing number of pulses per spot (i.e., dose) in an SF₆ atmosphere. The data suggest that an ablation and redeposition process occurs which increases the penitente height with number of pulses. The features shown in Figure 1(d) are similar in size and shape to those created using typical fsec lasers by other research groups. Processing in this manner creates high levels of sulphur doping near the surface which is believed to be responsible for the enhanced infrared absorption.

The optical absorption data are shown in Figure 3. These data show an increase in absorption through the visible spectral region as compared to untextured silicon which increases with the dose. Additionally, enhanced absorption extends well below the bandgap into the infrared region.

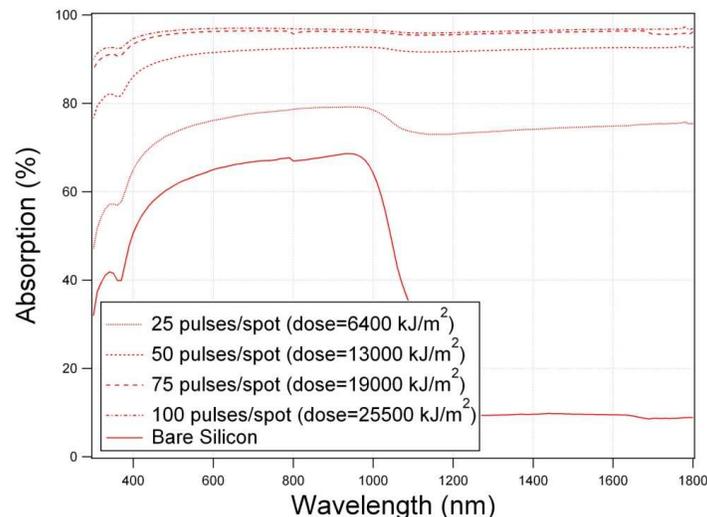


Figure 3: Absorption data for samples of microstructured silicon created by exposure to various doses. The solid line shows untextured silicon for comparison.

3. CONCLUSION

Adoption of an embedded Ag thin layer and a surface microstructure on a-Si thin film solar cells provides a capability to tailor and enhance absorption over visible to infrared wavelengths. Physical mechanisms involved are plasmonics, enhanced scattering, bandgap changing by dopant, etc. However, further studies to fully understand them are required.

ACKNOWLEDGMENT

The authors acknowledge the support from the United States Office of Scientific Research (AFOSR) and the United States Air Force Research Laboratories (AFRL).

REFERENCES

1. Rockstuhl, C., S. Fahr, and F. Lederer, *J. Appl. Phys.*, Vol. 104, 123102-7, 2008.
2. Lu, Y. L., L. Li, and K. Reinhardt, *Materials Science Forum*, Vols. 636–637, 855, 2010.
3. Lu, Y. L. and X. B. Chen, *Appl. Phys. Lett.*, Vol. 94, 193110, 2009.
4. Her, T., R. J. Finlay, C. Wu, S. Deliwala, and E. Mazur, *Appl. Phys. Lett.*, Vol. 73, 1673, 1998.
5. Wang, W., S. M. Wu, K. Reinhardt, Y. Lu, and S. C. Chen, *Nano Letters*, Vol. 10, 2012, 2010.
6. Yuan, H., V. E. Yost, M. R. Page, P. Stradins, D. L. Meier, and H. M. Branz, *Appl. Phys. Lett.*, Vol. 95, 123501, 2009.
7. Crouch, C. H., J. E. Carey, J. M. Warrender, M. J. Aziz, E. Mazur, and F. Y. Genin, *Appl. Phys. Lett.*, Vol. 84, 1850, 2004.

Layer-structured $\text{Bi}_5\text{F}_{0.5}\text{Co}_{0.5}\text{Ti}_3\text{O}_{15}$ Thin Films Grown by Pulsed Laser Deposition

Yalin Lu¹, Gail Brown², Gregory Kozlowski³, and Xiaobing Chen⁴

¹Laser and Optics Research Center, Department of Physics, U.S. Air Force Academy, CO 80840, USA

²AFRL/RXPSO, 3005 Hobson Way Bldg 651, Wright-Patterson AFB, OH 45433, USA

³Department of Physics, Wright State University, Dayton, OH 45435, USA

⁴College of Physics Science and Technology, Yangzhou University, Yangzhou 225002, China

Abstract— Layer-structured $\text{Bi}_5\text{Fe}_{0.5}\text{Co}_{0.5}\text{Ti}_3\text{O}_{15}$ (BFCTO) ceramics exhibits a large magnetic moment increase of ~ 3 times by substituting half Fe by Co. The material has a four-layer unit cell, and presents a remarkable coexistence of ferroelectricity (FE) and ferromagnetism (FM) above room temperature (RT). The measured $2P_r$ and $2M_r$ are $13 \mu\text{C}/\text{cm}^2$ and $7.8 \text{ memu}/\text{g}$, respectively, and the material's magnetic transition temperature is $\sim 345^\circ\text{C}$. This therefore raised an interest in further exploring the material's potential photonic applications when in thin films. In this research, an initial effort of thin film growth using the advanced pulsed laser deposition (PLD) approach was demonstrated. Such films' structural, morphological, and dielectric properties were studied.

1. INTRODUCTION

Aurivillius oxides are a family of materials built up by the stacking of $[\text{A}_{n-1}\text{B}_n\text{O}_{3n+1}]$ perovskite-like layers and $\alpha\text{-PbO}$ -type bismuth-oxygen $[\text{Bi}_2\text{O}_2]$ slabs [1, 2]. In this notation 'A' in general represents an alkaline earth or lanthanide metal and 'B' normally denotes a transition metal. Such compounds have received strong interest in the past as high-temperature piezoelectric devices, non-volatile memories and electro-optic devices, since they are usually ferroelectrics with high Curie temperature, anisotropic characteristics, relative easiness in making into thin films [3]. When substituted with magnetic ions such as Fe, Mn, Co, et al., a new Aurivillius material family can be created, which can be further written as $(\text{BiFeO}_3)_n(\text{A}_{m-1}\text{B}_m\text{O}_{3m+1})$. If 'A' and 'B' are Bi and Ti, respectively, it becomes $(\text{BiFeO}_3)_n(\text{Bi}_{m-1}\text{Ti}_m\text{O}_{3m+1})$. Because of a possible co-presence of FE and ME transitions, such materials may be considered as potential multiferroic materials [4]. Many studies on structural and electrical properties have been done in the past for $n = 1$ and $m = 2$ and 3 oxides, while little is known about the $n = 1$ and $m = 4$ composition, i.e., $\text{Bi}_5\text{FeTi}_3\text{O}_{15}$ (BFTO) [5], a typical four-layer subcategory in the family. In BFTO, A-sites are occupied by Bi^{3+} ions and B-sites can be shared by magnetic Fe^{3+} ions with d-electrons and ferroelectrically active Ti^{4+} ions with d^0 configurations. BFTO was reported to have a Curie temperature at 730°C and a fairly large ME effect. However, it has a low Neel temperature of T_N at $\sim 80\text{K}$ [6].

With a goal to increase the transition temperature, substituting half of the Fe ions by Co in BFTO was suggested in our team, and therefore a new layer-structured $\text{Bi}_5\text{Fe}_{0.5}\text{Co}_{0.5}\text{Ti}_3\text{O}_{15}$ (BFCTO-0.5) bulk ceramics can be formed. In coincidence to an expected coupling mechanism among Fe and Co ions, the material was found with a ~ 3 times magnetic moment increase and a coexistence of FE and FM well above RT. The measured $2P_r$ and $2M_r$ are $13 \mu\text{C}/\text{cm}^2$ and $7.8 \text{ memu}/\text{g}$, respectively [7]. The material's magnetic behavior below 275°C is relaxation-like, and most importantly, its magnetic transition temperature is raised to $\sim 345^\circ\text{C}$. Subsequently, interests in exploring such material's potential in photonic applications require a successful thin-film growth of it. This paper describes our preliminary effort on growing the BFCTO-0.5 thin films on (001)-cut $(\text{La,Sr})(\text{Al,Ta})\text{O}_3$ (LSAT) substrates using the pulsed laser deposition (PLD) method.

2. EXPERIMENTAL

BFCTO-0.5 thin films were grown by the pulsed laser deposition approach. (001)-cut LSAT substrates were selected in the experiment, considering a possible lattice matching between BFCTO-0.5 and LSAT. The used target is a $\sim 2\text{mm}$ thick and $\Phi\text{-15mm}$ pallet made directly from the BFCTO-0.5 ceramics. During the growth, we were able to change the substrate temperature, oxygen partial pressure, average time of flight, film thickness, etc., in order to optimize the growth conditions. In our discussion below, the BFCTO-0.5 thin film grown on LSAT substrate has a film thickness

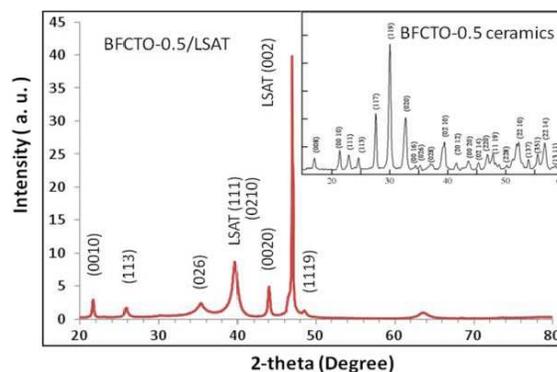


Figure 1: $\theta/2\theta$ XRD spectrum of BFCTO-0.5 thin film grown on (001)-cut LSAT. The inset is $\theta/2\theta$ XRD data from its bulk ceramics.

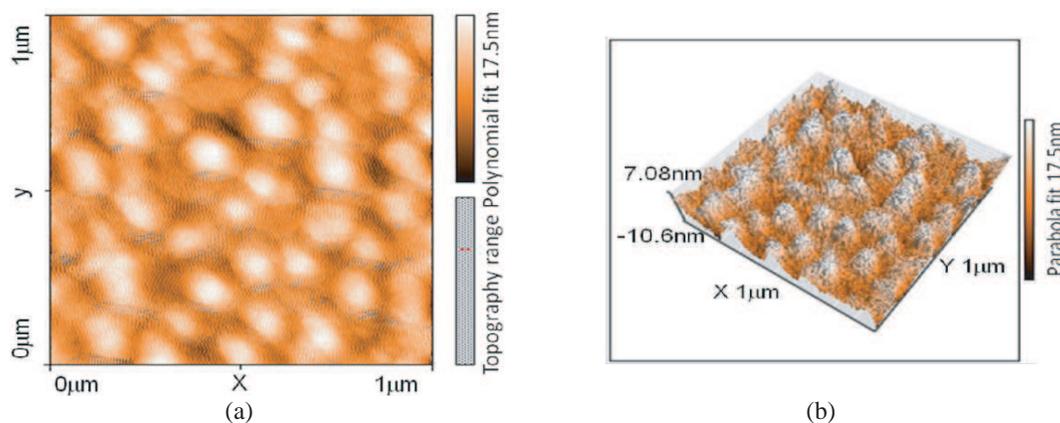


Figure 2: AFM image of the BFCTO-0.5. (a) Grain size. (b) Surface roughness.

around 100 nm. It was grown at the temperature of 600 °C and under 100 *mT* O₂ partial pressure. The average time of flight was set at 6.196 μ sec.

The film's crystalline phase analysis was performed by the standard X-ray diffraction method (XRD) using the Cu K_{α} radiation, and was identified according to both the standard Aurivillius structure and the comparison with diffraction data directly from the LSAT substrate. Surface morphology of the BFCTO-0.5 film was observed by an atomic force microscope (AFM) operating in the tapping mode. The compositional analysis was performed using both XPS and XPS mapping of each element inside the formula. As a preliminary effort on property characterization, we also measured the film's dielectric behavior at high frequency, using a scanning-tip microwave near-field (evanescent) microscope.

3. RESULT AND DISCUSSION

Figure 1 shows the XRD patterns of BFCTO-0.5 thin film grown on the (001)-cut LSAT substrate. Based on the orthorhombic crystalline structure, all observed X-ray diffraction peaks inside the XRD spectrum can be identified, indicating the existence of a single four-layer Aurivillius phase. The orthorhombic BFCTO-0.5 has a lattice parameters of $a = 0.5373$ nm, $b = 0.543$ nm, and $c = 4.1268$ nm. Comparing to the cubic LSAT ($a = 0.3868$ nm), lattice mismatch between them is about 28% (inside the a - b plane). This level of large lattice mismatch will hardly favor an epitaxial growth. However, comparing to the bulk ceramics' XRD spectrum (the inset in Fig. 1), the film shows a certain degree of the oriented growth. Apparently, this degree of orientation can be further optimized by changing growth parameters such as the substrate temperature, oxygen partial pressure, average time of flight, film thickness, etc., or by changing the thermal procedures to be used during the growth.

Surface morphology of the thin film was shown in Figs. 2(a) and 2(b). The two tapping mode AFM images over an area of $1 \mu\text{m}^2$ show clearly the grain size and the surface roughness of the film. The average grain size is about 100 nm (Fig. 2(a)) and the surface roughness is within 17 nm

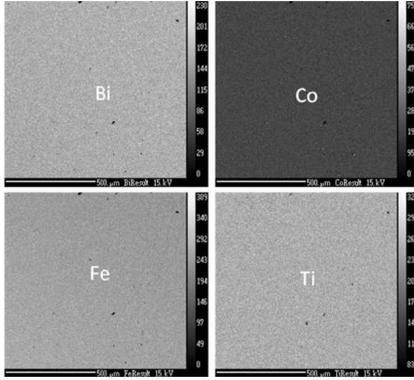


Figure 3: XPS imaging over an area of 1 mm^2 at Bi- $4f$, Fe- $2p$, Co- $2p$, and Ti- $2f$ lines.

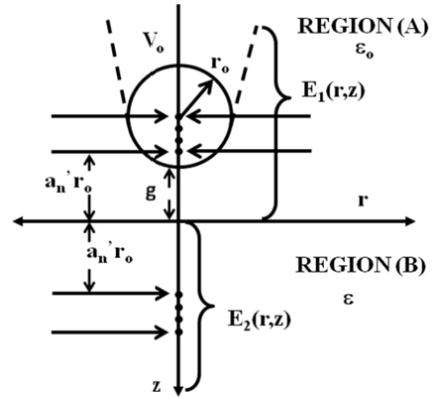


Figure 4: Schematic of the tip-dielectric environment for quantitative measurement in a scanning-tip microwave near-field microscope.

Table 1: XPS imaging analysis of the BFCTO-0.5 film's composition.

| ELEMENT | VALENCE | WEIGHT (%) | ATOM (%) |
|---------|---------|------------|----------|
| Fe | 3 | 1.95 | 2.16 |
| Ti | 4 | 9.47 | 12.25 |
| Co | 3 | 2.17 | 2.28 |
| Bi | 3 | 70.30 | 20.85 |
| O | 2 | 16.12 | 62.45 |

(Fig. 2(b)), which will be acceptable for possible optical applications. Both grain size and surface roughness can be affected by the lattice mismatch between the film and the substrate and various growth conditions. This indicates the potential to further improve the film's quality by optimizing those conditions.

X-ray photoelectron spectroscopy (XPS) imaging is a powerful analytical tool because it allows particular information on both elements and bonding to be recorded on a two-dimensional distribution map. Inside the XPS mapping system, optional deflectors are used. The deflector can shift the acceptance area by a certain lateral distance, independent of the kinetic energy set in the analyzer, and in the dispersion and the non-dispersion direction of the analyzer. Using a 2D detector, scanning the deflector voltage is only necessary in the dispersion direction of the analyzer. The chemical distribution over an extended area can be then determined by scanning the deflector voltage in two directions. Fig. 3 shows the XPS mapping results at Bi- $4f$, Fe- $2p$, Co- $2p$, and Ti- $2f$ lines over a 1 mm^2 area on the film. The mapping over the area is uniform. The measured quantitative results were listed in Table 1. According to the atomic percentage ratio, the as-grown film's composition is quite close to the designed composition inside the target.

A scanning tip microwave near-field microscope provides a unique capability to identify a material's linear and nonlinear dielectric properties at microwave frequencies and with a high spatial resolution [8, 9]. The design of the microscope involves a sharpened metal tip mounted on the center conductor of a high quality factor (high Q) coaxial $\lambda/4$ resonator. The tip extends beyond an aperture formed in the end wall of the resonator. The resonant frequency (f_r) and Q change as a function of the dielectric properties of the material placed in the vicinity of the tip. Ideally, the size of the aperture should be as small as possible to shield off the far-field propagating components of microwave energy so that only the near-field components, which stay on the tip surface, will interact with the sample and give rise to the imaging signals. More importantly, quantitative measurements can now be obtained through theoretical analysis and proper calibration (Fig. 4) [10]. On the BFCTO-0.5 thin film, the measured dielectric constant at 14.5 GHz resonance frequency is: real $\epsilon_1 = 7.7$, imaginary $\epsilon_2 = 0.02$ and the tangent loss ($\tan \delta$) = 0.00026.

4. CONCLUSION

BFCTO-0.5 thin films were prepared using the pulsed laser deposition method. X-ray diffraction analysis confirmed the single phase nature and a certain level of oriented growth. The flat surface morphology and small grain size were verified by AFM. XPS imaging shows the film's composition is very close to the target ceramic, and therefore, closes to the designed composition. Its dielectric constant at 14.5 GHz resonance frequency is: real $\varepsilon_1 = 7.7$, imaginary $\varepsilon_2 = 0.02$ and $\tan \delta = 0.00026$. The research is an on-going effort, further to optimize the film's quality via improving the growth conditions and to perform more detailed characterization over optical, ME, FE property are necessary.

ACKNOWLEDGMENT

The authors acknowledge the support from the United States Office of Scientific Research (AFOSR) and the United States Air Force Research Laboratories (AFRL).

REFERENCES

1. Aurivillius, B., *Ark. Kemi.*, Vol. 1, 463, 1949.
2. Aurivillius, B., *Ark. Kemi.*, Vol. 2, 519, 1950.
3. Desu, S. B., P. C. Joshi, X. Zhang, and S. O. Ryu, *Appl. Phys. Lett.*, Vol. 71, 1041–1043, 1997.
4. Srinivas, A., D. Kim, K. Hong, and S. V. Suryanarayana, *Appl. Phys. Lett.*, Vol. 83, 2217–2219, 2003.
5. Kubel, F. and H. Schmid, *Ferroelectrics*, Vol. 129, 101–112, 1992.
6. Singh, R. S., T. Bhimasankaram, G. S. Kumar, and S. V. Suryanarayana, *Solid State Commu.*, Vol. 91, 567–569, 1994.
7. Mao, X. Y., W. Wang, X. B. Chen, and Y. L. Lu, *Appl. Phys. Lett.*, Vol. 95, 082901, 2009.
8. Xiang, X. D., C. Gao, F. Duewer, H. T. Yang, and Y. L. Lu, US Patent 7550963, 2009.
9. Lu, Y., F. Duewer, N. B. Ming, P. G. Schultz, and X. D. Xiang, *Science*, Vol. 276, 2004, 1997.
10. Gao, C., F. Duewer, Y. L. Lu, and X. D. Xiang, *Appl. Phys. Lett.*, Vol. 72, 1146–1148, 1998.

Light Trapping within the Grooves of Diffraction Gratings

Mario M. Jakas and Francisco Llopis

Departamento de Física Fundamental y Experimental, Electrónica y Sistemas
Universidad de La Laguna, E-38205 La Laguna, Tenerife, Spain

Abstract— The possibility of using the rectangular-shaped grooves of a perfectly conducting diffraction grating as a light-trapping structure is analyzed. To this end a numerical code, which calculates the electromagnetic field produced by a well collimated, monochromatic light using the Rayleigh-Bloch method is developed. This code is then used to calculate the mean-square electric field amplitude over the groove volume, resulting from a beam of light which arrives to the grating within a cone, having all polarization states and within a small, but not negligible range of wavelengths. It is found that the average electromagnetic energy within the groove is considerably larger than that one finds far away above the grating. No such large peaks as those previously reported for well-collimated, monochromatic light are observed. However, figures in this paper show that diffraction gratings can be successfully used for producing light-confinement in photovoltaic and optoelectronic devices.

1. INTRODUCTION

In extending the sensitivity of Silicon-based photovoltaic (PV) cells to infrared (IR) light, one faces the problem arising from the low absorption Silicon has for wavelengths greater than that of visible light. A similar situation also appears in thin-films PV cells because, then, the attenuation length of light becomes comparable to the cell thickness. In both cases, designers have resorted to the concept of *light trapping* and, as a result, a variety of structures have been devised for such a purpose, cf. Ch. 9 in Ref. [1]. Among them, the diffraction gratings appear to be quite a promising alternative [2]. This approach is based on the works of Wirgin et al. [3, 4], which demonstrate that the grooves of a diffraction grating act as a resonant cavity where the electromagnetic field can be considerably *enhanced* due to resonant excitation of modal-waves. Accordingly, the mean-square amplitude of the electric field within the grooves can be several hundred times larger than that of the incident light. Furthermore, as is often the case with resonances, such enhancements take place within a fairly narrow range of wavelengths which, in turns, depends on the width and depth of the grooves, the polarization and angle of incidence of the light.

The aforementioned results, however, are obtained for well-collimated, monochromatic, either TE- or TM-polarized light, therefore, one cannot use them to assess the performance of the diffraction gratings under sunlight illumination. For that reason, this paper deals with calculating the enhancement for non-polarized, non-monochromatic and not well-collimated beams. To this end, the same procedure as those already described in Ref. [2], namely the Rayleigh-Bloch method [5–7], is used to find the electromagnetic fields along the grating. Likewise, a perfectly conducting grating and rectangular-shaped, b -wide h -deep grooves are assumed. According to the present study, beams of light arriving to the grating uniformly distributed over a 30-degree cone and having all states of polarization, give rise to field amplitudes within the grooves that, after taking an average over the groove volume, can be as large as three times that of the incident light. Furthermore, such enhancements occur for wavelengths ranging from $1b$ to $2b$ and appear to have little or nearly no dependence with the groove depth h . Results in this paper, suggest that diffraction gratings are promising as a structure to trap light within their grooves, on a relatively wide range of wavelengths, incidence angles and states of polarization.

In the following section, the model calculation is presented. Section 3 begins with validating the numerical code by comparing results obtained using the present model with those in previously published similar calculations. Secondly, the field enhancement within the grooves is calculated as a function of the wavelength of the incoming light, and for h/d ranging from 1 to 8, d being the spatial frequency of the grating. This section also include a brief discussion and analysis of the results presented therein. Finally, in Section 4 a summary and concluding remarks are offered.

2. FIELD CALCULATION

As sketched in Fig. 1, a linearly polarized light illuminates a perfectly conducting diffraction grating. The incoming light has a wave vector \mathbf{k}_i and its electric field is denoted as \mathbf{E}_i . Similarly, d , h and b represent the period, depth and width of the diffraction grating, respectively. According to the

assumptions schematically summarized in such a figure, one has $k_{ix} = k_i \sin \theta \cos \varphi$, $k_{iy} = -k_i \cos \theta$, and $k_{iz} = k_i \sin \theta \sin \varphi$. In the same fashion as $E_{ix} = E_i (\sin \phi \cos \theta \cos \varphi - \cos \phi \sin \varphi)$, $E_{iy} = E_i \sin \phi \sin \theta$, and $E_{iz} = E_i (\cos \phi \cos \varphi + \sin \phi \cos \theta \sin \varphi)$.

As is customary, for $y > 0$ the *Rayleigh expansion* is assumed,

$$\mathbf{E}^{(y>0)}(\mathbf{r}) = \mathbf{E}_i e^{i\mathbf{k}_i \cdot \mathbf{r}} + \sum_n \mathbf{R}_n e^{i\mathbf{q}_n \cdot \mathbf{r}} \quad \text{for } y > 0, \quad (1)$$

where \mathbf{q}_n is the wave vector of the n -th order reflected beam and $\mathbf{R}_n = (R_{nx}, R_{ny}, R_{nz})$ is the corresponding electric field vector. Similarly, $\mathbf{q}_n = (q_{nx}, q_{ny}, k_{iz})$, where $q_{xn} = k_{ix} + 2\pi n/d$, $n = 0, \pm 1, \pm 2, \dots$, and $q_{ny} = +\sqrt{k_i^2 - q_{nx}^2 - k_{iz}^2}$, where the plus sign in front of the square root means that the positive, either real or imaginary solution to the square root has to be used.

Within the grooves ($y < 0$), however, one has the so-called *modal waves*. These are solution to the wave equation which also satisfy the boundary conditions over the perfectly conducting walls of the grating, namely,

$$\begin{aligned} E_x^{(y<0)}(\mathbf{r}) &= \sum_{m,N} \mathcal{E}_{x,m} e^{i(k_{iz}z + Nk_{ix}x)} X'_{m,N}(x) Y_m(y) \\ E_y^{(y<0)}(\mathbf{r}) &= \sum_{m,N} \mathcal{E}_{y,m} e^{i(k_{iz}z + Nk_{ix}x)} X_{m,N}(x) Y'_m(y) \\ E_z^{(y<0)}(\mathbf{r}) &= i \sum_{m,N} \mathcal{E}_{z,m} e^{i(k_{iz}z + Nk_{ix}x)} X_{m,N}(x) Y_m(y) \end{aligned} \quad (2)$$

where

$$\begin{aligned} X_{m,N}(x) &= \begin{cases} \sin [K_{mx}(x - x_N)] & \text{for } x_N < x < x_N + b \\ 0 & \text{otherwise} \end{cases} \\ X'_{m,N}(x) &= \begin{cases} \cos [K_{mx}(x - x_N)] & \text{for } x_N < x < x_N + b \\ 0 & \text{otherwise} \end{cases} \\ Y_m(y) &= \sin [K_{my}(y + h)] \quad \text{and} \quad Y'_m(y) = \cos [K_{my}(y + h)], \end{aligned} \quad (3)$$

where $K_{mx} = m\pi/b$ for $m = 0, 1, 2, \dots$ and $K_{my} = +\sqrt{k_i^2 - k_{iz}^2 - K_{mx}^2}$, where, again, the positive solution to the square root must be taken. Similarly, x_N denotes the x -coordinate of the right wall in the N -th groove, i.e., $x_N = Nd$ for $N = 0, \pm 1, \pm 2, \dots$

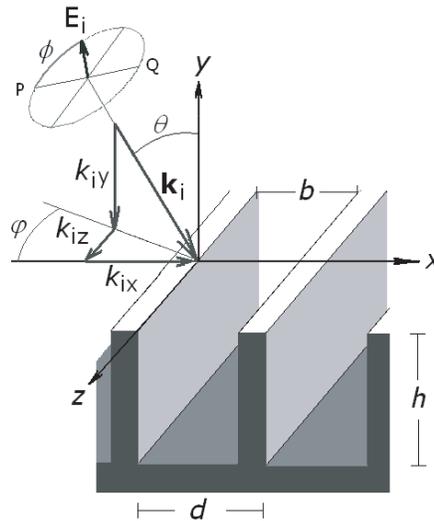


Figure 1: Sketch showing a diffraction grating illuminating with linearly polarized light. d , h and b denote the period, depth and width of the grating, while \mathbf{k}_i and \mathbf{E}_i are the wave vector and the electric field of the light, respectively. Line PQ is assumed to be parallel to $(y = 0)$ -plane. Reflected light as well as those within the grooves are not shown in this figure.

The coefficients in the Rayleigh and modal expansion can be obtained using the continuity of transversal fields along the groove aperture, the boundary conditions of electromagnetic fields on the surface of a perfect conductor (see Refs. [5, 8], and the fact that $\nabla \cdot \mathbf{E} = 0$ all above the grating surface. Accordingly, one has

$$\left. \begin{aligned} \frac{\partial}{\partial y} E_z^{(y>0)} - \frac{\partial}{\partial z} E_y^{(y>0)} &= \frac{\partial}{\partial y} E_z^{(y>0)} - \frac{\partial}{\partial z} E_y^{(y>0)} \\ \frac{\partial}{\partial x} E_y^{(y>0)} - \frac{\partial}{\partial y} E_x^{(y>0)} &= \frac{\partial}{\partial x} E_y^{(y>0)} - \frac{\partial}{\partial y} E_x^{(y>0)} \end{aligned} \right\} \begin{array}{l} E_{x,z}^{(y>0)} = E_{x,z}^{(y<0)} \\ 0 < x < b \end{array} \quad (4)$$

$$\left. \begin{aligned} \frac{\partial}{\partial x} E_z^{(y>0)} - \frac{\partial}{\partial z} E_x^{(y>0)} &= 0 \\ \frac{\partial}{\partial x} E_y^{(y>0)} - \frac{\partial}{\partial y} E_x^{(y>0)} &= 0 \end{aligned} \right\} \begin{array}{l} E_{x,z}^{(y>0)} = 0 \\ b < x < d \end{array} \quad (5)$$

$$\left. \begin{aligned} q_{nx} R_{nx} + q_{ny} R_{ny} + k_{iz} R_{nz} &= 0 \\ K_{mx} \mathcal{E}_{x,m} + K_{my} \mathcal{E}_{y,m} + k_{iz} \mathcal{E}_{z,m} &= 0 \end{aligned} \right\} \text{above metallic grating.} \quad (6)$$

It must be noticed that if first of the two equations in (5) holds, the second holds as well. This implies that the second equation in (5) is therefore redundant and can be discarded. Similarly, one can easily observed that the first equations in (4) and (5) can be written as a single one, running over the whole range, namely,

$$E_{ix} e^{ik_{ix}x} + \sum_{n=-\infty}^{+\infty} R_{nx} e^{iq_{nx}x} = \begin{cases} \sum_{m=0}^{+\infty} \mathcal{E}_{x,m} \cos(K_{mx}x) \sin(K_{my}h) & 0 < x < b \\ 0 & b \leq x < d \end{cases} \quad (7)$$

and

$$E_{iz} e^{ik_{iz}x} + \sum_{n=-\infty}^{+\infty} R_{nz} e^{iq_{nz}x} = \begin{cases} i \sum_{m=0}^{+\infty} \mathcal{E}_{z,m} \sin(K_{mx}x) \sin(K_{my}h) & 0 < x < b \\ 0 & b \leq x < d \end{cases} \quad (8)$$

By taking the Fourier transform of these equations in the x variable, one obtains

$$R_{nx} = -E_{ix} \delta_{n,0} + \frac{1}{d} \sum_{m=0}^{+\infty} \mathcal{E}_{x,m} \tilde{C}_{m,n} \sin(K_{my}h) \quad (9)$$

$$R_{nz} = -E_{iz} \delta_{n,0} + \frac{i}{d} \sum_{m=0}^{+\infty} \mathcal{E}_{z,m} \tilde{S}_{m,n} \sin(K_{my}h) \quad (10)$$

where,

$$\tilde{S}_{m,n} = \int_0^b dx e^{-iq_{nx}x} \sin(K_{mx}x) \quad \text{and} \quad \tilde{C}_{m,n} = \int_0^b dx e^{-iq_{nx}x} \cos(K_{mx}x). \quad (11)$$

After plugging Eqs. (9) and (10) into Eq. (4), and using Eq. (6) one finds,

$$\begin{aligned} & 2 \frac{(k_{iy}^2 + k_{iz}^2) E_{iz} + k_{ix} k_{iz} E_{ix}}{k_{iy}} e^{ik_{ix}x} \\ &= \sum_{m=0}^{+\infty} \left\{ \left[\frac{K_{mx} k_{iz}}{K_{my}} \sin(K_{mx}x) \cos(K_{my}h) - \sin(K_{my}h) U_m^{(xx)}(x) \right] \mathcal{E}_{x,m} \right. \\ & \quad \left. + \left[\frac{K_{my}^2 + k_{iz}^2}{K_{my}} \sin(K_{mx}x) \cos(K_{my}h) - \sin(K_{my}h) U_m^{(xz)}(x) \right] \mathcal{E}_{z,m} \right\}, \end{aligned} \quad (12)$$

and

$$\begin{aligned}
& 2 \frac{(k_{iy}^2 + k_{ix}^2) E_{ix} + k_{ix} k_{iz} E_{iz}}{k_{iy}} e^{ik_{ix}x} \\
&= \sum_{m=0}^{+\infty} \left\{ \left[\frac{K_{mx}^2 + K_{my}^2}{K_{my}} \cos(K_{mx}x) \cos(K_{my}h) - \sin(K_{my}h) U_m^{(zx)}(x) \right] \mathcal{E}_{x,m} \right. \\
&\quad \left. + \left[\frac{K_{mx} k_{iz}}{K_{my}} \cos(K_{mx}x) \cos(K_{my}h) + \sin(K_{my}h) U_m^{(zz)}(x) \right] \mathcal{E}_{z,m} \right\}, \tag{13}
\end{aligned}$$

where

$$\begin{aligned}
U_m^{(xx)}(x) &= \frac{1}{d} \sum_{n=-\infty}^{+\infty} \frac{q_{nx} k_{iz}}{q_{ny}} \tilde{C}_{m,n} e^{iq_{nx}x}, & U_m^{(xz)}(x) &= \frac{i}{d} \sum_{n=-\infty}^{+\infty} \frac{q_{ny}^2 + k_{iz}^2}{q_{ny}} \tilde{S}_{m,n} e^{iq_{nx}x} \\
U_m^{(zx)}(x) &= \frac{i}{d} \sum_{n=-\infty}^{+\infty} \frac{q_{nx}^2 + q_{ny}^2}{q_{ny}} \tilde{C}_{m,n} e^{iq_{nx}x}, & U_m^{(zz)}(x) &= \frac{1}{d} \sum_{n=-\infty}^{+\infty} \frac{q_{nx} k_{iz}}{q_{ny}} \tilde{S}_{m,n} e^{iq_{nx}x}
\end{aligned} \tag{14}$$

Equations (12) and (13) can be solved numerically using standard procedures for solving a system of linear algebraic equations. In this paper, the Gauss elimination algorithm in Ref. [9] was used and, it must be mentioned, this routine appeared to handle all calculations satisfactorily well.

3. RESULTS AND DISCUSSIONS

In the first place, the reflectance of a gold grating for TM-polarized light at 21 degree incidence is calculated. The results are plotted in Fig. 1 and compared with experiments and theoretical calculations from Ref. [10]. As one can see, a fairly good agreement between the present results and previous theoretical calculations is observed. Small discrepancies are nevertheless observed. This can be attributed to the fact that results in Ref. [10] were obtained using a different method, namely the *transfer matrix* formalism, and also due to the fact that rather than assuming a perfect conductor, as in this paper, a more realistic dielectric function for the Gold grating was employed in these calculations.

In order to further validate this method, the so-called *enhancement of the electric field* reported in Ref. [4] is studied. Accordingly, the square modulus of electric field along the center of the

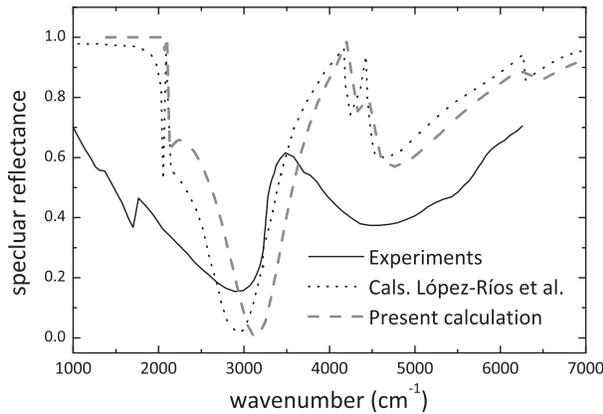


Figure 2: Reflectance coefficient for $\theta = 21$ degree incidence TM-polarized light scattered off a gold grating with grooves of a period $d = 3.5 \mu\text{m}$, width $b = 0.857 \mu\text{m}$, and depth $h = 0.171 \mu\text{m}$. Experiments (full line) and previous theoretical calculations (dotted) are from Ref. [10]. The present results are indicated as a dashed line.

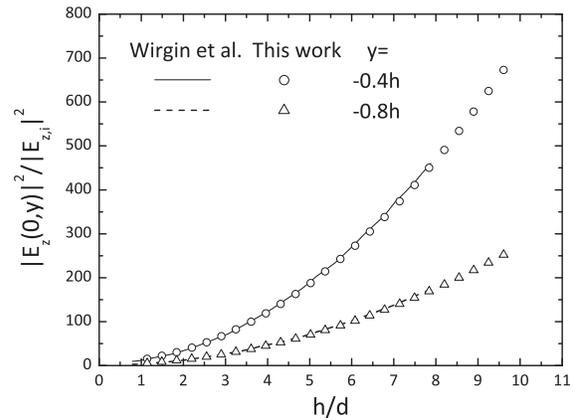


Figure 3: Electric field enhancement as a function of depth h . Field is normalized to that of the incident wave and calculated along the groove center at two depths, namely $y/h = -0.4$ and -0.8 . Light is TE-polarized and $\theta_i = 0^\circ$, $d = 0.38 \mu\text{m}$, $b = 0.35 \mu\text{m}$, and $\lambda = \lambda_1$ (see text). Lines are calculations in Ref. [4], whereas, symbols stand for the presents results.

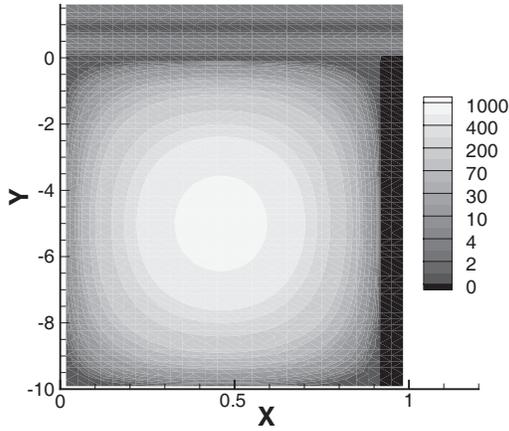


Figure 4: Electric field enhancement ($|\mathbf{E}|^2(x, y)/|\mathbf{E}_i|^2$) corresponding to the cases described in the caption of Fig. 4 for $h/d = 10$.

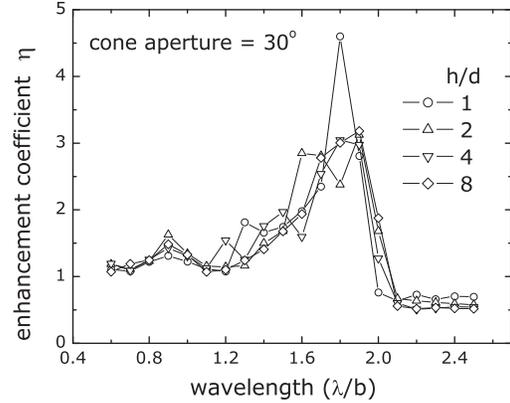


Figure 5: Mean enhancement coefficient for light arriving within a 30 degree cone and randomly polarized upon a $b/d = 0.95$ and $h/d = 1, 2, 4, 8$ diffraction grating.

groove is calculated and the results, which are previously divided by the square modulus of the incident wave, are plotted in Fig. 3. The dimensions of the grating are indicated in the caption of such a figure. The wavelengths used in these calculations, however, depend on the groove depth h and width b , since, according to theoretical prediction by the same Authors, the field enhancement should exhibit a series of peaks, at wavelengths given by expression,

$$\lambda_m = 2 [(m/h)^2 + (1/b)^2]^{-0.5} \text{ for } m = 1, 2, \dots$$

As one can readily observe, the present calculations reproduce the results in Ref. [4] remarkably well. It is worth noticing that the enhancement appears to be a monotonically increasing function of the groove depth and, within the conditions in Fig. 4, it can be as large as 600 times that of the incident beam.

The results of calculating the enhancement of the electric field within the grooves for the case of $h/d = 10$ is shown in Fig. 4. There, one can see that there is a single maximum in the center of the groove, where the electric field becomes of the order of 10^3 times that of the incoming light. It must be noticed that the gray-scale used in the contour plot is not linear and the enhancement is larger than 10^2 over nearly the entire volume of the groove. Above the grating, however, the enhancement is hardly larger than unity.

Having verified the numerical code, one can proceed with the problem of light trapping within the grooves. To commence with, it must be noticed that the electric field enhancement, as defined in Ref. [4], is not adequate for the purpose in this paper, since it is *local* and therefore, it may not necessarily denote an accumulation but a mere redistribution of the energy within the groove volume. Such a problem can be solved by taking an average of the electric field enhancement over the groove volume, one can thus introduce the *enhancement coefficient*, given by the expression

$$\eta = \frac{1}{2A|\mathbf{E}_i|^2} \int_0^b dx \int_{-h}^0 dy |\mathbf{E}(x, y)|^2, \quad (15)$$

where $|\mathbf{E}_i|$ is the electric field amplitude of the incoming light, $\mathbf{E}(x, y)$ is the electric field within the groove and A is the cross sectional area of the groove, i.e., $A = hb$. It is important to realize that since the field amplitude does not depend on z , integration along the z -direction is unnecessary. Moreover, note that η is normalized to twice the modulus squared of the electric field in the incident light. Therefore, a η unity means that, within the groove, the field would be, in an average, as intense as that one finds on top of perfectly reflecting surface. In more practical terms, it implies that the use of a diffraction grating as a light-trapping structure would be justifiable provided $\eta > 1$.

Figure 5 shows the results of calculating the mean enhancement coefficient for a beam of light, randomly polarized and uniformly distributed over a 30 degree wide cone. Diffraction gratings with $b/d = 0.95$ and h/b ranging from 1 to 8 are used and, similarly, the wavelength of the incoming

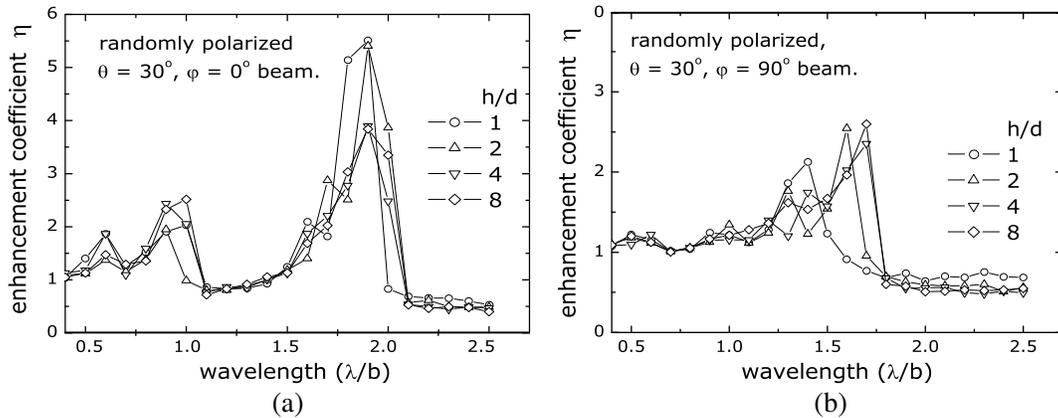


Figure 6: Mean enhancement coefficient for randomly polarized light arriving at 30 degree incidence angle on $b/d = 0.95$ and $h/d = 1, 2, 4, 8$ diffraction gratings. Incidence plane is (a) perpendicular and (b) parallel to the grating rules, respectively.

light are uniformly distributed over an interval of wavelength equal to the separation between the points, i.e., $\Delta\lambda/b = 0.1$. It must be mentioned that using these rather wider wavelength interval is a way of obtaining smoother curves. Otherwise the plot will look unnecessarily messy, due to a number of narrow peaks that will appear which, on the other hand, are of no relevance for the purpose of this paper.

As one can readily see, the enhancement coefficient becomes larger than unity for wavelengths ranging from b to $2b$. For wavelength smaller than b , however, the enhancement becomes nearly unity, indicating that, then, the diffraction grating introduces no advantages compared to a flat, perfectly reflecting surface. η reaches its maximum value around $\lambda/b \approx 1.8$, becoming smaller than one for $\lambda/b > 2$. A result that stems from the fact that for $\lambda/b > 2$ TE-polarized light can hardly penetrate within the grooves. It is worth noticing that the enhancement coefficient can be as large as 3 or 4 and it does not seem to depend on the groove depth h thus implying that total amount of energy within the grooves volume increases with an increase of h .

In order to analyze the influence of the azimuthal angle, namely φ in Fig. 1, η is recalculated for the same four groove depth used in the previous conical beams. This time, however, the beam is assumed to be non-polarized, arriving to the grating with $\theta = 30^\circ$ and both $\varphi = 0$ and 90° azimuthal angles, respectively. It must be mentioned that a non-polarized beam is realized by a set of randomly distributed, linearly polarized light. The results of these calculations, which are plotted in Fig. 6, show that light arriving perpendicular to the direction of the grating ($\varphi = 0$) produces η larger than that of light arriving along the grooves ($\varphi = 90^\circ$).

It can be also noticed that ($\varphi = 0$)-beams exhibit, apart from the main peak at $\lambda/b \approx 1.8$, a couple of smaller peaks around $\lambda/b \approx 1$ and 0.6 . For $\varphi = 90^\circ$, however, each curve has basically one maximum, and they are located at a smaller wavelength compared to those of the ($\varphi = 0$)-beams. Notably the maxima of the ($\varphi = 0$)-beams occur, approximately, in the same range of λ where ($\varphi = 90^\circ$)-curves become slightly smaller than unity. This results explain why non-polarized light yields η larger than unity over a relatively wide range of wavelengths, as observed in Fig. 5.

4. SUMMARY AND CONCLUDING REMARKS

The possibility of using a diffraction grating as a light trapping device is investigated. To this end, the electric field amplitude over the surface of a diffraction grating is calculated. This is performed by resorting to the Rayleigh-Bloch method [5], and assuming that the diffraction grating is made of a perfectly conducting material, containing rectangular-shaped grooves of width b and depth h , and having a spatial period d ($d > b$) as illustrated in Fig. 1. Results in this paper show that the mean-square electric field amplitude within the grooves of the diffraction grating can be up to fourfold larger than that one may have for a flat, perfectly reflecting surface. This seem to be the case, for beams of light arriving to the grating uniformly distributed over a 30-degree cone and all states of polarization. Such enhancements appear to take place for wavelengths ranging from b to $2b$, and they show nearly no dependence with the grooves depth. In short, it can be safely concluded that the grooves of diffraction gratings can be profitably used to build electronic devices where light is weakly absorbed and so, a light trapping structure is required.

ACKNOWLEDGMENT

This work has been supported in part by the European Union Program IBPOWER (211640) and the NUMANCIA II project funded by the Comunidad Autónoma de Madrid.

REFERENCES

1. Nelson, J., *The Physics of Solar Cells*, Imperial College Press, London, 2002.
2. Llopis, F., I. Tobías, and M. M. Jakas, *J. Opt. Soc. Am. B*, Vol. 27, 1198–2006, 2010.
3. Glass, N. E. and A. A. Maradudin, *Phys. Rev. B*, Vol. 27, 5150–5153, 1983.
4. Wirgin, A. and A. A. Maradudin, *Phys. Rev. B*, Vol. 31, 5573–5576, 1985.
5. Kong, J. A., *Electromagnetic Wave Theory*, John Wiley & Sons, New York, 1986.
6. Lochbihler, H. and R. A. Depine, *Applied Optics*, Vol. 32, 3459–3465, 1993.
7. Depine, R. A. and D. C. Skigin, *J. Opt. Soc. Am. A*, Vol. 11, 2844–2850, 1993.
8. Jackson, J. D., *Classical Electrodynamics*, John Wiley & Sons, New York, 1962.
9. Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran 77: The Art of Scientific Computation*, Cambridge University Press, Cambridge, 1992.
10. López-Ríos, T., D. Mendoza, F. J. García-Vidal, J. Sánchez-Dehesa, and B. Pannetier, “Surface shape resonances in lamellar metallic gratings,” *Phys. Rev. Letters*, Vol. 81, No. 3, 665–668, 1998.

Nondestructive Evaluation of Extended Scatterers Using Phaseless Data Subspace-based Optimization Method in the Framework of the Method of Moments

L. Pan, X. Chen, and S. P. Yeo

Department of Electrical and Computer Engineering, National University of Singapore, 117576, Singapore

Abstract— In this paper, we propose the Phaseless Data Subspace-based Optimization Method (PDSOM) in the framework of the Method of Moments (MoM) so as to reconstruct the relative-permittivity profiles of extended scatterers in Transverse Electric case, by utilizing only phaseless measurements (i.e., intensity data of the total field with no phase information). The numerical result demonstrates that the proposed method is capable of reconstructing complicated patterns with rapid rate of convergence and robust immunity to noise; and, more importantly, the comparison unravels the framework-dependence of PDSOM, i.e., PDSOM in the framework of MoM outperforms PDSOM in framework of coupled dipole method (CDM), which supplies a new way of improving the capability of reconstruction.

1. INTRODUCTION

Inverse scattering technique is one of the key approaches to reconstructing the characteristics of an object (shape and internal structure) from a set of measurement data of radiation scattered from the object. Recently, it has been of considerable interest, because of its ready applications in remote-sensing, military surveillance, biomedical diagnosis, and nondestructive testing.

To meet the challenges presented by the ill-posed problems encountered in inverse scattering, a number of iterative reconstruction solutions have been proposed over the past two decades, for instances, the distorted Born iterative method [1], the contrast source inversion method [2, 3], and the subspace-based optimization method (SOM) [4–6]. Nevertheless, one of the main drawbacks of these methods lies in the need to measure both the amplitude and the phase of the scattered fields, and that is why these methods are conventionally referred to as full-data inverse scattering approaches. It is commonly acknowledged that phase is generally more difficult to measure than amplitude. As a matter of fact, accurate knowledge of the phase information involves far more sophisticated measurement equipments, especially as the working frequency increases. Furthermore, researchers have reported that the accuracy of phase measurements cannot be guaranteed for operating frequencies approaching the millimeter-wave band and beyond, due to the fact that the phase data is more prone to noise corruption during measurement than the amplitude data. Consequently, the adoption of phaseless inverse scattering techniques is mandatory at optical frequencies, and strongly suggested at microwave frequencies.

Despite the lack of the phase information and the resultant even worse ill-posedness, the phaseless inverse scattering problem is still solvable, either by the indirect approach (phase retrieval) or by the direct approach, as reported in [7–12]. In the indirect method, the phaseless inverse scattering problem is first converted to a full-data problem, by retrieving the phase information from the measured intensities, usually in an iterative manner. After that, the full-data inverse scattering methods can be readily adopted to reconstruct from both amplitude and phase information. The disadvantage of this approach is that it requires *a priori* information concerning the scatterer (such as the compact support and limits on the spatial frequency bandwidth) and usually requires more computational cost for the phase retrieval. The direct approach generates the reconstruction directly from the measured amplitude, and yet requires naturally more incidences compared with the reconstruction from both amplitude and phase. The existing methods in the family of direct approach seem to be time-consuming and rely, to a great extent, on a good initial guess, which may further increase the computational cost.

In this paper, we shall extend the idea of original SOM in the MoM framework to handle inverse scattering problems without recourse to phase measurements. To differentiate between these two versions of the SOM, we borrow the prefix terminology coined by other researchers [7]: the original SOM (based on the full set of data) and the revised SOM (based only on amplitude data) will, for convenience, be referred to as full data SOM (FDSOM) and phaseless data SOM (PDSOM), respectively, in the paper. The original contributions of this paper may be summarized as follows:

- (a) It provides the solution of applying PDSOM in the framework of MoM.
- (b) It studies the framework-dependence of PDSOM by conducting a comparative study on PDSOM in the framework of MoM and PDSOM in the framework of CDM, which was proposed in [11]. The numerical experiments show that the former outperforms the latter, and this encouraging observation provides us with a new means of improving the reconstruction quality.
- (c) It evaluates the performance of PDSOM in the case of partial view. For previously reported experiments with PDSOM, the circular antenna array is adopted, as shown in Fig. 1(a). Nevertheless, this perfect interrogation is not always plausible in the real applications. In integrated circuit failure analysis or bomb detection, for instance, only the one-sided antenna array is possible, see Fig. 1(b), as a result, the number of receiving antennas reduces accordingly. Mathematically, the ill-posedness of the inverse problem increases as the number of linearly independent measurements reduces. The experiment results not only exhibit the practicality and good reconstruction using partial view, but also demonstrate that the presented approach is able to deal with considerably ill-posed reconstruction problems.

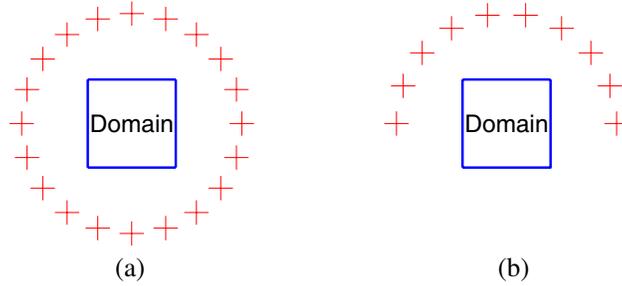


Figure 1: Antenna arrays (represented by red crosses) and the domain. (a) Circular antenna array. (b) Semicircular antenna array

2. FORMULATION OF THE INVERSE SCATTERING PROBLEM

2.1. Formulation of Forward Scattering

We consider a two-dimensional setting under transverse electric (TE) incidence for the inverse problem where the cylindrical dielectric scatterers (which are inhomogeneous in the x - y plane but invariant in the z direction) are located in the domain $\mathbf{D} \subset \mathbf{R}^2$ in the x - y plane. N_t unit magnetic current line sources (TE polarization) are arranged at \mathbf{r}_p^t ($p = 1, 2, \dots, N_t$) in a circle around the scatterers, which are illuminated successively by the sources with the incident waves expressed as $\mathbf{H}_p^{inc}(\mathbf{r}) = \hat{\mathbf{z}}H_{z,p}^{inc}(\mathbf{r}) = \hat{\mathbf{z}}\frac{i}{4}H_0^{(1)}(k_0 \cdot |\mathbf{r}_p^t - \mathbf{r}|)$, and $\mathbf{E}_p^{inc}(\mathbf{r}) = \hat{\mathbf{x}}E_{x,p}^{inc}(\mathbf{r}) + \hat{\mathbf{y}}E_{y,p}^{inc}(\mathbf{r})$ can be obtained by Ampère's Law. The scattered waves are received by N_r receivers located at \mathbf{r}_q^r ($q = 1, 2, \dots, N_r$), which are also placed in a circle surrounding the scatterers. The inverse problem consists of determining the relative permittivity profile of the scatterers, given a set of $N_t N_r$ scattering data, $\mathbf{E}_p^{scat}(\mathbf{r}_q^r)$, where $p = 1, 2, \dots, N_t$; $q = 1, 2, \dots, N_r$. In practice, the domain \mathbf{D} is discretized into a total number of N_d subunits, with the centers of the subunits located at \mathbf{r}_m ($m = 1, 2, \dots, N_d$), and the problem then reduces to determining $\epsilon(\mathbf{r}_m)$.

The eclectic field integral equation (EFIE) can be expressed as the relationship between the incident electric field $\mathbf{E}^{inc}(\mathbf{r}_m) = \hat{\mathbf{x}}E_x^{inc}(\mathbf{r}_m) + \hat{\mathbf{y}}E_y^{inc}(\mathbf{r}_m)$ and the equivalent electric polarization current densities $\mathbf{J}(\mathbf{r}_m) = \hat{\mathbf{x}}J_x(\mathbf{r}_m) + \hat{\mathbf{y}}J_y(\mathbf{r}_m)$, which is known as the MoM method, and referred to as the *state equation* in this paper,

$$E_{p,x}^{inc}(\mathbf{r}_m) = \sum_{n=1}^{N_d} \frac{A_n \eta_0}{4k_0} \left[\frac{\partial^2 H_0^{(1)}(k\rho_{mn})}{\partial x_n \partial y_n} J_{p,y}(\mathbf{r}_n) - \frac{\partial^2 H_0^{(1)}(k\rho_{mn})}{\partial y_n^2} J_{p,x}(\mathbf{r}_n) \right] + \frac{J_{p,x}(\mathbf{r}_m)}{\theta(\mathbf{r}_m)}, \quad (1a)$$

$$E_{p,y}^{inc}(\mathbf{r}_m) = \sum_{n=1}^{N_d} \frac{A_n \eta_0}{4k_0} \left[\frac{\partial^2 H_0^{(1)}(k\rho_{mn})}{\partial x_n \partial y_n} J_{p,x}(\mathbf{r}_n) - \frac{\partial^2 H_0^{(1)}(k\rho_{mn})}{\partial x_n^2} J_{p,y}(\mathbf{r}_n) \right] + \frac{J_{p,y}(\mathbf{r}_m)}{\theta(\mathbf{r}_m)}, \quad (1b)$$

($m = 1, 2, \dots, N_d$), where the subscript p means that the incident field is due to the p th source antenna; the subscripts x and y denote the x - and y -components, respectively; the superscript *inc* refers to the incident field radiated by the line sources; k_0 and η_0 are the free-space wave number and impedance respectively; A_n denotes the area of the n th subunit. and ρ_{mn} is the distance between two locations \mathbf{r}_m and \mathbf{r}_n , $\rho_{mn} = \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2}$, and $\theta(\mathbf{r}_m) = \frac{jk_0}{\eta_0} \cdot \frac{\epsilon(\mathbf{r}_m) - 1}{\epsilon(\mathbf{r}_m)}$.

Eq. (1) can be expressed in a compact form as $\bar{\mathbf{J}}_p^d = \bar{\theta}(\bar{\mathbf{E}}_p^{inc} + \bar{\mathbf{G}}_D \cdot \bar{\mathbf{J}}_p^d)$, where the exact definition of $\bar{\mathbf{G}}_D$ can be found in [5, 6]. After obtaining the current densities $\mathbf{J}(\mathbf{r}_m)$, the scattered field can be determined by using the *field equation*,

$$E_{p,x}^{scat}(\mathbf{r}_q^r) = \sum_{n=1}^{N_d} \frac{\eta_0 A_n}{4k_0} \left[\frac{\partial^2 H_0^{(1)}(k\rho_{qn})}{\partial y_n^2} J_{p,x}(\mathbf{r}_n) - \frac{\partial^2 H_0^{(1)}(k\rho_{qn})}{\partial x_n \partial y_n} J_{p,y}(\mathbf{r}_n) \right], \quad (2a)$$

$$E_{p,y}^{scat}(\mathbf{r}_q^r) = \sum_{n=1}^{N_d} \frac{\eta_0 A_n}{4k_0} \left[\frac{\partial^2 H_0^{(1)}(k\rho_{qn})}{\partial x_n^2} J_{p,y}(\mathbf{r}_n) - \frac{\partial^2 H_0^{(1)}(k\rho_{qn})}{\partial x_n \partial y_n} J_{p,x}(\mathbf{r}_n) \right], \quad (2b)$$

($q = 1, 2, \dots, N_r$). Similarly, the Eq. (2) can be expressed in compact form as $\bar{\mathbf{E}}_p^{scat} = \bar{\mathbf{G}}_S \cdot \bar{\mathbf{J}}_p^d$, where the exact definition of $\bar{\mathbf{G}}_S$ can be found in [5, 6].

We use the variable $\bar{\mathbf{E}}_p^{rad}$ to represent the field radiated by the p th transmitting antenna and received by the N_r receiving antennas, which can be expressed as a column vector of size $2N_r$ in the following manner: $\bar{\mathbf{E}}_p^{rad} = [E_{p,x}^{rad}(\mathbf{r}_1^r), E_{p,x}^{rad}(\mathbf{r}_2^r), \dots, E_{p,x}^{rad}(\mathbf{r}_{N_r}^r), E_{p,y}^{rad}(\mathbf{r}_1^r), E_{p,y}^{rad}(\mathbf{r}_2^r), \dots, E_{p,y}^{rad}(\mathbf{r}_{N_r}^r)]^T$. Without any loss of generality, the phaseless data may conveniently be taken to be the square of the intensity of the total received field, and denoted by $\bar{\mathbf{F}}_p(\mathbf{r}_q^r)$, where $p = 1, 2, \dots, N_t$; $q = 1, 2, \dots, N_r$. It is known that the square of the intensity of the total received field due to the p th transmitting antenna can be recast as a column vector of size $2N_r$ defined by the following *intensity equation*:

$$\bar{\mathbf{F}}_p = \left(\bar{\mathbf{E}}_p^{rad} + \bar{\mathbf{E}}_p^{scat} \right)^* \circ \left(\bar{\mathbf{E}}_p^{rad} + \bar{\mathbf{E}}_p^{scat} \right), \quad (3)$$

where \circ denotes the Hadamard product (or Schur product), which is an element-wise product, and the superscript $*$ denotes the complex conjugate. Eq. (1), Eq. (2), and Eq. (3), constitute the foundation for the PD inverse scattering problem.

2.2. Reconstruction with Phaseless Data

In the first step of PDSOM, we perform the SVD of $\bar{\mathbf{G}}_S$, and obtain $\bar{\mathbf{G}}_S = \bar{\mathbf{U}} \cdot \bar{\mathbf{\Sigma}} \cdot \bar{\mathbf{V}}^*$. And then the equivalent electric polarization current density ($\bar{\mathbf{J}}_p^d$) on scatterers due to the p th incidence can be divided into two orthogonally complementary portions (viz., the deterministic portion $\bar{\mathbf{J}}_p^s$ and the ambiguous portion $\bar{\mathbf{J}}_p^n$) in the following manner: $\bar{\mathbf{J}}_p^d = \bar{\mathbf{J}}_p^s + \bar{\mathbf{J}}_p^n = \bar{\mathbf{V}}^s \cdot \bar{\alpha}_p^s + \bar{\mathbf{V}}^n \cdot \bar{\alpha}_p^n$. In this equation, L is the number of leading singular values in the $\bar{\mathbf{\Sigma}}$ matrix; $\bar{\mathbf{V}}^s$ is the basis of the signal subspace, and consists of the first L right singular columns of the $\bar{\mathbf{V}}$ matrix); $\bar{\mathbf{V}}^n$ is the basis of the noise subspace, and consists of the remaining $2N_d - L$ right singular columns of the $\bar{\mathbf{V}}$ matrix); and $\bar{\alpha}_p^s$ and $\bar{\alpha}_p^n$ are two vectors of coefficients for the deterministic and ambiguous portions, respectively.

In the second step of PDSOM, we calculate the deterministic portion $\bar{\mathbf{J}}_p^s$ (or, more specifically, its coefficient vector $\bar{\alpha}_p^s$) by solving a nonlinear optimization problem. On basis of this idea, an objective function is proposed as follows,

$$\bar{\alpha}_p^s = \arg \min_{\bar{\alpha}_p^s} \left\| \bar{\mathbf{F}}_{mea,p} - \left(\bar{\mathbf{E}}_p^{rad} + \bar{\mathbf{G}}_S \cdot \bar{\mathbf{V}}^s \cdot \bar{\alpha}_p^s \right)^* \circ \left(\bar{\mathbf{E}}_p^{rad} + \bar{\mathbf{G}}_S \cdot \bar{\mathbf{V}}^s \cdot \bar{\alpha}_p^s \right) \right\|, \quad (4)$$

where $\bar{\mathbf{F}}_{mea,p}$ denotes the square of the measured intensity vector due to the p th incidence. The objective function in this optimization problem is a quartic polynomial and the gradient of the function can be computed in a straightforward manner.

The final step of the PD-SOM is to determine the ϵ_r profile of the scatterers by minimizing the sum of relative mismatches in both *state equation* and in *intensity equation*.

$$\bar{\epsilon}_r = \arg \min_{\bar{\epsilon}_r} \sum_{p=1}^{N_t} \frac{\|\bar{\bar{A}} \cdot \bar{\alpha}_p^n - \bar{B}_p\|^2}{\|\bar{J}_p^s\|^2} + \frac{\|\bar{F}_{\text{mea},p} - \bar{C}_p^* \circ \bar{C}_p\|^2}{\|\bar{F}_{\text{mea},p}\|^2}, \quad (5)$$

where $\bar{\bar{A}} = \bar{V}^n - \bar{\xi} \cdot \bar{G}_D \cdot \bar{V}^n$, and $\bar{B}_p = \bar{\xi} \cdot (\bar{E}_p^{\text{inc}} + \bar{G}_D \cdot \bar{J}_p^s) - \bar{J}_p^s$, $\bar{C}_p = \bar{E}_p^{\text{rad}} + \bar{G}_S \cdot \bar{I}_p^s + \bar{G}_S \cdot \bar{V}^n \cdot \bar{\alpha}_p^n$. It is recommended that $\bar{\alpha}_p^n$ (which is also a function of $\bar{\epsilon}_r$) be obtained by the least squares solution [4], i.e., $\bar{\alpha}_p^n = (\bar{\bar{A}}^* \cdot \bar{\bar{A}})^{-1} \cdot (\bar{\bar{A}}^* \cdot \bar{B}_p)$.

3. NUMERICAL SIMULATION

The numerical result is based on the following set of values assigned to their configuration parameters. $N_t = 10$ sources (serving as transmitting antennas) are evenly distributed on a semicircle of radius 5λ , with their locations given by $(5\lambda \cdot \cos \frac{\pi p}{N_t}, 5\lambda \cdot \sin \frac{\pi p}{N_t})$ $p = 1, 2, \dots, N_t$. A total number of $N_r = 20$ receivers are also evenly distributed on a circle of radius 5.5λ , with their locations given by $(5.5\lambda \cdot \cos \frac{\pi q}{N_r}, 5.5\lambda \cdot \sin \frac{\pi q}{N_r})$ $q = 1, 2, \dots, N_r$. The scattered field is then recorded in the format of the multi-static response (MSR) matrix $\bar{\bar{K}}$ with size of $2N_r \times N_t$. White Gaussian noise $\bar{\bar{\kappa}}$ is thereafter added to the MSR matrix, and the resultant noise-corrupted matrix $\bar{\bar{K}} + \bar{\bar{\kappa}}$ is treated as the measured MSR matrix which we utilized to reconstruct the scatterers. The noise level is quantified in percentage form as $\frac{\|\bar{\bar{\kappa}}\|_F}{\|\bar{\bar{K}}\|_F} \cdot 100\%$, where $\|\cdot\|_F$ denotes the Frobenius norm

of a matrix, and in our simulations, we have chosen the noise level to be 10%. We have found from preliminary trials that 10 iterations are sufficient for the second and third steps, respectively, of the PDSOM. To initiate the optimization process, it is convenient to choose free space with $\epsilon_r(\mathbf{r}_m) = 1$, $m = 1, 2, \dots, N_d$ as the default for the initial guess.

The test structure is depicted in Fig. 2(a) where the annular scatterer (with relative permittivity of 1.5 and inner and outer radii of 0.15λ and 0.3λ respectively) is centered at the origin of the x - y coordinate system and the selected domain (which is a square of 0.8λ for each side) has been discretized into a grid of 25×25 subunits. When applying the SOM in the presence of noise, choosing an appropriate value for the number of leading singular values (L) from the spectrum of the the matrix $\bar{\bar{G}}_S$ [4] not only improves the rate of convergence and the quality of the reconstructed patterns, but also strengthens the immunity to noise; we infer from the singular value decomposition of $\bar{\bar{G}}_S$ that L should be chosen as 9. As can be seen from Fig. 2(c), the reconstructed pattern satisfactorily exhibits the scatterer's key features, and is regarded to be rather successful, in view of the following facts: firstly, the noise level is as high as 10%; secondly, data obtained from semicircular measurement is more limited than data available from circular detector array; thirdly and most importantly, such reconstruction is impossible for PDSOM in the framework of CDM, whose result is illustrated in Fig. 2(b). This interesting contrast indicates the importance fact that the performance of PDSOM is not framework-independent. Consequently, choosing a

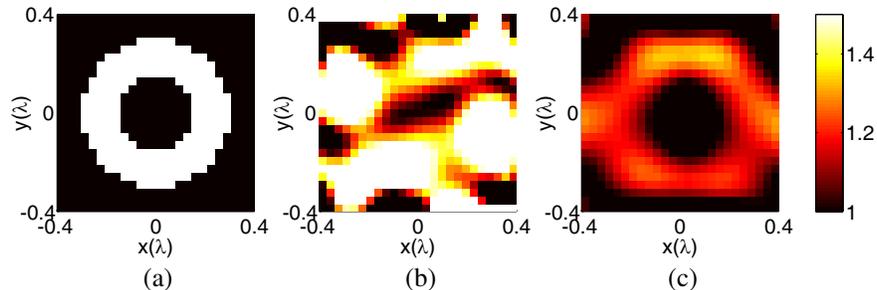


Figure 2: An annulus with inner radius 0.15λ and outer radius 0.3λ . (a). Exact relative permittivity. (b) Reconstructed relative permittivity using PDSOM in the framework of CDM. (c) Reconstructed relative permittivity using PDSOM in the framework of MoM.

suitable framework should be taken into account as an important factor influencing the quality of reconstruction, in the study of inverse scattering problem.

4. CONCLUSION

In this paper, we have proposed the PDSOM in the framework of MoM, for the reconstruction of the relative permittivity pattern from the measured phaseless electromagnetic signal received by the semicircular antenna array. The numerical experiment shows fast convergence rate (10 iterations) and robustness against noise (10%). The reconstructed pattern shows good resolution (about 0.15 wavelength), and clearly exhibits the key subwavelength features of the exact pattern. More significantly, the comparative study shows the framework-dependence of PDSOM, and our numerical simulation proves that MoM framework surpasses CDM framework in the sense of reconstructing capability, especially when only the partial view, the semicircular detector array, for instance, is available. This observation supplies a new way to improve the capability of reconstruction. Although our numerical results are for scatterers under TE incidence (since the corresponding case of TM incidence has been widely considered by others), it should be pointed out that the PD-SOM can be readily extended to 2-D TM and 3-D cases as well.

ACKNOWLEDGMENT

The authors wish to acknowledge the financial support from the US Department of the Air Force AOARD R&D Grant 094130.

REFERENCES

1. Chew, W. C. and Y. M. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted born iterative method," *IEEE Transactions on Medical Imaging*, Vol. 9, 218–225, 1990.
2. Van Den Berg, P. M. and R. E. Kleinman, "A contrast source inversion method," *Inverse Problems*, Vol. 13, 1607–1620, 1997.
3. Abubakar, A., P. M. van den Berg, and J. J. Mallorqui, "Imaging of biomedical data using a multiplicative regularized contrast source inversion method," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, 1761–1771, 2002.
4. Chen, X., "Application of signal-subspace and optimization methods in reconstructing extended scatterers," *J. Opt. Soc. Am. A*, Vol. 26, 1022–1026, 2009.
5. Pan, L., K. Agarwal, Y. Zhong, S. P. Yeo, and X. Chen, "Subspace-based optimization method for reconstructing extended scatterers: Transverse electric case," *J. Opt. Soc. Am. A*, Vol. 26, 1932–1937, 2009.
6. Pan, L., X. Chen, and S. P. Yeo, "Application of the subspace-based optimization method in the framework of the method of moments: Transverse electric case," *Asia-Pacific Microwave Conference*, 980–983, Singapore, Dec. 2009.
7. Li, L., H. Zheng, and F. Li, "Two-dimensional contrast source inversion method with phaseless data: TM case," *IEEE Trans. Geosci. Remote Sens.*, Vol. 47, No. 6, 1719–1736, 2009.
8. Crocco, L., M. D'Urso, and T. Isernia, "Inverse scattering from phaseless measurement of the total field on a closed curve," *J. Opt. Soc. Am. A*, Vol. 21, 622–631, 2004.
9. Bucci, O., L. Crocco, M. D'Urso, and T. Isernia, "Inverse scattering from phaseless measurements of the total field on open lines," *J. Opt. Soc. Am. A*, Vol. 23, 2566–2577, 2006.
10. D'Urso, M., K. Belkebir, L. Crocco, A. Litman, and T. Isernia, "Phaseless imaging with experimental data: facts and challenges," *J. Opt. Soc. Am. A*, Vol. 25, 271–281, 2008.
11. D'Uroso, M., L. Crocco, and T. Isernia, "Faithful non-linear imaging from only-amplitude measurements of incident and total fields," *Optical Express*, Vol. 15, No. 7, 3804–3815, 2007.
12. D'Uroso, M., L. Crocco, and T. Isernia, "Quantitative imaging from diffracted fields intensities: An inversion method and its experimental validation," *Journal of Modern Optics*, Vol. 57, No. 9, 777–782, 2010.

Theoretical Model for Optical Sensing of a Random Monolayer of Particles

A. García-Valenzuela¹, C. Sánchez-Pérez¹, E. Gutiérrez-Reyes², and R. G. Barrera²

¹Centro de Ciencias Aplicadas y Desarrollo Tecnológico
Universidad Nacional Autónoma de México, Apartdo Postal 70-186, Distrito Federal 04510, México

²Instituto de Física, Universidad Nacional Autónoma de México
Apartado Postal 20-364, Distrito Federal 01000, México

Abstract— We summarize the derivation of a simple model for the coherent-reflectance of light from a random monolayer of particles using a multiple scattering formalism and an effective field approximation. The resulting model is applicable for low surface-coverage fractions of the monolayer but for all angles of incidence and is suitable for sensing applications.

1. INTRODUCTION

Random monolayers of particles play a fundamental role in diverse physical or chemical processes and devices. For instance, chemical sensors consisting of chemically active particles interacting with their environment may be devised and interrogated using a light beam. An example of such devices is nicely described in Ref. [1].

When light is incident on a random monolayer of particles we may divide the scattered field in two components: a coherent and a diffuse component. The coherent component corresponds to the average scattered field, whereas the diffuse component relates to the fluctuations of the scattered field around the average. The power carried by the average field is commonly referred to as the coherent component. We may then define the coherent reflectance or transmittance of a random monolayer as the ratio between the power carried by the reflected or transmitted component and the incident power. The amount of power carried by the coherent and the diffuse components depend on the particle size, refractive index and angle of incidence. A few models for the coherent reflectance of light from a random monolayer of particles have been devised in the past. The most robust one, valid for larger surface-coverage fractions of the monolayer, is the so called Quasi-Crystalline Approximation [2]. However it is also the most complicated one. There is no formula in closed-form for such model making it rather unsuitable for sensing applications. The simplest model was developed in Refs. [3] to simulate the reflectance of a random monolayer of large particles adsorbed on a glass substrate at an angle of incidence near Brewster's angle of the supporting interface. However this model is valid only for low angles of incidence and ignores the multiple reflections between the substrate and the monolayer and between the particles. Finally, some of us presented a simple heuristic model for the coherent reflectance of a random monolayer of particles which remains well behaved at large angles of incidence in Ref. [4]. The results of this latter model were found to be consistent with some experimental data for reflectivity measurements in an internal reflection configuration from a sparse monolayer of large particles adsorbed at the base of a glass prism. Here we present a formal derivation of this model based on a multiple scattering formalism.

2. MULTIPLE SCATTERING FORMALISM

Let us first consider a free standing random monolayer of particles. That is, let us suppose the particles are embedded in homogeneous, non-magnetic medium of refractive index n_m . We will refer to this medium surrounding the particles as the matrix. For simplicity let us consider all particles are spherical and have the same radius a and refractive index n_p . We will indicate the position of any particle in space with the coordinates of its center, \vec{r}_n . In a monolayer of particles, the centers of all particles lie on a plane. Let us place the origin of our coordinate system on the monolayer plane with the x and y axes on the plane as shown in Fig. 1.

Let us suppose a plane wave, $\vec{E}^i(\vec{r}) = E_0 \exp(i\vec{k}^i \cdot \vec{r} - i\omega t)\hat{e}_i$, is incident to the monolayer, where $\vec{k}^i = k_y^i \hat{a}_y + k_z^i \hat{a}_z$ with $k_y^i = k_m \cos \theta_i$, $k_z^i = k_m \sin \theta_i$, $k_m = 2\pi n_m / \lambda$ and λ is the vacuum wavelength. The electric field scattered by the n th particle may be written as,

$$\vec{E}_n^S(\vec{r}) = \int \vec{G}_m(\vec{r}, \vec{r}') \cdot \vec{J}_{ind}^{(n)}(\vec{r}') d^3 r', \quad (1)$$

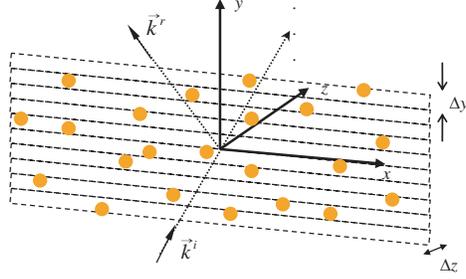


Figure 1: Coordinate system. Strips along the x axis are indicated.

where $\vec{G}_m(\vec{r}, \vec{r}')$ is the Green function of the vector wave-equation satisfied by the electric field in the matrix medium, and $\vec{J}_{ind}^{(n)}(\vec{r}')$ is the excess current induced inside the particles, that is, the difference between the current induced inside the particles and the current that would be induced in the matrix without the particle. $\vec{J}_{ind}^{(n)}(\vec{r}')$ includes all excess currents, whether they are polarization, conduction or magnetizations currents. The total scattered field is the sum of the scattered fields by all particles. The current induced in any particle may be related to the exciting field, that is, the net field incident on the particle through the transition operator of a particle centered at the origin, $\vec{T}(\vec{r}', \vec{r}'')$ (see Ref. [5] and references therein). This operator may be expressed in terms of its momentum representation as,

$$\vec{T}(\vec{r}', \vec{r}'') = \frac{1}{(2\pi)^6} \int d^3 p' \int d^3 p'' \exp(i\vec{p}' \cdot \vec{r}') \vec{T}(\vec{p}', \vec{p}'') \exp(-i\vec{p}'' \cdot \vec{r}''). \quad (2)$$

We can translate this operator to that for a particle centered at \vec{r}_n by replacing \vec{r}' and \vec{r}'' with $\vec{r}' - \vec{r}_n$ and $\vec{r}'' - \vec{r}_n$. Then, the scattered field from N particles may be written formally as,

$$\vec{E}^S(\vec{r}) = \sum_{n=1}^N \int \vec{G}_m(\vec{r}, \vec{r}') \cdot \vec{T}(\vec{r}' - \vec{r}_n, \vec{r}'' - \vec{r}_n) \cdot \vec{E}_n^{exc}(\vec{r}'') d^3 r' d^3 r'', \quad (3)$$

where \vec{E}_n^{exc} is the exciting field to the n th particle. For a given configuration of N particles in the monolayer we may set up a system of N equations for the exciting fields for each particle. Solving this system of equations, introducing the calculated exciting fields in Eq. (3), performing the indicated integrals and taking the configurational average, yields the coherent scattered fields from the system of particles. Clearly such calculation is limited to a finite number of particles, is time consuming and does not provide physical insight into the process. It would not be suitable, for instance, for inverting the parameters of a monolayer from experimental data in a sensing experiment. To this end, it is desirable to have simple approximate models in closed form.

3. DEVELOPING SIMPLE MODELS

Here we will derive the model reported in Ref. [4] using a multiple-scattering formalism, stating clearly the approximations used, and pointing out a way for improvements while keeping its simplicity.

The idea is to use an effective field approximation previously used for a half-space of a random system of particles [5] but adapted to the monolayer problem. The basic idea is to approximate the exciting field to all the particles as an effective plane wave and setting up a consistency equation from which we may derive the parameters of the exciting field.

Let us assume the center of the particles are confined to a slab of space between $\Delta z/2$ and $-\Delta z/2$, and suppose the exciting field is a plane wave given by $\vec{E}^{exc}(\vec{r}) = E_{exc} \exp(i\vec{k}^{exc} \cdot \vec{r} - i\omega t) \hat{e}_{exc}$. We want to calculate the scattered field from the monolayer on the monolayer. To this end, we divide the monolayer in thin strips of width Δy and calculate the average field scattered by each strip. Then, we may calculate the total average scattered field at a point within the slab containing the monolayer by adding the contributions from all strips.

Considering strips of the slab containing the monolayer along the y -axis (see Fig. 1) it is convenient to work with the following plane-wave expansion of the Green function,

$$\vec{G}_m(\vec{r}, \vec{r}') = \frac{i}{8\pi^2} \int \int dk_x^s dk_z^s \frac{1}{k_y^s} (\vec{1} - \hat{k}_\pm^s \hat{k}_\pm^s) \cdot \exp[i\vec{k}_\pm^s \cdot (\vec{r} - \vec{r}')], \quad (4)$$

where $\vec{k}_\pm^s = k_x^s \hat{a}_x \pm k_y^s \hat{a}_y + k_z^s \hat{a}_z$ and $k_y^s = \sqrt{k_m^2 - (k_x^s)^2 - (k_z^s)^2}$ (this is the positive root). Where there is a choice of sign, we should use the upper one when $y > y'$, and the lower one when $y < y'$. The field scattered by all particles may be written as,

$$\vec{E}^S(\vec{r}) = \frac{i}{8\pi^2} E_{exc} \sum_{n=1}^N \iint dk_x^s dk_z^s \frac{1}{k_y^s} (\vec{1} - \hat{k}_\pm^s \hat{k}_\pm^s) \cdot \exp[i(\vec{k}^{exc} - \vec{k}_\pm^s) \cdot \vec{r}_n] \vec{T}(\vec{k}_\pm^s, \vec{k}^{exc}) \exp[i\vec{k}_\pm^s \cdot \vec{r}] \hat{e}_{exc}, \quad (5)$$

where we used Eqs. (2) and (4) in Eq. (3) and performed the integrals on d^3r' , d^3r'' , d^3p' and d^3p'' . Now, let us split the thin layer in thin strips of width Δy along the y -axis of the slab containing the monolayer.

Let us take the configurational average of the scattered field from each strip. We will assume the probability density of finding any sphere at \vec{r}_n is constant within the strip and is given by $1/V$, where V is the volume of the strip. Ignoring the exclusion volume between spheres, the average scattered field is N times the average of that for any given sphere, where N is the number of particles in the strip of volume V . The average of the scattered field from any sphere located in a strip around y_n corresponds to integrating on x_n from $-L/2$ to $L/2$ (with $L \rightarrow \infty$), on z_n from $-\Delta z/2$ to $\Delta z/2$ and on y_n from $y_o - \Delta y/2$ to $y_o + \Delta y/2$, and multiplying by $1/V$ where $V = \Delta y \Delta z L$. Then, the average scattered field from all particles on a strip around y_o yields,

$$\left\langle \vec{E}^S(\vec{r}) \right\rangle_{strip} = \frac{i}{4\pi} E_{exc} \rho \Delta y \Delta z \int dk_z^s \frac{1}{k_y^s} (\vec{1} - \hat{k}_\pm^s \hat{k}_\pm^s) \cdot \vec{T}(\vec{k}_\pm^s, \vec{k}^{exc}) \exp[i\vec{k}_\pm^s \cdot \vec{r}] \exp[i(\vec{k}_y^{exc} \mp \vec{k}_y^s) y_o] \hat{e}_{exc} \quad (6)$$

where $\rho = N/V$ is the volume number-density of particles and now $k_y^s = \sqrt{k^2 - (k_x^{exc})^2 - (k_z^s)^2}$. Now, for a monolayer of particles we may take the limit $\Delta z \rightarrow 0$ but $\rho \rightarrow \infty$ so that the product $\Delta z \rho \rightarrow \rho_s$ is the surface number-density of particles. The surface density of particles in the monolayer may be related to the surface coverage fraction, Θ , as $\rho_s = \Theta/(\pi a^2)$.

Now, we add up the scattered field from all strips at a point \vec{r}_m at, or near the monolayer plane and take the limit $\Delta y \rightarrow 0$. This corresponds to integrating on dy_o from $-\infty$ to y_m and from y_m to ∞ . In the first integral we choose \hat{k}_+^s, \vec{k}_+^s and $-k_y^s$ in the integrand, whereas for the second integral we choose \hat{k}_-^s, \vec{k}_-^s and k_y^s . We then assume that k_y^s has a small imaginary part and thus, the integrated terms evaluated at $\pm\infty$ are zero. The new integrand becomes, $\exp(i k_x^{exc} x_m + i k_y^{exc} y_m + i k_z^s z_m)$ times, $(\vec{1} - \hat{k}_+^s \hat{k}_+^s) \cdot \vec{T}(\vec{k}_+^s, \vec{k}^{exc}) / i(k_y^{exc} - k_y^s) - (\vec{1} - \hat{k}_-^s \hat{k}_-^s) \cdot \vec{T}(\vec{k}_-^s, \vec{k}^{exc}) / i(k_y^{exc} + k_y^s)$. We may perform the remaining integral on dk_z^s by the method of residues. Note that the first term in the integrand has two poles since $1/i(k_y^{exc} - k_y^s) = -i(k_y^{exc} + k_y^s)/(k_z^s + k_z^{exc})(k_z^s - k_z^{exc})$, whereas there are no singularities in the second term inside the integrand (a physical \vec{T} should not have poles). Now, let us suppose k_z^{exc} has a small imaginary part. Then, there is one pole in the upper half-plane and another in the lower half-plane (of the k_z complex plane). If $z_m > 0$ we may close the contour of integration in the upper half-plane (counter clockwise) and if $z_m < 0$ we may close the contour of integration in the lower half-plane (clockwise). We get,

$$\left\langle \vec{E}^S(\vec{r}_m) \right\rangle = \frac{i E_{exc} \rho_s}{2 k_z^{exc}} \begin{cases} (\vec{1} - \hat{k}^{exc} \hat{k}^{exc}) \cdot \vec{T}(\vec{k}^{exc}, \vec{k}^{exc}) \exp[i\vec{k}^{exc} \cdot \vec{r}_m] \hat{e}_{exc} & \text{for } z_m > 0 \\ (\vec{1} - \hat{k}_r^{exc} \hat{k}_r^{exc}) \cdot \vec{T}(\vec{k}_r^{exc}, \vec{k}^{exc}) \exp[i\vec{k}_r^{exc} \cdot \vec{r}_m] \hat{e}_{exc} & \text{for } z_m < 0 \end{cases}, \quad (7)$$

where $\hat{k}^{exc} = k_x^{exc} \hat{a}_x + k_y^{exc} \hat{a}_y + k_z^{exc} \hat{a}_z$ and $\hat{k}_r^{exc} = k_x^{exc} \hat{a}_x + k_y^{exc} \hat{a}_y - k_z^{exc} \hat{a}_z$. This result is nearly the same one obtained in [5] and used in [4], except that here, the scattered field is specified for all values of z_m whereas before the result was obtained formally only for $|z_m| > a$. Note that the scattered field in this approximation is discontinuous at $z_m = 0$. Clearly, this result comes from taking into account, in the dilute limit, the multiple scattering for the average wave among strips of the monolayer.

Now, in the so called effective field approximation, we take the exciting field felt by the particles as the average field. Let us use a slightly different approximation. Let us take the exciting field as given by the average of the fields outside the particles, that is, we do not include the internal field within any given particle as part of the exciting field to the other particles. Then we may write, $\vec{E}^{exc}(\vec{r}) = \vec{E}^i(\vec{r}) + \left\langle \vec{E}^S(\vec{r}) \right\rangle$. Using Eq. (7) in the latter consistency equation requires $\vec{k}^{exc} = \vec{k}^i$ and

$\hat{e}_{exc} = \hat{e}_i$. Now if we consider points with $z_m > 0$, we get, $E_{exc} = E_0[1 - \frac{i\Theta}{4\pi a^2 k_z}(\vec{1} - \hat{k}^i \hat{k}^i) \cdot \vec{T}(\vec{k}^i, \vec{k}^i) \cdot \hat{e}_i]^{-1}$ which completely specifies the exciting field in this approximation. Using again Eq. (7) we get the average reflected and transmitted fields by the monolayer, and from these, we may obtain a coherent transmission and reflection coefficients, t_{coh} and r_{coh} :

$$t_{coh} = \frac{1}{1 + \alpha S(0)} \quad \text{and} \quad r_{coh} = \frac{\alpha S_j(\pi - 2\theta_i)}{1 + \alpha S(0)}, \quad (8)$$

with $\alpha = 2\Theta/(k_m a)^2 \cos \theta_i$, $j = 1$ or 2 for a TE or a TM polarized incident wave and we used the results, $(\vec{1} - \hat{k}^i \hat{k}^i) \cdot \vec{T}(\vec{k}^i, \vec{k}^i) \cdot \hat{e}_i = i \frac{4\pi}{k_m} S(0) \hat{e}_i$, and $(\vec{1} - \hat{k}^r \hat{k}^r) \cdot \vec{T}(\vec{k}^r, \vec{k}^i) \cdot \hat{e}_i = i \frac{4\pi}{k_m} S_j(\pi - 2\theta_i) \hat{e}_i$ previously demonstrated in [5], where $S(0)$ is the forward scattering amplitude of an isolated sphere (embedded in the matrix medium), and S_1 and S_2 are the diagonal elements of the amplitude scattering matrix as defined by Bohren and Huffman in [6].

If we now suppose the monolayer is supported by a flat interface, we must take into account the multiple reflections of the average wave between the monolayer and the flat interface. We may construct the coherent reflection coefficient using the well known formula for the compound reflection coefficient of two parallel interfaces, where the reflection coefficient of one of the interfaces is that of the monolayer given in Eq. (9), as shown in Ref. [4]. This is valid whether the monolayer is on one side of the interface or the other.

4. FINAL REMARKS

We obtained Eq. (7) from a multiple scattering formalism in the monolayer and approximating the exciting field to the particles as a simple plane wave. We were able to set up a consistency equation for points in the transmission half-space, from which we obtained t_{coh} and r_{coh} in Eq. (8). With the approach presented here, there is room for some improvements while keeping the same degree of simplicity. For points on the other side of the monolayer (the reflection half-space), no meaningful consistency equation was obtained. A way to improve the present model is to include in the exciting field a plane wave in the direction of coherent reflection. In this way we may be able to set up a consistency equation for all points in space which may yield a more realistic exciting field for the monolayer problem. Also, we may include the internal fields within the particles in the effective field approximation, and thus, in the consistency equation. We believe that such improved models may still be attractive for optical sensing applications and will be published elsewhere in the future.

ACKNOWLEDGMENT

We acknowledge financial support from Consejo Nacional de Ciencia y Tecnología (Mexico) through grant 49482F and to Dirección General de Asuntos del Personal Académico from Universidad Nacional Autónoma de México through grant IN-120309.

REFERENCES

1. Okamoto, T., I. Yamaguchi, and T. Kobayashi, "Local plasmon sensor gold colloid monolayers deposited upon glass substrates," *Opt. Lett.*, Vol. 25, No. 6, 372–374, 2000.
2. Loiko, V. A., V. P. Dick, and A. P. Ivanov, "Features in coherent transmittance of a monolayer of particles," *J. Opt. Soc. Am. A*, Vol. 17, 2040–2045, 2000.
3. Van Der Zeeuw, E. A., L. M. Sagis, G. J. M. Koper, H. M. T. Mann, and D. Bedeaux, "The suitability of angle scanning reflectometry for colloidal particle sizing," *J. Chem. Phys.*, Vol. 105, 1646–1653, 1996.
4. Peña-Gomar, M. C., J. J. F. Castillo, A. García-Valenzuela, R. G. Barrera and E. Pérez, "Coherent optical reflectance from a monolayer of large particles adsorbed on a glass surface," *Applied Optics*, Vol. 45, No. 4, 626–632, 2006.
5. Barrera, R. G. and A. García-Valenzuela, "Coherent reflectance in a system of random Mie scatterers and its relation to the effective medium approach," *J. Opt. Soc. Am. A*, Vol. 20, No. 2, 296–311, 2003.
6. Bohren, C. F. and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, Wiley-Interscience, New York, 1983.

CAROLS SMOS CAL/VAL Campaigns

M. Zribi^{1,2}, M. Pardé², J. Boutin³, P. Fanise², D. Hauser², M. Dechambre², Y. Kerr¹,
M. Leduc-Leballeur², G. Reverdin³, N. Skou⁴, S. S. Søbjaerg⁴, C. Albergel⁵, C. Calvet⁵,
J. P. Wigneron⁶, E. Lopez-Baeza⁷, K. Saleh⁷, A. Ruis⁸, and J. Tenerelli⁹

¹CESBIO (CNRS/IRD/CNES/UPS), Toulouse, France

²LATMOS, Guyancourt, France

³LOCEAN, Paris, France

⁴DTU-Space, Lyngby, Denmark

⁵CNRM/GAME, Toulouse, France

⁶INRA-Bordeaux, Cestas, France

⁷Facultat de Fisica, Valencia, Spain

⁸IEEC/ICE-CSIC, Bellaterra, Spain

⁹National Institute of Marine Research, Brest, France

Abstract— The CAROLS “Cooperative Airborne Radiometer for Ocean and Land Studies” L band radiometer was designed and built as a copy of the EMIRAD II radiometer constructed by the Technical University of Denmark team. It is a fully polarimetric and direct sampling correlation radiometer. CAROLS is installed on board a dedicated French ATR42 research aircraft, in conjunction with other airborne instruments. Following initial laboratory qualifications and other scientific campaigns, SMOS CAL/VAL campaigns involving 22 flights were carried out over South West France, the Valencia site and the Bay of Biscay (Atlantic Ocean) in spring 2010, in coordination with *in situ* field campaigns.

1. INTRODUCTION

Passive microwave remote sensing of Soil moisture and ocean salinity has been at the center of attention of many research programs, for several decades. SMOS satellite mission is based on an aperture synthesis L-band radiometer [1], designed and developed by the European Space Agency (ESA). In the context of the validation activity for the SMOS mission, the authors proposed to design, build and operate the CAROLS L-Band radiometer from an aircraft. Because the sensitivity of L-band brightness temperature to salinity is very small (-0.45 K/psu at a physical temperature equal to 105 K), it was necessary to build a very accurate, sensitive, and stable system. The radiometer was installed in the French research ATR-42 aircraft. In coordination with *in situ* field campaigns, it is used together with a suite of other airborne instruments: a scatterometer — or a C-Band radar — used to measure ocean roughness and winds, a GPS receiver to estimate ocean surface roughness, an infra-red radiometer to measure sea surface temperature, and a visible wavelength wide angle camera to provide a qualitative description of the surface.

This paper presents in Section 2 a detailed description of the CAROLS instrument with the other the instruments used in our campaigns. In Section 3, we describe the SMOS CAL/VAL campaigns. The CAROLS data quality and first results are discussed in Section 4. Our conclusions are given in Section 5.

2. CAROLS RADIOMETER AND AUXILIARY INSTRUMENTS

2.1. Carols Radiometer

CAROLS is a total power radiometer and has a simple structure and high theoretical sensitivity. The receiver was developed as a copy of the EMIRADII radiometer [2], in collaboration between the TUD (Technical University of Denmark) and the LATMOS laboratory (Laboratoire Atmosphères, Milieux, Observations Spatiales). It is a fully polarimetric correlation radiometer using direct sampling [2]. The antenna system comprises two large, identical Potter horns and two bulky waveguide orthomode transducers (OMT). These provide dual-incidence measurements.

All of the analog components are placed inside a double insulated and thermally regulated box and their physical temperatures are recorded in real time by means of temperature sensors. The accuracy of the receiver is highly dependant on the stability of the noise source and other components used in the noise injection circuitry. Table 1 illustrates performance of CAROLS instrument.

Table 1: Main technical specifications of the CAROLS radiometer.

| CAROLS: correlation radiometer with direct sampling | |
|---|----------------------------------|
| Receiver type | Total power receiver |
| Operating frequency | [1.4–1.427 GHz] |
| Bandwidth | 24 MHz @ –3 dB |
| Polarization | Vertical and Horizontal |
| Sampling rate | 139.4 MHz |
| Integration time | 1 ms and 1.8 s |
| Internal calibration | Load and noise diode |
| Sensitivity | 0.1 K for 1 sec integration time |
| Stability | 0.1 K |
| System noise temperature | 150 K |
| Physical temperature of AFE | 45°C |
| Physical temperature of DFE | 90°C |

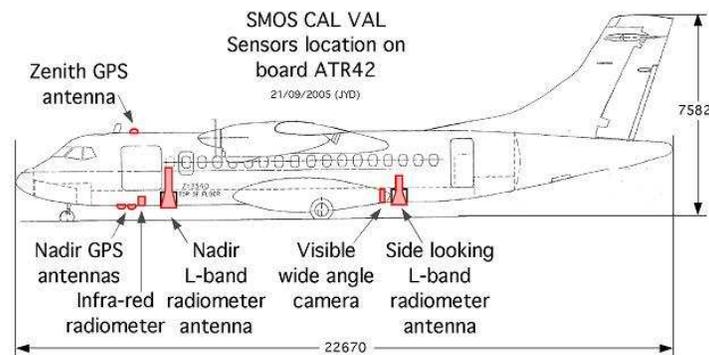


Figure 1: CAROLS instrument with two antennas (one slant and one nadir).

2.2. GPS GOLD-RTR System

The GPS system was supplied by the IEEC (Institut d'Estudis Espacials de Catalunya or Institute for Space Studies of Catalonia) [3]. It is composed of 3 antennas and a GPS Open Loop Differential Real Time Receiver (GOLD-RT). The aim is to retrieve information concerning sea surface roughness, in the form of the effective Probability Density Function (PDF) of the sea surface slopes.

2.3. Infrared Radiometer (IRR)

The infrared radiometer is part of the standard equipment of the research ATR42. This instrument points to nadir, and has a 3° field of view. It measures the radiance and the brightness temperature of the Earth's surface in three channels, namely 8.7, 10.8 and 12 μm , and is used to provide surface temperature estimations.

3. DESCRIPTION OF THE CAROLS CAMPAIGNS

3.1. Description of the Interfaces with the Research ATR-42 Aircraft

The research ATR42 has been available for research experiments since early 2006 (<http://www.safire.fr>). It is an original ATR42-320, specifically modified for scientific use. Two configurations were proposed for the CAROLS campaign flights. In the first configuration, providing dual-incidence data, the CAROLS radiometer was operated with two antennas, one pointing to nadir for vertical measurements, and the other pointing to the right side of the aircraft (slant antenna), at an approximately 30° incidence angle (Figure 1).

3.2. Flight Descriptions

All measurements were made from altitudes ranging between 600 m and 3000 m. Different flight patterns were carried out during the oceanic flights, in addition to straight horizontal flight portions: wing wags with a $\pm 25^\circ$ roll; nose wags with a $\pm 5^\circ$ pitch; and circular flights with 15° positive and

negative roll angles. These movements are useful for the qualification of the instrument's behavior and validation.

3.3. Measurement Sites

We selected three different test sites for the CAROLS data acquisitions, over land and ocean surfaces. Simultaneously to the CAROLS measurements, different *in situ* measurements were acquired in order to qualify the CAROLS data, the direct emissivity models, and finally the inversion algorithms.

3.3.1. Bay of Biscay

The ocean measurements were made over the Bay of Biscay. In order to support the Airborne CAROLS experiment, ship campaigns were organized by the LOCEAN team for each CAROLS campaign, along the same track as that used by the aircraft. Measurements of SSS, sea surface temperature (SST), wind speed, foam, wave spectrum, heat and momentum flux were collected on the ship. Different drifters measuring SSS and SST were deployed.

3.3.2. SMOSMANIA Site

This site is situated in the South West of France. SMOSMANIA (Soil Moisture Observing System — Meteorological Automatic Network Integrated Application) determines soil moisture values, based on a portion of the Météo-France automatic ground station network (the RADOME network) [4].

3.3.3. Valencia Site

The Valencia Site is located in South East Spain, about 80 km inland, to the West of Valencia. Within the Valencia validation site, an area of 30 km × 50 km was selected for the experiment.

4. CAROLS AIRBORNE DATA ANALYSIS

4.1. Carols Calibration Data

During any given flight, periodic calibrations effectively decrease the influence of gain fluctuations, and increase the radiometer's sensitivity. For any calibration step, the gain and the receiver noise temperature can be estimated with help of two calibrated points (load and load+noise diode). In order to retrieve absolute values of natural emissions, absolute calibration of the measured brightness temperatures is needed. Absolute calibrations were carried out over the ocean. For the two antennas, the measurements were compared to calculated temperatures using the TRAP software [5]. For these comparisons, measurements recorded under conditions of only limited variations in roll and pitch, and also SSS, were used.

4.2. CAROLS Measurements over Ocean

Figure 2 shows the data collected over the ocean, influenced by wing wag movements with strongly varying incidence angles, for both antennas. The results indicate a strong degree of coherence between the data collected by the two antennas. The upper TB values are for the X polarization, whereas the lower values correspond to the Y polarization. The nadir antenna data is plotted in red, and that with the slant antenna is plotted in black. For the same incidence angle, the variations

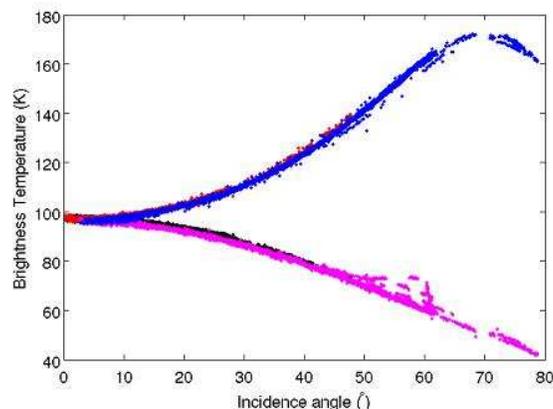


Figure 2: Tb (K) measured for the X and Y polarizations, for both nadir (red) and slant (black) antennas, as a function of incidence angle.

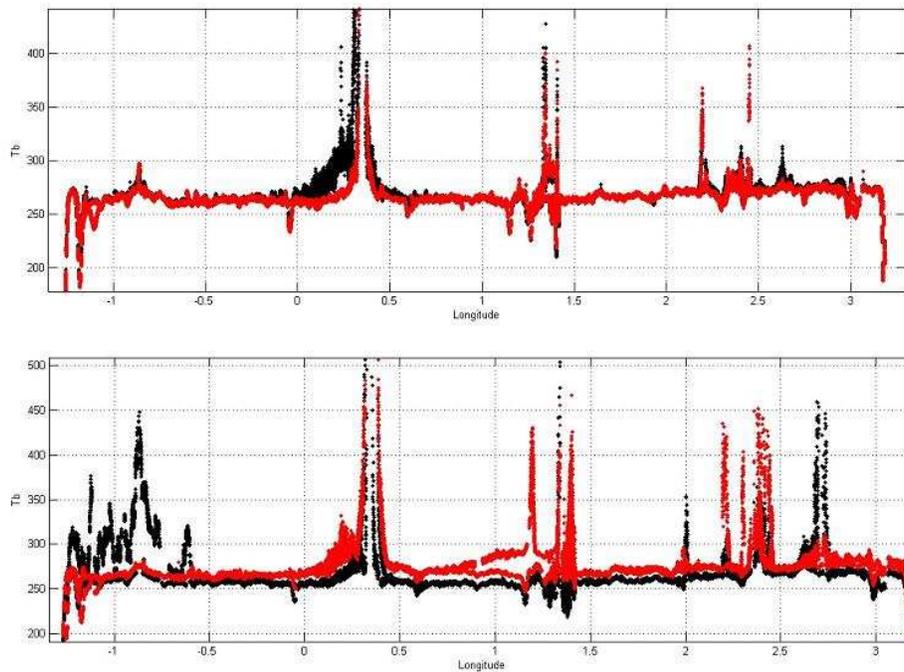


Figure 3: Time variations of T_b during the SMOSmania transect, in 2010, for the Nadir antenna (upper figure, red and black lines respectively for the X and Y polarizations), and Slant antenna (lower figure, red and black lines respectively for the X and Y polarizations). A mixed Kurtosis and thresholds to the T_b standard deviation algorithm [6] was applied for the data used here.

between the nadir and slant antenna measurements are less than 0.2 K for the X polarization and less than 1 K for the Y polarization.

4.3. CAROLS Measurements over Land

Figure 3 shows the measured temperatures recorded over the SMOSMANIA site, acquired with both the nadir and slant antennas after RFI elimination [6]. For the nadir antenna, we find a good agreement between the T_x and T_y measurements. For the slant antenna, we verify that the T_x (close to T_V) values are higher than the T_Y (close to T_H) values. For CAROLS measurements, we qualitatively validate the agreement between Brightness temperatures and ground moisture measurements, with an increasing value of T_B on the driest flight days. The detailed inversion results will be published in the near future.

5. CONCLUSION

In the context of the SMOS mission, the CAROLS L-Band radiometric instrument has been built and tested. A large database was acquired during SMOS CAL/VAL campaigns over three studied sites (SMOSMANIA, VAS and the Gulf of Biscay). The retrieved brightness temperatures are found to be in good agreement and coherence with the simulated values over ocean. These measurement campaigns have been successful, and the detailed science results will be published in the near future.

ACKNOWLEDGMENT

This work is a SMOS/ESA Cal/Val project and was funded by the TOSCA/CNES program and ESA. The authors would like to thank the French CNES for the considerable support it gave to the CAROLS project, as well as to all of the CAROLS campaigns. The authors would also like to thank the technical teams from CESBIO, CNRM, LATMOS, LOCEAN, SAFIRE, DT-INSU, DTU, INRA, and the SAFIRE pilots for their support and contributions to the successful airborne and ground campaigns.

REFERENCES

1. Kerr, Y., P. Waldteufel, J. P. Wigneron, Martinuzzi, J. Font, and M. Berger, "Soil moisture retrieval from space: The soil moisture and ocean salinity (SMOS) mission," *IEEE Transaction*

- on Geoscience and Remote Sensing*, Vol. 39, 1729–1735, 2001.
2. Rotbøll, J., S. S. Søbjerg, and N. Skou, “A novel L-band polarimetric radiometer featuring subharmonic sampling,” *Radio Science*, Vol. 38, No. 3, 11-1–11-7, 2003.
 3. Cadarech, E. and A. Rius, “A new technique to sense non-Gaussian features of the sea surface from L-band bi-static GNSS reflections,” *Remote Sensing of Environment*, Vol. 112, No. 6, 2927–2937, June 2008.
 4. Calvet, J. C., N. Fritz, F. Froissard, D. Suquia, A. Petitpa, and B. Piguet, “In situ soil moisture observations for the CAL/VAL of SMOS: the SMOSMANIA network,” *International Geoscience and Remote Sensing Symposium, IGARSS*, 1196–1199, Barcelona, Spain, July 23–28, 2007, doi:10.1109/IGARSS.2007.4423019.
 5. Reul, N., J. Tenerelli, B. Chapron, S. Guimbard, V. Kerbaol, and F. Collard, “CoSMOS OS campaign, scientific data analysis report, ifremer report,” November 2006.
 6. Pardé, M., M. Zribi, P. Fanise, and M. Dechambre, “Analysis of RFI issue using the CAROLS L-Band experiment,” *IEEE Transaction on Geoscience and Remote Sensing*, 2010, in press.

Radio Frequency Interferences Investigation Using the Airborne L-band Full Polarimetric Radiometer CAROLS

M. Pardé¹, P. Fanise¹, M. Zribi², and M. Dechambre¹

¹UVSQ-CNRS, LATMOS, Guyancourt, France

²CESBIO-IRD, Toulouse, France

Abstract— In the present paper, different methods are proposed for the detection and mitigation of the undesirable effects of radio frequency interference (RFI) in microwave radiometry. The first of these makes use of kurtosis to detect the presence of non-Gaussian signals, whereas the second imposes a threshold on the standard deviation of brightness temperatures, in order to distinguish natural emission variations from RFI. The third approach is based on the use of a threshold applied to the third and fourth Stokes parameters. These methods have been applied and tested with a CAROLS radiometer operating in the L-band. We also used of a spectral analyser in the view of dividing the spectrum into subbands for subband filtering. These instruments were operated during airborne campaigns made in spring 2009 over the South West of France. The performance of different algorithms is analyzed for both of them. We showed that the kurtosis method is well adapted to pulsed RFI, whereas the method based on the second moment is well adapted to continuous-wave RFI. Subband filtering allow us to mitigate RFI effects, the efficiency of these method is better using algorithm based on Euclidian distance estimation.

1. INTRODUCTION

In recent years, various studies have revealed the substantial influence of Radio-Frequency Interference (RFI) in microwave radiometry. In fact, this type of noise can corrupt measurements, thereby introducing a significant deterioration to the database quality. In some regions of the world, it can in some cases render the data simply unusable. In an effort to circumvent this significant problem, RFI detection and mitigation has been studied in a small number of papers [2, 4].

In [4], the authors have shown the ability of an ADD (Agile Digital Detector) to detect and remove RFI from microwave measurements. The ADD performance was experimentally verified under controlled laboratory conditions and in the field, near to a commercial air traffic control radar. In [2], the authors tested a double detector based on kurtosis, under equivalent conditions (on ground and in the laboratory).

With the advent of various satellite radiometers operating in the L band (1400–1427 MHz, a protected frequency band), which are designed mainly for soil moisture estimation over the full terrestrial surface, this particular problem of measurement perturbations is becoming increasingly significant. Small levels of RFI, of only a few Kelvin, can cause the soil to appear to be dryer than in reality.

The objective of the present study is to enable natural surface emissions, interpreted for the purpose of soil moisture measurements, to be correctly estimated. For this, we made use of airborne measurements acquired in the L-band with CAROLS (Cooperative airborne radiometer for Ocean and Land Studies) [6], over different terrestrial areas including the South West of France, in which we identified the presence of strong RFI signals. We also tested the use of a Spectral Analyser (SA) to mitigate RFI by dividing the spectrum in several subbands. We propose to test and compare different approaches, using or the SA or the radiometer, in order to identify and mitigate both forms of RFI, pulsed and continuous-wave, present in the CAROLS data.

The present paper is organized as follows: Section 2 presents our database, with a description of the CAROLS radiometer, Spectral Analyser and CAROLS flight patterns. Section 3 discusses the approaches proposed to detect and eliminate data contaminated with RFI. Finally, our conclusions are presented in Section 4.

2. DATA PRESENTATION

2.1. CAROLS Flight Plans

CAROLS radiometer flights [6] were carried out with the French ATR42 operational research aircraft, operated by SAFIRE. The onboard instrumentation also included an inertial navigation unit combined with a GPS receiver, for aircraft attitude and position recording. The Slant antenna was pointing at an angle of 32° (Slant antenna).

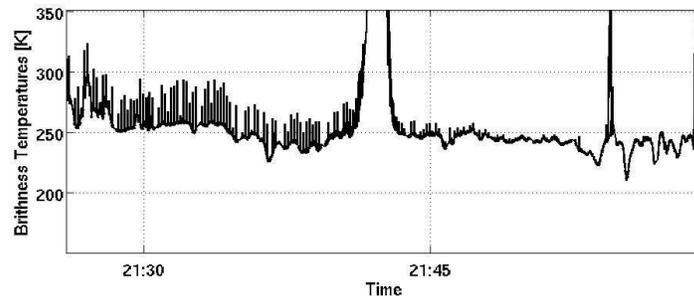


Figure 1: $T_b H$ measurements during Flight 07 using the slant antenna. Values averaged over 40 ms. The time axis corresponds to local time ($= UTC + 1$).

2.2. CAROLS Radiometer

The receiver, designed and built as a copy of the EMIRAD II radiometer from the DTU team, and adapted to the French ATR42 research aircraft, is a fully polarimetric correlation radiometer with direct sampling [5]: all four Stokes parameters describing the electric field are measured. Calibration is carried out using internal loads and a noise diode, which adds approximately 120 K to the radiometer input. This setup ensures frequent internal calibration capability, and allows the radiometer to calibrate any phase difference between the two input channels preceding the digital correlator. More information about the CAROLS radiometer is available in [6].

For the purposes of the present study we selected one specific flights over the SMOSmania sites in the South West of France, flown on 27 and April, 2009, of which we chose to analyze only one part: the return of the aircraft from Biscarosse to Francazal airport near Toulouse. We selected measurements corresponding to the Slant antenna only. All measurements (4 Stokes parameters and kurtosis values) were averaged over 40 ms periods.

All of the above-ground transects were disturbed by interference, in both polarizations (see Figure 1 for the H measurements); pulsed sources appear clearly throughout the flight. Moreover, two point sources led to a strong perturbation of the measurements, in the vicinity of Auch and Toulouse (after 21 : 40 and near 21 : 55). The measurements in the H polarization were also disturbed near to the ocean, probably due to military emissions from the Cazeaux base (right side of Figure 1).

The following comparisons are developed and illustrated, for the quantity T_b measured by the Slant antenna in the H polarization only, because the RFI effects were more significant on this antenna.

2.3. Spectral Analyser

For the purpose of analysing the L-band spectrum in the south of France, we dedicated a flight to the use of a Spectral Analyser (SA). The digital signal processing consists in the computation of a Fast Fourier transform (FFT), followed by the power spectrum estimation and accumulation. When connected to the H channel of the CAROLS radiometer slant antenna, the SA allows to record integrated spectra.

Figure 2 represents the data collected during the spectral analysis flight performed the 29 of May. To plot this figure, each frequency channel was first calibrated using samples acquired with the internal load source, with and without an internal noise source. The spectral variations were plotted as a function of time from the Ocean to Toulouse (from the left to the right of the figure).

During land survey several peaks corresponding to the presence of three distinct types of RFI can be identified: pulsed RFI sources in the frequency range 1.418–1.419 GHz, emitted by military and civil radars, CW RFI in a narrow band (CEL), and CW RFI occupying the full bandwidth (Auch). Because the latter form of RFI is distributed over the full bandwidth of interest, it was not possible to restore the correct values for T_b .

3. RFI FILTERING

Naturally occurring thermal emissions, and self-generated receiver noise measured by a radiometer, have well modeled Gaussian distributions. Interference-free Gaussian signals have a constant relationship between the power of their second and fourth central moments, μ_2 and μ_4 . This

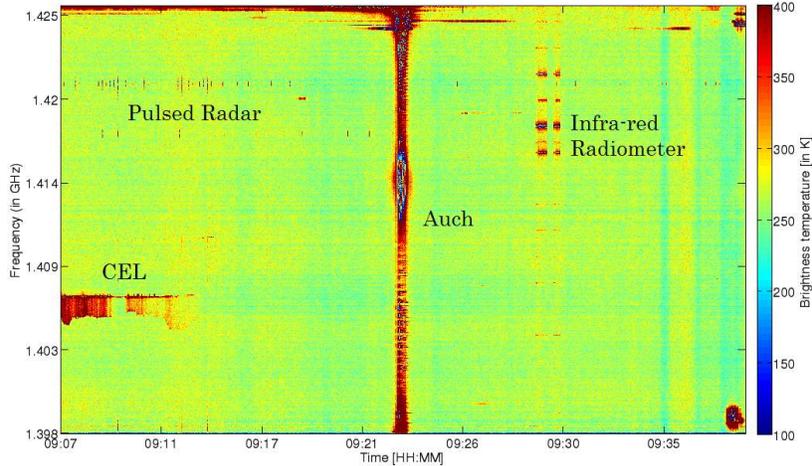


Figure 2: Time-Frequency representation of calibrated Tb values, during the flight from the Atlantic Ocean to Toulouse city (France).

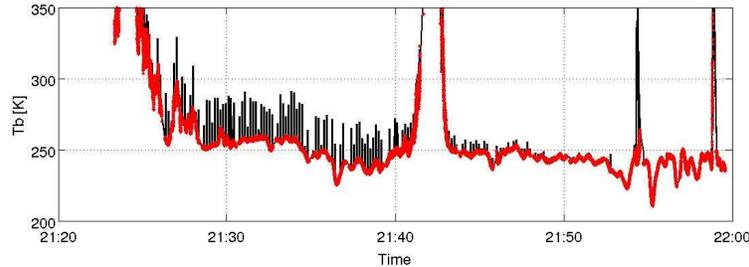


Figure 3: Tb values measured in H polarization along the flight track. The black line represents the unfiltered data, and the red points correspond to the values after application of the F_K3 filter.

relationship is the kurtosis, which is defined as the ratio:

$$\beta_4 = \frac{\mu_4}{\mu_2^2} \quad (1)$$

and is equal to 3 if the input is Gaussian. The deviation of this quotient from its constant value of three is an indicator of the presence of non-Gaussian RFI. A pulsed sinusoidal waveform of low duty cycle (radar-type signal) will have a β_4 greater than 3, and this property can be used to flag RFI-contaminated signals.

This last condition applied to the CAROLS data respects the theoretical. As a general rule, the F_K conditions are considered in kurtosis algorithm applications. However, as shown in Table 1, the application of this condition is restrictive, since only 72% of the analyzed data remains unflagged. Moreover, the mean of the Tb measurements in H polarization is greater in this case than that resulting from F_K3.

Figure 3 illustrates RFI elimination using F_K filter. Under these conditions, data exhibiting radar-pulse contamination is found to be completely eliminated. On the other hand, the continuous wave noise present in the vicinity of Auch city is not correctly eliminated: many Tb values as high as 300K are not removed. One very simple method for discriminating between RFI data and non-polluted data could be the use of a flag, based on the estimation of the brightness temperature standard deviation (second moment, filter named F_STD). We propose to analyze different options, making use of the Stokes parameters to filter RFI-contaminated data. We considered an empirical threshold for both T_3 and T_4 , chosen from RFI-free data, and then applied the filter to the Tb values in H polarization, based on this threshold. The two filters are: F_T3: threshold of $\pm 3K$ and F_T4: threshold of $\pm 3K$.

From the results illustrated in the three last sections, the kurtosis and standard deviation approaches are seen to be complementary. In fact, the former method correctly detects radar pulses, but has difficulties to eliminate continuous-wave RFI noise, such as that observed in the vicinity of

Table 1: RFI percentages and mean Tb values in H polarization, depending on the filter applied to the data set.

| Filtering method | Percentage of kept data (%) in H pol | H (K) |
|----------------------|--|---------|
| Without modification | 100 | 356,6 |
| F_K | 79,9 | 250,7 |
| F_STD | 77.5 | 246,8 |
| F_STD_K | 70.4 | 246,7 |
| F_T3 | 75,5 | 248,8 |
| F_T4 | 62,5 | 247,8 |

Table 2: % of deleted data and mean Tb values, for 3 different masking algorithms, and division of the spectrum into 8 or 16 subbands.

| | number of subbands | % of deleted data | Mean Tb value (K) |
|------------------|--------------------|-------------------|-------------------|
| Without blanking | | | 270.7 |
| β_4 | 8 | 32.2 | 266.3 |
| ED | 16 | 8.3 | 261.4 |
| Shapiro-Wilk | 16 | 18.8 | 266,4 |

Auch. In the other hand, standard deviation approaches are well adapted to the detection of such continuous wave RFI sources, characterized by a strong increase in the value of this parameter. For these reasons, we propose the identification (and removal) of RFI based on the combination (referred to as F_STD_K) of these two methods: in this case, the F_K filter is first applied to the data, followed by the F_STD filter. We also tested the possibility of making simultaneous use of Kurtosis filters (F_K), and T_3 and T_4 (F_T3 and F_T4).

Table 1 provides a summary of the outcome of various RFI filtering strategies, using the different combinations of the aforementioned filters. For each case, we compute the percentage of data retained, together with the mean brightness temperature over the whole transect for the H polarization. It can be seen that in the absence of RFI mitigation, the mean value of Tb is very high: 356.6 K. The application of the Kurtosis filters F_K allows the mean value of Tb to be significantly reduced. It can be seen that the F_STD filter produces a very low mean value of Tb (246.7 K).

4. RFI MITIGATION ALGORITHM USING SUBBAND FILTERING

In this section we present three algorithms, which we applied in order to recover Tb values free from the influence of RFI. In order to identify the presence of RFI, an easy approach consist in dividing the spectrum into different subbands, and then applying different criteria to each subband so as to classify the samples according to whether or not they are affected by RFI. A threshold needs to be defined for each criteria.

We compare 3 different RFI mitigation algorithms. One based on the Kurtosis (β_4) parameter, computed for each subband, one based on the Shapiro Wilk (SW) parameter [1] and one based on the Euclidian distance (ED). Two factors are used in this analysis: the number of deleted data, and the number of mean Tb computed for only one small part of the transect, corresponding to a path length of approximately 40 km (close to the dimensions of the central field of view of a SMOS pixel). These values are presented in the following Table 2, for 8 or 16 subbands.

In the case of the kurtosis test, the results are given for 8 subbands only, a large number of psd (power spectral density) values is needed to compute β_4 ; Kurtosis values with less psd are incorrect.

The most efficient mask was that based on the ED criterion, using 16 subbands. This mask allowed only 8.3 % of the data to be deleted. This low percentage of deleted data together with the low Tb value ensures a small number of false alarms. The β_4 mask and the SW test led to a higher mean Tb value, with a larger amount of deleted data. This outcome can be interpreted as these algorithms also mask some uncorrupted parts of the spectrum.

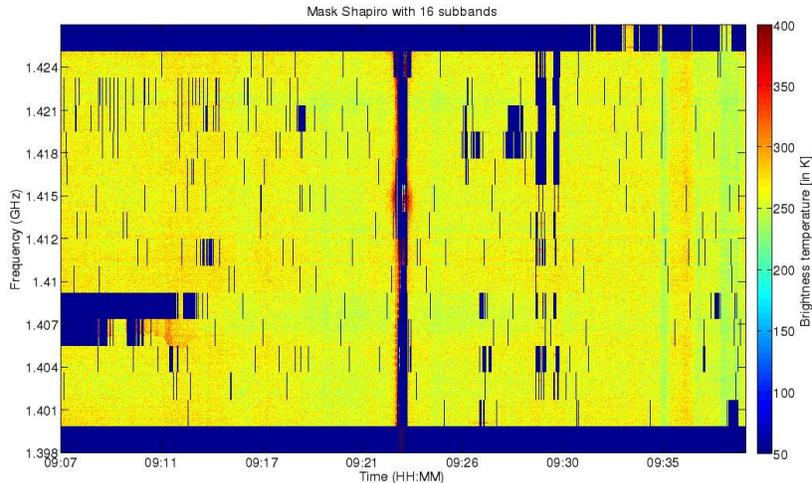


Figure 4: Tb after correction using the S - W filter, when the spectrum is divided into 16 subbands.

5. CONCLUSION

In the present study, different methods have been evaluated for the detection and elimination of RFI from CAROLS radiometric data. We applied this analysis to the South West of France, characterized by the presence of strong RFI emissions due to two types of source in particular: pulsed radars and antennas emitting a continuous and very high level of RFI. The first filtering techniques are applied on foul-band radiometric data, they are based on an algorithm using kurtosis parameter, standard deviation of Tb or on third and fourth Stokes parameters. Our analysis of the resulting RFI corrections reveals that the two methods (Kurtosis and STD) are complementary. In fact, the kurtosis parameter is well adapted to the correction of radar pulses. On the other hand, standard deviation analysis is better adapted to continuous noise sources characterized by an approximately Gaussian distribution, which are not easily detected with kurtosis. A FFT analyser has been developed and added to the radiometer receiver in order to retrieve and remove RFI from radiometric measurements. After correction, a new “clean” Tb without corrupted subband is calculated. Results from L-band observations have demonstrated clearly the effectiveness of frequency domain blanking for RFI mitigation using Euclidian distance. Pulsed and continuous wave RFI are removed and the Tb can be correctly retrieved. This is an important result because the ocean salinity and the surface soil moisture will be derived even if RFI corrupt the measurements. In the future, the Euclidian distance could be replaced by the Mahalanobis distance, calculated with information concerning subband and polarization diversity and polarimetric information; that could improve the classification between RFI and non RFI samples.

REFERENCES

1. Güner, B., M. T. Frankford, and J. T. Johnson, “A study of the Shapiro-Wilk test for the detection of pulsed sinusoidal radio frequency interference,” *IEEE Trans. Geosci. Remote Sens.*, Vol. 47, 1745–1751, 2009.
2. Piepmeier, J. R., P. N. Mohammed, and J. J. Knuble, “A double detector for RFI mitigation in microwave radiometers,” *IEEE Trans. Geosci. Remote Sens.*, Vol. 46, No. 2, 458–464, 2008.
3. Pardé, M., M. Zribi, P. Fanise, and M. Dechambre, “Analysis of RFI issue using the CAROLS L-band experiment,” *IEEE Trans. Geosci. Remote Sens.*, 2010, accepted.
4. Ruf, C., S. M. Gross, and S. Misra, “RFI detection and mitigation for microwave radiometry with an agile digital detector,” *IEEE Trans. Geosci. Remote Sens.*, Vol. 44, 694–706, 2006.
5. Rotbøll, J., S. S. Søbjærg, and N. Skou, “Novel L-band polarimetric radiometer featuring subharmonic sampling,” *Radio Science*, Vol. 38, No. 3, 1–7, 2003.
6. Zribi, M., M. Pardé, J. Boutin, P. Fanise, D. Hauser, M. Dechambre, Y. Kerr, M. Leduc-Leballeur, G. Reverdin, N. Skou, S. S. Søbjærg, C. Albergel, C. Calvet, J. P. Wigneron, E. Lopez-Baeza, K. Saleh, A. Ruis, and J. Tenerelli, “CAROLS: A new airborne L-Band radiometer for ocean surface and land observations,” *IEEE Trans. Geosci. Remote Sens.*, submitted.

Interpretation of CAROLS L-band Measurements in the Gulf of Biscay (September 2007)

J. Boutin¹, M. Leduc-Leballeur², M. Pardé², M. Zribi³, P. Fanise²,
G. Reverdin¹, J. Tenerelli⁴, and N. Reul⁵

¹LOCEAN (CNRS/UPMC/IRD/MNHN), France

²LATMOS (UVSQ/CNRS/UPMC), France

³CESBIO (CNRS/IRD/CNES/UPS), France

⁴CLS, France

⁵LOS/IFREMER, France

Abstract— The L-band Cooperative Airborne Radiometer for Ocean and Land Studies (CAROLS) radiometer flew four times over the Gulf of Biscay between September 24 to September 28, 2007 around 20UTC. These flights were the first ones over the ocean of this new instrument. Concomitant ship campaign and drifter deployments provide in situ ground truths for sea surface salinity (between 34.6 and 35.8 pss) and temperature (between 15°C and 17°C). Wind speed (between 2 and 10 m/s) and direction are estimated from the QSCAT scatterometer.

Observed variations of CAROLS brightness temperatures in horizontal polarization are very consistent with the ones simulated using a direct modeling of L-band radiometric signals over the ocean, demonstrating a good sensitivity of CAROLS instrument and the realism of the modeling. During circle flights during which variations of the galactic signal scattered by the sea surface was moderate (less than 0.5 K), observed azimuth variations at 15° incidence angle are consistent with the galactic noise signal scattered by the sea surface plus a small signal due to rough sea asymmetry as simulated by a two-scale model. On September 28, a more than 1 K increase of T_b over 2° longitude is observed, linked to an increase of wind speed from 3 to 8 m/s.

1. INTRODUCTION

Since the 90th, measuring sea surface salinity (SSS) from space using the L-band microwave radiometry saw renewed interest (1) (2). The Soil Moisture and Ocean Salinity (SMOS) satellite was launched in November 2009 and the Aquarius/SAC-D satellite is planned to be launched in Spring 2011. In order to refine the modeling of the L-band radiometric signal in preparation of the SSS retrieval from these new missions, the L-band Cooperative Airborne Radiometer for Ocean and Land Studies (CAROLS) radiometer has been built at LATMOS (3) and has been mounted onboard ATR42 Meteo France airplane. In this paper we analyze the measurements performed by CAROLS during the first four flights over the ocean and we compare them with simulations performed using L-band radiometric modeling developed before launch to interpret SMOS ocean measurements.

We focus 1) on azimuth variations observed during circle flights and related to galactic noise scattered towards CAROLS antenna and to wind direction and 2) on wind speed signature across a spatial 3–8 m/s wind speed variation.

2. DATA AND METHOD

The four flights were performed over the Gulf of Biscay between September 24 and September 28, 2007 around 20UTC (Figure 1). Brightness temperatures (T_b) were measured by one antenna looking at 33 degrees on the right hand side of the aircraft and optionally by a nadir antenna. In order to reduce T_b noise, T_b have been averaged over 1 s. Due to an imperfect decoupling of the X polarisation with the Y polarisation, only the Y polarisation (corresponding to the horizontal polarisation in the sea surface frame) was usable.

Concomitant ship campaign and drifter deployments provide in situ ground truths for sea surface salinity (between 34.6 and 35.8 pss, Figure 1) and temperature (between 15°C and 17°C). Wind speed and direction are estimated from the QSCAT scatterometer descending orbits which flew within 2 hours from the CAROLS flight. Except on 28 Sept. (see Figure 1), wind speed was between 7 and 10 m/s.

CAROLS measurements are compared with simulations conducted with the Terrestrial Radiometry Analysis Package (TRAP) software (4) run for CAROLS geometry and observed geophysical

conditions. TRAP uses the physical modelling of atmospheric radiative transfer (5), sea surface emissivity (flat sea model of Klein and Swift and rough sea model either from the two scale model that uses the Durden and Vesecki wave spectrum multiplied by a factor two (6) (2scale/DV2) or from the sea surface slope approximation model (7) that uses the Kudryatsev wave spectrum (8) (SSA/Kudr.) and galactic glint (9) that were implemented for the first processing of the Soil Moisture and Ocean Salinity satellite data. The simulations are integrated over the CAROLS lobe antenna pattern measured in anechoic chamber.

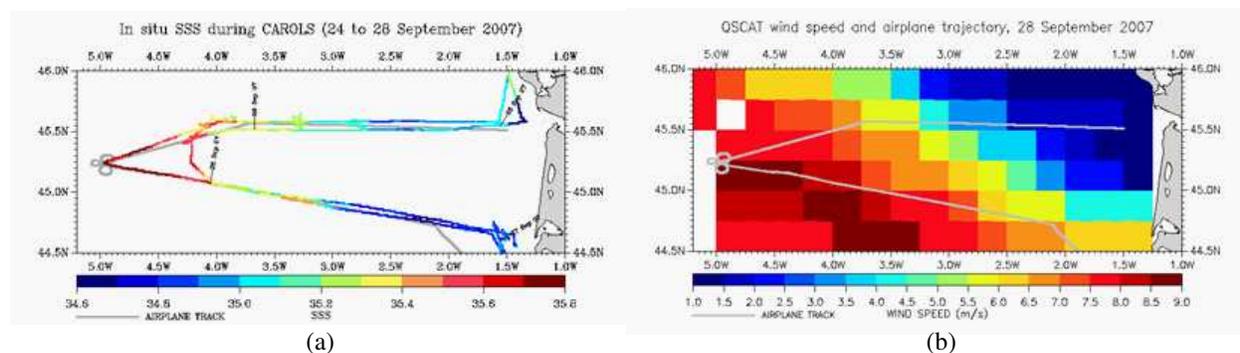


Figure 1: Geophysical conditions during CAROLS 2007 campaign. (a) SSS (color coded) measured onboard the Cote de La Manche ship and by sea surface drifters; (b) Wind speed from QSCAT scatterometer. The plane track is indicated as a grey line.

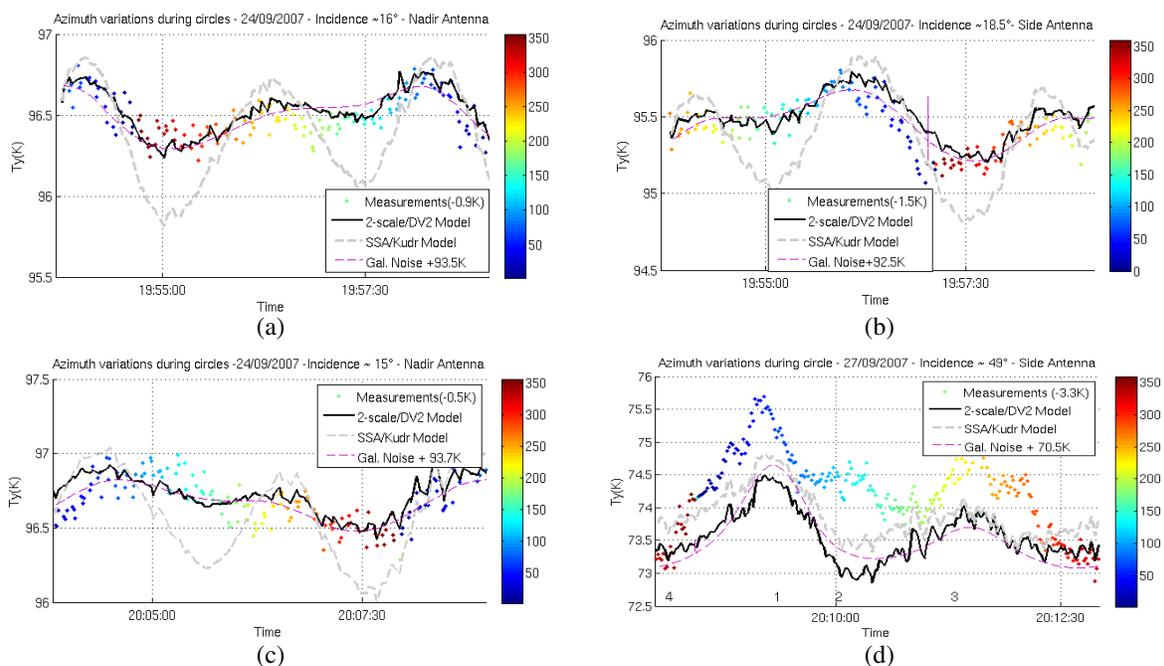


Figure 2: Variations of T_y as observed (color points) and modeled using 2scale/DV2 (black line) and SSA/Kudr. (grey dashed line) during 4 circle flights. Azimuth angle is color coded; it is defined anticlockwise, with an origin pointing westward. The scattered galactic noise shifted by a constant amount to fit within the simulated T_y scale is represented as a purple dashed line. (a) (c) nadir antenna at an incidence angle close to 15° on 24/09/07 (wind speed of 7.6 m/s). Side antenna pointing at (b) about 18.5° on 24/09/07 (wind speed of 7.6 m/s) and (d) about 49° on 27/09/07 (wind speed of 6.2 m/s); numbers refer to the periods shown on Figure 3. Measurements have been corrected for an arbitrary bias (see text and legend on the figures).

3. RESULTS

A variable shift between measurements and simulations is observed all along the flights. Investigations after the flights, indicate that a connector was imperfectly isolated, leading to imperfect

calibration and that the radiometer bandwidth was too large, leading to large radio interferences, especially over land. As a consequence, in the following, only portions of flights during which no sudden jump in measurements of a few tenth of Kelvins is observed are analyzed. Arbitrary biases have been removed from the measurements for an easier reading of the figures.

Circles above the Gascogne buoy (45.3°N , 5°W) allow to explore all the azimuth angles. In the simulations, the azimuth variation is created by the azimuth variation of the scattered galactic noise and of the sea surface emissivity. Figure 2 (Left and top right) illustrates the azimuth variations of the observed and simulated signal close to 15° incidence angle by both antennas. For these incidence angles, the part of the sky affecting the measurements always contains the Cygnus A source and a variable part of the milky way which creates about 0.4 K azimuth variation of the scattered galactic signal (see purple line on the Figures). At first order, the azimuth variation of the measurements is driven by the azimuth variation of the scattered galactic noise. The additional azimuth variation coming from the rough sea surface emissivity as simulated with the 2scale/DV2 model is small and the total azimuth variation simulated with this model is within the noise of measurements. On another hand, the additional azimuth variation coming from the rough sea surface emissivity simulated the SSA/Kudr. model at 15° incidence angle appears to be too large.

At large incidence angle (Figure 2 right bottom), the modeled scattered galactic noise is much more variable than at 15° and it varies by 1.5 K. Depending on the azimuth, the part of the sky specularly reflected by the surface towards the radiometer includes or not the milky way (see

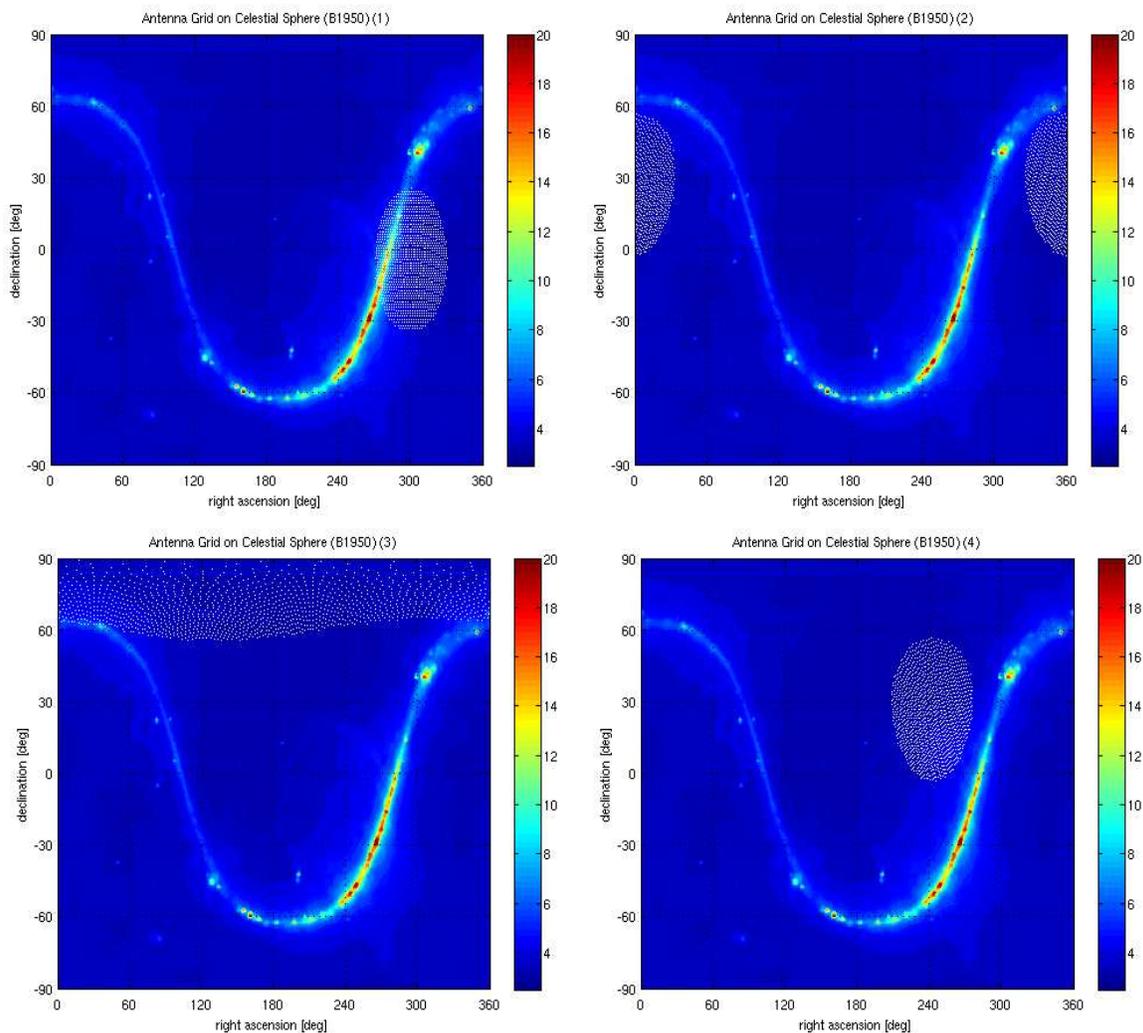


Figure 3: Points in the sky reflected (in specular direction) into the lobe of the CAROLS side antenna superimposed onto maps of L-band sky emission during four periods of a circle performed on 27/09/07 (the four periods are indicated on Figure 2 bottom, right).

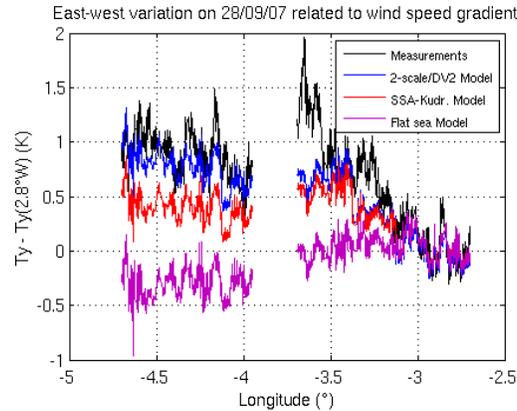


Figure 4: 28 September, northern transect, Tb anomalies referred to the values at 2.8°W (wind speed equal to 3 m/s): measurements at 33° incidence angle (black), simulations with the complete model using for the rough emissivity 2-scale/DV2 model (blue) or SSA/Kudr. model (red), simulations without taking into account the rough sea emissivity (purple). No data are shown between 3.7°W and 3.9°W corresponding to strong attitude variations due to the change in the direction of the aircraft.

Figure 3). The observed variations are even larger (up to 2.5 K) and neither the additional signal coming from the rough surface emissivity modeled with the 2 scale/DV2, nor the SSA/Kudr. is enough to reproduce the observation. In addition strong signal around 150° azimuth angle cannot be (even qualitatively) explained by our models. This is suspected to be due to radio interferences coming from the Spanish Basque coast as unexplained large signals have often been observed in this azimuth direction on CAROLS flights and as strong radio interferences have been detected on the first SMOS images in this region (the sources are now off).

On September 28, a large east-west wind speed gradient occurs (Figure 1, right). Over the northern transect, no connection problem was suspected between 2.7°W and 4.7°W in the Tb measured by the side antenna (no sudden jump on Tb was observed) so that it was possible to compare modeled and measured Tbs (Figure 4). In the simulations, the increase of Tb associated with the increase of wind speed ($\sim 0.2 \text{ K/m/s}$) is partly compensated by the east-west increase of SSS ($\sim -0.5 \text{ K/pss}$) (Figure 4). The best agreement between the measurements and the simulations is obtained using the 2-scale/DV2 model except between 3.5°W and 3.7°W: during that period, the STORM scatterometer measurements onboard the aircraft suggested a roughness increase (not shown, (3)) that was not observed by QSCAT that sample this area 1h45mn before the aircraft.

4. DISCUSSION AND CONCLUSIONS

Although they reveal some instrumental flaws, the first flights of CAROLS instrument over the ocean were very encouraging. Over selected parts of the flights, the variability of CAROLS measurements is very consistent with the variability expected from model simulations. It allows to draw some conclusions about the models used in the pre-launched version of the SMOS salinity processor:

- at 15° incidence angle and 7 m/s wind speed, under moderate azimuth variations of the galactic signal, the azimuth variation simulated with the Reul galactic glint model and the 2-scale-DV2 rough emissivity model is consistent with the observations, contrary to that obtained with the SSA/Kudr. rough emissivity model.
- at 49° incidence angle and 6 m/s wind speed, under strong azimuth variations of the galactic signal, the azimuth variation simulated with the Reul galactic glint model and either the 2-scale-DV2 rough emissivity model or the SSA/Kudr. model is smaller than the observed one.
- at 33° incidence angle, under wind speed varying between 3 and 8 m/s, the best agreement between measurements and simulations is found when the 2-scale-DV2 rough emissivity model is used.

These conclusions are only valid for the horizontal polarization and need to be confirmed with other flights as wind speed may have locally varied between the time of QuikScat and CAROLS measurements separated by 1h45mn.

ACKNOWLEDGMENT

This work is a SMOS/ESA Cal/Val project and was funded by the TOSCA/CNES program. The authors would like to thank the French CNES for the considerable support it gave to the CAROLS project. The authors would also like to thank the technical teams from LATMOS, LOCEAN, SAFIRE, DT-INSU, DTU, and the SAFIRE pilots for their support and contributions to the successful airborne and ground campaigns. We are indebted to S. Zine and N. Martin for their advice and support with the computing aspects of this study. QuikScat data are produced by Remote Sensing Systems and sponsored by the NASA Ocean Vector Winds Science Team. Data are available at www.remss.com.

REFERENCES

1. Lagerloef, G. S. E., C. T. Swift, and D. M. Le Vine, "Sea surface salinity: the next remote sensing challenge," *Oceanography*, Vol. 8, 44–50, 1995.
2. Font, J., A. Camps, A. Borges, M. Martín-Neira, J. Boutin, et al., "SMOS: The challenging sea surface salinity measurement from space," *Proceedings of the IEEE*, Vol. 98, 649–665, 2010.
3. Zribi, M., M. Pardé, J. Boutin, P. Fanise, D. Hauser, et al., "CAROLS: A new airborne L-band radiometer for ocean surface and land observations," *IEEE Trans. Geosc. Rem. Sens.*, 2010, submitted.
4. Reul, N., J. Tenerelli, B. Chapron, S. Guimbard, V. Kerbaol, et al., "Cosmos OS campaign: Scientific data analysis report," 2008.
5. Liebe, H. J., G. A. Hufford, and M. G. Cotton, "Propagation modeling of moist air and suspended water/ice particles at frequencies below 1000 GHz," *AGARD 52nd Specialists' Meeting of the Electromagnetic Wave Propagation Panel*, presented, Palma de Mallorca, Spain, 1993.
6. Dinnat, E. P., J. Boutin, G. Caudal, and J. Etcheto, "Issues concerning the sea emissivity modeling at L band for retrieving surface salinity," *Radio Sci.*, Vol. 38, 8060, 2003.
7. Johnson, J. T. and M. Zhang, "Theoretical study of the small slope approximation for ocean polarimetric thermal emission," *IEEE Trans. Geosci. Remote Sens.*, Vol. 97, 2305–2316, 1999.
8. Kudryavtsev, V. N., V. K. Makin, and B. Chapron, "Coupled sea surface-atmosphere model. 2. Spectrum of short wind waves," *J. Geophys. Res.*, Vol. 104, 7625–7639, 1999.
9. Tenerelli, J. E., N. Reul, A. A. Mouche, and B. Chapron, "Earth-viewing l-band radiometer sensing of sea surface scattered celestial sky radiation — Part I: General characteristics," *IEEE TGRASS*, Vol. 46, 2008.

Retrievals of Soil Moisture and Optical Depth from CAROLS

M. Pardé², J.-P. Wigneron¹, M. Zribi³, Y. Kerr³, P. Fanise²,
 J.-C. Calvet⁴, C. Albergel⁴, A. Albitar³, F. Cabot³, F. Demontoux³, E. Jacqueline³,
 E. Lopez-Baeza⁵, A. Mialon³, C. Moisy³, N. Novello¹, P. Richaume³,
 K. Saleh⁶, M. Schwank³, P. Waldteufel³, E. Zakharova⁴, and M. Dechambre²

¹INRA, EPHYSE, Bordeaux, France

²UVSQ-CNRS, LATMOS, Guyancourt, France

³CESBIO, Toulouse, France

⁴Météo-France, CNRM, Toulouse, France

⁵Universitat de Valencia, Valencia, Spain

⁶Cambridge University, Cambridge, England, UK

Abstract— We propose in this paper to evaluate a method to retrieve soil moisture (SM) and vegetation optical thickness, in areas of unknown roughness and unknown vegetation water content in view of operational applications, by using airborne Tb measurements acquired in South-West of France. Results are compared to *in situ* measurements, manual and automatic ones included in SMOSmania network, in the South-West of France.

1. INTRODUCTION

The European Space Agency has launched the SMOS (Soil Moisture and Ocean Salinity) satellite during the spring of 2009. The generation of land products from SMOS brightness temperature (Tb) measurements relies on the inversion of the microwave forward model L-MEB (L-band Microwave model of the Biosphere) [7]. The L-MEB model assembles a set of equations describing the emission and scattering of the surface, vegetation and atmosphere at L-band (1.4 GHz). The calibration of L-MEB for different surfaces has been addressed by numerous studies over the last ten years, most of them based on the analysis of ground-based L-band data. The study presented here is based on results of the CAROLS study [8], designed in support of the CAL-VAL of SMOS and of the evaluation of the L-MEB model over land surfaces.

We propose in this paper to evaluate a method to retrieve soil moisture (SM) in areas of unknown roughness and unknown vegetation water content in view of operational applications, by using airborne Tb measurements acquired at two incidence angles and two polarizations. We made the assumption that the roughness parameter was dependent on SM; this roughness correction was used in the 2-Parameter retrievals of soil moisture and optical depth for all flights. Results are compared to *in situ* measurements, manual and automatic ones included in SMOSmania network, in the South-West of France. Very good agreement between these sources of SM estimation was obtained which is very promising for future SMOS validation activities.

We will first present CAROLS data (II), the radiative transfert model (III) and finally comparison between measurements and model inversion (IV).

2. CAROLS DATA PRESENTATION

In the context of the SMOS Cal/Val project, an airborne radiometer was developed and installed in an aircraft since 2007. Airborne campaigns were conducted in 2007, 2008 and 2009 for scientific and technical purposes and a final campaign was conducted in 2010 after the launch of the SMOS satellite. Details of the campaigns are available in the paper of Zribi et al. [8] and we summarize here the main informations about these campaigns.

2.1. CAROLS Radiometric Data

In 2007, we developed with the collaboration of TUD (Technical University of Denmark) a radiometer as a copy of EMIRAD [1]. This radiometer was tested onboard in 2007. Its main characteristics are:

- Frequency: 1400–1427 MHz at -3 dB, Digital I/Q demodulation and correlation for accurate estimation of the 3rd and 4th Stokes parameters,
- Data integrated for 1 ms, and recorded on a primary storage PC,
- The raw data pre-integrated to $1.8 \mu\text{s}$ is recorded on a dedicated PC.

Two antennas were installed in the plane, one nadir looking and one right side looking antenna (view angle around 33 degrees). The radiometer was programmed to switch on each antenna alternatively every 5 seconds.

The radiometer was calibrated using laboratory measurements, automatic calibration with two internal sources and additional losses are estimated using a well known ocean target. Far from the coast, salinity and temperature are stable during the campaign and we were able to compute Tb values with a sufficient accuracy. The radiometer performances were estimated in laboratory and we showed that stability and accuracy were around 0.1 K during the whole 2009 campaign.

2.2. CAROLS 2009 Flights

In 2009 13 flights were conducted during Spring (23/04/2009–28/05/2009). The plane was taking off in Toulouse (France) and different sites were overlighted: Gulf of Biscay, South West of France and Valencia site (VAS).

In this paper we only used Tb measured in the South West of France, overflying the SMOSmania test sites operated by Météo-France [2]. In Figure 1, we show the flight path and the location of the SMOSmania sites.

2.3. Brightness Temperature Preprocessing

Before estimating the soil moisture, different steps were applied to correct Tb.

- The first step is the projection of Tb measured from the antenna frame (X and Y polarization) into the earth frame (H and V polarization).
- Combination of Tb values measured with the nadir antenna with the more closed Tb values measured with the slant antenna.
- Filtering data that are polluted by Radio Frequency Interferences (RFI) was based on Kurtosis and Tb standard deviation thresholding [4].

Then radiative model inversion was applied to these corrected Tb data.

2.4. Ground Based Measurements

In the meantime of the flights, different automatic and manual measurements were done *in situ*:

- Soil moisture (SM) and soil temperature (ST) were estimated every 30 mn using TDR (Time Domain Reflectometry) measurements on the 12 sites within the SMOSmania network [2]. In this study, we used TDR measurements at 5 cm depth to evaluate the Tb inversion process.
- Specific fields were chosen in 2009 for *in situ* gravimetric SM measurements, temperature and theta probes measurements too. One part of these fields are located near Bordeaux, on the Western part of the flight and the other part is located near Toulouse.

Location of the 12 SMOSmania ground stations are shown on Figure 1.

On this figure, we showed the flight path and we can see the distance between some of the stations and the flight path. To estimate the soil moisture values under the path, we had to extrapolate the values measured by the automatic stations.

All SM *in situ* measurements were used to evaluate the ability of the model inversion to estimate surface SM.

3. RADIATIVE TRANSFERT MODEL

The radiative model we used to compute surface Tb values is the L-Meb model [7] described and used in many studies. This model represents the soil as a surface in contact with the atmosphere, and the vegetation as a homogeneous layer. This model computes soil emission based on the Dobson [3] or Mironov model, vegetation contribution to the surface Tb is computed based on the τ - ω model where the ω parameter is supposed to be close to zero. Then the only vegetation parameter is the optical thickness τ which, in this study was supposed not to be independent on the view angle. The general equation describing the effect of soil and vegetation on the surface Tb is:

$$Tb(\theta, P) = (1 - \omega_{\theta, P})(1 - \gamma_{\theta, P})1 + \Gamma_{\theta, P} \gamma_{\theta, P} \cdot T_{veg} + (1 - \Gamma_{\theta, P}) \gamma_{SM, P} T_{eff} \quad (1)$$

where T_{eff} and T_{veg} are the effective soil and vegetation temperatures, ω and γ are respectively the single scattering albedo and the transmissivity of the vegetation layer, varying with the polarization



Figure 1: Position of the SMOSmania stations in the South-West of France. Red line show the position of the Flight path.

P and the view angle θ . The transmissivity is linked to the optical thickness:

$$\gamma_{\theta,P} = \exp - \frac{\tau P}{\cos(\theta)} \quad (2)$$

The reflectivity of a rough soil, Γ is controlled by the soil roughness parameters. T_{eff} was estimated based on (i) the surface temperature (T_s) (ii) the in-depth physical temperature (T_{depth}), using the equation:

$$T_{eff} = T_{depth} + (T_s - T_{depth} \cdot (SM/wa))^{wb} \quad (3)$$

where $wa = 0.739$ and $wb = 0.2585$. We used airborne infra-red temperature measurements to estimate T_s . In-depth temperature were estimated using SMOSmania measurements at 50 cm depth extrapolated under the flight path (temperature variations are supposed to be weak).

Roughness parameter is also a major problem in the soil surface Tb modelisation. We made the assumption in this paper that the roughness parameter H depends on surface SM as described in [6]:

$$H = 1 - 1.13 \times SM \quad (4)$$

We finally have two unknown parameters to estimate: SM and the optical depth, τ . For that purpose we used non linear regression based on Gauss-Marquardt algorithm to minimize a cost function between the 4 simulated Tb values and the 4 measured ones (the method is described in [5]).

4. RESULTS OF SURFACE SOIL MOISTURE ESTIMATION

For each 4 Tb measurements we estimated the two SM and τ parameters. SM values were compared with two pools of *in situ* data: SM measured in SMOSmania stations and SM measured manually under the flight path. SMOSmania SM measurements were spatial interpolations of measurements mad at each each CAROLS sites. These results are presented in the next 2 figures: Figures 2 and 3.

In Figures 2, we present SM spatial variations as estimated by CAROLS during two flights compared to *in situ* measurements. We show that despite some strong SM variations, probably

due to heterogeneous landscapes, and a light underestimation in SM estimation, a general good agreement was found between these two ways of estimating SM. Temporal and Spatial variations are well mapped by CAROLS retrievals. In Figure 3, we compared local SM values to the mean remote-sensely retrieved SM value. Data concerning all flights where plotted on the same figure and good correlation is found: $R = 0.69$, the RMSE is $0.08 \text{ m}^3/\text{m}^3$.

Figure 4 shows estimated τ spatial variations for the flights 5 and 13. Values are between 0 and 0.2 in general, but in the Western part, it seems that this estimated value is larger, may be due to the presence of ‘Les Landes’ forest.

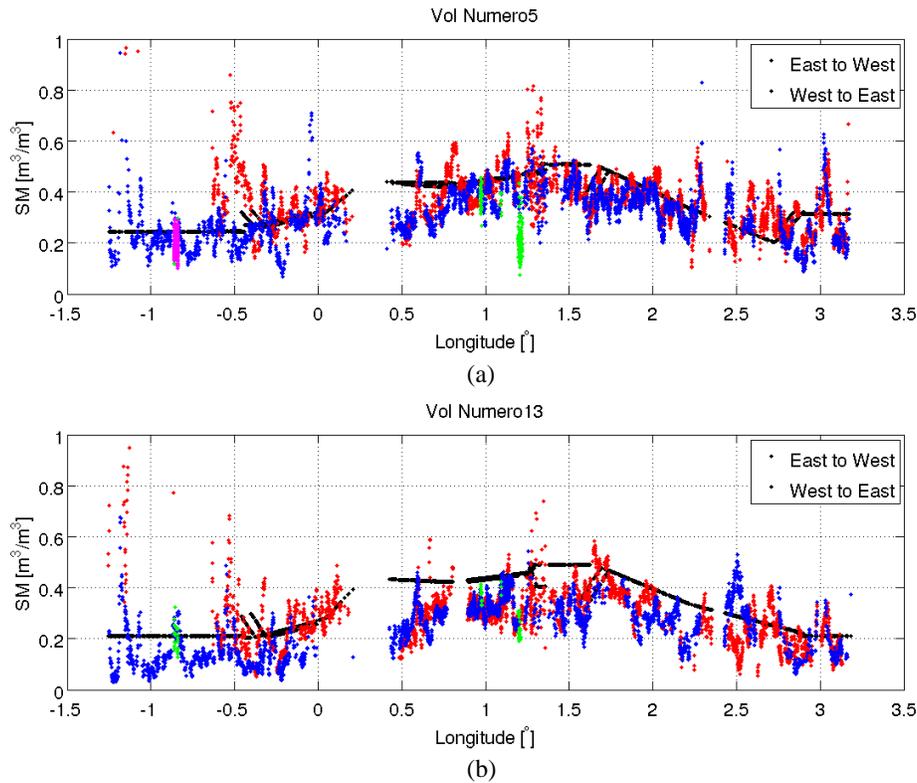


Figure 2: Comparison between SM spatial variation estimated with CAROLS and with in situ measurements, concerning flight 5 (a) and 13 (b). Red and Blue points are related to remote sensed SM, for out and home travel; Black points are interpolation of SMOSmania SM automatic measurements under the flight track; green and mauve are undertrack SM measurements.

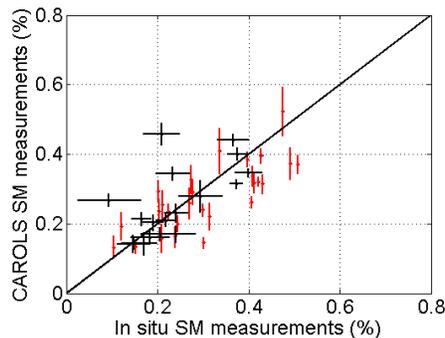


Figure 3: SM retrievals from CAROLS flight compared to local SM measurements. Black points concern SMOSmania measurements, red one concern handmade measurements under the flight track. Bars indicate the standard deviation of the field soil moisture and the retrieved soil moisture.

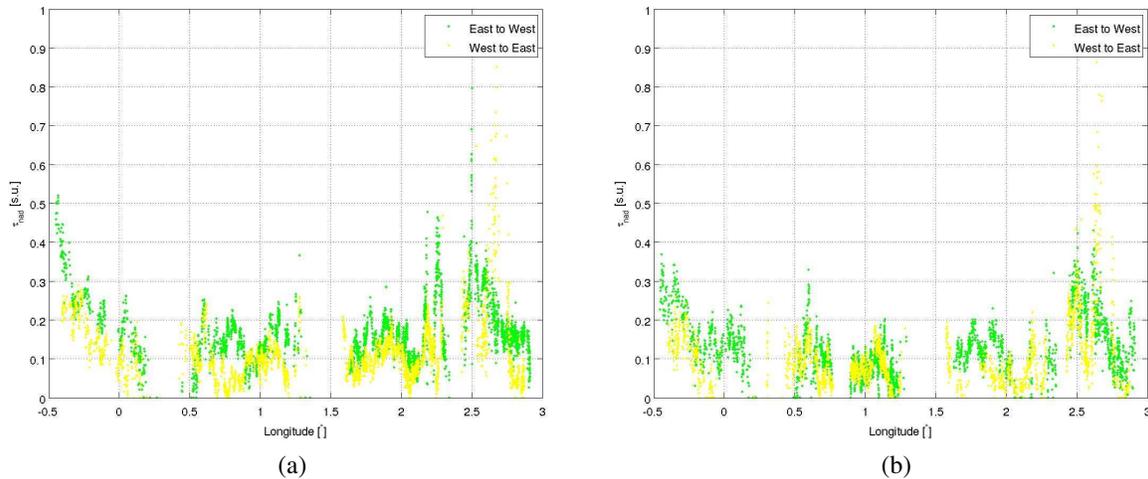


Figure 4: Vegetation optical thickness τ parameter variations with space for flight (a) 5 and (b) 13. The two colors are for the two parts of the flight (out and home).

5. DISCUSSION AND CONCLUSION

This paper investigated L-band measurements over forests and low crops obtained from an airborne L-band radiometer (CAROLS) during the CAROLS'09 campaign in South-West France. The main objective of the study was to evaluate the performance of the microwave radiative transfer model which sits at the core of ESA's SMOS soil moisture algorithm. The approach developed here is based on a two-parameter inversion of the radiative transfer model (L-MEB) using non linear regression using 4 brightness temperature measurements: two view angles and two polarizations. We compared SM estimations with *in situ* SM measurements and we showed a very good agreement between these two data sets. The next steps of this study are (i) correction of Tb measurements using DEM (ii) comparison between τ and remotely sensed vegetation parameters like LAI or NDVI.

REFERENCES

1. Rotbøll, J., S. S. Søbjaerg, and N. Skou, "Novel L-band polarimetric radiometer featuring subharmonic sampling," *Radio Science*, Vol. 38, No. 3, 1–7, 2003.
2. Calvet, J. C., N. Fritz, F. Froissard, D. Suquia, A. Petitpa, and B. Piguet, "In situ soil moisture observations for the CAL/VAL of SMOS: The SMOSMANIA network," *International Geoscience and Remote Sensing Symposium, IGARSS*, 1196–1199, Barcelona, Spain, July 23–28, 2007, doi:10.1109/IGARSS.2007.4423019.
3. Dobson, M. C., F. T. Ulaby, M. T. Hallikainen, and M. A. El-Rayes, "Microwave dielectric behavior of wet soil. Part II: Dielectric mixing models," *IEEE Trans. Geosci. Remote Sens.*, GE-23, 35–46, 1985.
4. Pardé, M., M. Zribi, P. Fanise, and M. Dechambre, "Analysis of RFI issue using the CAROLS L-band experiment," *IEEE Trans. Geosci. Remote Sens.*, accepted, 2010.
5. Pardé, M., J. P. Wigneron, A. Chanzy, P. Waldteufel, Y. Kerr, and S. Huet, "Retrieving surface soil moisture over a wheat field: Comparison of different methods," *Remote Sensing of the Environment*, Vol. 87, Vol. 2–3, 334–344, 2003.
6. Saleh, K., J.-P. Wigneron, P. de Rosnay, J.-C. Calvet, M.-J. Escorihuela, Y. Kerr, and P. Waldteufel, "Impact of rain interception by vegetation and mulch on the L-band emission of natural grass," *Remote Sensing of Environment*, Vol. 101, 127–139, 2006.
7. Wigneron, J. P., Y. Kerr, P. Waldteufel, K. Saleh, M. J. Escorihuela, P. Richaume, P. Ferrazzoli, P. de Rosnay, R. Gurney, J. C. Calvet, M. Guglielmetti, B. Hornbuckle, C. Mätzler, T. Pellarin, and M. Schwank, "L band microwave emission of the biosphere (LMEB) model: Description and calibration against experimental data sets over crop fields," *Remote Sensing of the Environment*, Vol. 107, 639–655, 2007.
8. Zribi, M., M. Pardé, J. Boutin, P. Fanise, D. Hauser, M. Dechambre, Y. Kerr, M. Leduc-Leballeur, G. Reverdin, N. Skou, S. S. Søbjaerg, C. Albergel, C. Calvet, J. P. Wigneron, E. Lopez-Baeza, K. Saleh, A. Ruis, and J. Tenerelli, "CAROLS: A new airborne L-Band radiometer for ocean surface and land observations," *IEEE Trans. Geosci. Remote Sens.*, submitted.

A Compact Single Feed, Low Cost Broadband Switched-beam Antenna for Mobile Wimax Applications

C. D. Nikolopoulos, C. I. Tsitouri, T. D. Dimousios, and C. N. Capsalis

Division of Information Transmission Systems and Material Technology
School of Electrical and Computer Engineering, National Technical University of Athens, Greece

Abstract— A broadband and low cost switched beam antenna suitable for WIMAX applications is presented in this paper. Present mobile terminal antennas are expected to show increased bandwidth as well as low-loss impedance matching while maintaining low-profile geometry. A custom Genetic Algorithm (GA) is used in order to optimize the proposed array's parameters regarding resonance frequency and radiation pattern. The optimized antenna exhibits small size (dimensions 12×12 cm), satisfactory directivity for mobile terminal applications of 6.2 dB over an operation bandwidth of 120 MHz around the central frequency of 3.5 GHz. Simulation results and parameters range are exposed below followed by the according radiation pattern and variation of the simulated VSWR.

1. INTRODUCTION

With the rapid growth of wireless mobile communication technology and the demand for broadband applications the world is looking a new, potentially disruptive, technology. The mobile WIMAX technology is an ideal mean for a new generation of mobile web applications that are being supplemented by a simultaneous shift in consumer behavior. The major change in the consumer behavior can be summarized as a strong move towards mobility. In that case more and more attention is being paid to the design of new handsets offering wider bandwidth, interference cancellation, multipath fading mitigation, direction of arrival destination etc [1, 2].

In this paper, a Cross-PIFA array is proposed for mobile-WIMAX application in the frequency of 3.5 GHz aiming to combine the Switched-Parasitic Arrays (SPAs) [3, 4], with the wideband and low profile Planar Inverted F Antennas (PIFAs).

2. DESIGN AND OPTIMIZATION OF THE CROSS PIFA

2.1. Architecture of the Proposed Structure

The proposed structure consists of a conductive top plate (four planar inverted F antennas-PIFA's connected forming a cross) lying over a finite sized ground plane which is interconnected through four wires (feeding wires-only one active at the time as explained later in this paper), and four other elements as shorting strips, in a configuration designed for symmetrical coverage of the horizontal plane (Fig. 1). The radiation characteristics and performance of the array can be adjusted by altering the dimensions of the structure. Also due to the structures symmetry, by electrically controlling the active (feeding wire) and passive (shorting strips) elements, the array's beam pattern may be steered through the azimuth plane.

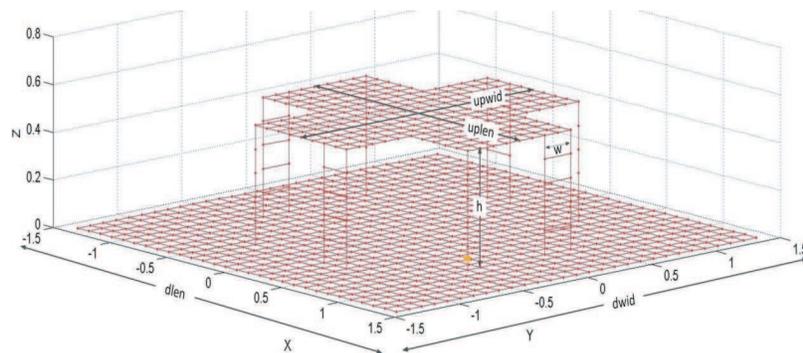


Figure 1: Architecture and dimensioning of Cross-Pifa.

It is noted that the switch between the active and parasitic state of the wire may be achieved using a p.i.n. diode and a two-state RF switches. When the diode is conducting, the corresponding

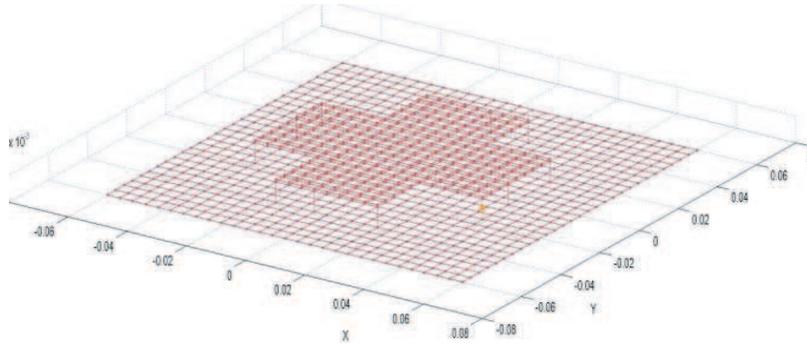


Figure 2: Implementation of the Switched-Beam array using the SNEC platform.

element is short-circuited and acts as a reflector regardless of the state of the RF switch. Else, in the case where the diode is not conducting, there are two cases: If the RF switch is open, the element is open circuited and is virtually transparent to the array, else (when the RF switch is closed) the element becomes active. Beam steering is achieved by appropriately setting the array's elements to active, parasitic or idle (transparent) state [5, 6].

2.2. Implementation and Analysis of the Cross-PIFA

The structure under consideration is depicted in Fig. 1. By altering the length of the active element (l_a), the length of the parasitic elements (l_p), the dimensions of top plate (uplen, upwid) (cross), and the dimensions of ground plate (dlen, dwid), a structure with adequate operational bandwidth and desired azimuth radiation patterns at this frequency band will be derived (Fig. 2). The structure is generated with the use of SNEC simulation software package and its performance is optimized via utilizing the GAs technique. This optimization tool is a stochastic search technique that uses the mechanisms of natural selection and genetics (crossover, mutation) to explore non-linear and discontinuous solution spaces [7]. SNEC is a hybrid MoM-UTD (Uniform Theory of Diffraction) antenna and electromagnetics simulation program. The MoM primitives available in the code are wire segments, whereas the UTD primitives supported are dielectrically coated plates and elliptical cylinders [7, 8]. The MoM is a numerical electromagnetic technique used to compute the radiation pattern and input impedance of wire-structured antennas [7]. For this reason, the ground plane, the top plates and the shorting strips of the Cross-PIFA were modeled as wire grid plates. In Fig. 2 the SNEC implementation of the cross-PIFA using wire segments is depicted. A GA is suitable for facing multi-parametric problems, such as the design of antennas, where a set of performance criteria, such as gain and SWR, should be satisfied [8]. In this paper, the GA module incorporated in SNEC is utilized.

The objective function is the driving force behind the GA. It is called from the GA to determine the fitness of each solution string generated during the search. As already being pointed out, the initial purpose is to offer a radiation pattern with certain characteristics (maximum gain at the direction of 0° , 3-dB beamwidth of 90° and relative sidelobe levels below -10 dB), while attaining input impedance at 3.5 GHz [9].

The objective function that satisfies the aforementioned demand is the following [7]:

- i. One set of 360 points is used to form the desired directivity pattern. Each point $D(\varphi)$ represents the desired normalized directivity pattern value at angular position φ . $D(\varphi)$ is formed with angular step of 1° and $D_{calculated}(\varphi)$ is the normalized directivity pattern calculated by the software. The peak gain of $D(\varphi)$ is obtained at $\varphi = 0^\circ$. Furthermore, $D(\varphi)$ exhibits a 3dB beamwidth of 90° and relative sidelobe level less than -10 dB. The first relative error term is:

$$e_1 = \frac{1}{360} \sum_{\phi=0^\circ}^{359^\circ} \left(\frac{D_{calculated}(\phi) - D(\phi)}{D(\phi)} \right)^2 \quad (1)$$

- ii. Input impedance matching is also required. So, a relevant error term is taken into consideration:

$$e_2 = abs \left(\frac{R_{in} - 50}{50} \right)^2 + abs \left(\frac{X_{in}}{50} \right)^2 \quad (2)$$

Table 1: Genetic Algorithm parameters range and results. (Wavelength equal to $\lambda_o = 0,0857$ m corresponding to an operating freq. of 3.5 GHz).

| Element | Range of Variation | Step | Results | Physical Dimentions |
|-------------------------------------|--|-----------------|-------------------------|---------------------|
| Length of top plate (uplen) | $0.15\lambda_o - 1.5\lambda_o$ | $0.05\lambda_o$ | $0.9\lambda_o$ | 7.7 cm |
| Width of top plate (upwid) | $0.15\lambda_o - 1.5\lambda_o$ | $0.05\lambda_o$ | $0.9\lambda_o$ | 7.7 cm |
| Length of ground plate (dlen) | $UpLen + 2 * (0.05\lambda_o - 0.5\lambda_o)$ | $0.05\lambda_o$ | $UpLen + 0.25\lambda_o$ | 12 cm |
| Width of ground plate (dwid) | $UpWid + 2 * (0.05\lambda_o - 0.5\lambda_o)$ | $0.05\lambda_o$ | $UpWid + 0.25\lambda_o$ | 12 cm |
| Height of wires/shorting strips (h) | $0.05\lambda_o - 0.2\lambda_o$ | $0.05\lambda_o$ | $0.05\lambda_o$ | 0.42 cm |
| Width of shorting strips (w) | $0.05\lambda_o - 0.4\lambda_o$ | $0.05\lambda_o$ | $0.25\lambda_o$ | 2.14 cm |

where R_{in} and X_{in} represent the real and imaginary part of the input impedance respectively and a characteristic impedance of 50Ω is taken into account.

- iii. The cumulative error is given by the summation of the terms described above after they have been multiplied by a properly selected weight factor each.

$$err = w_1e_1 + w_2e_2 \quad (3)$$

- iv. The fitness function is defined as:

$$OF = \frac{1}{1 + \sqrt{err}} \quad (4)$$

The simulation frequency was set to 3.5 GHz, since the spatial parameters of the antenna under examination are expressed in terms of λ_o , the electrical dimensions of the antenna remain constant and the derived data is suitable for application in other frequencies as well.

3. NUMERICAL RESULTS

A lot of GA runs took place for different values of the weight factors incorporated in the fitness function. Regarding the radiation pattern with the maximum gain at 0° best results were obtained when $W_1 = 1$ and $W_2=2$. Table 1 describes the variation of the parameters that took part in the GA optimization procedure and the respective radiation pattern is depicted in Fig. 2. The proposed array dimensions are expressed in terms of the number of segments. Each segment length was selected to be equal to $0.05 * \lambda$. The results of the optimization implementation are exhibited also in Table 1. The desired impedance bandwidth is determined by the band of frequencies where the value of the reflection coefficient at the feed point is less than -10 dB, corresponding to a VSWR with a value of no more than 2, when a characteristic impedance of 50Ω is considered. In Fig. 3, the variation of simulated VSWR around the frequency band of 3500 MHz is plotted.

As shown in the Fig. 3, the optimized array exhibits impedance matching at 3.4 GHz (VSWR = 1.51) and an extended operational bandwidth of 120 MHz compatible for mobile WIMAX applications. Also, in Fig. 2, it is depicted that the simulated array demonstrates a main beam towards 0° , with a 3 dB-beamwidth of 60° , a gain of 6.2 dB and a front to back ratio of 6 dB at 3.5 GHz. It should be noted that these values are maintained through the entire 120 MHz operating bandwidth of the proposed structure, thus making the achieved beamwidth, gain and front-to-back ratio quite satisfactory.

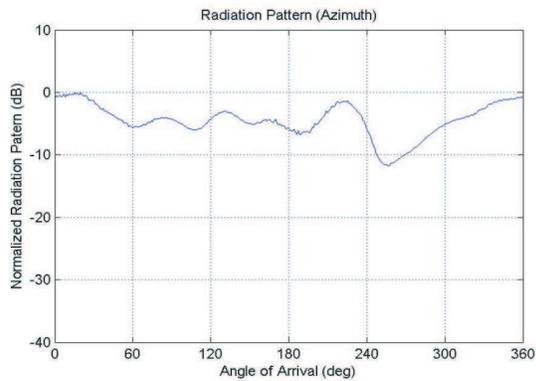


Figure 3: Radiation pattern of the antenna at xy plane at 3500 MHz

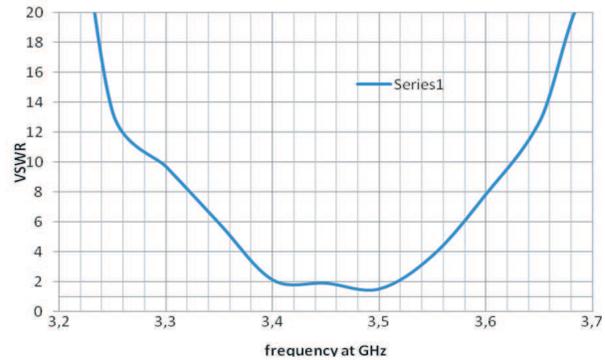


Figure 4: VSWR of the optimized switched-beam array around the frequency band of 3.5 GHz.

4. CONCLUSION

A low-profile Cross-Pifa was presented in this paper to add new perspectives in mobile terminals offering wider bandwidth compared to simple PIFA's and switched-beam capabilities with a significant directivity of 6.2 dB. The operational bandwidth performance was optimized using the technique of the genetic algorithms and the central aim in design is to broaden the bandwidth in limited height so that can be adjusted in any mobile-portable handsets.

REFERENCES

1. Waterhouse, R., *Printed Antennas for Wireless Communications*, John Wiley & Sons, Inc., 2007.
2. Kumar, A., *Mobile Broadcasting with Wimax*, Focal Press, 2008.
3. Kamarudin, M. R. B., P. S. Hall, F. Colombel, and M. Himdi, "Electronically switched beam disk-loaded monopole array antenna," *Progress In Electromagnetics Research*, Vol. 101, 339–347, 2010.
4. Dimousios, T. D., C. D. Nikolopoulos, S. A. Mitilineos, and C. N. Capsalis, "A new low-profile and cost SPA-PIFA for mobile 2.4 GHz ISM applications," *Journal of Electromagnetic Waves and Applications*, Vol. 24, No. 7, 881–891, 2010.
5. Preston, S. L., D. V. Thiel, T. A. Smith, S. G. O'Keefe, and J. W. Lu, "Basestation tracking in mobile communications using a switched parasitic antenna array," *IEEE Transactions on Antennas and Propagation*, Vol. 46, No. 6, 841–844, June 1998.
6. Mitilineos, S. A., C. A. Papagianni, G. I. Verikaki, and C. N. Capsalis, "Design of switched-beam planar arrays using the method of genetic algorithms," *Progress In Electromagnetic Research*, Vol. 46, 105–126, 2004.
7. "SuperNec v. 2.4 MOM technical reference manual," <http://www.supernec.com/manuals/snm-omtrm.htm>.
8. Fourie, A. and D. Nitch, "SuperNEC: Antenna and indoor-propagation simulation program," *IEEE Antennas and Propagat. Mag.*, Vol. 42, No. 3, 31–48, June 2000.
9. Rahmat-Samii, Y. and E. Michielssen, *Electromagnetic Optimization by Genetic Algorithms*, John Wiley & Sons, Inc., 1999.

Compact MIMO Microstrip Antenna with Defected Ground for Mutual Coupling Suppression

Fitri Yuli Zulkiffi and Eko Tjipto Rahardjo

Antenna Propagation and Microwave Research Group (AMRG)
Center for Information and Communication Engineering Research (CICER)
Electrical Engineering Department, Faculty of Engineering
Universitas Indonesia, Depok 16424, Indonesia

Abstract— A compact four triangular shape microstrip antenna with defected ground for multiple input multiple output (MIMO) systems is presented in this paper. The multi-feed four triangular patches are placed uniquely without any distance to each other; however, the mutual coupling reduction is achieved by adding slots on the antenna's ground.

The antenna is designed to operate at 3.3 GHz WiMAX application and studied experimentally regarding mutual coupling reduction. Details of the antenna design, simulated and measured results on the return loss and mutual coupling reduction of the proposed antenna are presented.

1. INTRODUCTION

The wireless multimedia technology has increased the capacity and reliability requirements of wireless communication systems. Therefore, the multiple input multiple output (MIMO) system has been one of the most suitable and promising technology for this, because it is well suited for high capacity, high data rate and high reliability requirements [1–3]. One main advantage of the MIMO system is it can improve capacity and reliability requirements, without increasing transmitted power or bandwidth [2]. MIMO system can also utilize several antennas for different diversity schemes. However multiple antennas bring drawbacks in the increasing size of the system and worsen the isolation between them, therefore cause distorted radiation pattern and decrease channel capacity.

Several works have been conducted to reduce the mutual coupling effect between antenna elements in MIMO systems. Some works have used spacing between antennas [1] and [4], slits added in the ground plane [3] and the ground plane is divided by slots [2].

In [1] and [3], the mutual coupling effect has been suppressed to -20 dB for a spacing of around $\lambda/4$, while for [3], by adding the slit, the mutual coupling can be suppressed also to -20 dB. Overall, the mutual coupling effect achieved in these recent works is to -20 dB, however these results were achieved by increasing the distance between elements. Therefore, in this paper a compact MIMO antenna without using any spacing between the elements is proposed. The antenna is designed to work at frequency 3.3 GHz for WiMAX application.

2. ANTENNA DESIGN

The antenna design proposed in this paper consists of four element triangular shape patches. The design of the single element patch is depicted in Fig. 1. The antenna consists of two layer substrate FR4 with dielectric constant $\epsilon_r = 4.4$ and substrate thickness of 1.56 cm. On the upper substrate, a triangular shape patch is designed. On the top of the lower substrate, a slot is designed and on the bottom of it, a microstrip line is designed as the feedline to the antenna, which is shown by dashed lines.

The feeding technique used is aperture coupling. The width of the $50\ \Omega$ microstrip line is 3.06 mm. The dimensions of the single patch is $W \times L = 30\ \text{mm} \times 30\ \text{mm}$, with the length of the triangle patch $a = 25\ \text{mm}$. This single element patch antenna is placed side by side (without no spacings between) with the same single element design to make a configuration of the compact MIMO 2×2 antenna. Therefore, the total dimension of the ground plane of the MIMO antenna is $60\ \text{mm} \times 60\ \text{mm}$.

This four element antenna design which has no spacing between each antenna, causes the impact of mutual coupling towards the antenna performance is no longer negligible. To suppress the mutual coupling effect and give isolation between each element, the slot on the ground is designed. The configuration of the MIMO antenna which is a compact four triangular shape patch microstrip antenna with defected ground is depicted in Fig. 2. Fig. 2(a) shows the top view, while Fig. 2(b) shows the exploded view of the antenna design. The dashed lines indicates the slot (which divides

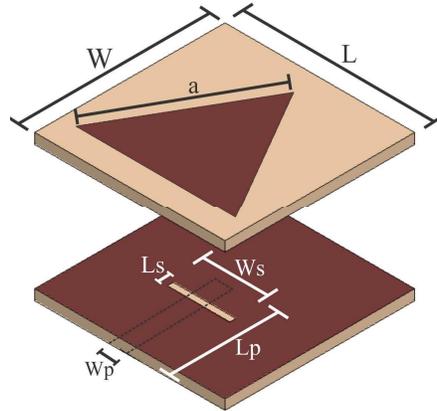


Figure 1: Exploded view of the single element patch antenna.

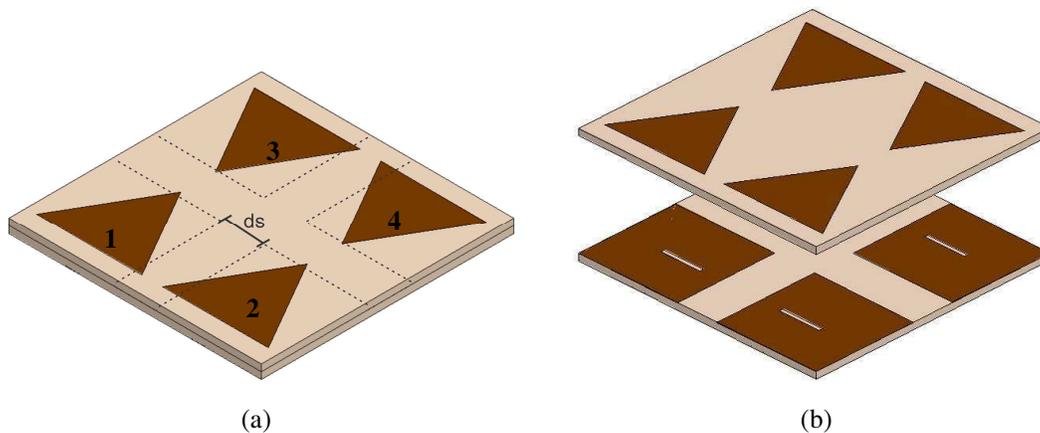


Figure 2: Multi-feed 4-element patch antenna design. (a) Top view. (b) Exploded view.

the ground) with width $ds = 10$ cm. This slot is added for the isolation between the four elements. In this paper, the impact of the width of the slot on the ground (ds) towards the mutual coupling effect is studied.

3. SIMULATION RESULTS

The antenna design was simulated using the commercial software Ansoft HFSS. Simulation results shown in Fig. 3 is observed by varying the width ds from no slots added ($ds = 0$ mm) to $ds = 10$ mm. Fig. 3(a) shows the simulated reflection coefficient of the antenna design with varied ds . It is shown that only slight shifts of resonant frequencies occurred. The impedance bandwidth of the proposed antenna ($S_{11} \leq -10$) is around 80 MHz–100 MHz.

Figure 3(b) shows the mutual coupling suppression due to the increase width (ds) of the slot on the ground plane. The mutual coupling between each antenna element can be obtained by measuring the isolation between each element [5]. Fig. 3(b) shows, as the width of the slot on the ground plane increases, the mutual coupling effect between the elements is suppressed.

By observing Fig. 3, it is shown that before adding the slot which divides the ground planes ($ds = 0$ mm), the mutual coupling effect S_{12} is -15.5168 dB for the frequency at 3.35 GHz. The impedance bandwidth is 80 MHz which occurs from 3.32–3.4 GHz.

After adding the slot which divides the ground plane, the most suppressed mutual coupling effect occurs when $ds = 9$ mm with $S_{12} = -21.9987$ dB. This is slightly better than when $ds = 10$ mm with $S_{12} = -21.789$ dB, however the bandwidth when $ds = 9$ mm is only 80 MHz from 3.29–3.37 GHz, while for $ds = 10$ mm is 100 MHz, from 3.3–3.4 GHz. Therefore, the proposed antenna uses $ds = 10$ mm, which has increased the isolations between the ports by about 6 dB. This design is further fabricated and measured.

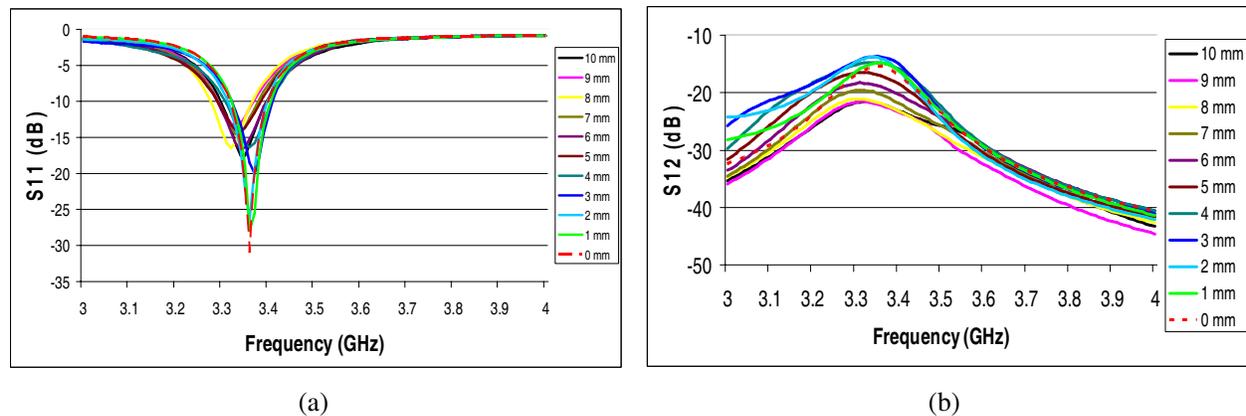


Figure 3: Simulation results with varied ds . (a) Coefficient reflection. (b) Mutual coupling S_{12} .

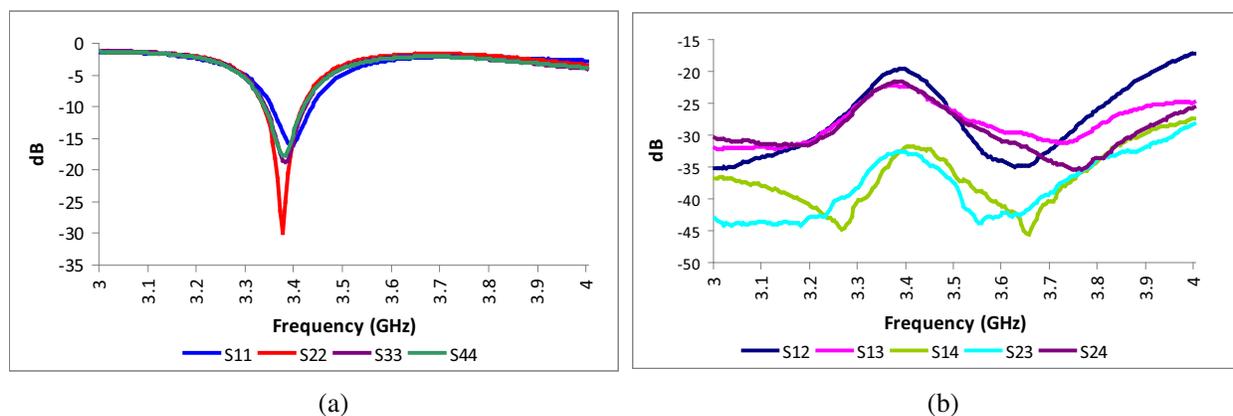


Figure 4: measurement results with $ds = 10$ mm. (a) Coefficient reflection. (b) Mutual coupling S_{12} .

4. MEASUREMENT RESULTS

The antenna design after fabrication is then measured in anechoic chamber at Electrical Engineering Department, Faculty of Engineering, Universitas Indonesia. The measurement result of the coefficient reflection is shown in Fig. 4(a), while the measured mutual coupling result is shown in Fig. 4(b).

The measurement shows in Fig. 4(a) that the measured S_{11} has a bandwidth from 3.35 GHz–3.435 GHz (85 MHz), for S_{22} from 3.335 GHz–3.415 GHz (80 MHz), for S_{33} from 3.34 GHz–3.42 GHz (80 MHz) and for S_{44} from 3.335 GHz–3.42 GHz (85 MHz). Therefore, the overall percentage bandwidth of the MIMO antenna is around 4%. The best impedance matching occurs at S_{22} with return loss of -29.939 dB.

For the mutual coupling, it can be seen from Fig. 4(b) that the mutual coupling effect at frequency 3.35 GHz for $S_{12} = -20.6672$ dB, for S_{13} is -22.4875 dB, for S_{14} is -35.5126 dB, for S_{24} is -22.4828 dB and for S_{23} is -34.6546 dB. The measurement result shows that the lowest isolation occurs at S_{14} and S_{23} because the ports are placed apart the farthest compared to the others.

From the measured mutual coupling results, the mutual coupling is relatively weak, although the antenna elements are placed relatively near to each other (without no spacing). The low coupling between the elements is caused of the small antenna elements and the aperture antenna is inherent in low coupling between adjacent elements. The slight difference between simulation and measurement results are due to the imperfect fabrication condition of the antenna.

Furthermore, both the E - and H -plane radiation pattern of the compact MIMO antenna for all of the ports was also measured as shown in Fig. 5. The H -plane shows good broadside radiation pattern while the E -plane shows directional radiation pattern.

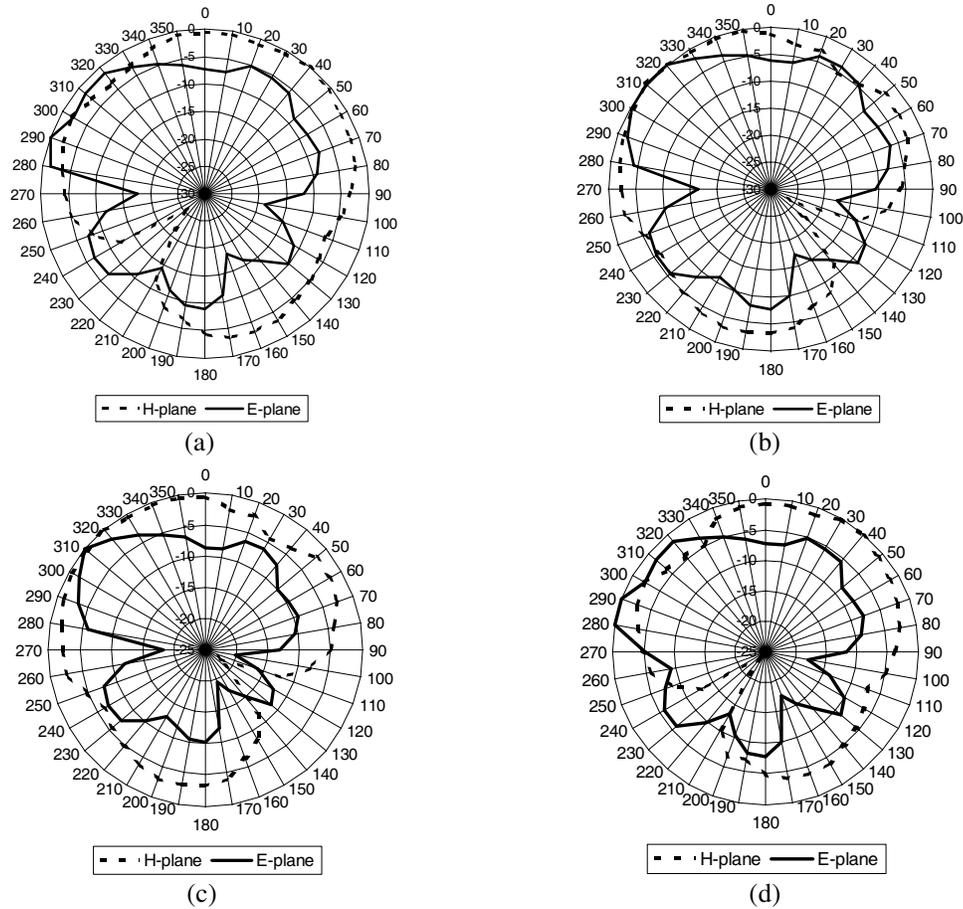


Figure 5: Measured radiation pattern of the MIMO antenna. (a) Port 1. (b) Port 2. (c) Port 3. and (d) Port 4.

5. CONCLUSIONS

A compact MIMO microstrip antenna with defected ground for WiMAX application at frequency 3.3 GHz is designed, fabricated and measured. Measurement result shows that the mutual coupling is suppressed to -20.6672 dB with impedance bandwidth of 85 MHz. The separation of the ground plane increased the isolations between the ports by about 6 dB.

ACKNOWLEDGMENT

This work is partly supported by The Ministry of Education under national strategic research 2010 with contract number: 2622/H2.R12/PPM.00.01 Sumber Pendanaan/2010.

REFERENCES

1. Fakhr, R. S., "Compact size and dual band semicircle shaped antenna for MIMO applications," *Progress In Electromagnetics Research C*, Vol. 11, 147–154, 2009.
2. Zhang, H., Z. Wang, J. Yu, and J. Huang, "A compact MIMO antenna for wireless communication," *IEEE Antennas and Propagation Magazine*, Vol. 50, No. 6, December 2008.
3. Li, H., J. Xiong, and S. He, "A compact planar MIMO antenna system of four elements with similar radiation characteristics and isolation structure," *IEEE Antennas and Wireless Propagation Letters*, Vol. 8, 1107–1110, 2009.
4. Chung, K. and J. H. Yoon, "Integrated MIMO antenna with high isolation characteristic," *Electron. Lett.*, Vol. 43, 199–201, 2007.
5. Karaboikis, M., C. Soras, G. Tsachtsiris, and V. Makios, "Four-element printed monopole antenna systems for diversity and MIMO terminal devices," *Proceedings of 17th International Conference on Applied Electromagnetics and Communications*, Croatia, October 1–3, 2003.

Performance and Capacity Analysis of Compact MIMO Aided OFDM-SDMA Systems

Mostafa Hefnawi

Department of Electrical and Computer Engineering, Royal Military College of Canada, Kingston, Canada

Abstract— This paper investigates antenna array coupling effects on the symbol error rate (SER) performance and system capacity of compact MIMO aided multiuser OFDM-SDMA systems. The results show that at close antenna separation the mutual coupling degrades the SER performance and reduces the system capacity.

1. INTRODUCTION

Multiple Input Multiple Output Orthogonal Frequency Division Multiplexing (MIMO-OFDM) system with Space Division Multiple Access (SDMA) has recently attracted significant interests among researchers and developers because of its potentials for realizing spectrally-efficient wireless services at low complexity/costs [1–7]. This costeffectiveness aspect presents a strong case for the deployment of MIMO OFDM-SDMA technologies in upcoming wireless systems, such as the WiMAX, and the upcoming 4G networks [8, 9]. These technologies will be used in handheld devices such as cellphones or PDAs where the separation between the antenna elements becomes small and the effect of mutual coupling between the antennas becomes important, and thus it should be taken into account in addition to multipath fading and multiuser interference. This effect has been widely studied in conventional antenna array and MIMO systems [10, 13] but has not previously been investigated in OFDM-SDMA systems incorporating multiuser detection algorithms such as maximum ratio combining (MRC) combining, minimum mean-squared error (MMSE), successive interference cancelation (SIC), parallel interference cancellation (PIC), and maximum likelihood sequence estimation/ maximum likelihood detection (MLSE/MLD [1–6]. MMSE detection scheme [3], though slightly inferior to the MLD, is known to exhibit very good error rate performance at low complexity, and has thus been widely deployed. Therefore, the focus in this paper is on MMSE based OFDM-SDMA systems where the performance degradation due to antenna mutual coupling, multipath fading and multi-access interference (MAI), are taken into account. We consider an uplink OFDM-SDMA system with multiuser access interference, where multiple antenna-aided base station (BS) receiver detects independent OFDM data streams from multiple SDMA users simultaneously on the same frequency band. Each of the users is equipped with antenna array as well for its data transmission. The transmitter (T_x) employs eigen beamforming weights, while the receiver (R_x) computes the receive weights using the MMSE criterion. Using Monte-Carlo simulations, we evaluate the impact of antenna mutual coupling on the SER performance and system capacity of the MMSE based MIMO OFDM-SDMA systems over Rayleigh fading environment.

2. SYSTEM MODEL AND ANALYSIS

Figure 1 displays the model we assumed for the uplink scenario. Let $\mathbf{x}[k] = \{x[k]_1, x[k]_2, \dots, x[k]_L\}$ denote the set of L user signals transmitted on each sub-carrier, k , $k = 1, \dots, N_c$, where N_c denotes the number of subcarriers per OFDM symbol in the system. The expression for the array output in Fig. 1 can be written for each sub-carrier as

$$\mathbf{y}[k] = \sum_{i=1}^L \mathbf{H}[k]_i \mathbf{w}[k]_i^t x[k]_i + \mathbf{n}[k] \quad (1)$$

where $\mathbf{y}[k] = [y[k]_1, y[k]_2, \dots, y[k]_{N_r}]^T$ is the $N_r \times 1$ vector containing the outputs of the N_r -element array at the BS, with T denoting the transpose operation, and $\mathbf{H}[k]_i$ is the $N_r \times N_t$ channel matrix representing the frequency-domain fading coefficients from user i 's N_t -element antenna array to the BS's N_r -element antenna array, on each subcarrier. $\mathbf{w}[k]_l^t = [w[k]_{l1}^t, w[k]_{l2}^t, \dots, w[k]_{lN_t}^t]^T$ is the complex transmit weighting vector at the transmitter for user l , $l = 1, \dots, L$, and $\mathbf{n}[k] = [n[k]_1, n[k]_2, \dots, n[k]_{N_r}]^T$ is the $N_r \times 1$ complex additive white Gaussian noise (AWGN) vector. The transfer functions from the users' devices to the BS antenna array (the cascade of $\mathbf{H}[k]_i$ and $\mathbf{w}[k]_i^t$), result in a unique spatial signature for each user, which can be exploited to effect the

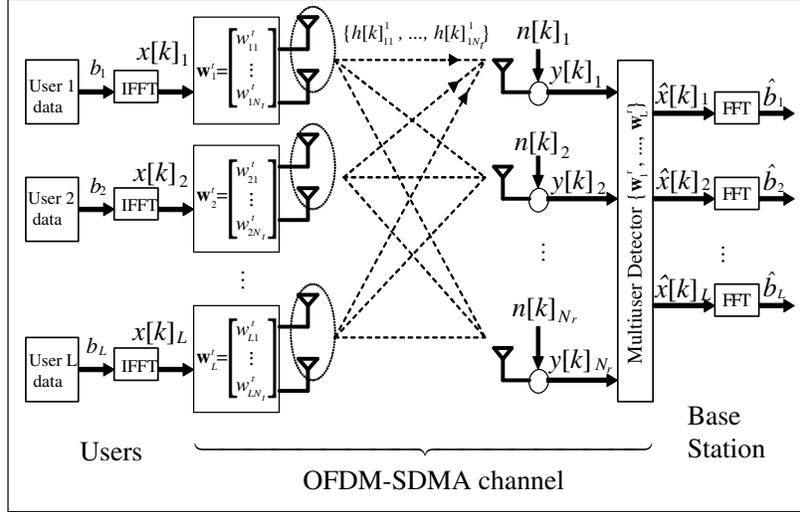


Figure 1: Multiple antenna aided OFDM-SDMA system.

separation of the user data at the BS using appropriate multiuser detection techniques analogous to the CDMA system. In the linear detection technique, the BS detects all L users simultaneously at the multiuser detection module of the OFDM-SDMA system, by multiplying the output of the array with the respective receiving weight vectors for each user l , $\mathbf{w}[k]_l^r = [w[k]_{l1}^r, w[k]_{l2}^r, \dots, w[k]_{lN_r}^r]^T$, $l = 1, \dots, L$. The detection of user l out of L users (with $L-1$ interfering users) can thus be depicted as

$$\hat{x}[k]_l = \mathbf{w}[k]_l^{rH} \mathbf{y}[k] = S_d + S_I + N \quad (2)$$

where $S_d = \mathbf{w}[k]_l^{rH} \mathbf{H}[k]_l \mathbf{w}[k]_l^t x[k]_l$ is the desired signal for the detection of user l 's signal on subcarrier k , $S_I = \mathbf{w}[k]_l^{rH} \sum_{i=1, i \neq l}^L \mathbf{H}[k]_i \mathbf{w}[k]_i^t x[k]_i$ is the multiple-access interference (MAI) contributed by the $L-1$ other users sharing the same broadband channel with the desired user, and $N = \mathbf{w}[k]_l^{rH} \mathbf{n}[k]$ is the noise signal.

2.1. MIMO-MMSE in OFDM-SDMA Systems

At the transmitter side of MIMO-MMSE scheme, eigen beamforming is used, where the transmit beamforming vector for the l th user $\mathbf{w}[k]_l^t$ is given by

$$\mathbf{w}[k]_l^t = \sqrt{P_l} \mathbf{u}_{\max, l} \quad (3)$$

where P_l is the total power transmitted by user l in a subcarrier and $\mathbf{u}_{\max, l}$ denotes the eigen vector corresponding to $\lambda_{\max, l}$, the maximum eigen value of $\mathbf{H}[k]_l^H \mathbf{H}[k]_l$ on each subcarrier k .

At the receiver side, a linear filter with coefficient $\mathbf{w}[k]_l^{r, mmse} = [w[k]_{l1}^{r, mmse}, w[k]_{l2}^{r, mmse}, \dots, w[k]_{lN_r}^{r, mmse}]^T$ combines the received signals, $\mathbf{y}[k]$, to obtain soft estimates

$$\hat{x}[k] = \mathbf{w}[k]_{MMSE}^H \mathbf{y}[k] \quad (4)$$

where $\hat{\mathbf{x}}[k] = [\hat{x}[k]_1, \hat{x}[k]_2, \dots, \hat{x}[k]_L]$ is the estimate of the transmitted user data $x[k]$ at the output of the multiuser detector module in Fig. 1, and $\mathbf{W}[k]_{MMSE} = [\mathbf{w}[k]_1^{r, mmse}, \mathbf{w}[k]_2^{r, mmse}, \dots, \mathbf{w}[k]_L^{r, mmse}]^T$ is the matrix whose l th column consists of the MMSE weights (filter coefficients) for user l , $l = 1, 2, \dots, L$. The receive weight vector for the l th user in the MMSE solution, $\mathbf{w}[k]_l^{r, mmse}$, is chosen as the unique vector minimizing the mean squared error (MSE) for the l th user as

$$MSE = E [(x[k]_l - \hat{x}[k]_l)^* (x[k]_l - \hat{x}[k]_l)] \quad (5)$$

For a given noise power σ_n^2 , and with $E[x[k]_l^* x[k]_l] = 1$, $l = 1, 2, \dots, L$, this means that $\mathbf{W}[k]_{MMSE}$ has to obey the L sets of linear equations as [1]

$$[\mathbf{H}(k)\mathbf{H}(k)^H + \sigma_n^2 \mathbf{I}] \mathbf{W}[k]_{MMSE} - \mathbf{H}(k) = 0 \quad (6)$$

yielding the MMSE weights for all users as

$$\mathbf{W}[k]_{MMSE} = [\mathbf{H}[k]\mathbf{H}[k]^H + \sigma_n^2\mathbf{I}]^{-1} \mathbf{H}[k] \quad (7)$$

The receive weights for the MIMO-MMSE solution for the l th user in the presence of OFDM-SDMA multiuser access interference, are given by

$$\mathbf{w}[k]_l^{r,mmse} = [\mathbf{H}[k]\mathbf{H}[k]^H + \sigma_n^2\mathbf{I}]^{-1} \mathbf{H}[k]\mathbf{w}[k]_l^t, \quad (8)$$

with σ_n^2 denoting the variance of the AWGN.

2.2. Symbol Error Rate (SER) Performance of OFDM-SDMA Systems

- *SINR Expression*: The signal-to-interference-plus-noise ratio (SINR), $\gamma[k]_l$, for user l , is given by

$$\gamma[k]_l = \frac{\|\mathbf{w}[k]_l^{r,mmse} \mathbf{H}[k]_l \mathbf{w}[k]_l^t x[k]_l\|^2}{\mathbf{w}[k]_l^{r,mmse} \mathbf{B}[k]_l \mathbf{w}[k]_l^t} \quad (9)$$

where $\mathbf{B}[k]_l$ is the covariance matrix of the interference-plusnoise, given by

$$\mathbf{B}[k]_l = \sum_{i=1, i \neq l}^L \mathbf{H}[k]_i \mathbf{w}[k]_i^t x[k]_i (x[k]_i \mathbf{w}[k]_i^t)^H \mathbf{H}[k]_i^H + \sigma_n^2 \mathbf{I}_{N_r} \quad (10)$$

- *Probability of Error Expression*: Let $P_r(E, k)_l$ denotes the probability of error associated with the k th subcarrier for user l , $l = 1, \dots, L$. In this analysis, we consider modulation formats for which $P_r(E, k)_l$ can be expressed as [6]

$$P_r(E, k)_l = E_{\gamma[k]_l} \left[aQ\sqrt{2b\gamma[k]_l} \right] \quad (11)$$

where $E[\cdot]$ denotes the expectation operator, $Q(\cdot)$ denotes the Gaussian Q -function, and a and b , are modulation-specific constants. For binary phase shift keying (BPSK), $a = 1$, $b = 1$, for binary frequency shift keying (BFSK) with orthogonal signaling $a = 1$, $b = 0.5$, while for M -ary phase shift keying (M-PSK) $a = 2$, $b = \sin^2(\pi/M)$. The average SER performance for user l , can be estimated as [6]

$$SER_l = \frac{1}{N_c} \sum_{k=0}^{N_c-1} P_r(E, k)_l \quad (12)$$

2.3. OFDM-SDMA Channel Capacity

Assuming equal power allocation to users, with total power transmitted by the l th SDMA user in an OFDM subcarrier equal to P , and assuming a large number of users so that the interference can be considered a Gaussian process, then the Ergodic capacity of the OFDM-SDMA channel, per subcarrier, for each user l is given by [7]

$$C(\mathbf{H}[k]_l, \mathbf{w}[k]_l^t) = E \left[\log_2 \left\{ 1 + \frac{P}{N_t} \left| \mathbf{B}[k]_l^{-\frac{1}{2}} \mathbf{H}[k]_l \mathbf{w}[k]_l^t \right|^2 \right\} \right] \quad (13)$$

where $l = 1, \dots, L$.

2.4. Mutual Coupling

In order to include the effect of mutual coupling, the channel matrix $\mathbf{H}[k]_l$ can be modified as follows

$$\mathbf{H}_C[k]_l = C_r \mathbf{H}[k]_l C_t$$

where $\mathbf{H}_C[k]_l$ is the new channel matrix, C_r is the $N_r \times N_r$ transmitter coupling matrix, and C_t is the $N_t \times N_t$ receiver coupling matrix. C_t and C_r can be written using fundamental electromagnetic and circuit theory [11] as

$$C_{r,t} = (Z_A + Z_T)(\mathbf{Z} + Z_T \mathbf{I}_N)^{-1} \quad (14)$$

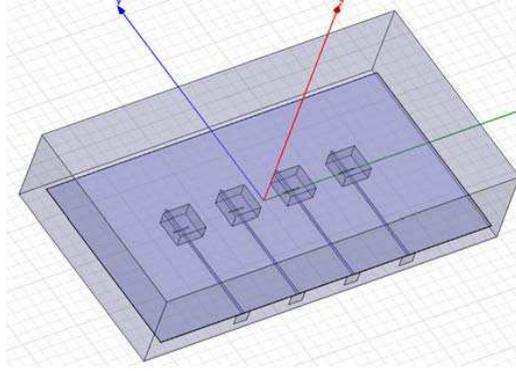
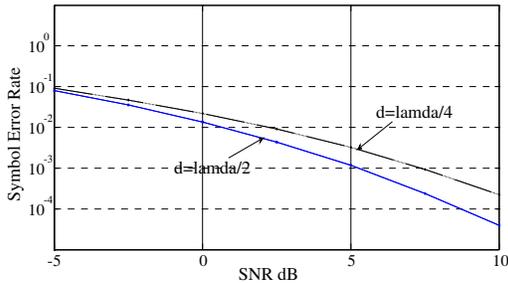
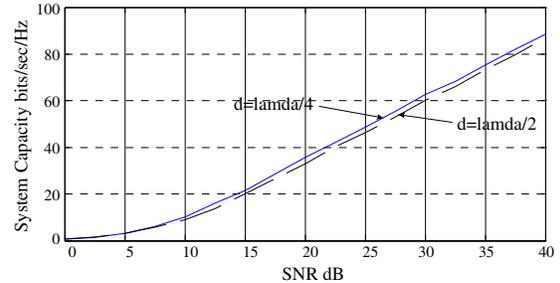


Figure 2: Four-element DRA array using HFSS.

Figure 3: SER of OFDM-SDMA systems for 4×4 MIMO-MMSE with two cases of antenna spacing.Figure 4: Capacity of OFDM-SDMA system for two cases of antenna spacing: $d = 0.5\lambda$ and $d = 0.25\lambda$.

where Z_A is the element's impedance in isolation, Z_T is the impedance at each antenna termination and is chosen as the complex conjugate of Z_A to obtain an impedance match for maximum power transfer, \mathbf{I}_N is the identity matrix, and \mathbf{Z} is an $n \times n$ mutual impedance matrix given by

$$\mathbf{Z} = \begin{bmatrix} Z_A + Z_T & Z_{12} & \dots & Z_{1n} \\ Z_{21} & Z_A + Z_T & \dots & Z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{n1} & Z_{n2} & \dots & Z_A + Z_T \end{bmatrix} \quad (15)$$

In general, commercial software packages such as HFSS and CST based on different numerical techniques can be used to obtain the mutual impedance matrix \mathbf{Z} .

3. SIMULATION RESULTS

In our simulation setup we consider an OFDM-SDMA systems with: $N_t = N_r = 4$, $L = 2$ users. We used $N_c = 256$ in the simulations and we assumed BPSK modulation for all the subcarriers. We consider two cases of inter-element spacing, $d = 0.5$ and $d = 0.25$ wavelengths apart. HFSS was used to obtain the mutual impedance matrix of a four-element dielectric resonator antenna (DRA) array fed by a microstripslot transition as shown in Fig. 2.

Figure 3 compares the SER performance of the 4×4 MIMOMMSE scheme for the two cases of inter-element spacing. It is noted that the SER at $d = \lambda/2$ outperforms the BER at $d = \lambda/4$, especially at high SNR. Thus, when the separation between the antenna elements becomes small the effect of mutual coupling becomes important.

Figure 4 displays the system capacity for both cases, $d = \lambda/2$ and $d = \lambda/4$. It is shown that at high SNR (≥ 15 dB), at least 3 bits/sec/Hz of capacity impairment is observed with $d = 0.25\lambda$ compared to $d = 0.5\lambda$. Although this impairment seems negligible, it might be of significant impact in broadband wireless services like the WiMAX.

4. CONCLUSION

This paper investigated the impact of mutual coupling between antenna elements on the SER performance and the capacity of the MIMO-MMSE beamforming scheme for OFDM-SDMA systems in presence of multiuser access interference. We found that at close antenna separation the

mutual coupling degrades the SER performance and reduces the system capacity. Therefore, the consideration of mutual coupling is indispensable in the analysis, design and implementation of MIMO-beamforming in OFDM-SDMA systems.

REFERENCES

1. Vandenameele, P., L. Der Perre, M. Engels, B. Gyselinckx, and H. De Man, "A combined OFDM/SDMA approach," *IEEE J. Sel. Areas Commun.*, Vol. 18, No. 11, 2312–2321, November 2000.
2. Jiang, M., S. Ng, and L. Hanzo, "Hybrid iterative multiuser detection for channel coded space division multiple access OFDM systems," *IEEE Transactions on Vehicular Technology*, Vol. 55, No. 1, January 2006.
3. Thoen, S., L. Deneire, L. V. Der Perre, M. Engels, and H. De Man, "Constrained least squares detector for OFDM/SDMA-based wireless networks," *IEEE Trans. Wireless Commun.*, January 2003.
4. Munster, M. and L. Hanzo, "Performance of SDMA multi-user detection techniques for walsh-hadamard-spread OFDM schemes," *IEEE-VTC'01*, Vol. 4, 2319–2323, October 2001.
5. Wong, K.-K., R. Cheng, K. B. Letaief, and R. D. Murch, "Adaptive antennas at the mobile and base stations in an OFDM/TDMA system," *IEEE Trans. Commun.*, Vol. 49, No. 1, January 2001.
6. Sulyman, A. I. and M. Hefnawi, "Adaptive MIMO beamforming algorithm based on gradient search of the channel capacity in OFDM-SDMA systems," *IEEE Commun. Lett.*, September 2008.
7. Blum, R. S., J. H. Winters, and N. R. Sollenberger, "On the capacity of cellular systems with MIMO," *IEEE Commun. Lett.*, Vol. 6, No. 6, 242–244, June 2002.
8. Fujitsu Microelectronics joins Cisco to demonstrate first multi-vendor interoperability using mobile WiMAX with beamforming, MIMO, WiMAX world exhibition, Chicago, September 30–October 2, 2008 [information available at: <http://us.fujitsu.com/micro>].
9. "Beamforming boosts the range and capacity of WiMAX networks," *White paper by Fujitsu Microelectronics America, Inc.*, July 2008 [available at: <http://us.fujitsu.com/micro>].
10. Yuan, H., K. Hirasawa, and Y. Zhang, "The mutual coupling and diffraction effects on the performance of a CMA adaptive array," *IEEE Transactions on Vehicular Technology*, Vol. 47, No. 3, August 1998.
11. Svantesson, T., "Modeling and estimation of mutual coupling in a uniform linear array of dipoles," *Proc. ICASSP'99*, 2961–2964, Phoenix, AZ, March 1999.
12. Wyglinski, A. M. and S. D. Blostein, "On uplink CDMA cell capacity: Mutual coupling and scattering effects on beamforming," *IEEE Transactions on Vehicular Technology*, Vol. 52, No. 2, March 2003.
13. Durrani, S. and M. E. Bialkowski, "Effect of mutual coupling on the interference rejection capabilities of linear and circular arrays in CDMA systems," *IEEE Transactions on Antennas and Propagation*, Vol. 52, No. 4, April 2004.

Wide-Band Rectangular Dielectric Resonator Antenna for Wireless Applications

Achraf Jaoujal, Noura Aknin, and Ahmed El Moussaoui

Information and Telecommunications Systems Laboratory, Faculty of Sciences
Abdelmalek Essaadi University, Tetouan, Morocco

Abstract— In this letter, we proposed a dual-band rectangular DRA using Alumina (Al_2O_3 , $\epsilon_r = 9.8$) as radiating element and fed by a simple network technique, named the strip-fed method (coax + conducting strip). As a result, this work presents the characteristics of a new dual-band antenna operating at 2.879 (2.81–3.09 GHz)/3.506–3.941 (3.09–4.69 GHz) GHz providing higher gain values (more than 8 dBi), and a fractional bandwidth of 50.19%.

1. INTRODUCTION

In the last decade, the demand of antennas miniaturization and the wide band had increased exponentially for that many antenna engineers have been tried to design compact size antennas with wide bandwidth performance. The DRA is a good candidate for these requirements. DRAs was proposed for the first time by the professor S. A. Long in 1983 [1], and it has received increasing attention in the last two decades. The DRA has many common advantages to micro strip antennas (MSA) such small size, low cost, low loss, light weight, and ease of excitation. Additionally, DRAs have wider bandwidth ($\sim 10\%$ for dielectric constant ~ 10) than conventional MSA [2] since they do not use metallic radiators. As consequence, DRAs radiation efficiency (more than 98%) [5], radiation pattern performances and antennas directivity become higher compared to those of microstrip antennas operating in the same frequencies in millimeter wave band [6] or in microwave range [7].

A lot of works were carried out to achieve wide bandwidth enhancements using different techniques and different geometries such as conical [8], elliptical [9,10], tetrahedral [11], well [12], stair [13] and H-shaped [14]. In this work we present a new shape which provides a UWB ($> 25\%$). The studded structure consists in a rectangular DRA which contains two symmetrical and rectangular gaps, and fed by a simple network technique named the strip-fed method (coax + conducting strip) [3–15]. The analyses in this paper is performed using the commercial codes HFSS and CST [16,17].

2. ANTENNA DESIGN

Figure 1 shows the evolution of the studded DRA. Fig. 1(a) shows the rectangular DRA excited by a simple network technique named the strip-fed method (coax + conducting strip) [3–15]. At the back face we have placed an identical strip in front of the first one (Fig. 1(b)) to excite more than mode [15]. Fig. 1(c) shows the rectangular DRA with two symmetrical and rectangular gaps the first in the left and the second in the right side. Adding gaps to the DRA provides a wider bandwidth and a good gain compared to the structure in Fig. 1(a).

The dimensions of the proposed antenna are 14.3 mm width, 34 mm length, and 15 mm height with a dielectric constant of 9.8, and it is supported by a $104.3 \times 124 \text{ mm}^2$ ground plane. The rest

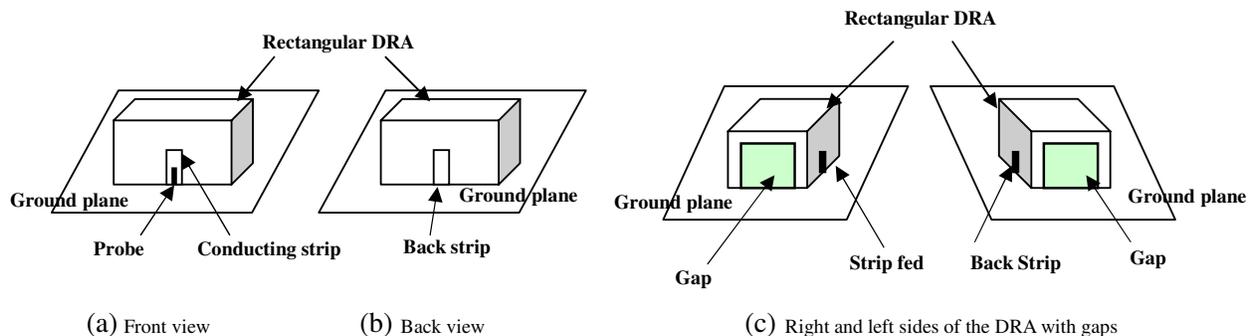


Figure 1: Evolution of the studded DRA.

of the antenna parameters are shown in Fig. 2 with $W_D = 14,3$ mm, $L_D = 34$ mm, $H_D = 15$ mm, $g = 45$ mm, $W = 1,15$ mm and $H_{gap} = 11$ mm.

3. RESULTS

Figure 3 compares reflection coefficients of the rectangular DRA and the modified DRA, which has two rectangular gaps in right and left sides. While the first rectangular DRA model provides an 17.71% impedance matching bandwidth, the DRA with two rectangular gaps provides much wider impedance matching bandwidth of around 50.19%.

In the Figure 4, we can find a comparison between the reflection coefficients of the modified rectangular DRA simulated in the CST and HFSS simulators.

So, we can notice that the result is verified for the two simulators whose are based in two different methods (HFSS in finite element method and CST in FDTD method). The simulated matching frequency band of the proposed antenna for -10 dB reflection coefficient is from 2.81 GHz to 3.09 GHz and from 3.09 GHz to 4.69 GHz with a total fractional bandwidth of 50.19%. The obtained gains are 9.04 dBi at 2.879 GHz, 8.61 dBi at 3.506 GHz and 6.90 at 3.941 GHz.

Figures 3, 5 and 7 represent the radiation patterns of the radiation patterns of the proposed antenna in the x - y , y - z and x - z planes at 3.06, 2.879 and 3.941 GHz respectively. As shown in

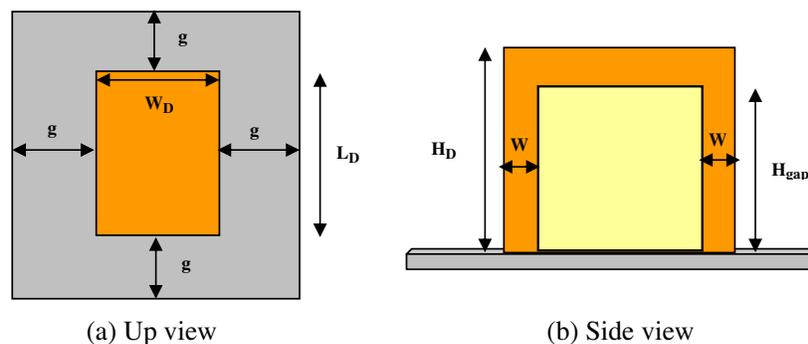


Figure 2: Geometry of the rectangular DRA, (a) up and (b) side view.

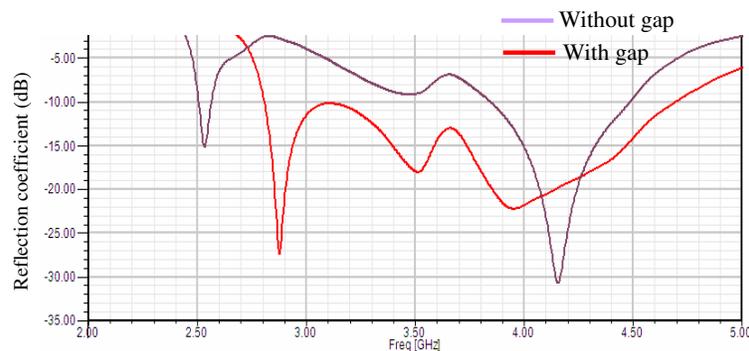


Figure 3: Reflection coefficients for rectangular DRA with and without gaps.

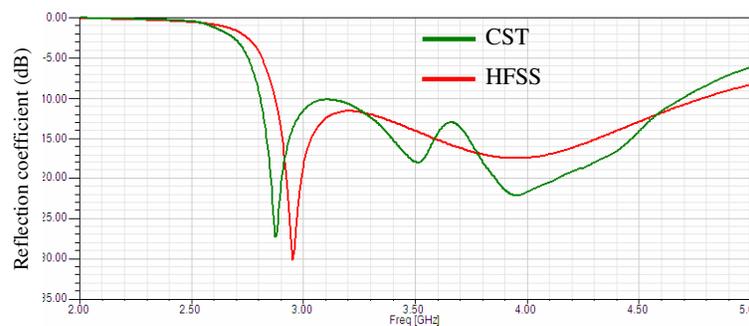


Figure 4: CST and HFSS simulated impedance matching bandwidth.

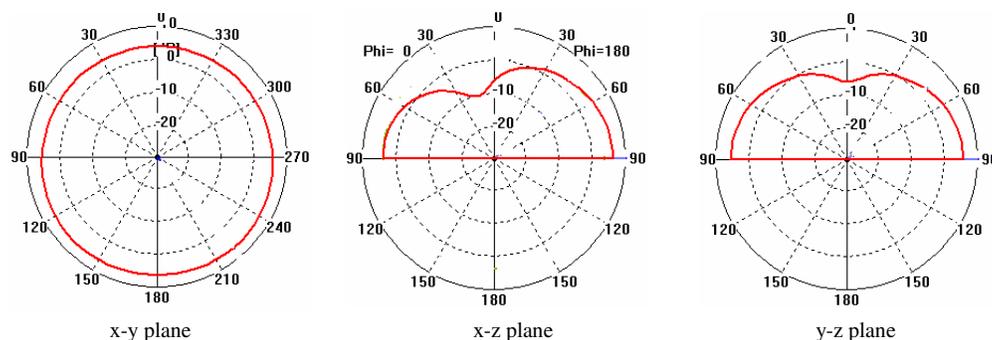


Figure 5: Simulated radiation patterns of the proposed antenna at 3.506 GHz.

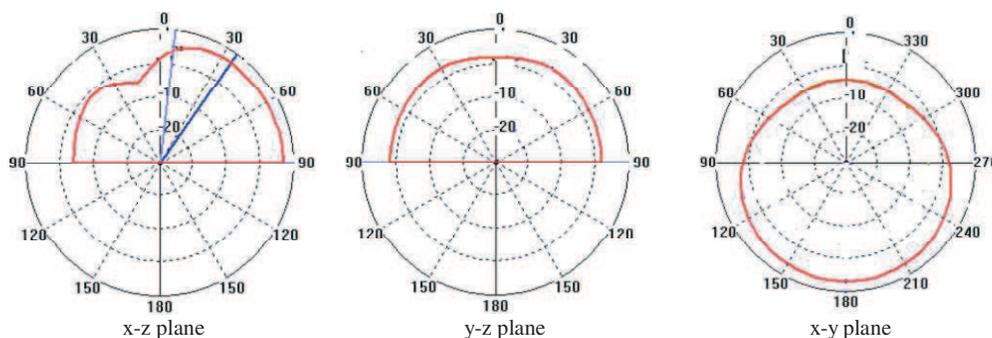


Figure 6: Simulated radiation patterns of the proposed antenna at 2.879 GHz.

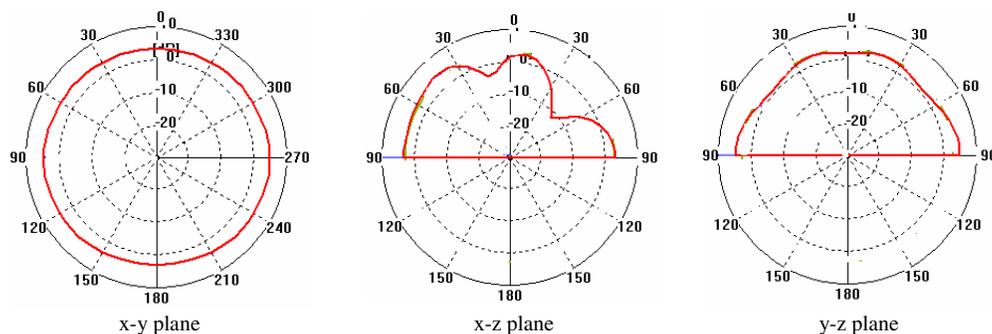


Figure 7: Simulated radiation patterns of the proposed antenna at 3.941 GHz.

In the figures above, the radiation patterns at 3.06, 2.879 and 3.941 GHz in the $x-y$ plane are very similar to those of the monopole antenna, while in the $y-z$ plane and $x-z$ the radiation patterns are directional.

4. CONCLUSION

A rectangular DRA with two lateral gaps and excited by a simple feeding network, named the strip-fed method was presented. The structure provided height bandwidth as compared to the simple rectangular DRA. We presented the return losses coefficient and the radiation patterns of the proposed antenna structure. As a result, this antenna has two bands, a narrow bandwidth between 2.81 and 3.09 GHz and a wide bandwidth between 3.09 and 4.69 GHz having a fractional impedance bandwidth of 50.19%. The maximum radiation gains obtained are 9.04 dBi at 2.879 GHz and 8.61 dBi at 3.506 GHz.

REFERENCES

1. Long, S. A., M. W. McAllister, and L. C. Shen, "The resonant cylindrical dielectric cavity antenna," *IEEE Trans. Antennas Propag.*, Vol. 31, No. 5, 406–412, May 1983.
2. Kishk, A. A., "Experimental study of broadband embedded dielectric resonator antennas excited by narrow slot," *IEE Antennas and Wireless Propagation Letters*, Vol. 4, 2005.

3. Li, B. and K. W. Leung, “Strip-fed rectangular dielectric resonator antennas with/without a parasitic patch,” *IEEE Transactions on Antennas and Propagation*, Vol. 53, No. 7, Jul. 2005.
4. Moon, J.-I. and S.-O. Park, “Small chip antenna for 2.4/5.8-GHz dual ISM-band applications,” *IEEE Antennas and Wireless Propagation Letters*, Vol. 2, 2003.
5. Mongia, R. K., A. Ittipiboon, and M. Cuhaci, “Measurement of radiation efficiency of dielectric resonator,” *IEEE Microwave and Guided Wave Letters*, Vol. 4, No. 3, 80–82, Mar. 1994.
6. Lai, Q., G. Almpanis, C. Fumeaux, H. Benedickter, and R. Vahldieck, “Comparison of the radiation efficiency for the dielectric resonator antenna and the microstrip antenna at ka band,” *IEEE Transactions on Antennas and Propagation*, Vol. 56, No. 11, 3589–3592, Nov. 2008.
7. Drossos, G., Z. Wu, and L. E. Davis, “A comparative study of circular microstrip and cylindrical dielectric resonator antennas,” *Tenth. International Conference on Antennas and Propagation*, Vol. 1, 38–42, Apr. 14–17, 1997.
8. Kishk, A. A., Y. Yin, and A. W. Glisson, “Conical dielectric resonator antennas for wideband applications,” *IEEE Transactions Antenna and Propagations*, Vol. 50, 469–474, Apr. 2002.
9. Kishk, A. A., “Elliptic dielectric resonator antenna for circular polarization with single feed,” *Microwave and Optical Technology Letters*, Vol. 37, No. 6, 454–456, Jun. 2003.
10. Vijumon, P. V., S. K. Menon, M. N. Suma, B. Lehakumari, M. T. Sebastian, and P. Mohanan, “Broadband elliptical dielectric resonator antenna,” *Microwave and Optical Technology Letters*, Vol. 48, No. 1, 65–67, Jan. 2006.
11. Kishk, A. A., “Wideband dielectric resonator antenna in a truncated tetrahedron form excited by a coaxial probe,” *IEEE Transactions Antenna and Propagations*, Vol. 51, No. 10, 2907–2912, Oct. 2003.
12. Chang, T.-H. and J.-F. Kiang, “Broadband dielectric resonator antenna with an offset well,” *IEEE Antennas and Wireless Propagation Letters*, Vol. 6, 564–567, 2007.
13. Chair, R., A. A. Kishk, and K.-F. Lee, “Wideband stair-shaped dielectric resonator antennas,” *IET Microwaves, Antennas & Propagation*, Vol. 1, No. 2, 299–305, Apr. 2007.
14. Liang, X. L. and T. A. Denidni, “H-shaped dielectric resonator antenna for wideband applications,” *IEEE Antennas and Wireless Propagation Letters*, Vol. 7, 163–166, 2008.
15. Jaoujal, A., N. Aknin, and A. El Moussaoui, “Dual-band rectangular dielectric resonator antenna for ISM and 4 GHz bands applications,” *2009 Mediterranean Microwave Symposium (MMS)*, 1–3, 2010.
16. “HFSS: High frequency structure simulator based on finite element method,” Vol. 11, Ansoft Corporation, 2007.
17. “CST: Computer simulation technology based on FDTD method,” CST Computer Simulation Technology AG, 2008.

A Very Small UWB Dielectric Resonator Antenna for Mobile and Wireless Communications Systems

M. Aoutoul¹, N. Healey¹, J. Kiwan¹, F. Bourzeix¹, B. Lakssir¹, and M. Essaaidi²

¹Moroccan Foundation for Advanced Science, Innovation and Research — Microelectronics
Rabat, Morocco

²Faculty of Science, Abdelmalek Essaadi University, Tetuan, Morocco

Abstract— The proposed dielectric resonator antenna (DRA) structure has a very low profile, a very small size, $10 \times 10 \times 2.5 \text{ mm}^3$ DRA and $10 \times 35 \text{ mm}^2$ substrate size, and low permittivity constant (10.2). The simulated impedance bandwidth, achieved is about 73%, from 5.59 GHz to 12 GHz. The considered DRA radiation pattern is quasi-omni directional and has a simulated gain values up to 5 dB, which makes it suitable for higher band Direct Sequence Ultra Wideband (DS-UWB) communication systems and X band imaging systems applications.

1. INTRODUCTION

UWB antennas, the key elements of any ultra wideband systems, has become a competitive academic and industrial topic after 2002 date of the first standards ruling the commercial use of ultra wideband technology reported by the Federal Communications Commission (FCC).

Recently, DRAs have attracted a steady increasing attention during the last decade due to their main advantages such as wider bandwidth, better radiation efficiency [1–4] compared to those of their microstrip antennas counterparts and due to other additional characteristics such as light weight and low size, and their ability to support harsh environment conditions (i.e., high temperature degree). In the last few years, UWB DRAs [5–10] have been designed for high data rate local wireless communication systems, radars and imaging systems, in microwave and millimeter wave frequency ranges, responding to the extensive demand of wideband application and benefiting from their higher efficiency performance. However broad band and ultra wideband DRA structures are still suffering from some designing disadvantages making them less competitive such as complex geometries, higher dielectric constant material involved and/or the incompactness of proposed DRAs.

UWB DRAs have been designed for high data rate local wireless communication systems, radars and imaging systems, in microwave and millimeter wave regions, responding to the high demand of wideband application and benefiting from their higher efficiency and gain performance. However broad band and ultra wideband DRA structures are still suffering from some designing disadvantages making them less competitive such as complex geometries, relative bigger size, higher dielectric constant material involved and/or the incompactness of proposed DRAs.

In addition to its usefulness for UWB systems, the designed antenna structure can be integrated in portable devices for applications requiring higher frequencies band superior than 5 GHz. Recently new higher frequencies spectrum allocations have been established responding to a relevant increased demand because the lower frequencies spectrum, less than 5 GHz, witnessed a full employment and congestion by many standards.

In this work, we propose an extremely compact UWB DRA having a simple rectangular geometry, simple feeding technique and wider bandwidth (more than 5 GHz) which makes it low cost and, hence, practical for fabrication and integration into hand held devices. The working principle of this antenna is similar to that we have reported in [5] but the actual topology size is enormously reduced by about 74% compared to the original one. A parametric study has been done to reach the optimum structure to meet exactly the characteristics of the DS-UWB higher band, involving the principal geometrical and physical parameters, using the commercial electromagnetic simulator CST MWS and Ansoft HFSS to validate the optimum results obtained by CST MWS.

2. STRUCTURE DESIGN

The proposed antenna structure has a very low profile, a very small size of the radiator element, $10 \times 10 \times 2.5 \text{ mm}^3$, and low permittivity constant (10.2). The structure design is similar to that which has been developed in our previous work in 2009 [11]. The ground plane on the bottom antenna is half truncated and has three slots while the feeding mechanism is a stepped microstrip line, thereby

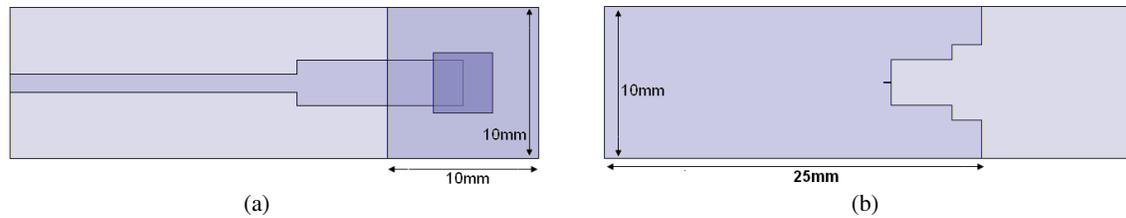


Figure 1: Antenna structure geometry: (a) top view and (b) bottom view.

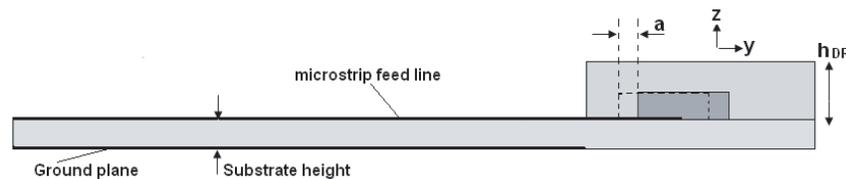


Figure 2: Antenna structure geometry: side view.

ensuring a good impedance matching. All geometrical values of the three slots and the stepped microstrip line are optimized using CST MWS and validated by Ansoft HFSS electromagnetic commercial simulators. The dielectric resonator has been made up by ceramic material with 10.2 permittivity constant, while the whole structure is printed on the Rogers RT/Duroid 3010 material with $25 \times 10 \times 1.27 \text{ mm}^3$ dimensions.

3. PARAMETRIC STUDY AND RESULTS

The scope of this study is designing a more compact DRA, while preserving good UWB performances. To achieve these aims, two design approaches have been followed in this work: the first one consists of minimizing extremely the substrate and ground plane to reduce the total antenna size, and in the second approach we have made investigation on the positioning effect of the placed metallic layer inside the dielectric resonator element. As result, the entire antenna size has been decreased by about 74% (from $30 \times 45 \text{ mm}^2$ to $10 \times 35 \text{ mm}^2$) with the ability of the new structure to operate in frequency range between 5.59 GHz and 12 GHz; this makes it a very small UWB dielectric resonator DRA for higher band DS-UWB frequency spectrum (6 to 10.6 GHz) as defined by FCC and for X band applications (IEEE Radar Designation).

4. RESULTS AND DISCUSSION

In microwave, the precision is key word of the game. After cutting the extended substrate and ground plane, surrounding the feeding line and radiator element, the positioning of the metallic layer, placed inside the dielectric resonator element, has been investigated to see how it influences on the antenna performances. An iterative process of simulations has lead to the CST simulated results shown by Figure 3 which illustrates that metallic object' position alters significantly the results obtained; hence, the optimum position, to achieve a large impedance bandwidth, is that when the metallic layer element is moved toward negative y -axis direction by a value of ' a ' parameter equals to 0.5 mm. Figure 4 shows the return loss parameters result simulated by Ansoft HFSS and CST MWS where the fractional antenna bandwidth is 73% (from 5.59 to 12 GHz); the two curves show a good agreement in the lower and upper band frequencies and slight difference between the two curves behavior which means that the fabricated prototype performances will be closed to the simulated ones.

The antenna radiation pattern has been plotted for three frequencies peaks: 6.15, 8 and 10 GHz in the two planes XY and YZ (Figures 5 and 6). The simulated results show that this antenna parameter increases versus frequency and can reach a maximum value about 5 dB at 10 GHz, the antenna radiation pattern is quasi omni-directional allowing it to be used for indoor LAN purposes for higher DS-UWB band.

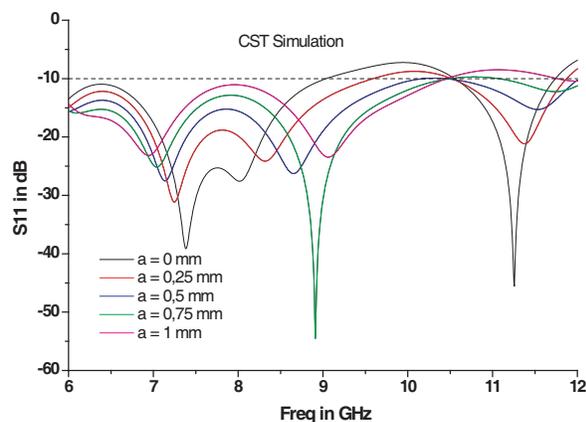


Figure 3: CST MWS Simulated Return Loss for different values of 'a' parameter.

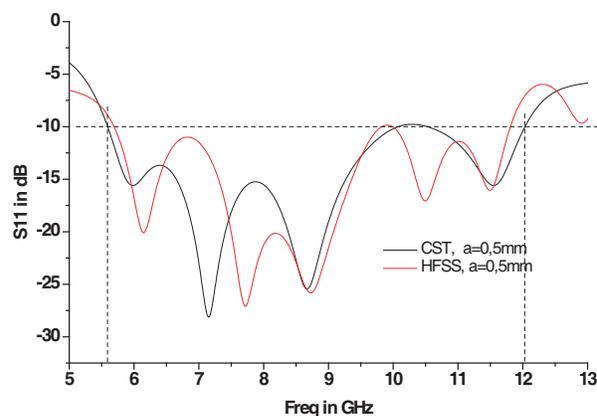


Figure 4: DRA simulated Return Loss of the optimised structure.

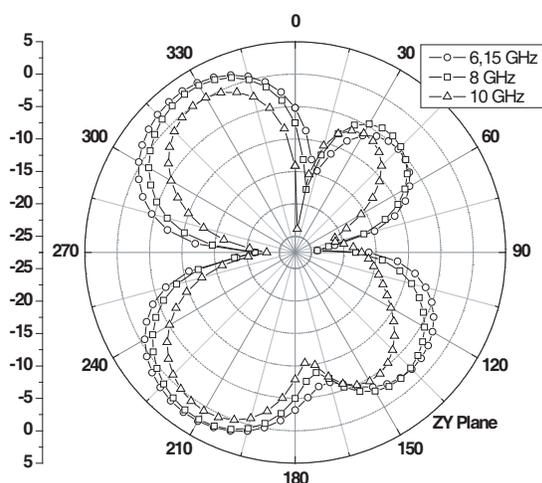


Figure 5: HFSS simulated gain at YZ plane.

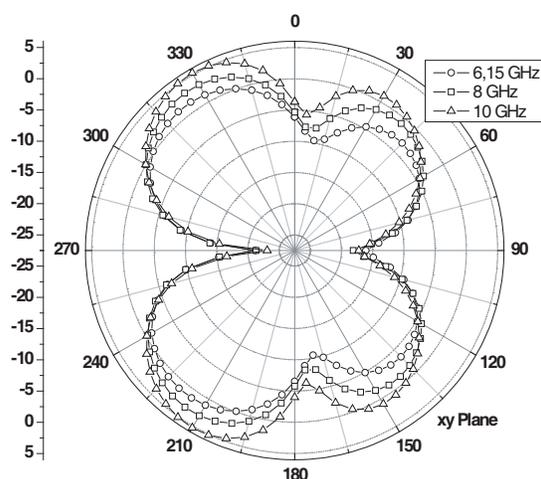


Figure 6: HFSS simulated gain at XY plane.

5. CONCLUSION

The extensive demand for higher debit applications and UWB system for communication and imaging systems has taken more attention of academic and industrial researchers. The use of the metallic layer and its positioning has helped to design this UWB DRA antenna which has an extremely small size, relatively higher gain (up to 4.5 dB) and large bandwidth from 5.6 to 12 GHz (fractional BW of 73%), this will make it suitable for higher UWB band (6–10.6 GHz) and X band (IEEE Radars band designation) applications.

ACKNOWLEDGMENT

The authors of this paper thank Dr. M. M'Hamed Drissi from INSA of Rennes, France, for his help in simulation with HFSS software.

REFERENCES

1. Long, S. A., M. W. Mcallister, and L. C. Shen, "The resonant cylindrical dielectric cavity antenna," *IEEE Trans. Antennas Propagat.*, Vol. 31, 406–412, 1983.
2. Petosa, A., A. Ittipiboon, Y. M. M. Antar, D. Roscoe, and M. Cuhaci, "Recent advances in dielectric resonator antenna technology," *IEEE Antennas and Propagation Magazine*, Vol. 40, No. 3, 35–48, 1998.
3. Mongia, R. K. and P. Bhartia, "Dielectric resonator antennas — A review and general design relations for resonant frequency and bandwidth," *International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering*, Vol. 4, No. 3, 230–247, 1994.

4. LuK, K. M. and K. W. Leung, *Dielectric Resonator Antennas*, Research Studies Press LTD., 2002.
5. Lapierre, M., Y. M. Antar, A. Ittipiboon, and A. Petosa, “Ultra wideband monopole/dielectric resonator antenna,” *IEEE Microwave and Wireless Components Letters*, Vol. 15, No. 1, 7–9, January 2005.
6. Guha, D., Y. M. M. Antar, A. Ittipiboon, A. Petosa, and D. Lee, “In search of ultra wideband performance of the monopole-dielectric resonator antenna,” *First European Conference on Antennas and Propagation, 2006. EuCAP 2006*, Digital Object Identifier: 10.1109/EUCAP.2006.4584581, 1–4, 2006.
7. Yuehe, G. and K. P. Esselle, “A UWB probe-fed dielectric resonator antenna,” *IEEE 69th Vehicular Technology Conference, 2009. VTC Spring 2009*, Digital Object Identifier: 10.1109/VETECS.2009.5073351, 1–2, April 26–29, 2009.
8. Morsy, M. M., M. R. Khan, and F. J. Harackiewicz, “Ultra wideband hybrid dielectric resonator antenna (DRA) with parasitic ring,” *2010 IEEE International Conference on Wireless Information Technology and Systems (ICWITS)*, Digital Object Identifier: 10.1109/ICWITS.2010.5611809, 1–4, 2010.
9. Huitema, L., M. Koubeissi, C. Decroze, and T. Monediere, “Handheld dielectric resonator antenna for ultra wideband applications,” *2010 International Workshop on Antenna Technology (iWAT)*, 1–4, 2010.
10. Guha, D., B. Gupta, and Y. M. M. Antar, “Hybrid monopole-DRA: New geometries for improved ultra-wideband operation,” *2010 IEEE International Symposium Antennas and Propagation Society (APSURSI)*, Digital Object Identifier: 10.1109/IWAT.2010.5464875, 1–4, 2010.
11. Aoutoul, M., O. El-Mrabet, M. Essaaidi, and A. El Moussaoui, “A compact rectangular dielectric resonator antenna for UWB wireless communication systems,” *Microwave and Optical Technology Letters*, Vol. 51, No. 10, 2281–2286. October 2009.

A Super-miniaturized Low Profile Antenna on a Substrate of Rose Curve Resonators

A. Kabiri¹, L. Talbi¹, and O. M. Ramahi²

¹Universite du Quebec Outaouais, Canada

²University of Waterloo, Canada

Abstract— Artificial magnetic materials (AMMs) are used as a substrate in design of miniaturized low profile antennas. We demonstrate the design of a microstrip patch antenna on a substrate composed of Rose curve resonators (RCRs). The RCR, a novel rose shape broken metallic inclusion, is a generic inclusion for design of an AMM with desired properties. It is shown that an AMM composed of RCRs allows the design of a patch antenna operating at desired frequency while occupying specific dimensions. In fact, the proposed substrate enables design of highly miniaturized low profile antennas. A prototype patch antenna is designed and fabricated, and its performance is compared with counterpart designs in the literature.

1. INTRODUCTION

Shrinking the size, simplifying the interface and encapsulating the multi-functionality in wireless hand-held devices are of the highest priority in the communications industry. This kind of demand is very challenging for antenna designers because the smaller the devices are, the smaller the antennas should be. Of course, the smaller size antennas should retain their capability to fulfill the ever shrinking system requirements. For example, wide bandwidth and high performance are demanding parameters in the antenna design. Recently, an influx of research has been devoted to antenna miniaturization using artificial magnetic materials [1–3]. Artificial magnetic materials are used as an antenna substrate of low profile antennas to reduce the size of antennas while maintaining good performance [2].

An artificial magnetic material (AMM) known as a magnetic metamaterial is a composition of electrically small metallic broken-loop inclusions, periodically aligned in parallel planes perpendicular to the direction of incident magnetic field. Impinging with an electromagnetic radiation, an AMM exhibits an enhanced homogenous magnetic response inside the medium at the vicinity of an LC resonance of the rings, and an effective homogenous electric response normal to the plane of the ring resonators [4].

Let consider a half-wavelength resonant patch antenna operating at an angular frequency ω_{op} . The size of the antenna ℓ_a is approximately equal to the half of the wavelength in the substrate; hence,

$$\ell_a \lesssim \frac{\lambda}{2} = \frac{\lambda_0/2}{\sqrt{\epsilon_{eff}\mu_{eff}}} \quad (1)$$

where λ_0 is the wavelength of the radiation in the free space, and ϵ_{eff} and μ_{eff} are the relative permittivity and permeability of the substrate. Increasing the effective medium parameters results in antenna size reduction. Note that if the antenna is miniaturized using a high permittivity dielectric, the antenna suffers from significant reduction in bandwidth and deterioration in the impedance matching [5–7]. Therefore, alternatively, for miniaturizing the size of the antenna, one can use magneto-dielectric materials ($\epsilon_{eff} > 1$, $\mu_{eff} > 1$) or high permeability materials for the substrate. This hypothesis has been examined by using a transmission-line analysis, and it was verified that a substrate with $\mu_{eff} \simeq \epsilon_{eff}$ and $\mu_{eff} \gg 1$ retains the antenna bandwidth and radiation quality factor after miniaturization [8]. Since an AMM effectively acts as a homogenous magneto-dielectric medium, it is a suitable candidate for patch antenna miniaturization. One can determine desired properties of the substrate, and then, design the substrate with the design methodology proposed in [9].

In this work, the performance of a patch antenna while lying on an AMM substrate composed of Rose curve resonators (RCRs) is analyzed. The RCR is a generic inclusion for AMMs, and can be formulated in polar coordinates in the following parametric form [10]

$$R_n(r_0, a) : r(\theta) = r_0 + a \cos(n\theta) \quad (2)$$

where $r(\theta)$ represents the position of an inclusion's contour in the polar coordinate, and θ is the polar angle which sweeps the contour aside the slit. The parameters r_0 and a are constants and n is an integer representing the order of the curve. These parameters are calculated in a design procedure so that the produced AMM fulfills desired magnetic properties [9]. An AMM composed of RCRs can be designed in order to maintain a desirable frequency dispersion over a specific frequency range, while the AMM meets a specific magneto-dielectric properties. Therefore, the designer, by choosing a suitable artificial magneto-dielectric substrate, has the opportunity of designing an antenna with a desired performance. Next, the designer can configure the AMM to respond based on the substrate used in the antenna design procedure. A significant advantage of such a substrate composed of RCRs is that it can be used in design of highly miniaturized microstrip patch antennas. In fact, increasing the effective substrate parameters results in the antenna size reduction.

2. MICROSCRIP ANTENNA DESIGN

A resonant patch antenna mounted on a magneto-dielectric substrate is designed to work at 600 MHz. The patch size is desired to be less than $40 \text{ mm} \times 40 \text{ mm}$. The ground plane is roughly taken three times larger than the patch size in each side to avoid the restriction of the gain, and the antenna is fed by a coaxial line from the bottom. The substrate is an AMM slab composed of Rose curve inclusions with height of 16 mm. Metallic Rose curve inclusions are periodically aligned on printed circuit boards (PCBs), and PCBs are stacked parallel to the E -plane of the microstrip patch antenna to form the substrate (see Fig. 1(a)). Therefore, while the antenna is excited, the magnetic field produced in the substrate is normal to inclusions, and induces an electric current on inclusions to enhance the magneto-dielectric properties of the substrate. The AMM was separately designed to provide $\mu_{eff} = 10 \pm 6\%$ and $\epsilon_{eff} = 18 \pm 6\%$ over a frequency bandwidth of 2 MHz and with a magnetic loss tangent (MLT) and electric loss tangent (ELT) of less than 0.05 (see [5] for design procedure). The suitable inclusion which can fulfil the required properties is a 13th order Rose curve with $r = 5.86 \text{ mm}$, $a = 0.32 \text{ mm}$, and trace width of 1.47 mm (see Fig. 1(b)). Figs. 2(a) and (b) show the real effective permeability and permittivity of the designed AMM, and the electric and magnetic loss tangent in the slab medium, respectively. Moreover, Fig. 2(c) shows the real and imaginary part of the refractive index. Indeed, the refractive index corresponds to the miniaturization factor (\mathcal{M} -factor). The values of the effective parameters of the substrate were extracted through plane wave analysis by the scattering parameters of the AMM simulated with a commercial EM simulator at 600 MHz. The results are summarized in Table 1.

The schematic of the designed structure is demonstrated in Fig. 3(a). To check the performance of the antenna, the patch antenna was simulated, and the operating parameters were extracted using a full-wave antenna simulator. However, due to the large feature size ratio between the antenna and the AMM, a huge number of inclusions are needed to fill up the substrate. Thus, embedding

Table 1: The effective parameters of designed AMM for an antenna substrate at 600 MHz.

| $\epsilon_{eff} \text{ Re}$ | $\tan \delta_\epsilon$ | $\mu_{eff} \text{ Re}$ | $\tan \delta_\mu$ | n_{eff} |
|-----------------------------|------------------------|------------------------|-------------------|------------------|
| 18.52 | 18.52 | 10.08 | -0.374 | $13.68 + j0.088$ |

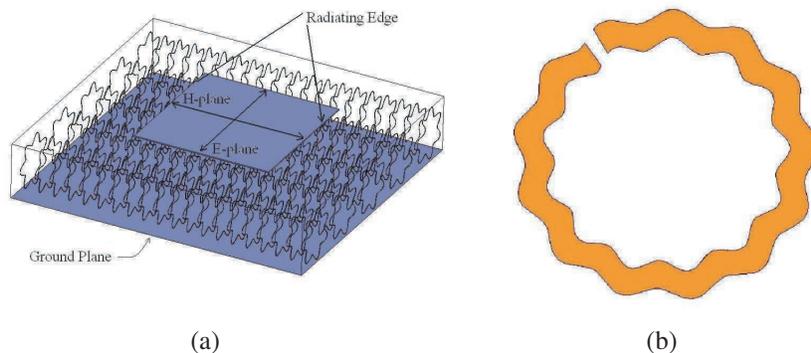


Figure 1: (a) A schematic of a coaxial-fed patch antenna lying over an AMM composed of rose curve inclusions. (b) A 13th order rose curve resonator.

the metallic inclusions in the design setup consumes a huge, and practically inefficient, resources for simulation. As a consequence, the simulation was alternatively done with a homogeneous, yet anisotropic, substrate with permittivity and permeability functions defined based on the AMM electromagnetic characteristics.

The simulated S_{11} is presented in Fig. 3(b), and it shows a relative bandwidth ($S_{11} < -10$ dB) of 0.67%. The bandwidth of the antenna is relatively low compared to other counterparts proposed in the literature, and make the antenna useful for narrow band applications such as sensor applications. The radiation pattern of the antenna at 600 MHz is plotted in Fig. 3(c). The numerical calculations show a maximum gain of 0.5 dB and 59% radiation efficiency for the miniaturized antenna. According to the design dimensions, the size of the patch antenna is equal to $L = 31$ mm which is $1/16.2$ of $\lambda_0 = 50$ cm, the wavelength at the resonance frequency of 600 MHz. Thus, a

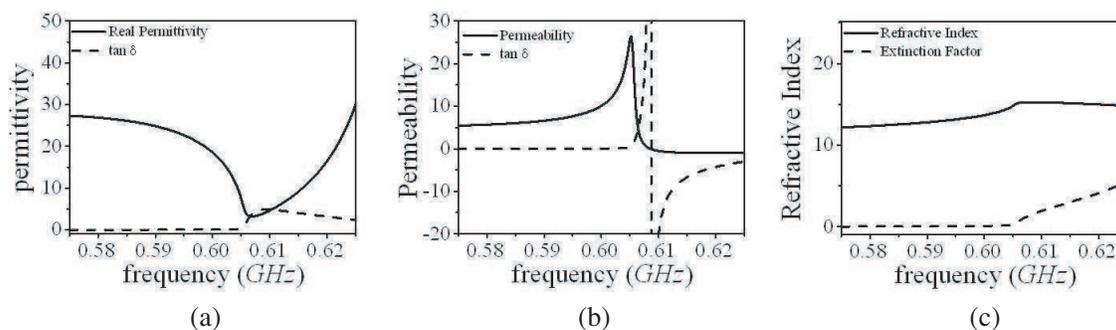


Figure 2: (a) The permittivity and electric loss tangent functions of the substrate composed of rose curve inclusions. (b) The permeability and magnetic loss tangent functions of the substrate composed of rose curve inclusions. (c) The refractive index and the extinction factor functions of the substrate composed of rose curve inclusions.

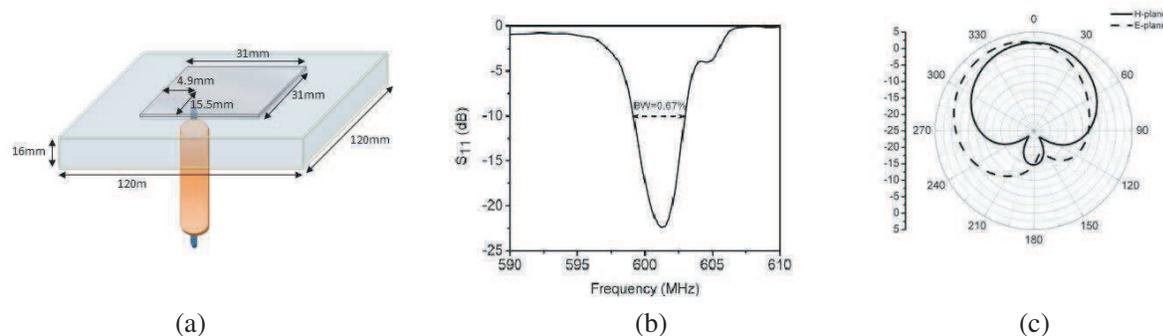


Figure 3: (a) The design schematic of the coaxial-fed microstrip patch antenna. (b) The simulated return loss of the miniaturized patch antenna. (c) The simulated radiation pattern of the miniaturized patch antenna.

Table 2: Comparison table of different miniaturized antenna designs. sim = simulation results, mes = measurement results, filling = a typical AMM dispersion function.

| Source | inclusion | f_{res} MHz | Gain | BW | Efficiency | \mathcal{M}_{factor} | sim/mes |
|-----------|--------------|---------------|---------|------|------------|------------------------|---------|
| [6] | Hilbert | 615 MHz | -3.4 dB | 2.1% | 50% | 5.8 | mes |
| [1] | Metasolenoid | 860 MHz | - | 1.3% | 70% | 1.2 | mes |
| [2] | CSR | 3.0 GHz | 1.3 dB | 3.3% | - | 1.4 | sim |
| [2] | CSR | 2.5 GHz | -0.5 dB | 3.2% | - | 1.5 | mes |
| [5] | Spiral | 250 MHz | -3.9 dB | 0.8% | 20% | 6.5 | mes |
| [3] | filling | 1.58 GHz | - | 6.0% | 89% | 1.3 | sim |
| [7] | Using EBG | 2.4 GHz | 5.2 dB | - | 90% | 1.3 | mes |
| This work | Rose curve | 600 MHz | 0.5 dB | 0.7% | 59% | 8.1 | sim |

M -factor of 8.1 is achieved using artificial magnetic substrate. Note that Chu's theoretical limit of the relative bandwidth for the antenna with provided dimensions is about 0.79% in comparison with 0.67% relative bandwidth achieved in this work. This work shows a remarkable attempt to approach the Chu's limit for small size antennas. In the calculation of Chu's theoretical limit, the patch length was considered as the antenna's dimension. For completeness, the simulation results were compared with a similar patch antenna design with substrates composed of various AMM's inclusions' geometry. A comparison of different experiments and simulations reported in the literature were summarized in Table 2.

The antenna is fabricated, and the simulation results are validated with experimental measurements. Although the results of measurements are not provided in this manuscript, they will be presented in the symposium.

3. CONCLUSION

We presented an enabling technique for design of an ultra miniaturized patch antenna. In the proposed technique, we used an artificial magnetic material composed of Rose curve resonators as the substrate of a patch antenna. AMMs composed of Rose curve resonators can be designed to meet a desired magneto-dielectric properties. Therefore, a deft designer can order the antenna substrate needed for achieving specific antenna performance and size. The substrate is an engineered medium composed of Rose curve resonators. In an earlier work, it was shown that such a medium can be designed to meet specific electromagnetic properties; i.e., a specific permittivity and permeability over a frequency band while showing a maximum loss tangent at a desired frequency [9]. Of course, some fundamental limitations are applied on design of AMMs which already discussed in [11]. In fact, design of a desired AMM enables a significant potential for design of ever shrinking low profile antennas in the wireless communication and biomedical industry.

REFERENCES

1. Ikonen, P., S. I. Maslovski, C. R. Simovski, and S. A. Tretyakov, "On artificial magnetodielectric loading for improving the impedance bandwidth properties of microstrip antennas," *IEEE Transaction on Antenna and Propagation*, Vol. 54, No. 6, 1654–1662, June 2006.
2. Lee, Y. and Y. Hao, "Characterization of microstrip patch antennas on metamaterial substrates loaded with complementary split-ring resonators," *Microwave and Optical Technology Letters*, Vol. 50, No. 8, 2131–2135, August 2000.
3. Karkkainen, M. K., S. A. Tretyakov, and P. Ikonen, "Numerical study of pifa with dispersive material fillings," *Proceeding of the International Symposium of Antennas and Propagation Society*, 1885–1888, June 2007.
4. Marqus, R., F. Medina, and R. Rafii-El-Idrissi, "Role of bianisotropy in negative permeability and left-handed metamaterials," *Physical Review B*, Vol. 65, 144440/1C6, April 2002.
5. Buell, K., H. Mosallaei, and K. Sarabandi, "A substrate for small patch antennas providing tunable miniaturization factors," *IEEE Transaction on Microwave and Technology*, Vol. 54, No. 12, 134–146, January 2006.
6. Yousfi, L. and O. Ramahi, "Miniaturized wideband antenna using engineered magnetic materials with multi-resonator inclusions," *Proceeding of the International Symposium of Antennas and Propagation Society*, 1885–1888, June 2007.
7. Korkontzila, E. G., D. B. Papafiliippou, and D. P. Chrissoulidis, "Miniaturization of microstrip patch antenna for wireless applications by use of multilayered electromagnetic band gap substrate," *Proceeding of the International Symposium of Antennas and Propagation Society*, 1–6, November 2006.
8. Hansen, R. C. and M. Burke, "Antennas with magneto-dielectrics," *Microwave Optical Technology Letters*, Vol. 26, No. 2, 75–78, 2000.
9. Kabiri, A. and O. M. Ramahi, "A simple approach for synthesizing of multipurpose metamaterials," *Proceeding of the International Symposium of the Antenna and Propagation Society*, Toronto, Canada, July 2010.
10. Kabiri, A. and O. M. Ramahi, "Metamaterials composed of rose curve inclusions," *Proceeding of the International Symposium of the Antenna and Propagation Society*, Toronto, Canada, July 2010.
11. Kabiri, A., L. Yousefi, and O. M. Ramahi, "On the fundamental limitations of artificial magnetic materials," *IEEE Transaction on Antenna and Propagation*, Vol. 58, No. 7, 2345–2353, July 2010.

Thermographic Analysis of Swiss Albino Mice Exposed to 1.8 GHz GSM Frequency

A. D. Usman¹, W. F. Wan Ahmad¹, M. Z. A. Ab Kadir¹, M. Mokhtar², and A. Rusnani³

¹Centre of Excellence on Lightning Protection (CELP), Faculty of Engineering
Universiti Putra Malaysia, UPM Serdang, Selangor 43400, Malaysia

²Wireless and Photonic Networks (WiPNET), Research Center of Excellence
Faculty of Engineering, Universiti Putra Malaysia, UPM Serdang, Selangor 43400, Malaysia

³Faculty of Electrical Engineering, Universiti Teknologi Mara (UiTM) Pulau Pinang
Permatang Pauh 13500, Pulau Pinang, Malaysia

Abstract— Exposure to radio frequency electromagnetic field has been a subject of debate for many years. Currently, related regulations provided were meant to safeguard occupational and general public for short term thermal effect only. Due to that reason, this study was designed to investigate the long term thermal effect of electromagnetic field exposure of 1.8 GHz GSM frequency. 72 unrestrained Swiss Albino mice were used as surrogate for this study and were divided into two groups of 1.8 GHz and control groups. The cage for the 1.8 GHz group was placed at far field distance from the directional antenna, and the exposure were for 7 hours/day, 7 days/week and for 4 weeks. While the cage for control group is sham exposed. Thermographic pictures were taken daily to obtain average temperature of the mice brain before and after being exposed to the 1.8 GHz GSM down-link frequency. The highest and lowest mice brain temperatures recorded for 1.8 GHz group are on day 20 and day 6 with temperature of 37.0°C and 33.9°C, respectively. While, the highest and lowest mice brain temperatures recorded for control group are on day 1 and day 24 with temperature of 36.9°C and 32.7°C, respectively. Daily rise in temperature for 1.8 GHz and control groups range from 0.3°C to 2.1°C and 0.3°C to 2.8°C, respectively. The mice weight taken on weekly basis shows a decrease for 1.8 GHz expose group as compared to the mice in control group. The death rate was also found to be higher for mice in 1.8 GHz group compared to the mice in control group.

1. INTRODUCTION

Microwave frequencies used in our daily lives has attracted a lot of concern on the possible health consequences which may likely result with time. The widespread usage is more prominent in telecommunication industries. Multipurpose usages of mobile phones and its proximity to humans could go on and on with time. According to International Telecommunication Industries (ITU), the demands for communication services were not affected even during the world economic recess. It was reported that after reaching 4.6 billion subscribers at the end of 2009, the number of mobile subscribers globally is estimated to reach 5 billion by 2010 [1].

Concerns on the possible health consequences during short and long term usage of this technology have given birth to a lot of researches on possible biological effects. However, the results are inconclusive to date. Though several guidelines exist from various organizations for decades, these guidelines have gone through harmonization where the ones that are widely adopted are provided by IEEE and ICNIRP [2, 3]. Though most of the exposure from communications instruments are reported to be below the recommended values, there are complains of discomfort and fears due to cumulative effects that might arise as a result of long time usages. This however, leads to a lot of complicating researches with many pointing to possible biomarkers of effect that will likely lead to secondary consequences. Among which are; increased incidence of cancer as a result of exposure to radio frequency electromagnetic field radiation (RF EMR) as reported in the researches done by [4–7], DNA damage as in [8–14] and general discomfort in [15, 16]. Some reports are however, in contrary as in [17–21].

Investigation on RF EMR exposure as it relates to temperature changes on the body at long time basis are still not able to be understood until now. When EM waves hit a body, part of the waves are reflected while other are absorbed and the penetration orientation changes at every surface between elements of the body. Part of EM energy absorbed is converted into heat, provoking an elevation of temperature. The energy transformations into the tissues create an increase of the molecular kinetic energy, thereby absorbing the EM waves. Researches aimed at measuring the heat generation effect on the body of animal and humans on long term basis are currently rather

scanty. Insights into some of these studies are provided in [22–26], where in these studies though it was reported as long term, the active exposure period were either short or the animals used as surrogate were put in restrained condition. On the other hand, this study is aimed at using mice under unrestrained condition placed at far field from the antenna. Average field strength of 1.5 V/m was generated for the 1.8 GHz exposed group, while the control group is sham exposed, i.e., placed away from any RF EMF source. The research was also designed to mimic long term whole body exposure effect due to RF EMR of a typical base station where the exposure period was for 7 hours/day, 7 days/week and for 4 weeks. The main aim of this study is to determine possible indicators that may lead to possible biological effect as a result of long term exposure to RF EM field of GSM frequencies under similar environmental condition. During this study, the increase of kinetic energy which translates into heat generation is to be measured directly using an infra-red thermographic camera (IRTC). Infra-red thermal imaging was chosen as it is non-invasive and capable of measuring local temperature directly as oppose to other methods. Nowadays, IRTCs are used in many application ranging from industries, human and veterinary medicine. Above all it is capable of measuring object temperature without any direct physical contact while continuously monitor the surface of interest.

2. EXPERIMENTAL PROTOCOL

The study was designed to measure the whole body surface temperature of mice before and after exposure to 1.8 GHz RF EMF each day for a period of 4 weeks. Animal care and handling was carried out according to the guidelines set by World Health Organization (WHO) in Geneva, Switzerland, Malaysian Animal Handling Code of Conduct and National Research Council guide for the care and use of laboratory animals [27]. The RF EMF design and exposure set-up were done following Electronic Communication Commission (EEC) protocol [28]. 72 male Swiss Albino mice (*Mus Musculus*) aged about 4 weeks on arrival and weighing between 34g to 45g were obtained from Veterinary Research Institute in Ipoh, Perak, Malaysia. A total of 72 mice were used for this study in which the mice were divided into 2 groups of 36 mice each. The groups were named as 1.8 GHz and control groups. 36 mice in each group were randomly selected, housed in 2 cages of 18 mice each and were allowed to acclimatize for a week before the exposure started according to suggestion by [29–31]. The cages were made of metal with plastic coating and the dimensions were 0.6 m × 0.42 m × 0.24 m. The floor of the cage was made of removable plastic covered with wooden chips. This cage is capable of accommodating 25 unrestrained mice according to [27]. Hence, with only 18 mice in each cage, all the mice are able to move around freely. Both the 1.8 GHz and control groups' ambient temperature and humidity throughout the experiment were maintained at 27°C ± 2°C and 65% ± 5%, respectively. For both the 1.8 GHz and control groups, ferrite tile floor absorbers were used to prevent ground reflection and interference from other source. The water bottle and the food container were made of plastics and assorted mixed food and water were available to the mice ad libitum. Wooden chips beddings were used to supply comfort and absorb mice waste. The beddings were changed regularly to avoid infection.

IFR AM/FM signal generator model 2023B with frequency range of 9 kHz–2.05 GHz and a resolution of 1 Hz was used and set at +13 dBm power to provide the required exposure signal for 1.8 GHz group. FSL spectrum analyzer with frequency range of 9 kHz–6 GHz was used to measure the received signal strength from different position in the cage as well as to make sure there is no interference from any neighboring signal. An SMP RF survey meter with frequency range of 3 kHz–3 GHz was used to validate the field strength generated inside each cage. The cage were placed at far field distance using far field equation parameters shown in (1), where R is the distance from the source to the far field, D is the largest dimension of the source antenna and λ is the wavelength of the transmitted signal.

The maximum couplings of the field generated with the mice body were ensured and the average field strength generated and measured inside the cages with both the spectrum analyzer and the SMP field meter is 1.5 V/m. This is correspond to the typical field strength obtained during the GSM base stations site surveys conducted at far field distances in some residential areas in Malaysia. The exposure was conducted for 7 hours/day, 7 days/week and for 4 weeks including the weekend and public holidays. Measurements of temperature rise in the body of the mice were done each day before and after exposures using an IRTC of IR Snap shot Model 525. The camera resolution is 0.1°C and Snap View Pro 2.1 software is used for the thermo graphic analysis. This IRTC has a wavelength of 8–14 μm which is quite suitable for measurements of the mice temperatures. Note that the measurements using IRTC resulting thermographic pictures taken in the morning and

evening each day for before and after exposure cases, respectively.

Specific absorption rate (SAR) value calculated using (2) and (3) was found to be 0.2W/kg. These equations considered the worst case relation between specific energy and temperature provided the effect of cooling is neglected [32], where t is the temperature rise in $^{\circ}\text{C}$, J is the specific energy absorption in Jkg^{-1} , $c = 0.85$ and T is the exposure time in s.

$$R = \frac{2D^2}{\lambda} \quad (1)$$

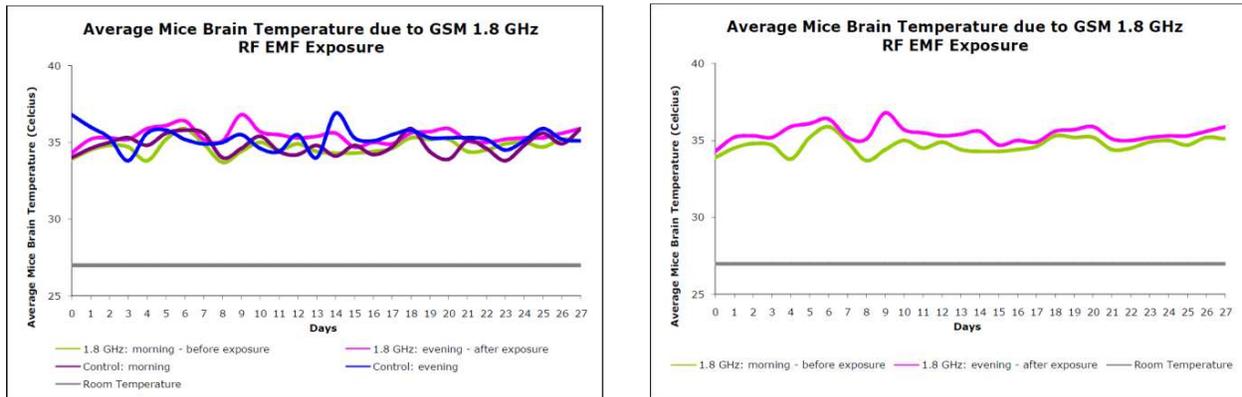
$$t = \frac{J}{c \times 4180} \quad (2)$$

$$J = SAR \times T \quad (3)$$

3. RESULTS AND DISCUSSIONS

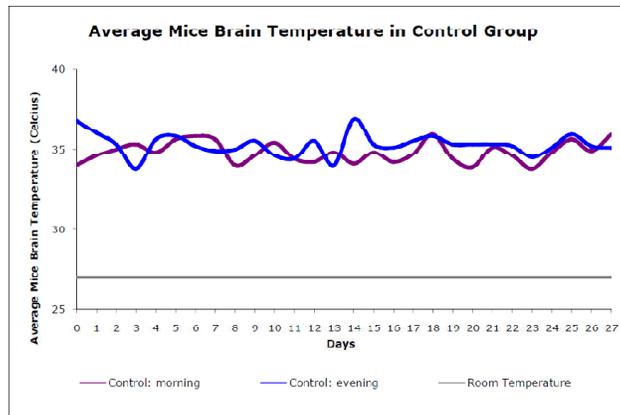
The IRTC was used to capture the body temperature of three randomly chosen mice from each group before the exposure in the morning and after 7 hours exposure in the evening where the average temperature of the mice brain for both groups are as displayed in Figure 1. Weights of all the mice in each group were taken on day 1 and subsequently on weekly basis where the average values are tabulated in Figure 2. The dead mice during the study from each group was sent for postmortem examination, with the total death against time of each group is recorded.

It is shown that for mice in 1.8 GHz group, the average brain temperature increases after the 7 hours exposure, while for the mice in control group, there is no definite pattern of temperature changes. This is evidently seen on each day until day 27. It is found that the highest daily temperature observed after the 7 hours exposure was 36.8°C on day 10 and the lowest being 34.3°C



(a)

(b)



(c)

Figure 1: (a) Average mice brain temperature due to GSM 1.8 GHz RF EMF exposure in 1.8 GHz and control groups. (b) Average mice brain temperature in 1.8 GHz group. (c) Average mice brain temperature in control group.

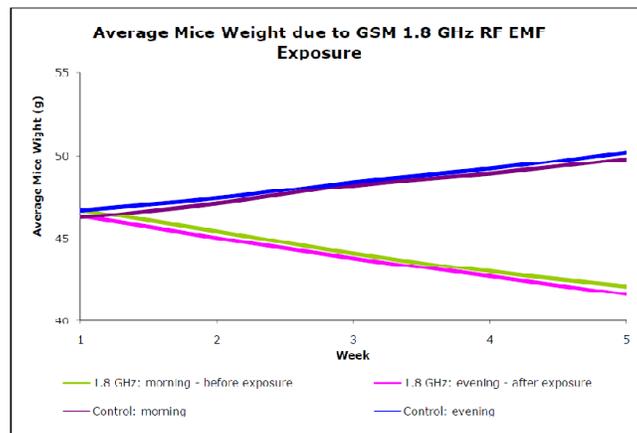


Figure 2: Average mice weight due to GSM 1.8 GHz RF EMF exposure.

on day 1 for the mice in 1.8 GHz group. The incremental temperature difference between the reading before and after the exposure on the mice in 1.8 GHz group was high on day 5, recorded at 2.1°C and low on days 8, 18, 19, 24, and 25 with 0.3°C each. In this group, rise in body temperature as a result of exposure was demonstrated. On the hand, the average mice brain temperature in control group fluctuates where results are highly dependent upon daily temperature changes. The average high temperature was observed on day 15 with temperature readings of 36.9°C , and the lowest temperature is 34°C .

A summary of the paired sample T-test carried out on 1.8 GHz group is as listed in Table 1. It shows that $t(27) = 7.460$, $p = 0.000$. Since $\text{sig-}t(0.000) > \alpha(0.05)$, it is therefore, concluded that the temperature rise after the exposure during the evening (after 7 hours exposure) is significant compared to the temperature rise before the exposure at 0.05 level of significance. Also, an independent sample T-test conducted for 1.8 GHz and control groups is listed in Table 2. From the results $t(53.999) = 2.104$, and $p = 0.040$. Since $\text{sig-}t(0.040) < \alpha(0.05)$, it is concluded that there is significant difference in temperatures between 1.8 GHz and control groups at 0.05 level of significance.

From Figure 2, it is shown that the starting average weight of the mice in 1.8 GHz group was 46.7g (day 1). This weight decreases with prolong exposure and at the end of week four where it reduces to 41.6g (day 27). This however, is contrary to mice in control group where their weights increased with time due to the normal growth. In the control group, the average weight of the mice on day 1 was recorded as 46.2g and at the end of week four (day 27), the average weight was measured to be to 50.2g. The loss in weight in 1.8 GHz group within 4 weeks of exposure is 5.1g. On the other hand, the average weight gained in the control group within 4 weeks is 4g. Also, the death rate indicated that the total deaths recorded were 3 and 8 mice for control and 1.8 GHz groups, respectively. This demonstrated that the survival rate on mice in 1.8 GHz group is much lower compare to the mice in the control group.

As seen from the results, sharp temperature increases were noticed after the exposure in each day and the statistical analysis conducted corroborate the results shown in Figure 1. The analysis shows that there is a significant difference in temperatures recorded before and after the exposure as well as between the 1.8 GHz and the control groups. The results obtained in this experiment are quite comparable to [22–26].

A decrease in body weight with prolong 1.8 GHz RF EMF exposure was noticed after 4 weeks time. This may likely be as a result of lack of appetite as well as some other hidden effect that will likely be due to long exposure duration. Weight loss was demonstrated by mice in 1.8 GHz group, while mice in the control group gain weight as the days increases, with the highest record at week 4. Similarly the survival rate of mice in 1.8 GHz group is rather low compare to the mice in control group. It was also observed that mice in the 1.8 GHz group were not as active as mice in the control group. An increase urine secretion, clustering of the mice in one place and non free movement of the mice in the 1.8 GHz group were observed where vice versa behaviour indicated by mice in the control group.

Table 1: Paired sample T-test for 1.8 GHz exposed group.

| Temperature | Mean | SD | t | p |
|---------------------|---------|---------|-------|-------|
| Post test (Evening) | 35.4357 | 0.51368 | 7.460 | 0.000 |
| Pre test (Morning) | 34.7036 | 0.49177 | | |

Table 2: Summary of Independent sample T-test for 1.8 GHz exposed group.

| Group | n | Mean | SD | t | p |
|---------|----|---------|---------|-------|-------|
| 1.8 GHz | 28 | 35.4357 | 0.51368 | 2.104 | 0.040 |
| Control | 28 | 35.1464 | 0.51531 | | |

4. CONCLUSIONS

RF EMF exposure for long duration of time was observed on unrestrained male Swiss Albino due to 1.8 GHz GSM frequencies. An increase in body temperature was observed and recorded using IRTC. The statistical analysis carried out on temperature before and after exposure of 1.8 GHz group and temperatures between the control and 1.8 GHz groups were found to be significant at 0.05 level of significance. Despite the thermoregulatory mechanism of animal body when subjected to heat stress, it is evident that the temperature effect is cumulative as this gives rise to manifestation of some changes in both eating habit which lead to loss of weight and increased urine secretion as well as non active nature of the animal as compared to the control group. The death rate is another pointer to likely secondary effect caused by long term exposure, with 22% death rate is recorded for mice in the 1.8 GHz group, compared to only 6% of the mice in the control group died within 4 weeks. From the results it is evident that, the pattern of energy absorption inside irradiated body is non uniform and the biological responses are dependent on the exposure durations, frequency, intensity as well as ambient temperature.

REFERENCES

1. Hamadoun, T., "Strong global mobile cellular growth across all regions," *International Telecommunication Union Mobile World Congress*, Barcelona, 2010.
2. Institute of Electrical and Electronics Engineering (IEEE), "Standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz," New York, NY, 2005.
3. International Commission on Non-Ionizing Radiation Protection (ICNIRP), "Guidelines for limiting exposure to time-varying electric, magnetic and electromagnetic field (up to 300 GHz)," *Health Physics*, Vol. 74, 494–522, 1998.
4. Moulder, J. E., L. S. Erdreich, R. S. Malyapa, J. Meritt, W. F. Pickard, and Vijayalaxmi, "Cell Phones and Cancer: What is the Evidence for Connection?," *Radiation Res.*, Vol. 151, 513–531, 1999.
5. Moulder, J. E., K. R. Foster, L. S. Erdreich, and J. P. McNamee, "Mobile phones base stations and cancer: A review," *Int. J. Radiat. Biol.*, Vol. 81, 189–203, 2005.
6. Fritze, K., C. Sommer, B. Schmitz, G. Mies, K. A. Hossmann, M. Kiessling, and C. Wiessner, "Effect of global system for mobile communication (GSM) microwave exposure on blood-brain barrier permeability in rat," *Acta Neuropathol.*, Vol. 94, 465–470, 1997.
7. Norton, W. T., D. A. Aquino, I. Hozumi, F. C. Chiu, and C. F. Brosnan, "Quantitative aspects of reactive gliosis: Review," *Neurochem Res.*, Vol. 17, 877–885, 1992.
8. Sakar, S., S. Ali, and J. Behari, "Effect of low power microwave on the mouse genome: A direct DNA analysis," *Mutation Research*, Vol. 320, 141–147, 1994.
9. Lai, H. and N. P. Sing, "Acute low-intensity microwave exposure increases DNA single-strand breaks in rat cells," *Bioelectromagnetics*, Vol. 16, 207–210, 1995.
10. Lai, H. and N. P. Sing, "Single and double strand DNA breaks in rat brain cells after acute exposure to radiofrequency electromagnetic radiations," *Int J. Radiat. Biol.*, Vol. 69, 513–521, 1996.
11. Lai, H. and N. P. Sing, "Melatonin and a spin-trap compound block radiofrequency electromagnetic radiation — Induced DNA strand break in rat brain cells," *Bioelectromagnetics*, Vol. 18, 446–454, 1997.

12. Malyapa, R. S., E. W. Ahern, W. L. Straube, E. G. Moros, W. F. Pickard, and J. L. Roti Roti, "Measurement of DNA damage after exposure to electromagnetic radiation in the cellular phone communication frequency band (835.62 and 847.74 MHz)," *Radiat Res*, Vol. 148, 618–627, 1997.
13. Malyapa, R. S., E. W. Ahern, W. L. Straube, E. G. Moros, W. F. Pickard, and J. L. Roti Roti, "Measurement of DNA damage after exposure to 2450 MHz electromagnetic radiation," *Radiat. Res.*, Vol. 148, 608–617, 1997.
14. Philips, J. L., O. Ivaschuk, T. Ishida-Jones, R. A. Jones, M. Campbell-Beachler, and W. Haggren, "DNA damage in molt-4T lymphoblastoid cells exposed to cellular telephone radiofrequency fields in vitro," *Bioelectrochemistry and Bioenergetics*, Vol. 45, 103–110, 1998.
15. Mild, K. H., G. Oftedal, M. Sandstorm, J. Wilen, T. Tynes, B. Haugsdal, and E. Hauger, "Comparison of symptoms experienced by users of analogue and digital mobile phones," National Institute of Working Life, 1998.
16. Kumar, A., "Review on effects of electromagnetic radiation based on published papers," Quebec, Canada, 2000.
17. Inskip, P. D., R. E. Tarone, E. E. Hatch, T. C. Wilcosky, W. R. Shapiro, R. G. Selker, H. A. Fine, P. M. Black, J. S. Loeffler, and L. M. S. Linet, "Cellular-telephone use and brain tumors," *N. Engl. J. Med.*, Vol. 344, 79–86, 2001.
18. Joubert, V., P. Leveque, M. Cueille, S. Bourthoumieu, and Y. C. Yardin, "No apoptosis is induced in rat cortical neurons exposed to GSM phone fields," *Bioelectromagnetics*, Vol. 28, 115–121, 2007.
19. Muscat, J. E., M. G. Malkin, S. Thompson, R. E. Shore, S. D. Stellman, D. McRee, A. I. Neugut, and E. L. Wynder, "Handheld cellular telephone use and risk of brain cancer," *JAMA*, Vol. 284, 3001–3007, 2000.
20. Zook, B. C. and S. J. Simmens, "The effects of pulsed 860 MHz radiofrequency radiation on the promotion of neurogenic tumors in rats," *Radiat. Res.*, Vol. 165, 608–615, 2006.
21. Kim, T. H., T. Q. Huang, J. J. Jang, M. H. Kim, H. J. Kim, J. S. Lee, J. K. Pack, J. S. Seo, and W. Y. Park, "Local exposure of 849 MHz and 1763 MHz radiofrequency radiation to mouse heads does not induce cell death or cell proliferation in brain," *Exp. Mol. Med.*, Vol. 40, 440–447, 2008.
22. Ivana, K., K. Petr, A. K. G. Gurkan, P. Yunus, and C. S. Kemal, "Application of infrared thermography in animal production," *J. of Fac. of Agric.*, Vol. 22, 329–339, 2007.
23. Alla, Z., D. Samuel, V. Oleg, and B. Rafeal, "Dyanamic thermography: Analysis of hand temperature during exercise," *Annals of Biomedical Engineering*, Vol. 26, 988–993, 1998.
24. Christian, K., "Infrared thermal imaging to measure local temperature rises caused by handheld mobile phones," *IEEE Transaction on Instrumentation and Measurement*, Vol. 54, 1513–1519, 2005.
25. Ludwig, N., M. Gargano, F. Luzi, C. Carenzi, and M. Verga, "Technical note: Applicability of infrared thermography as non invasive measurement of stress in rabbit," *World Rabbit Science*, Vol. 15, 199–206, 2007.
26. Mikokajczyk, H., K. Maria, R. Ingeburg, and S. Eggert, "Thermal effect in rats exposed to heat at 40°C or to electromagnetic field," *J. Therm. Biol.*, Vol. 16, 115–119, 1991.
27. National Research Council (NRC), "Guide for the care and use of laboratory animals," Institute of Laboratory Animal Resources, Ed., National Academy Press, Washington DC, 1996.
28. EEC, "Measuring nonionizing electromagnetic radiation (9 kHz–300 GHz)," Recommended Practice, (02)04, 2003.
29. Warn, P. A., et al., "Infrared body temperature measurement of mice as an early predictor of death in experimental fungal infections," *Laboratory Animals*, Vol. 37, 126–131, 2003.
30. Shahryar, H. A., A. Lotfi, M. B. Ghodsi, and A. R. K. Bonary, "Effect of 900 MHz electromagnetic fields emitted from cellular phone on the T3, T4, and cortisol levels in Syrian hamsters," *Veterinary Institue in Pulawy*, Vol. 53, 233–236, 2009.
31. Weinert, D. and J. Waterhouse, "Daily activity and body temperature rhythms do not change simultaneously with age in laboratory mice," *Physiology and Behaviour*, Vol. 66, 605–612, 1999.
32. Saunders, R. D., et al., "Biological effects of exposure to non-ionising electromagnetic fields and radiation: Radiofrequency and microwave radiation," *NRPB*, London, 1991.

Influence of a Dielectric Insert of High Permittivity on the Transmit Performance of a 300 MHz Multi-channel MRI Loop Array

M. Kozlov and R. Turner

Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Abstract— Several recent reports have suggested that the transmit performance of a given MRI RF coil can be enhanced by incorporating suitable dielectric materials. To analyse this proposed approach more deeply, we investigated the effect of a dielectric material insert on loop array power balance, transmit performance and slice homogeneity. To obtain a useful database in a reasonable time, we varied only the insert permittivity, the load diameter, the separation between array and load, and the length of loop elements. For a range of parameter combinations, use of the insert increases the mean transmit RF field over the central transverse slice \mathbf{B}_{1+s} , and the homogeneity of this slice, but not the transmit performance, expressed as ratio of the mean RF field over the field of view to the square root of power deposited in the FOV, $\mathbf{B}_1 +^A / \sqrt{P_I}$. An array with the insert cannot provide better \mathbf{B}_{1+s} and homogeneity than an array without the insert in which the geometry is optimized for a given target region. We have found no general recipe for achieving the best array performance and homogeneity using the insert in an array with arbitrary geometry.

1. INTRODUCTION

Recently several reports [1, 2] have suggested that the transmit performance of a given MRI RF coil can be enhanced by incorporating suitable dielectric materials. However, these reports, which are mostly based on experimental results, do not provide detailed analysis of the power balance of the coil with and without an insert composed of such high permittivity dielectric materials, nor data regarding the transmit field homogeneity for different slices of the object. To analyse this proposed approach more deeply, we investigated the effect of the insert on loop array power balance, transmit performance and slice homogeneity. To obtain a useful database in a reasonable time, we varied only the insert permittivity, the load diameter, the separation between array and load, and the length of loop elements.

2. METHOD

We investigated 8-channel loop-type array coils with loop angular width of 37.5 degrees, having a range of diameters (from 200 mm to 280 mm) and lengths (80, 100 and 120 mm), mounted on a cylindrical acrylic former (not shown on Fig. 1) and loaded by cylinder with diameters 80, 120, 155 and 190 mm and length 375 mm. The latter minimizes the influence of load length on transverse slice homogeneity for arrays with different lengths. For all loads, electrical properties were close to those of average human tissue at 300 MHz — permittivity 52 and conductivity 0.55 S/m. A high permittivity hollow dielectric insert (in red color on Fig. 1), with permittivity ranging from 10 to 79, and with cylindrical geometry, was placed between the coil and load. In all cases, the thickness of the cylindrical acrylic former (5 mm) separated the array and the insert, and there was no air gap between the insert and load. The field of view (FOV) was taken as the part of the load located inside the array. In this arrangement, the cylindrical symmetry of the array, the insert and the load obviates performance optimization by RF shimming (adjustment of amplitude and phase for excitation signals). All arrays were excited in circular polarization mode, applying 1 W power to each port (array transmit power — $P_{transmit} = 8$ W), with a sequential 45 degree phase increment.

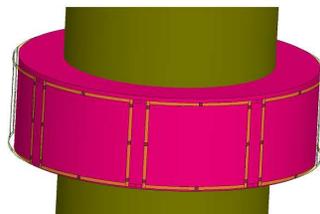


Figure 1: 3D EM setup.

For all geometries the array was tune/match/decoupled using capacitor and inductor based decoupling networks, or tuned only by the effective method of minimization of the power reflected by the entire array. We analyzed the effect of the insert on array transmit performance (calculated as the ratio of mean \mathbf{B}_{1+} over the entire FOV (\mathbf{B}_{1+}^A) to the square root of power deposited in the FOV (\mathbf{P}_l)) and on transmit magnetic field inhomogeneity (calculated as the ratio of standard deviation to mean \mathbf{B}_{1+} over a given transverse slice (\mathbf{B}_{1+s})) at three transversal slices (center and ± 25 mm from the array/load center). The Q factor of all capacitors was set equal to 1000.

3. RESULTS

For all arrays investigated, increasing the array diameter from the closest to the load to 280 mm (the largest array diameter used in the current investigation) results in decreased \mathbf{P}_l , \mathbf{B}_{1+s} and \mathbf{B}_{1+}^A . However, if the shape of FOV is the same (keeping the same array length and load diameter) the ratio of $\mathbf{B}_{1+}^A + \sqrt{\mathbf{P}_l}$ remains nearly constant (with only $\pm 4\%$ variation) for a given FOV for all arrays investigated without the insert. There is an intrinsic variation of \mathbf{B}_{1+s} and homogeneity, showing a rather complicated multi-parametric dependency on array and load setup, which we report separately.

An obvious approach to improve array performance is to find a way to increase \mathbf{P}_l for a given array and load geometry. Taking into account power balance and the knowledge that \mathbf{P}_l , the power deposited in the entire load (\mathbf{P}_{load}), the radiated power ($\mathbf{P}_{radiated}$) and the power delivered to entire coil (\mathbf{P}_{coil}) are all closely interrelated, increasing \mathbf{P}_l requires: a) minimization of the power reflected by entire array ($\mathbf{P}_{array-refl}$); b) maximization of the effective load resistance (\mathbf{R}_{eff}) seen by each loop coil; c) maximization of the ratio $\mathbf{P}_l/\mathbf{P}_{load}$.

Data obtained for some array/load geometries show that if the insert is significantly longer than the array (e.g., the same length as the load), the magnetic field propagates more strongly in parts of the load outside the array FOV. This results in a reduced ratio of $\mathbf{P}_l/\mathbf{P}_{load}$ and a decrease by 5–30% of both \mathbf{B}_{1+s} and the slice homogeneity. For this reason, most of our investigation was focused on an insert with length equal to the array length plus 5 mm, which in most cases increases \mathbf{B}_{1+s} at least for the central transverse slice.

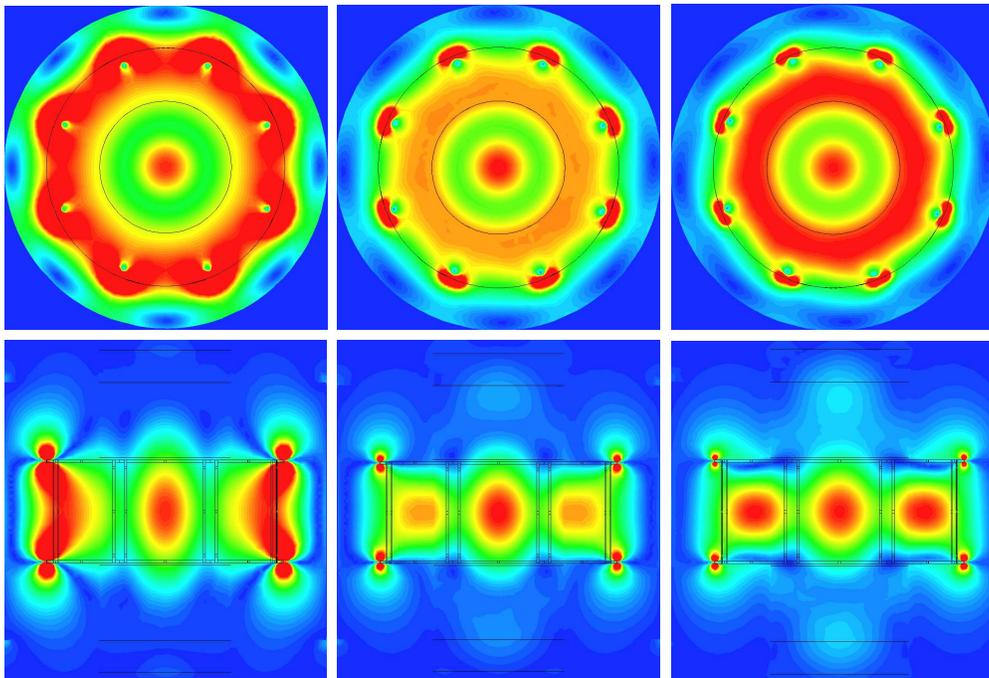


Figure 2: \mathbf{B}_{1+} profiles rescaled to individual maximum. 280 mm diameter 120 mm long array with 155 mm diameter load. From left to right: permittivity 0, 30, 52.

For this geometrical setup, the insert increases both the coupling between elements, and \mathbf{R}_{eff} . The first effect makes array decoupling (and minimization of $\mathbf{P}_{array-refl}$) more complicated, while the second makes it easier. Which effect dominates depends on the array and geometries, and the insert permittivity. The higher the permittivity of the dielectric material, the larger the increase of

both coupling and \mathbf{R}_{eff} . The smaller the load diameter and the higher the ratio of array diameter to load diameter, the higher the probability of an undesirable increase in \mathbf{P}_{array_refl} . \mathbf{P}_{array_refl} can be as high as 35 % of $\mathbf{P}_{transmit}$, for example, for a 240 mm diameter array of 80 mm length, loaded by an 80 mm diameter cylindrical load and insert permittivity of 52.

If the power reflected by the entire coil can be maintained at the same level or be further minimized (relative to original setup without the insert), use of the insert increases \mathbf{P}_{load} but can change the ratio $\mathbf{P}_l/\mathbf{P}_{load}$ in either direction. Furthermore, increased \mathbf{P}_l is a necessary but not sufficient condition for array performance improvement (Tables 1–4). While the ratio $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ remains stable or increases only slightly (not more than 4%) for a range of insert permittivity this value can decrease (or even drops significantly) if the insert permittivity is further increased. As a result the array performance depends on the behavior of both \mathbf{P}_l and $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ and can decrease, despite increasing \mathbf{P}_l . The value of insert permittivity at which the slope of $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ changes depends on the array/load geometry.

Our study suggests that for an 80 mm long array and 155 mm diameter mm load, no configuration

Table 1: Data for 280 mm diameter 120 mm long array with 155 mm diameter load.

| insert permittivity | | 0 | | | | 20 | | | | 30 | | | | 52 | | | |
|---------------------|-----------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|
| tuning | slice, mm | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l |
| C | 0 | 1.39 | 20 | 0.54 | 4.33 | 1.75 | 17 | 0.54 | 5.29 | 1.75 | 16 | 0.51 | 5.36 | 1.41 | 15 | 0.39 | 5.07 |
| | +25 | 1.25 | 21 | | | 1.47 | 19 | | | 1.42 | 19 | | | 1.06 | 18 | | |
| R | 0 | 1.39 | 20 | 0.54 | 4.30 | 1.74 | 17 | 0.54 | 5.17 | 1.80 | 16 | 0.51 | 5.60 | 1.47 | 15 | 0.40 | 5.13 |
| | +25 | 1.25 | 21 | | | 1.46 | 19 | | | 1.46 | 19 | | | 1.12 | 18 | | |
| I | 0 | 1.37 | 20 | 0.54 | 4.18 | 1.79 | 17 | 0.54 | 5.46 | 1.79 | 16 | 0.51 | 5.56 | 1.51 | 15 | 0.41 | 5.18 |
| | +25 | 1.24 | 21 | | | 1.50 | 19 | | | 1.46 | 19 | | | 1.17 | 18 | | |

Table 2: Data for 280 mm diameter 80 mm high array with 155 mm diameter load.

| insert permittivity | | 0 | | | | 20 | | | | 30 | | | | 52 | | | |
|---------------------|-----------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|
| tuning | slice, mm | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l |
| C | 0 | 1.49 | 18 | 0.76 | 2.43 | 1.85 | 17 | 0.76 | 3.76 | 1.97 | 17 | 0.76 | 4.07 | 2.03 | 18 | 0.71 | 4.36 |
| | +25 | 1.26 | 20 | | | 1.44 | 19 | | | 1.48 | 18 | | | 1.44 | 17 | | |
| R | 0 | 1.48 | 18 | 0.76 | 2.48 | 1.87 | 17 | 0.76 | 3.82 | 1.99 | 17 | 0.75 | 4.16 | 2.04 | 18 | 0.71 | 4.41 |
| | +25 | 1.24 | 20 | | | 1.45 | 19 | | | 1.49 | 18 | | | 1.44 | 18 | | |
| I | 0 | 1.46 | 18 | 0.76 | 2.42 | 1.84 | 17 | 0.76 | 3.73 | 1.97 | 17 | 0.75 | 4.06 | 2.04 | 18 | 0.71 | 4.41 |
| | +25 | 1.25 | 20 | | | 1.43 | 19 | | | 1.47 | 18 | | | 1.43 | 18 | | |

Table 3: Data for 200 mm diameter 80 mm high array with 155 mm diameter load.

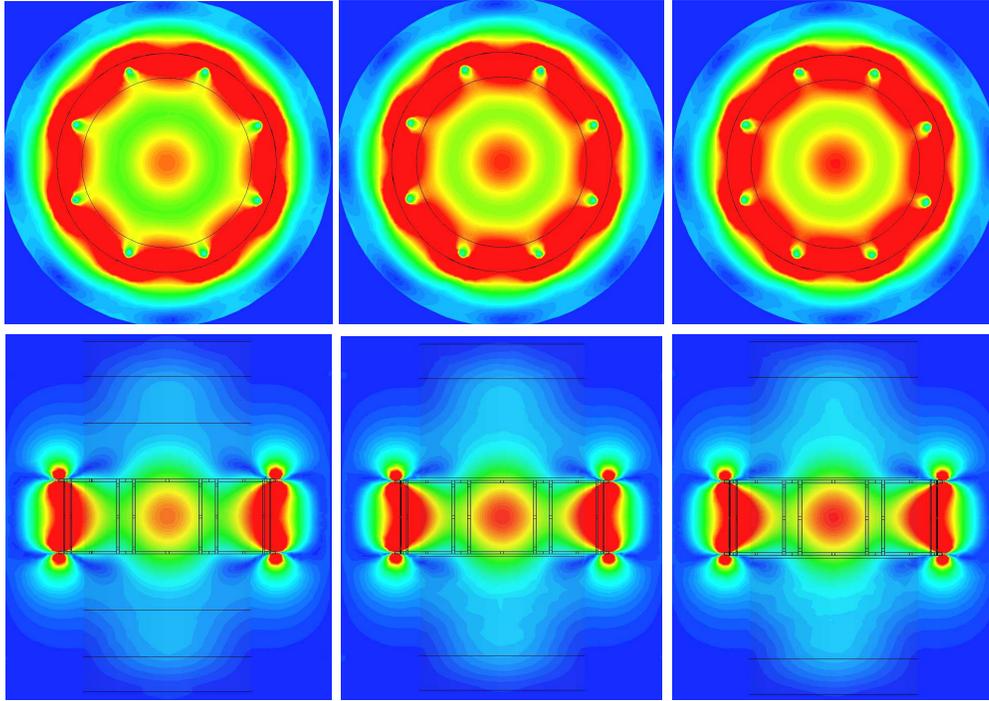
| insert permittivity | | 0 | | | | 20 | | | | 30 | | | | 52 | | | |
|---------------------|-----------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|
| tuning | slice, mm | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l |
| C | 0 | 1.88 | 15 | 0.74 | 3.96 | 1.95 | 15 | 0.74 | 3.98 | 1.98 | 16 | 0.75 | 4.00 | 1.98 | 16 | 0.73 | 4.00 |
| | +25 | 1.45 | 19 | | | 1.46 | 19 | | | 1.47 | 19 | | | 1.45 | 20 | | |
| R | 0 | 1.87 | 15 | 0.74 | 3.98 | 1.96 | 16 | 0.74 | 3.99 | 2.0 | 16 | 0.75 | 4.00 | 2.02 | 17 | 0.75 | 3.96 |
| | +25 | 1.44 | 20 | | | 1.48 | 19 | | | 1.48 | 19 | | | 1.45 | 19 | | |
| I | 0 | 1.87 | 15 | 0.74 | 3.98 | 1.97 | 16 | 0.74 | 4.00 | 2.01 | 16 | 0.76 | 4.01 | 2.04 | 17 | 0.75 | 4.00 |
| | +25 | 1.44 | 20 | | | 1.48 | 19 | | | 1.48 | 19 | | | 1.46 | 19 | | |

Table 4: Data for 240 mm diameter 80 mm high array with 80 mm diameter load.

| insert permittivity | | 0 | | | | 20 | | | | 30 | | | | 52 | | | |
|---------------------|-----------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|
| tuning | slice, mm | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + / \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l |
| C | 0 | 3.69 | 19 | 1.74 | 3.61 | 4.31 | 14 | 1.77 | 3.71 | 4.07 | 12 | 1.70 | 3.34 | 4.42 | 10 | 1.62 | 3.74 |
| | +25 | 3.27 | 20 | | | 3.67 | 17 | | | 3.30 | 16 | | | 3.05 | 15 | | |
| R | 0 | 3.56 | 19 | 1.74 | 3.35 | 5.00 | 14 | 1.81 | 4.76 | 5.24 | 12 | 1.76 | 5.06 | 5.22 | 10 | 1.63 | 5.18 |
| | +25 | 3.16 | 20 | | | 3.91 | 17 | | | 3.96 | 16 | | | 3.56 | 15 | | |
| I | 0 | 3.49 | 19 | 1.74 | 3.23 | 4.14 | 14 | 1.81 | 3.26 | 4.38 | 12 | 1.77 | 3.53 | 4.54 | 10 | 1.63 | 3.93 |
| | +25 | 3.09 | 20 | | | 3.19 | 17 | | | 3.27 | 16 | | | 3.12 | 15 | | |

Table 5: Data for 200 mm diameter 80 mm high array with 120 mm diameter load.

| permittivity | 0 | | | | 20 | | | | 30 | | | | 52 | | | | 79 | | | | |
|--------------|-------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|--------------------|----|--|----------------|
| Tuning | slice | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l | \mathbf{B}_{1+s} | % | $\mathbf{B}_1^A + \sqrt{\mathbf{P}_l}$ | \mathbf{P}_l |
| C | 0 | 2.41 | 18 | 0.95 | 4.34 | 2.58 | 15 | 0.98 | 4.34 | 2.63 | 14 | 0.98 | 4.32 | 2.65 | 13 | 0.96 | 4.28 | 2.63 | 11 | 0.93 | 4.26 |
| | +25 | 1.94 | 22 | | | 2.01 | 21 | | | 1.99 | 21 | | | 1.92 | 21 | | | 1.78 | 20 | | |
| R | 0 | 2.43 | 18 | 0.95 | 4.44 | 2.66 | 15 | 0.98 | 4.57 | 2.70 | 14 | 0.98 | 4.53 | 2.71 | 13 | 0.97 | 4.39 | 2.65 | 11 | 0.93 | 4.25 |
| | +25 | 1.97 | 22 | | | 2.04 | 21 | | | 2.03 | 21 | | | 1.95 | 21 | | | 1.82 | 20 | | |
| I | 0 | 2.36 | 18 | 0.95 | 4.21 | 2.58 | 15 | 0.98 | 4.32 | 2.63 | 14 | 0.98 | 4.31 | 2.66 | 13 | 0.97 | 4.26 | 2.64 | 11 | 0.93 | 4.23 |
| | +25 | 1.91 | 22 | | | 1.98 | 21 | | | 1.98 | 21 | | | 1.93 | 21 | | | 1.82 | 20 | | |


 Figure 3: \mathbf{B}_{1+} profiles rescaled to individual maximum. 200 mm diameter 80 mm high array with 155 mm diameter load. From left to right: permittivity 0, 30, 52.

exists where the use of the insert improves homogeneity by more than 10%. Usually we obtained the same or slightly poorer homogeneity. But for a 120 mm long array and all loads, or an 80 mm long array and 80 and 120 mm diameter loads, use of the insert can improve homogeneity by more than 45%.

Comparing data from Table 1 with the corresponding data ($\mathbf{B}_{1+s} = 1.58$ for the central slice and $\mathbf{B}_{1+s} = 1.46 \mu\text{T}$ for the ± 25 mm slices; inhomogeneity = 32 for the central slice and 36 for the ± 25 mm slices) for an array with the same length and diameter that is the closest to load, one can conclude that use of the insert essentially increases \mathbf{B}_{1+s} for the central slice and reduces inhomogeneity for all slices. Taking into account that $\mathbf{B}_1^A + \sqrt{\mathbf{P}_l}$ is not increased, this results in a smaller \mathbf{B}_{1+} for peripheral parts of the FOV.

If one is interested only in a volume ± 25 mm from load center, then Table 1 data would be best compared with that for a 200 mm diameter 80 mm long array that generates $\mathbf{B}_{1+s} = 1.87/1.44 \mu\text{T}$ and inhomogeneity = 16/20 (for central and ± 25 mm slices respectively). Including in our consideration data from Table 2 and Table 3 (as well as other data available, but not presented due to limited space) we conclude that for a given insert geometry: a) an array with an insert cannot provide significantly better \mathbf{B}_{1+s} and homogeneity than an array without an insert for which the geometry has been optimized for a given target region; b) there is no point in using the insert if array length is less than half the load diameter, and the separation between array and load is relative small (less than 1/4 of load diameter). In the latter case, use of the insert has minimal influence on both $\mathbf{B}_1^A + \sqrt{\mathbf{P}_l}$ and \mathbf{P}_l , but the ratio $\mathbf{P}_l/\mathbf{P}_{load}$ decreases as the insert permittivity increases.

4. CONCLUSIONS

Use of a high permittivity dielectric material insert is not a general approach for improving array performance and homogeneity. In some conditions use of the insert significantly increases \mathbf{B}_{1+s} and transverse slice homogeneity, but not the transmit performance, considered as the ratio $\mathbf{B}_{1+A}/\sqrt{\mathbf{P}_I}$. An array with an insert cannot provide significantly better \mathbf{B}_{1+s} and homogeneity than an array without an insert once the geometry has been optimized for a given target region. We have found no general recipe for achieving the best array performance and homogeneity using an insert in an array with a wide range of geometries. Investigation of array using only central transverse slice data is an unreliable guide for optimization, if performance and homogeneity over most of the FOV is important.

REFERENCES

1. Yang, Q. X., W. Mao, J. Wang, M. B. Smith, H. Lei, X. Zhang, K. Ugurbil, and W. Chen, "Manipulation of image intensity distribution at 7.0 T: Passive RF shimming and focusing with dielectric materials," *Journal of Magnetic Resonance Imaging*, No. 24, 197–202, 2006.
2. Haines, K., N. B. Smith, and A. G. Webb, "New high dielectric constant materials for tailoring the B_{+1} distribution at high magnetic fields," *Journal of Magnetic Resonance*, No. 203, 323–327, 2010.
3. Kozlov, M. and R. Turner, "Fast MRI coil analysis based on 3-D electromagnetic and RF circuit co-simulation," *Journal of Magnetic Resonance*, No. 200, 147–152, 2009.

Electromagnetic Compatibility between Implantable Cardiac Pacemakers and RFID Systems: Experimental Set-up, Test Protocol and Preliminary Results

E. Mattei¹, G. Calcagnini¹, F. Censi¹, M. Triventi¹, C. Desantis², P. Menna², and P. Bartolini¹

¹Italian National Institute of Health (ISS), Italy

²“Sapienza” University of Rome, Italy

Abstract— An experimental set-up for the evaluation of the electromagnetic interference (EMI) between implantable cardiac pacemakers (PMs) and radiofrequency identification (RFID) readers operating in the low-frequency (LF — 125 kHz) and high frequency (HF — 13.56 MHz) range is described. Two development kits were used to generate the RFID signal at 125 kHz and 13.56 MHz, and the EMI tests were performed with and without the presence of the RFID tag (passive). Eight PMs from 5 manufactures were tested inside a human torso simulator filled with a saline solution appropriate for each frequency. No significant degradations of the basic performances of the device were observed. Noteworthy, the field strength generated by the RFID readers used in this study is lower than the maximum values allowed by the RFID international standards: the magnetic field generated by the two RFID emitters during the EMI tests reached a peak value at 1 cm from the RFID antenna of 35.3 A/m at 125 kHz and 0.024 A/m at 13.56 MHz. The adopted experimental set-up and test protocol is suitable for further investigations, using other types of RFID readers and higher values of field.

1. INTRODUCTION

Radiofrequency identification (RFID) emissions have the potential to affect electronic devices. Particular care has to be paid for the electromagnetic interference (EMI) to implantable pacemakers (PMs) and implantable cardioverter-defibrillators (ICDs). A recent study conducted by the Food and Drug Administration in collaboration with major implantable PM and ICD manufacturers [1], demonstrates the effects of emissions from RFID readers on common implantable cardiac devices. Significant effects were observed especially for low frequency RFID readers, and the authors concludes that “are concerned that the continued proliferation of RFID without taking electromagnetic interference into consideration could cause clinically significant events for patients”, even if they “do not believe the current situation reveals an urgent public health risk”. Other comprehensive studies on the electromagnetic compatibility between implantable devices and RFID systems are scarce. The effects of the RFID signal have been wider investigated for not-implanted medical devices. In 2008, Van der Togt et al. [2] published a study on the interference between RFID readers and critical care medical equipment: in 123 tests, RFID induced 34 EMI incidents, among which 22 were classified as hazardous, 2 as significant, and 10 as light. The median distance between the RFID reader and the medical device in all incidents was 30 cm (range, 0.1–600 cm). In the same year, another group [3] performed EMI tests using an experimental set-up similar to the one adopted by Van der Togt et al., but, over more than 1500 tests, no significant effects on the performances of the medical devices were observed. A recent study [4] has also shown that the effect of the RFID signal depends on the presence or not of the RFID tag. This implies that the EMI tests must be also performed with a close-loop transmission between the RFID reader and tag. It seems thus clear how, given the rapid expansion of RFID technology in many settings of our every-day life and the lack of exhaustive data on the potential effects on electronic medical devices, further studies in this field are needed.

In this paper, an experimental in-vitro set-up for the evaluation of the electromagnetic interference (EMI) between implantable cardiac PMs and RFID readers, operating in the low-frequency (LF — 125 kHz) and high frequency (HF — 13.56 MHz) range, is proposed. The results of EMI tests on 8 PMs from 5 manufactures (Sorin, Medico, Biotronik, Medtronic, S. Jude) are also presented.

2. METHODS AND MATERIALS

The set-up and the test protocol adopted for the evaluation of the electromagnetic interferences between RFID systems and PMs are derived from the standard ANSI/AAMI PC69:2007 [5]. However, since this standard specifically addresses electromagnetic compatibility issues at high frequencies

(> 400 MHz), proper changes were needed, to better take into account the coupling mechanisms with electromagnetic fields at lower frequencies.

A human torso simulator (PVC rectangular box $-60 \times 40 \times 15 \text{ cm}^3 \sim 28.61$ inner volume) was used to host the PM and its leads. The torso simulator was filled with a saline solution whose concentration was properly chosen in order to account for the dielectric properties of the human tissues at the frequencies of interest. In particular, a conductivity of 0.2 S/m and 0.35 S/m was chosen for the tests in the LF and HF ranges, respectively.

The PM and the lead were fixed over a PVC grid ($20 \times 38 \text{ cm}^2$) and the lead path was arranged to form a loop with an area similar to the one of a realistic worst-case PM implant (about 225 cm^2 for a unipolar stimulation). The torso simulator was equipped with two couples of stainless steel plates mounted at the center of each of the 4 inner walls of the rectangular box. The first couple was used to monitor the PM activity; the second couple was used to simulate the electrical activity of the heart, in such a way that the PM recognizes it as an inhibition signal (Figure 1). For each PM, two tests were performed: in the first one, the inhibition signal was OFF and EMI was considered to occur if, after the activation of the RFID signal, the deviation in pace-to-pace interval exceeded 10% of the programmed rate; in the second test, the simulated heart signal was ON and no EMI was registered if the PM did not exhibit any pace pulse during application of the inhibition signal and RF signals. The monitoring of the PM activity as well as the generation of the inhibition signal was obtained by an acquisition/generation card (NI DAQ Card PCI 6052E — National Instruments, US) connected to a PC. A software interface developed in Labview (National Instruments, US) allowed to monitor in real-time the activity of the PM, to generate the simulated heart signal and to verify the correct behavior of the devices during the EMI tests. The programmed parameters of the PMs were kept as set by the manufactures. Six PMs were tested in the unipolar pacing/sensing modality, whereas 2 in the bipolar modality. As RFID sources, two development kits operating at the frequency of 125 kHz (Melexis, DVK90109) and 13.56 MHz (Texas Instruments, TRF7960) were used. Both the systems can communicate with passive tags; a PC-based interface was used to set the desired transmission protocol and to establish a close-loop communication between the reader and the tag.

The flow chart of the test protocol is reported in Figure 2. The initial distance between the reader and the PM was 10 cm. If no EMI phenomena occurred, the distance was reduced by a step of 2.5 cm, till a minimum distance of 2.5 cm from the device. The EMI tests were first performed with just the reader transmitting (open-loop signal); then with a tag placed near the PM and sending its ID to the reader (close-loop signal). In order to correlate possible EMI effects to the field strength generated by the RFID emitters, the magnetic fields generated by the two RFID emitters were measured at 1 cm and 2 cm from the RFID antenna, with and without the presence of the tag. Measurements at 13.56 MHz were performed by a near field H-probe (Rohde & Schwarz, HZ-14) connected to a spectrum analyzer (Advantest U3641). At 125 kHz, the magnetic field was evaluated by measuring the voltage induced on a custom-made loop (single turn — diameter = 1.8 cm) connected to a high-impedance oscilloscope (LeCroy WavePro 7000A). The magnetic

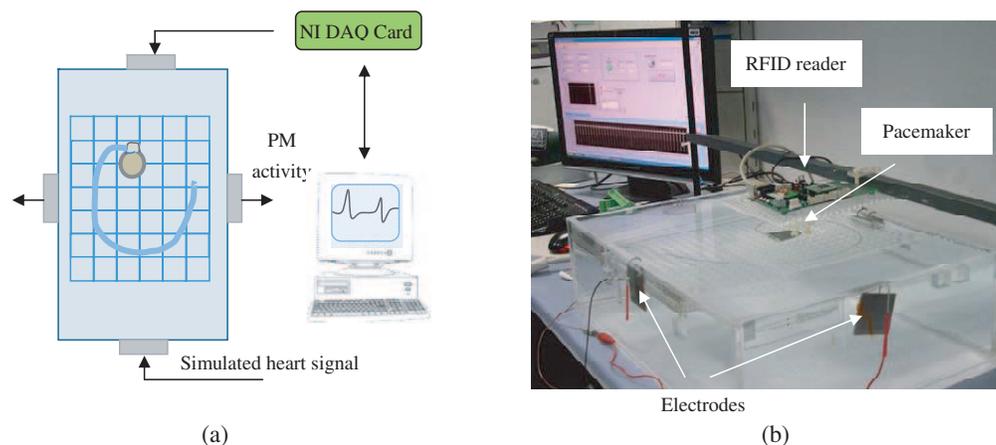


Figure 1: Experimental set up for the EMI test: (a) schematic representation and (b) picture of the actual set-up.

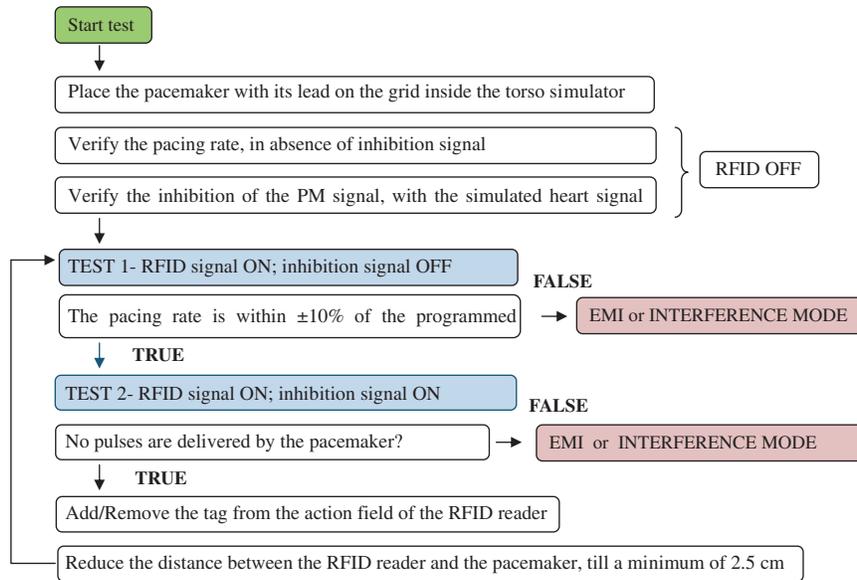


Figure 2: Flow chart of the test protocol.

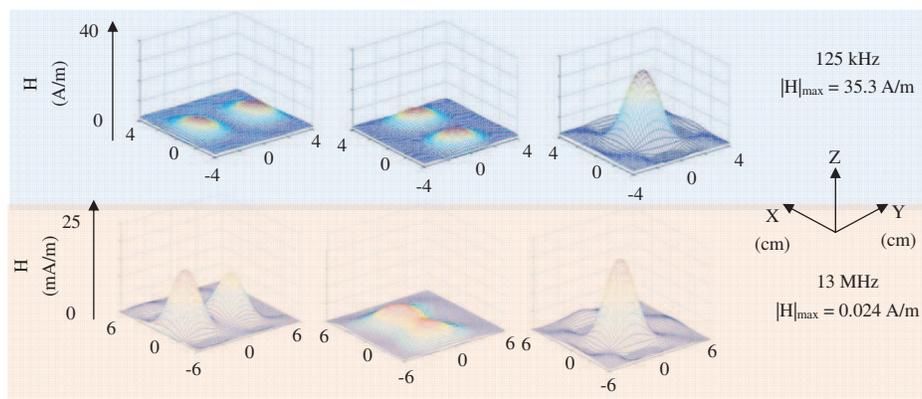


Figure 3: Spatial components of the magnetic field distribution at 1 cm from the RFID antennas: 125 kHz (upper panel); 13.56 MHz (lower panel).

field was obtained according to the relation:

$$H_{rms} = \frac{B_{rms}}{\mu_0} = \frac{V_{rms}}{2\pi f S} \quad (1)$$

where H_{rms} is the RMS value of the magnetic field strength, B_{rms} the magnetic field, μ_0 the magnetic permeability of free space ($4\pi \cdot 10^{-7} \text{ N/A}^2$), V_{rms} the RMS voltage induced in the loop, f the frequency of the time-varying signal (i.e., 125 kHz) and S the area of the loop.

3. RESULTS

The magnetic field distribution at 1 cm from the RFID antenna is reported in Figure 3: a peak value of 35.3 A/m at 125 kHz and 0.024 A/m at 13 MHz was observed. The presence of the tag does not substantially modify the magnetic field strength around the antenna. The magnetic field peak values with the tag placed next to the RFID reader become 34.9 A/m at 125 kHz and 0.023 A/m at 13.56 MHz. These differences are lower than the uncertainty of the measurement itself ($\pm 0.3 \text{ A/m}$ for the LF; $\pm 2 \text{ dB}$ for the HF). In all the EMI tests, no significant degradations of the basic performances of the PMs were observed, even when the RFID transmitter was placed in close proximity of the device. During the tests without the simulated heart signal, the pulse rate of the PM was not affected by the activation of the RFID signal, both for the LF and the HF reader. When the inhibition signal was applied to the torso simulator, the PM correctly recognized it and stopped delivering the pulses, even with the RFID readers ON. The presence of the tag was not

related to any particular effects caused by the RFID transmitters on the devices.

4. DISCUSSION

The definition of a proper experimental set-up and test protocol for the evaluation of the possible EMI phenomena induced by RFID systems on implantable PMs is a crucial aspect for the assessment of immunity levels/safety distances to be adopted to ensure the safety of the patient. Our proposal is based on the standard ANSI/AAMI PC69:2007, which provides indications on how to perform EMI tests in order to verify the electromagnetic compatibility of implantable PMs and ICDs in the frequency range between 450 MHz and 3 GHz. The LF and HF RFID systems transmit over lower frequency ranges and the coupling mechanism they used to communicate with the tag is much more an inductive coupling than an electromagnetic wave propagation. Appropriate modification of the standard are thus needed. First, the electrical conductivity of the saline solution has to be modified to account for the frequency dependences of the different dielectric of the human tissues. Another important aspect is the implant configuration during the test: since the RFID transmission at LF and HF generally provides the data transfer through a time-varying magnetic field, the PM lead path must be chosen to maximize the area covered by the implant. In addition, since the antennas of LF and HF RFID reader is generally a coil that generates the time-varying magnetic field, the EMI test has to be performed with the axes of the antenna coil perpendicular to the plane where the implant is reproduced. There is no need to perform additional tests with different orientations of the antenna. In the preliminary EMI tests that were performed no significant degradations of the basic performances of the PMs were observed. For the 8 PMs tested, the pulse rate as well as the sensing activity was not affected by the activation of the RFID signal. However, the field strength generated by the RFID readers used in these tests is not the maximum output level allowed. This may explain the differences respect to the results found by other groups [1].

5. CONCLUSION

The experimental set-up and test protocol described in this paper is suitable for the evaluation of the electromagnetic compatibility between PM and RFID system operating in the LF and HF range. In the preliminary tests that were performed, no EMI phenomena were observed. Noteworthy, the field strength generated by the RFID readers used in this study is lower than the maximum values allowed by the RFID international standards. The same experimental set up and test method can be adopted for further investigations, using other types of RFID readers and higher values of field.

REFERENCES

1. Seidman, J. S., R. Brockman, B. M. Lewis, J. Guag, M. J. Shein, M. J. Clement, J. Kippola, D. Digby, C. Barber, and D. Huntwork, "In vitro tests reveal sample radiofrequency identification readers inducing clinically significant electromagnetic interference to implantable pacemakers and implantable cardioverter-defibrillators," *Heart Rhythm*, Vol. 7, 100–107, 2010.
2. Van der Togt, R., E. J. van Lieshout, R. Hensbroek, E. Beinat, J. M. Binnekade, and P. J. M. Bakker, "Electromagnetic interference from radio frequency identification inducing potentially hazardous incidents in critical care medical equipment," *JAMA*, Vol. 299, 2884–2890, 2008.
3. Christe, B., E. Cooney, G. Maggioli, D. Doty, R. Frye, and J. Short, "Testing potential interference with RFID. Usage in the patient care environment," *Management & Technology*, 479–484, 2008.
4. Houlston, B., "Interference with the operation of medical devices resulting from the use of radio frequency identification technology," *NZMJ*, Vol. 122, No. 1297, 2009.
5. ANSI/AAMI PC69:2007, "Active implantable medical devices — Electromagnetic compatibility – EMC test protocols for implantable cardiac pacemakers and implantable cardioverter defibrillators," Association for the Advancement of Medical Instrumentation, 2007.

Impedance Variation of an Equipment under Test in a GTEM Cell

David Pouhè

Technical University Berlin, Einsteinufer 25, Berlin 10587, Germany

Abstract— This paper investigates the influence of the reactive effects of the GTEM cell on an EUT using the MoM. The approach used here is based on a framework recently developed. Emphasis is put on the impedance variation of the EUT which is a measure par excellence of the coupling between cell and EUT. The appeal of the method resides in its simplicity and its applicability.

1. INTRODUCTION

In a recent paper [1], the interaction between an equipment under test (EUT) and the cell reflecting metal walls as represented by images of the EUT has been investigated by restricting the scope of the study to the total incident field and its subsequent deviation from the primary excited TEM field, to the current on the EUT and its deviation when compared to the induced current in free space. The impedance variation of the EUT was set arbitrary as being constant and equal to 10% the value of the impedance the EUT would have in free space. It was essentially shown that strong and weak test conditions may occur since the overall impinging field on the EUT generally deviates from the primary TEM field. In addition to that, the deviations in field and the errors in induced current were quantified. The role that excited evanescent modes plays was neglected. At high frequencies however, the role of evanescent modes in both the impedance variation and the change in current on the EUT might be of no less importance and can therefore no longer be ignored. Hence, it is expedient to know which influence these waves can have on the overall impedance and on the surface current on an EUT placed in the cell. We address this issue in the present work. Especially, the approach introduced in [1] is generalized upon using the method of moment (MOM).

To underline the phenomenological principles of the interplay between the EUT and the cell by imparting deep insights into the physics behind the phenomena being investigated but also for the sake of completeness and to assume a continuous and smooth reading from one Section to another, the second section on Modal Constituents in [1] is partially reproduced in Section 2. In Section 3, due attention is given to the impedance variation as a measure of the coupling between EUT and the cell.

2. EXCITED MODES IN A GTEM CELL

As in [1], let us consider an EUT placed at a distance r_c from the apex in a GTEM cell. An incident TEM wave propagating in the positive r -direction from the apex to the hybrid termination illuminates the EUT and induces currents on the EUT, which in turn act as sources of secondary fields. Generally, the number of these secondary modes is infinite but through double symmetry with respect to the geometry of the EUT and to its position in the cell, a large number of these modes will not be excited (see Table 1). If in addition to that, the frequency range of interest is restricted to the frequencies where, beside the TEM wave, only a small number (one or two) of higher order modes propagate within the cell, their influence can clearly be seen. For these reasons and also for the sake of more clarity in the development of the approach, we will increase the number of propagating secondary higher order modes excited by the scatterer progressively. First, it is assumed in the following that the propagating scattered fields only consist of the reflected TEM wave and the TE_{10} -mode. Next, the number of propagating modes is increased to three, consisting thereby of the TEM mode, the TE_{10} and TE_{30} modes. In each of the two cases, all other higher-order modes are evanescent. Nevertheless, excited nonpropagating modes in the vicinity of the EUT are considered in the analysis as they play a role of no less importance in the variation of the impedance and current density on the EUT and thereafter in the fulfilment of the boundary conditions on the scatterer.

Due to the tapered shape of the cell, the higher-order modes propagating in the negative r -direction are totally reflected at their respective cutoff point and re-illuminate the EUT with a certain phase delay. This indeed means that the total field incident on the EUT is the superposition of the TEM mode and the propagating secondary higher-order waves. The number of these secondary fields depends of course on the considered frequency range. The frequency range may

Table 1: Cutoff frequencies of first higher order modes at a distance $r_c = 15$ m from the apex. $a = 063$ m, $b_1 = 031$ m.

| n ↓ \ m → | f_c [GHz] | | |
|-----------|----------------|-------|----------------|
| | 1 | 2 | 3 |
| 0 | 0.209* | 0.373 | 0.592* |
| 1 | 0.512 | 0.606 | 0.753 |
| 2 | 0.894** | 1.005 | 1.024** |

Values marked bold are modes which may be excited by a double symmetrical arranged dipole.

$$f_c = \sqrt{\eta(\eta + 1)/(2\pi r \sqrt{\varepsilon_0 \mu_0})}, \text{ whereby } \eta = \nu + 1/2 \text{ is used [2].}$$

* Excited TE_{mn} modes. ** Excited TM_{mn} modes.

Table 2: Index ν and corresponding modes [3].

| Mode | TE_{10} | TM_{11} | TE_{20} | TE_{11} | TE_{30} | TE_{21} | TM_{21} | TM_{12} | TM_{31} |
|-------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| ν | 5.6120 | 11.5721 | 11.5986 | 15.6976 | 17.6382 | 23.0681 | 24.2899 | 27.0515 | 27.8518 |

be chosen so to have only one higher-order mode propagating beside the principal mode [1]; or two and more as illustrated in Fig. 2 for of three propagating modes within the cell.

For frequencies where only TEM and TE_{10} modes propagate within the cell, the total field impinging the EUT is

$$\mathbf{E}_T^i = \mathbf{E}_{TEM}^0 + \mathbf{E}_{TE_{10}}^{re} \quad (1)$$

If the frequency range is chosen such that three modes (TEM, TE_{10} and TE_{30}) propagate inside the cell, the total field incident on the EUT is then

$$\mathbf{E}_T^i = \mathbf{E}_{TEM}^0 + \mathbf{E}_{TE_{10}}^{re} + \mathbf{E}_{TE_{30}}^{re} \quad (2)$$

where $\mathbf{E}_{TE_{10}}^{re}$ and $\mathbf{E}_{TE_{30}}^{re}$ are the fields propagating in the reverse direction to their respective counterparts in the forward direction. They are related to each other as follow

$$\mathbf{E}_{TE_p}^{re} = \mathbf{E}_{TE_p}^{f-} \cdot e^{j\psi_p} \quad (3)$$

In (3), ψ_p is the frequency dependent phase difference between $\mathbf{E}_{TE_p}^{f-}$ and $\mathbf{E}_{TE_p}^{re}$ defined as

$$\psi_p = 2\Theta(k_0 r) \quad (4)$$

whereby

$$\Theta(k_0 r) = \arctan\left(\frac{N_\nu(k_0 r)}{J_\nu(k_0 r)}\right), \quad (5)$$

with J_ν and N_ν being Bessel functions of the first and second kind of index ν respectively. Table 2 gives the index ν and the corresponding mode. $k_0 = 2\pi f \sqrt{\varepsilon_0 \mu_0}$ is the wave number. For simplicity, the double subscript mn ($m = 1, n = 0$ for the TE_{10} mode and $m = 3, n = 0$ for the TE_{30} mode) has been replaced by p .

TE_{10} and TE_{30} modes have of course different cutoff points. But this is less relevant for our analysis, since we are primarily interested in their contribution to the total field impinging the EUT and the subsequent effect on the scatterer.

The forward direction for higher order modes generated by the EUT throughout this paper is the direction from the EUT to the cell's boundaries. Note that, propagating modes excited by the EUT and scattered back from the waveguide conducting boundaries are accounted in the reflected waves. It is also recalled that for the sake of simplicity, the explicit dependence of the field vectors \mathbf{E} and \mathbf{H} on cross-sectional coordinates as well as on the longitudinal coordinate is suppressed when writing the equations. The time harmonic dependence $\exp(j\omega t)$ is also omitted for the same reason.

Perfect termination of the cell being assumed, and restricting for example the bandwidth of interest to frequencies where the propagating modes consist of only the principal mode and the first higher-order excited mode, i.e., the TE_{10} mode, the same approach of subdividing the cell into

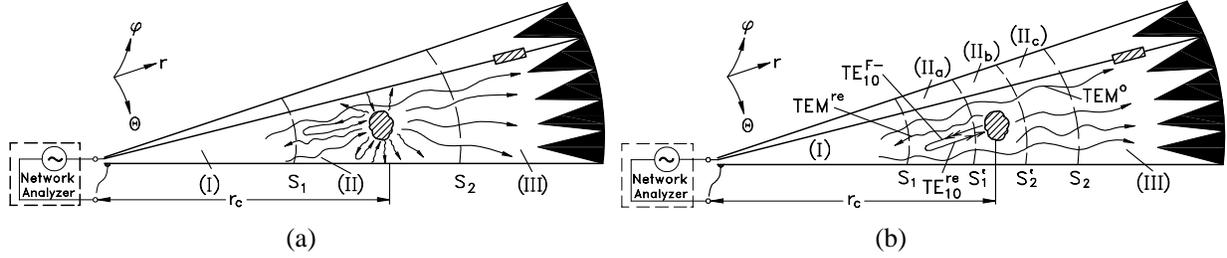


Figure 1: (a) Scattering of a small EUT in a GTEM cell and (b) the sketch of the calculating model.

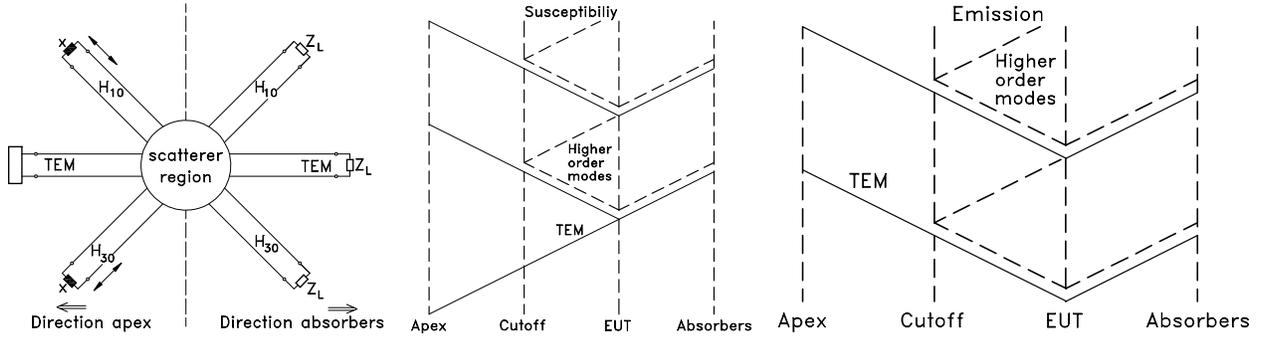


Figure 2: Schematic diagram showing three propagating modes within the cell. In the direction of the rear wall, all modes are absorbed. In the opposite direction, higher order modes are totally reflected at cutoff and re-illuminate the EUT. The phenomenon of illumination and re-illumination results in standing waves between EUT and cutoff point. Continuous lines denote the TEM wave and (non-vertical) dashed lines delineate higher-order modes.

regions (Fig. 1) developed in [1] can be followed. It leads to identical modal constituents described by Equations (1)–(3) in [1] for regions I, IIa and III¹.

We remind that no attempt was made in [1] to describe fields in region IIb, as it encompasses propagating and evanescent modes. But since in this paper we are merely interested in the impedance behavior of the EUT, we need to address the issue of describing fields in domain IIb.

In the vicinity of the EUT, the propagating modes are assumed dominant and nearly uniform over the source region. Nevertheless, the contribution of evanescent modes to the capacitive or inductive behavior of the EUT cannot be ignored. The overall field in this domain is therefore the sum of the incident field and all excited modes in Table 1. We then have

$$\mathbf{E}_T^{IIb} = \mathbf{E}^i + \mathbf{E}^s \quad (6)$$

where

$$\mathbf{E}^s = \mathbf{E}_{TEM}^{re} + \mathbf{E}_{TE_{10}} + \mathbf{E}_{TE_{10}} + \mathbf{E}_{TM_{12}} + \mathbf{E}_{TM_{32}} \quad (7)$$

Explicit expressions for the field components of the TEM and higher order TE and TM modes are found in [2].

3. IMPEDANCE VARIATION OF AN THE EUT

The phenomenon of illumination and reillumination of the EUT owing to the multiple reflections within the cell and to its tapering structure and generally known as the reactive effect of the cell implies that there is an interactive coupling between the GTEM cell and the EUT. The impact of this coupling can be captured by determining the variation of the impedance of the EUT when compared to the impedance it would have in free space.

The fields \mathbf{E}^s scattered by an EUT due to an incident field \mathbf{E}^i can be viewed as a reaction, between the test source and the actual source. A measure of this coupling is a reaction equation which in a MoM's approach may be written as [4, 5]

$$\sum_{n=1}^N I_n \langle \mathbf{F}_m, \mathcal{L}\mathbf{F}_n \rangle = \langle \mathbf{F}_m, \mathbf{E}^i \rangle \quad (8)$$

¹To avoid repetition, an explicit reproduction of the related mathematical expressions is omitted here. Readers are kindly reported to Section 2 in [1].

In a matrix form, (8) becomes

$$[Z_{mn}] \cdot [I_n] = [V_m] \quad (9)$$

where the generalized impedance matrix $[Z_{mn}]$, the generalized current vector $[I_n]$ and the generalized voltage $[V_m]$ are

$$[Z_{mn}] = \begin{pmatrix} \langle \mathbf{F}_1, \mathcal{L}\mathbf{F}_1 \rangle & \langle \mathbf{F}_1, \mathcal{L}\mathbf{F}_2 \rangle & \dots & \langle \mathbf{F}_1, \mathcal{L}\mathbf{F}_N \rangle \\ \langle \mathbf{F}_2, \mathcal{L}\mathbf{F}_1 \rangle & \langle \mathbf{F}_2, \mathcal{L}\mathbf{F}_2 \rangle & \dots & \langle \mathbf{F}_2, \mathcal{L}\mathbf{F}_N \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{F}_N, \mathcal{L}\mathbf{F}_1 \rangle & \langle \mathbf{F}_N, \mathcal{L}\mathbf{F}_2 \rangle & \dots & \langle \mathbf{F}_N, \mathcal{L}\mathbf{F}_N \rangle \end{pmatrix} = \begin{pmatrix} Z_{11} & Z_{12} & \dots & Z_{1N} \\ Z_{21} & Z_{22} & \dots & Z_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & Z_{N2} & \dots & Z_{NN} \end{pmatrix} \quad (10)$$

$$[I_n] = \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{pmatrix} \quad [V_m] = \begin{pmatrix} \langle \mathbf{F}_1, \mathbf{E}^i \rangle \\ \langle \mathbf{F}_2, \mathbf{E}^i \rangle \\ \vdots \\ \langle \mathbf{F}_N, \mathbf{E}^i \rangle \end{pmatrix}$$

respectively. \mathcal{L} is a linear operator while \mathbf{F}_n are basis functions. \mathbf{E}^i is the impressed field which may be the electric field of the primary TEM wave excited at the input of the cell or of the total incident field in the steady-state condition defined in (1) and/or (2). Recall that (8) and by extension (9) result from the boundary condition on the wire requiring tangential $\mathbf{E}_T = 0$. In obtaining (8) the surface current \mathbf{J} on the EUT has been expanded in a series of basis functions \mathbf{F}_n as

$$\mathbf{J} = \sum_n I_n \mathbf{F}_n \quad (11)$$

and the Galerkin method applied [4, 5]. The unknown complex coefficients I_n in (9) and (11), are determined through

$$[I_n] = [Z_{mn}]^{-1} \cdot [V_m] \quad (12)$$

The secondary modes arising from the induced surface current \mathbf{J}_n on the scatterer contribute to its impedance variation. This impedance variation is captured by the impedance matrix

$$[\Delta Z_{mn}] = \begin{pmatrix} \Delta Z_{11} & \Delta Z_{12} & \dots & \Delta Z_{1N} \\ \Delta Z_{21} & \Delta Z_{22} & \dots & \Delta Z_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta Z_{N1} & \Delta Z_{N2} & \dots & \Delta Z_{NN} \end{pmatrix} \quad (13)$$

which elements are

$$\Delta Z_{mn} = \iint_S \mathbf{J}_n \cdot \mathbf{E}_m^s dS \quad (14)$$

with \mathbf{E}^s being the scattered field given by (7). In the Galerkin representation, $\mathbf{J}_n = \mathbf{F}_n$.

From the approach developed here above, we can deduce that, in the GTEM cell, the overall impedance of an EUT is

$$[Z_{mn}] = [Z_{mn}] + [\Delta Z_{mn}] \quad (15)$$

where $[Z_{mn}]$ is the impedance the EUT would have in the free space (Eq. (10)) and $[\Delta Z_{mn}]$ the impedance variation due to the cell's reactive effect.

If the EUT is a thin wire as depicted in Fig. 3, Eq. (14) turns to become

$$\Delta Z_{mn} = - \int_{-sm/2}^{sm/2} \mathbf{J}_n(s) \cdot \mathbf{E}_m^s(s) ds \quad (16)$$

and the entries of the generalized impedance matrix $[Z_{mn}]$ are

$$Z_{mn}(f) = -j \frac{\mu\omega}{4\pi} \int_{\Delta s_n} F_n(s'_n) G(s, s') \vec{e}_s \cdot \vec{e}_{s'} ds' - \frac{j}{4\pi\epsilon\omega} \int_{\Delta s_n} \frac{\partial F_n(s'_n)}{\partial s'} \int_{\Delta s_m} \frac{\partial}{\partial s} G(s, s') ds ds' \quad (17)$$

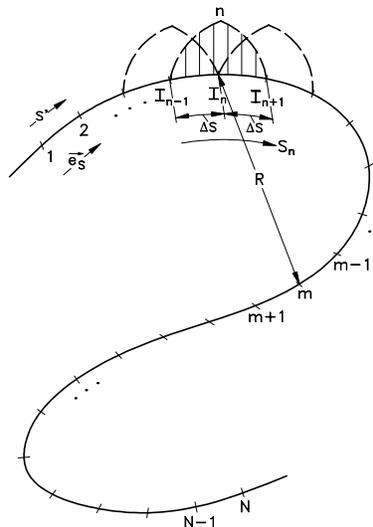


Figure 3: Segmented wire with overlapping piecewise expanded sinusoid functions.

where $\omega = 2\pi f$ and Δs_n denotes the integration over an element n . F_n are the expansion functions, which for the present case are taken to be piecewise sinusoidal functions

$$F_n(s') = \begin{cases} \frac{\sin(k_0(s' - s_{n-1}))}{\sin(k_0(s_n - s_{n-1}))} & \text{for } s_{n-1} \leq s' \leq s_n \\ \frac{\sin(k_0(s_{n+1} - s'))}{\sin(k_0(s_{n+1} - s_n))} & \text{for } s_n \leq s' \leq s_{n+1} \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

G is the free space Green function defined as

$$G(s, s') = \frac{e^{-jk_0 R(s, s')}}{R(s, s')} \quad (19)$$

$R(s, s')$ in (19) is the distance between the source point and the observation point.

The GTEM cell is generally used for susceptibility and emission measurements. The two cases are therefore of particular interest, especially for the excitation vector $[V_m]$. In both cases, $[V_m]$ depends on the excitation.

For the susceptibility case, independent from the re-illumination phenomenon, $[V_m]$ is always composed by the incoming fields. The aforementioned investigation applies without restriction.

For the emission, the excitation vector is respectively the voltage sources if the re-illumination is excluded or the superposition of the voltage sources and the higher-order modes totally reflected at the cut-off point if the re-illumination is considered. It is however worthwhile mentioning that, the variation of the input impedance is of crucial importance, when the EUT acts in the transmitting mode. This variation of the input impedance, indeed a key indicator of the mutual coupling between the cell and the EUT, is obtained as

$$\Delta Z_{in} = Z'_{in} - Z_{in} \quad (20)$$

where Z'_{in} and Z_{in} are the input impedance of the EUT in the cell and in free space respectively.

4. CONCLUSION

Based on a framework recently published, an approach for investigating the reactive effects of the GTEM cell is proposed. Emphasis has been put on the impedance variation. The method is general and can be applied to EUT of arbitrary geometry using the MoM. In particular, its application to thin dipoles as canonical EUTs may help capturing even small changes in the characteristics of the antenna and thereby providing a basic understanding of the effects of the mutual coupling between the cell and EUT. The appeal of the method resides in its simplicity and its applicability.

REFERENCES

1. Pouhè, D. and G. Mönich, "On the interplay between equipment under test and TEM cells," *IEEE Trans. EMC*, Vol. 50, No. 1, 3–12, Feb. 2008.

2. Pouhè, D., “Spherical waves in conical TEM cells,” *Progress In Electromagnetics Research*, Vol. 57, 209–236, 2006.
3. Pouhè, D., *Studies on Gigahertz Transverse Electromagnetic Cells*, Habilitation, TU Berlin, 2010.
4. Harrington, R. F., *Field Computation by Moment Methods*, R. E. Krieger Publishing Company, Inc., Krieger Drive, Malabar, Florida, 1968.
5. Stutzman, W. L. and G. A. Thiele, *Antenna Theory and Design*, 2nd Edition, John Wiley and Sons, New York, 1998.

Design Optimisation to Reduce the Magnetic Fields Propagated from DC Light Rail Transit Systems

A. Ogunsola^{1,2} and A. Mariscotti³

¹Department of Electrical and Electronics Engineering, Faculty of Engineering
University of Lagos, Lagos, Nigeria

²Rail Transit Division, Parsons Group International, London, United Kingdom

³Naval and Electrical Engineering Department, University of Genoa
Via all' Opera Pia 11A, 16145 Genoa, Italy

Abstract— The increasing implementation of Light Rail Transit System (LRTS) as an integrated form of public transportation, with the rail corridor running through city centers and connecting major commercial centers, presents a unique electromagnetic coupling scenario between sensitive installations and the dc magnetic field propagated from the LRTS. It is not surprising therefore, that there has been increasing interest in the prediction and mitigation of the magnetic fields propagated from such light rails. Numerous measurements have been performed, such as those by Lowes [1], Hsaio and Lin [2] and Yamazaki [3,4] with the view of quantifying the magnetic field from dc railways, however differences in Light Rail Vehicles restrict the use of measurement data across differing LRTS. Measurement data however has contributed to the development of numerical and computational analysis as applied to dc electrified railway. Waki et al. [5] investigated the use of magnetic shields, installed along the alignment, to attenuate the magnetic field propagated by the railway; this solution is feasible where the railway does not share the right of way with an existing road network. Schepper and Rabl [6], analysed a Compensation Conductor (CC) arrangement using a resistance network approach and reported a 42% improvement in the magnetic field due to the CC. Mariscotti and Ogunsola in [7] investigated the effectiveness of a mitigating technique based on the installation of CCs that can either be installed to be at the same potential of either the Overhead Contact Wire (OCW) or Running Rail (RR); the authors were able to demonstrate a 10dB attenuation of the propagating magnetic field. In this paper, the work of Mariscotti and Ogunsola is furthered by optimising the CCs arrangement with a view of maximising the attenuation of the propagated magnetic field. The approach to optimization is based on the minimization of either the mutual impedance term in the loop impedance expression or the weighted sum of magnetic field intensity in selected points. Computational results of the induction field computation is reported by means of contour curves, showing the achieved attenuation of the propagated magnetic field.

1. INTRODUCTION

Electrified traction systems are a source of electromagnetic field emissions; this is particularly true for large power systems but also for modern light rail, trolley bus and street car systems, that are embedded into cities, and thus closer to potential victims. The location of the alignment presents a unique scenario where the railway is in close proximity to sensitive installation [7]. The low frequency magnetic field may be computed easily by applying Biot-Savart's law to the whole set of the currents circulating in a given transversal section of the traction system, comprising the OCW, or third rail, and the return circuit, which for simply traction system comprises the RR only. Waki et al. [5] investigated the use of magnetic shields, installed along the alignment, to attenuate the magnetic field propagated by the railway; this solution is feasible where the railway does not share the right of way with an existing road network. Schepper and Rabl [6], analysed a Compensation Conductor (CC) arrangement using a resistance network approach and reported a 42% improvement in the magnetic field due to the CC. Mariscotti and Ogunsola in [7], investigated the effectiveness of a mitigating technique based on the installation of CCs that can either be installed to be at the same potential of either the OCW or RR; the authors were able to demonstrate a 10 dB attenuation of the propagating magnetic field. In this paper, the work of Mariscotti and Ogunsola is furthered by optimising the CCs arrangement with a view of maximising the attenuation of the propagated magnetic field.

2. COMPENSATION CONDUCTOR OPTIMISATION

The problem of external magnetic field minimization with respect to the geometry of Compensating Conductors (CC) may be formalized following two approaches. The CC geometry consists of their

position (x_i, y_i) and their cross section S_i with $i = 1 \dots N_C$. One approach builds an objective function J containing induction field values $B(x, y)$ for selected points computed with Biot-Savart's law and then numerically identifies the optimal solution (x_i, y_i, S_i) ; even if the calculation of B in the 2D surface containing the cross section is, in general, a valuable visualization tool for the expression of results, using it directly to drive the optimization poses the following additional issues: (1) the selection of points that correctly represent the problem and, (2) the possibility that closer locations with larger B values may "bias" the solution. The other approach considers the problem reformulated in terms of impedance terms expressions and minimization of the so-called loop impedance; in this approach the solution doesn't depend on the selected points, but accounts for the whole set of points on the 2D surface.

Before describing the two approaches in detail, it is worth defining more accurately the target of the analysis. If a set of points (x_k, y_k) with $k = 1 \dots K$ is considered, the requirement of minimization implies that each $B(x_k, y_k)$ value is minimized with respect to the $3N_C$ degrees of freedom of the problem given by the independent setting of (x_i, y_i, S_i) and to the fixed geometry of the existing traction line conductors (x_i, y_i, S_i) with $i = N_C + 1 \dots N$. Yet, with a more practical approach, the problem would be better expressed by the fact that each $B(x_k, y_k)$ is lower than a predetermined limit value B_k^* (given for instance by equipment susceptibility or human exposure), that may vary with the location — in real applications there are different limits for workers in technical rooms and along the line, separate limits for the general population at platforms, and separate limits for sensitive equipment in hospitals and laboratories.

3. B-FIELD MINIMIZATION

The B -field is computed for a set of reference points (x_k, y_k) and an objective function J is considered. For straight minimization, J may be in the form of a quadratic cost function, to allow the use of the first derivative for optimization, with the known issue of choosing the correct weighting factor w_k of each term, so that they are balanced and equally included in the optimal solution.

$$J = \sum w_k |B(x_k, y_k)|^2 \quad (1)$$

Alternatively, for the verification of compliance to limits, J needs to represent inequalities of the form $B(x_k, y_k) < q_k B_k^*$, where q_k is a coefficient for additional safety margin often required for human exposure and interference to safety related or safety critical equipment. The inequalities take the form of "sign" functions (output equal to 1 if positive, 0 if negative or null), possibly weighted by coefficients w_k , if a partial optimization with some limits violated is acceptable.

$$J = \sum w_k \text{sign}(B(x_k, y_k) - q_k B_k^*) \quad (2)$$

In all cases $B(x_k, y_k)$ is expressed by the Biot-Savart law.

4. LOOP IMPEDANCE MINIMIZATION

It is first observed that the CC compensating action is achieved by driving the current away from the compensated conductor(s) of the traction line and closer to the conductor(s) of opposite polarity. The total B -field in each point $B(x_k, y_k)$ is given by the sum of Biot-Savart terms extended to all the conductors of the traction line, including CCs. Taking a simple geometry with only two conductors (x_1, y_1) and (x_2, y_2) , one for each polarity, to form a rectangular loop, the total B -field at a point P is proportional to the separation of the two conductors $s = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$; the same term s appearing also at the denominator is in general negligible with respect to the distance of the same point P . The separation distance s appears also in the mutual inductance expression, calculated with a similar reasoning related to the linked flux. For this reason, the minimization of the mutual inductance and of the B -field intensity are considered equivalent facts. Secondly, while the optimization based on B -field minimization takes the currents I_i in the traction line conductors as an input to be determined externally (with an electrical simulator or some assumed worst case configuration), the use of the mutual inductance minimization and, in general, of the impedance matrix optimization is able to address the estimation of the current distribution among line conductors; of course the considered line is an ideal line with straight conductors, constant cross section and no transversal or lumped elements or connections. Thirdly, if the behavior at dc is also considered, then the current sharing between the traction line conductors is influenced by their cross section rather than the relative distance (as for the mutual inductance), so that

the mutual resistance terms need to be considered. The distribution of the B -field still depends on the geometrical properties and thus on the mutual inductance, but the driving currents I_i is distributed proportionally to the respective cross sections. The self and mutual terms that form the impedance matrix Z are calculated by means of Carson's equations [8], where the conductor's geometry (position and cross section) and electrical parameters (electrical resistivity and magnetic permeability) are the considered parameters. The Carson-Clem expressions are used to simplify the analysis with respect to Carson's equations; they hold when $x = \alpha r_i$ and $x = \alpha d_{ij}$ (defined for the self and mutual impedance terms respectively) are both smaller than 0.25, that is for small distances between conductors, as in the present case. Only the external (or geometrical) part of the impedance is calculated, leaving out internal terms to further simplify the analysis.

$$Z_{ii,ext} = 10^{-6} \left[0.987f + j\omega 2 \ln \frac{D_e}{r_i} \right] \quad Z_{ij,ext} = 10^{-6} \left[0.990f + j\omega 2 \ln \frac{D_e}{d_{ij}} \right] \quad (3)$$

where d_{ij} is the distance between conductors i and j ; D_e is the earth depth given by $D_e = 660 \sqrt{\frac{\rho}{f}} = 1.31 \sqrt{\frac{2\rho}{\mu\omega}}$; r_i is the radius of conductor i .

The impedance matrix Z thus has a general form of the kind

$$Z = \begin{bmatrix} R_{ii} + jX_{ii} & R_{ij} + jX_{ij} & \cdots & R_{im} + jX_{im} & R_{in} + jX_{in} \\ R_{ji} + jX_{ji} & R_{jj} + jX_{jj} & & R_{jm} + jX_{jm} & R_{jn} + jX_{jn} \\ \vdots & & \ddots & & \vdots \\ R_{mi} + jX_{mi} & R_{mj} + jX_{mj} & & R_{nn} + jX_{nn} & R_{mn} + jX_{mn} \\ R_{ni} + jX_{ni} & R_{nj} + jX_{nj} & & R_{nm} + jX_{nm} & R_{nn} + jX_{nn} \end{bmatrix}, \quad (4)$$

where (i, j) and (m, n) are two pairs of indices for the two circuit parts, and mutual terms below the diagonal are equal to their respective counterparts, since the Z matrix is symmetric. The terms of the same circuit part may be combined by the geometrical mean radius (GMR) rule [9] in order to obtain a two-conductor representation; the geometric mean radius for the i th conductor of a cable bundle is given by:

$$\text{GMR}_i = \sqrt[n]{n r_i R_{bundle}^{n-1}} \quad (5)$$

where R_{bundle} is the radius of the bundle and n the number of conductors in the bundle. The relation between GMR and the radius r of a circular solid nonmagnetic conductor, assuming uniform current distribution is given as $\text{GMR}_i = r e^{-1/4}$; however in the presence of skin depth, the electric current distribution is not uniform and it is therefore preferable to express GMR as:

$$\text{GMR}'_i = r e^{-\xi/4} \quad (6)$$

where ξ is a number between 0 and 1, and depends on frequency. When the frequency of the applied current is zero (i.e. dc current), then $\text{GMR}'_i = \text{GMR}_i$ and $\xi = 1$. For high frequency currents, the skin depth forces all the current to the surface of the conductor and $\text{GMR}'_i = r$ and $\xi = 0$. The resulting Z matrix for the equivalent two conductor representation is:

$$Z = \begin{bmatrix} R'_{11} + jX'_{11} & R'_{12} + jX'_{12} \\ R'_{12} + jX'_{12} & R'_{22} + jX'_{22} \end{bmatrix} \quad (7)$$

The mutual terms discussed above are the *off-diagonal* terms, resulting as the rightmost term of the loop impedance expression

$$Z_1 = (R'_{11} + jX'_{11}) + (R'_{22} + jX'_{22}) + Z_m \quad (8)$$

where $Z_m = -2(R'_{12} + jX'_{12})$.

The real and imaginary parts of Z_m are minimized separately: the mutual reactance term is mainly related to conductors position, while the resistive term takes into account the effect of conductors cross section. The solution for the resistive part is trivial since the optimum is reached by letting the cross sections increase to the limit fixed by practical and economical reasons: weight of the CC and mechanical suspension and/or fastening and cost of metal (especially in case of

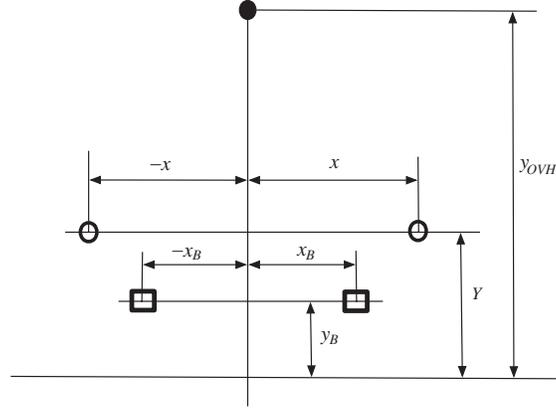


Figure 1: Cross section of the reference overhead railway system, with two CCs, one overhead and two running rails.

copper). The inductive part X'_{12} is evaluated via the application of Carson-Clem mutual impedance expression, by replacing the distance d_{ij} with the geometric mean distance of the set of conductors (that for the example with two CCs at overhead potential amount to five with $N_C = 2$, $N = 5$, as shown in Figure 1).

For each solution in terms of the optimal position of line conductors (x_i, y_i) the 2-D map of the B -field may be computed by Biot-Savart application as in [7].

$$z = \sqrt[6]{d_{13}d_{14}d_{15}d_{23}d_{25}d_{31}d_{32}d_{41}d_{42}d_{51}d_{51}} = \sqrt[6]{(d_{13})^2 (d_{14})^2 (d_{15})^2 (d_{24})^2 (d_{25})^2} \quad (9)$$

$$= \sqrt[3]{[x_R^2 + (y_R - y_{OVH})^2] [(x_R - X)^2 + (y_R - Y)^2] [(x_R + X)^2 + (y_R - Y)^2]}$$

By observing the structure of Carson-Clem expressions, minimizing Z_m corresponds to maximizing the GMD z . The first derivative of z with respect to the CC position (in general identified as (x_i, y_i) $i = 1 \dots N_C$, in the present case corresponding to $(-X, Y)$ and (X, Y)) corresponds to

$$\frac{\partial z}{\partial X} = 4X (X - Y + x_R + y_R) (X + Y - x_R - y_R)$$

$$\frac{\partial z}{\partial Y} = 4(x_R - Y + y_R) (X + Y + x_R + y_R) (X + Y - x_R - y_R) \quad (10)$$

that gives the trivial solution $X^* = x_R$ and $Y^* = y_R$, that is the optimal position of the CC is that of the running rails. The same set of equations, (9) and (10) may be analyzed by several types of optimization routines that may solve constrained optimization problems. The constraints limit the portion of space where the CC may be located: a minimum distance from the running rails ensures both electrical insulation and absence of mechanical interference with the sleepers and track ballast; considering that the CC in this example are at the OVH potential, electrical safety reasons impose that they are buried in a duct or trench, imposing thus that $Y < y_R$ with some margin, depending on infrastructure rules for trackside cableways.

5. RESULTS

Some sample results for a given current distribution in the line conductors corresponding to Test Case C (Figure 1(c) in [7]) as shown in Figure 2. The result from one of the optimization routines, *fminimax*, is shown in Figure 2. In Figure 2(a), all points are assigned an equal weight factor of unity, i.e., $w_k = 1$; with the resultant magnetic field shown in Figure 2(b). Figure 2(c), illustrates the case where the points are assigned linearly increasing weight factors starting from 1 to 10 towards the central position, 2 m above the rails and then decreasing to 1. The optimal CC position is thus a little higher for case Figure 2(c), resulting in lower $B(x_k, y_k)$ in the central points. The resultant magnetic field is shown in Figure 2(d).

In Figure 3, the B -field level obtained for one of the optimized solutions (unity weight factor of Figure 2(a)) and that of Test Case C in [7] are compared over a horizontal displacement of 30 m from the line axis (shorter than in [7] to appreciate smaller differences). The dashed contour curve

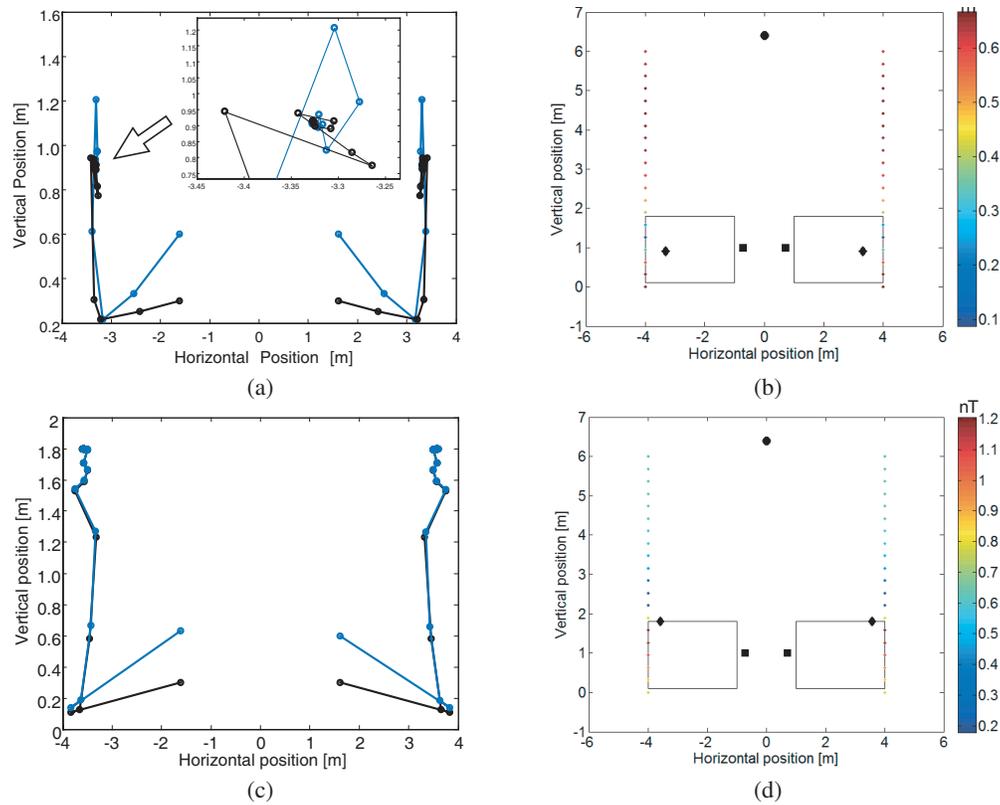


Figure 2: Sample results of the optimisation process: (a) and (b) minimax method, (c) and (d) least square non linear method; two initial CC positions at 0.3 and 0.6 m height.

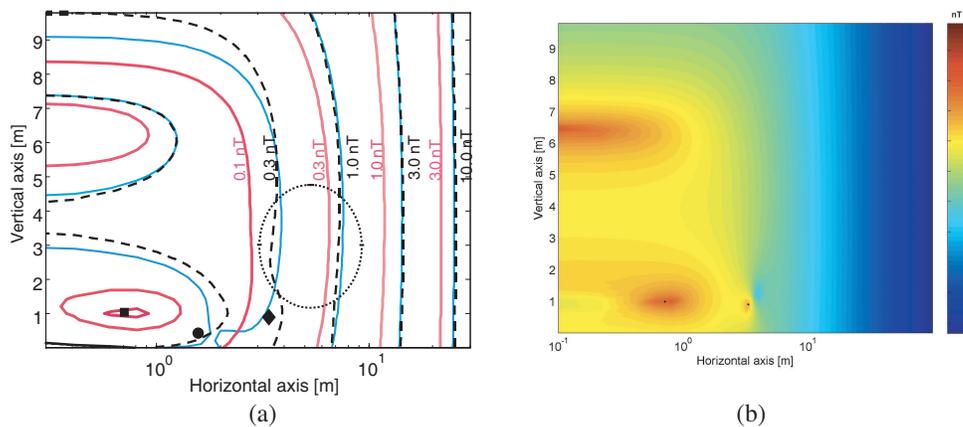


Figure 3: (a) comparison of B -field levels for sub-optimal [7] and optimal CC positions: non-compensated system (red), compensated system [7] (light blue), optimized system 2 (dashed black); (b) 2-D map of the calculated B -field.

indicates a reduction of the B -field intensity in the area subject to optimization (dotted ellipse in Figure 3) of about 1–1.5 dB, while they are approximately overlapped at larger distance. The advantage of the use of the considered class of optimization routines is that they scale up with the number of CC and the solution is always optimal in the assigned space for the assigned criteria (average field, maximum field, limit compliance, safety margins).

6. CONCLUSION

The problem of the reduction of B -field emissions from an electric traction systems equipped with overhead contact wire is formulated as a constrained optimization problem and the solution is attained by different approaches: a closed form calculation of the mutual impedance term Z_m (based on the calculation of the geometrical mean distance) and the definition of a suitable cost function of the B -field amplitude for numerical minimization. The results shown indicate that

depending on the problem requirements (for B -field limits and B -field values distribution) and constraints (accessible positions and minimum clearance) there are several options in terms of weighting, cost function and minimization routines.

REFERENCES

1. Lowes, F. J., "Magnetic monitoring of dc electric railways," *Physics in Technology*, Vol. 18, 209–214, 1987.
2. Ying-Tung, H. and K.-C. Lin, "Measurement and characterization of harmonics on the Taipei MRT DC system," *IEEE Trans. Industry Applications*, Vol. 40, 1700–1704, 2004.
3. Yamazaki, K., K. Kato, K. Kobayashi, Y. Uchikawa, Y. Kumagai, A. Haga, and K. Fujiwara, "Characteristics and prediction of magnetic noise due to DC electric railcars for biomagnetic measurements," *IEEE Trans. Magnetics*, Vol. 37, No. 4, 2884–2887, 2001.
4. Yamazaki, K., K. Kato, K. Kobayashi, K. Kawamata, A. Saga, N. Goto, S. Minegishi, and A. Haga, "Environmental low frequency magnetic field due to direct current electric railcars," *Electrical Engineering in Japan*, Vol. 137, No. 3, 10–21, 2001.
5. Waki, H., H. Igarashi, and T. Honma, "An analysis of shielding against dc magnetic fields generated by electric railways — Evaluation of effective permeability of fine structures," *Electrical Engineering in Japan*, Vol. 160, No. 4, 309–316, 2005.
6. Schepper, W. and C. R. Rabl, "Electromagnetic field emissions of electrified railways," *Computer Theoretikum und Praktikum für Physiker*, Vol. 11, 85–105, Fachbereich Physik der Martin-Luther-Universität Halle-Wittenberg, Hall, Germany, 1996.
7. Mariscotti, A. and A. Ogunsola, "Prediction and mitigation of magnetic fields propagated from dc light rail transit system," *Intern. Conf. on Electrical Systems for Aircraft, Railway and Ship Propulsion*, Bologna, Italy, Oct. 19–21, 2010.
8. Mariscotti, A. and P. Pozzobon, "Determination of the electrical parameters of railway traction lines: Calculation, measurement and reference data," *IEEE Trans. on Power Delivery*, Vol. 19, No. 4, 1538–1546, Oct. 2004.
9. CCITT (International Telegraph and Telephone Consultative Committee), "Directives concerning the protection of telecommunication lines against harmful effects from electric power and electrified rail-way lines," Volume 2, (ISBN 92-61-04031-4), Geneva, Switzerland.

Experimental Dynamical Evolution of Impulse and Delta Pulses through Dispersive Vegetation in Remote Sensing Frequency Bands

Ana V. Alejos^{1,2}, Luis Medina¹, Muhammad Dawood¹, Luis Rodríguez¹, and Paula Gomez²

¹EMAG and Wireless Group, New Mexico State University, Las Cruces, NM, USA

²University of Vigo, Campus Universitario, Maxwell St, Vigo 36310, Spain

Abstract— In this contribution, we introduce an approach to analyze and detect the formation of Brillouin precursor in a signal propagated through a dispersive vegetation for the frequency band of 2.4 GHz. We theoretically predict the evolution and formation of the Brillouin precursor for two pulses used in ultra wideband applications: impulse and delta.

1. INTRODUCTION

One important effect of the dispersive propagation is the evolution of signals called Brillouin precursor [1–5]. One main distinguishing property of the precursors is the algebraic peak amplitude attenuation compared to the conventional exponential decay of the original signal. The former is important for optimum transmitter and receiver design, demanding further research in that direction. The latter property leads to the optimal penetration depth in a given media and is of great interest for remote sensing applications involving soil, water, and foliage.

An experimental method has been developed in the frequency domain that gives as a result the frequency transfer function of the medium under study [6, 7]. This transfer function acts as a filter allowing as input signal any kind of waveform. In the outcome signal we can detect the formation of the Brillouin precursor superimposed to the rising and leading signal edges.

Herein, we introduce the experimental data obtained with the target of analyzing the time and spatial propagation of an ultra wideband input signal formed by an impulse and a delta pulse, as well as a sequence of the latter, through dispersive vegetation consisted of a single tree. We present the theoretical results achieved for both pulses and sequences as input waveforms, which modulate a sine carrier at 2.4 GHz.

The impulse waveform has been modeled by means of the first derivative of the Gaussian pulse. We observe that the pulse configuration affects the outcome, mainly in the impulse case.

2. FORMULATION

For the purposes of this paper, we used an FFT-based pulse decomposition technique to study the propagation of electromagnetic pulse $x(z, t)$ in through vegetation environment at distance z as given by (1) [5–7]:

$$x(z, t) = \frac{1}{N} \sum_{k=M_L}^{M_H} |X(0, k)| \cdot e^{-\alpha_k z} \cdot \cos(2\pi f_k t + \phi(0, k) - \beta_k z) \quad (1)$$

where $|X(0, k)|$ and $\phi(0, k)$ are the k -th component module and phase component of the discrete Fourier transform of the input signal $x(0, k)$; $\alpha(k)$ and $\beta(k)$ are the k -th frequency components of the attenuation and the phase constants derived from the dielectric permittivity $\epsilon_r(\omega)$ assumed for the dispersive medium of interest; and N is the number of frequency components considered $M_H - M_L$, and that should cover the frequency range of interest.

The dielectric model assumed to simulate the evolution of precursors is of vital importance. We have shown theoretically in [5] that a not accurate characterization of the dielectric properties can lead to a not exact estimation of the Brillouin precursor prediction. Specific dielectric models have been developed such as the given by Debye [5] considered suitable for modeling dispersive media as soils, foliage and sand. However, for this research, our experimental approach results valid to estimate the dielectric properties as we describe in the following section, and the achieved dielectric characterization has been used instead of theoretical models [6, 7].

3. EXPERIMENTAL METHODOLOGY

Our approach is based on the measurement of the transmission parameters $s_{21}(f)$ [6, 7]. This parameter was obtained using a vector network analyzer, in the range of 500 MHz to 3 GHz. The outcome $s_{21}(f)$ contains information about the effects of attenuation and phase for each frequency component of the signal travelling through the tree under study. Measurements have been conducted for vertical and horizontal polarizations.

The $s_{21}(f)$ obtained in presence of the tree is divided by the $s_{21}(f)$ measured in free space conditions. The result is not a transmission coefficient that we named general propagation factor. This filter $Ts_{21}(f)$ can be seen as a frequency filter so that we can analyze the evolution of any input signal $x(z, t)$ through the dispersive medium for any penetration depth z . It is enough to multiply in the frequency domain $X(z, f) \cdot Ts_{21}(f)$ and then apply an Inverse Fast Fourier Transform to observe the output signal in the time domain.

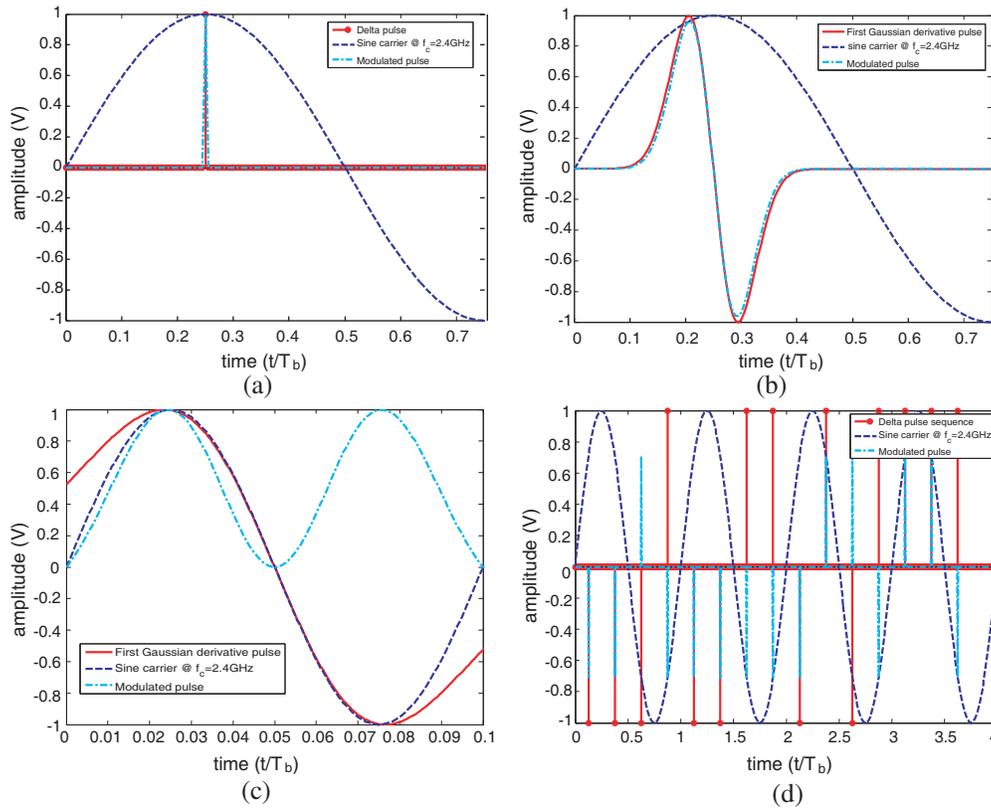


Figure 1: Different pulses utilized: (a) delta, (b) first derivative Gaussian ($0.025 \cdot T_b$, $0.00625 \cdot T_b$), (c) first derivative Gaussian ($0.05 \cdot T_b$, $0.0375 \cdot T_b$), and (d) a sequence of delta pulses.

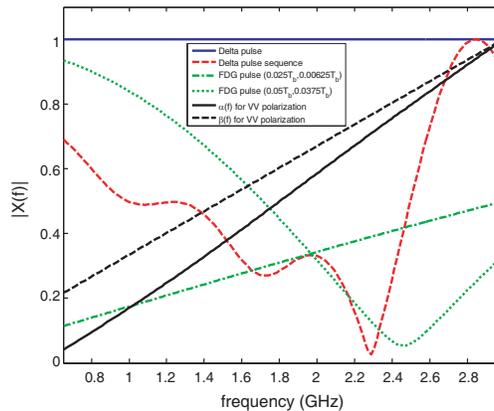


Figure 2: Frequency spectrum of the different pulses together with attenuation and phase constants experimentally obtained for VV polarization.

The ratio $T_{s21}(f)$ is also valid to estimate the complex dielectric permittivity $\varepsilon_r(\omega)$ of the medium under study [6, 7] with thickness d_m , since $T_{s21}(f) = \exp\{-[\alpha(f) + j\beta(f)] \cdot d_m\}$.

4. DYNAMICAL EVOLUTION

Once we obtain the attenuation $\alpha(f)$ and phase $\beta(f)$ constants of the involved medium, we can predict the spatial and time evolution of any input waveform by using (1). We have considered three signals: delta pulse, sequence of delta pulse and first derivative of Gaussian (FDG) pulse as as

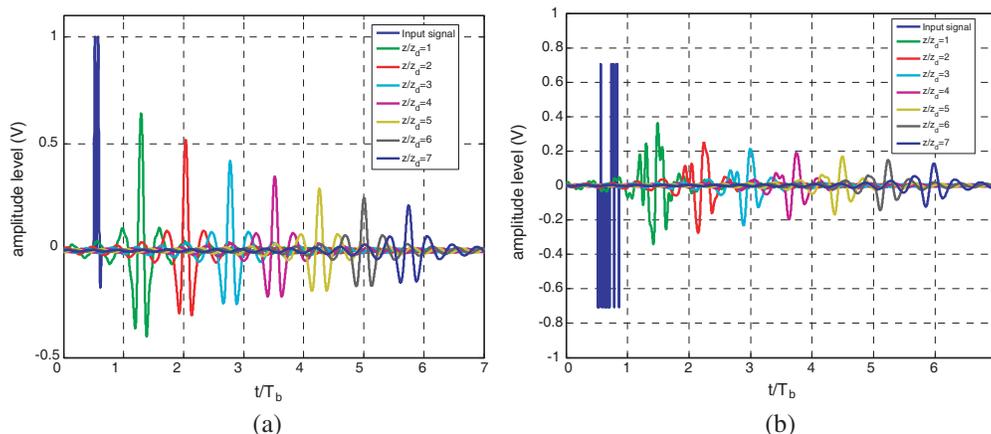


Figure 3: Evolution of the total propagated field of input modulated signals with sine carrier frequency $f_c = 2.4$ GHz into a vegetation medium characterized by the Maetzler model: (a) first derivative Gaussian ($0.05 \cdot T_b$, $0.0375 \cdot T_b$) and (b) sequence of delta pulses.

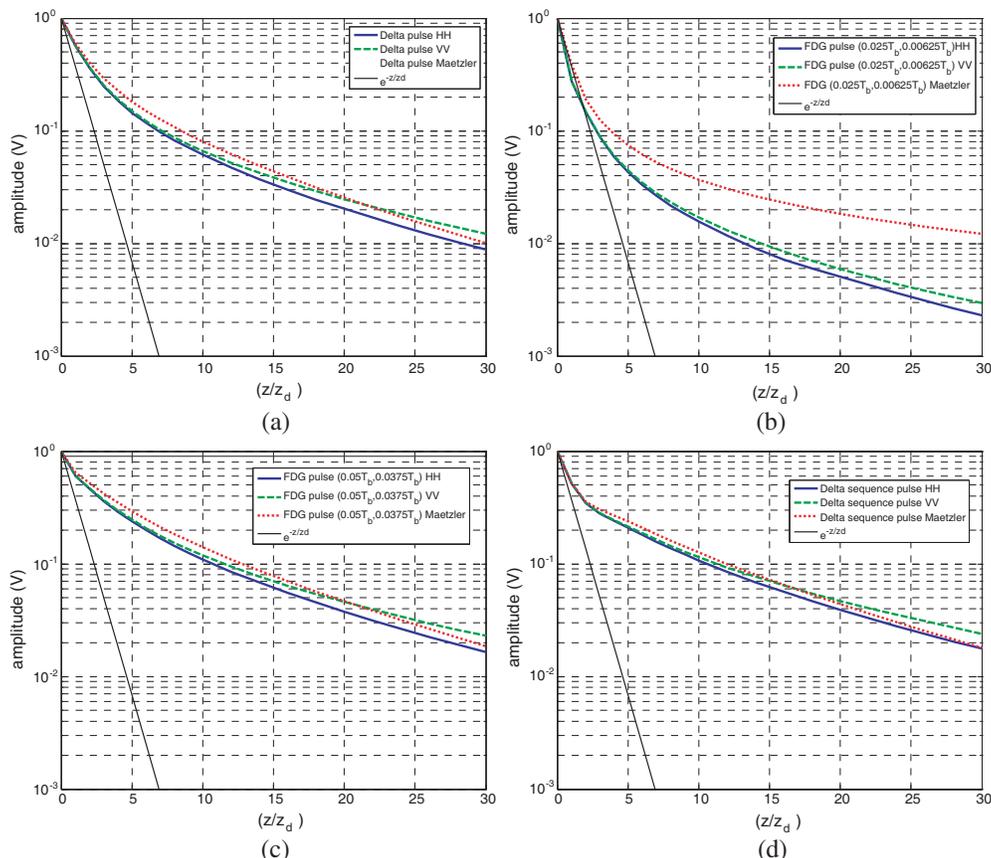


Figure 4: Power decay estimation with propagation distance for VV and HH polarizations and Maetzler model: (a) delta, (b) first derivative Gaussian ($0.025 \cdot T_b$, $0.00625 \cdot T_b$), (c) first derivative Gaussian ($0.05 \cdot T_b$, $0.0375 \cdot T_b$), and (d) a sequence of delta pulses.

an approach of impulse case. In (2)–(3) we indicate the settings used for the sequence of delta pulses and FDG:

$$x(0, t) = \sum_{k=1}^{15} a_k \cdot \delta(t - k \cdot 0.25 \cdot T_b) \quad (2)$$

$$x(0, t) = \frac{-2(t-a)}{b^2} \cdot \exp \left\{ - \left(\frac{t-a}{b} \right)^2 \right\} \quad (3)$$

We have considered two different settings for the FDG pulse: $(a, b) = (0.025 \cdot T_b, 0.00625 \cdot T_b)$ and $(a, b) = (0.05 \cdot T_b, 0.0375 \cdot T_b)$. In Fig. 1, we plot the resulting pulses used. In Fig. 2, we show their respective frequency spectrum together with $\alpha(f)$ and $\beta(f)$ to observe their effect of the different signal frequency components.

In Fig. 3, we observe the spatial and time evolution for the cases of FDG $(0.05 \cdot T_b, 0.0375 \cdot T_b)$ and the sequence of delta pulses as input pulses for different penetration depths into a vegetation model characterized by the experimental $\varepsilon_r(\omega)$ achieved for the VV polarization. The results considering the Maetzler model [5] has been also included. The peak amplitude decay of the analyzed pulses is shown in Fig. 4. We observe that all of them undergo an algebraic decay trend instead of the exponential one related to the narrowband component.

5. CONCLUSIONS

The evolution of ultra wideband pulses has been analyzed for the case of propagation through dispersive vegetation. The influence of the pulse configuration is explained by the different frequency components and the manner that the attenuation and phase constants affect them. All the pulses show an algebraically peak decay.

ACKNOWLEDGMENT

This work was partly supported by the New Mexico State University (NMSU), Las Cruces, NM, USA, and a NMSU patent is pending. We would also like to acknowledge the support of National Science Foundation (NSF) through the Bridge to Doctorate program, and the support received from the People Program of 7th FrameWork Programme (2008 Marie Curie IOF Action).

REFERENCES

1. Brillouin, L., *Wave Propagation and Group Velocity*, Academic, New York, 1960.
2. Oughstun, K. E., “Dynamical evolution of Brillouin precursor in rocard-powles-debye model dielectric,” *IEEE Transactions on Antennas and Propagation*, Vol 53, No. 5, 1582–1590, 2005.
3. Oughstun, K. E., *Electromagnetic and Optical Pulse Propagation 2*, Springer, New York, 2009.
4. Cartwright, N. A. and K. E. Oughstun, “Ultrawideband pulse propagation through homogeneous, isotropic, lossy plasma,” *Radio Science*, Vol. 44, RS4013, 2009.
5. Alejos, A. V., M. Dawood, and L. Medina, “Experimental dynamical evolution of the Brillouin precursor for broadband wireless communication through vegetation,” *Progress In Electromagnetics Research*, Vol. 111, 291–309, 2011.
6. Mohammed, H., M. Dawood, and A. V. Alejos, “Experimental detection of Brillouin precursors through tap water at microwave frequencies,” *IET Electronics Letters*, forthcoming 2011.
7. Alejos, A. V., M. Dawood, and H. U. R. Mohammed, “Analysis of Brillouin precursor propagation through foliage for digital sequences of pulses,” *IEEE Geoscience and Remote Sensing Letters*, Vol. 8, No 1, 59–63, 2011.

Investigation on a Ladder-shaped Frequency Selective Surface for Dual-band Operation

Keisuke Morimoto, Toshiaki Kitamura, and Daigo Yonetsu
Faculty of Engineering Science, Kansai University, Osaka, Japan

Abstract— A ladder-shaped frequency selective surface (FSS) is proposed. The proposed FSS, whose equivalent circuit is expressed as a combination of parallel and series LC circuits, shows dual stopband characteristics. Indoor shielding application for 2.4/5.2 GHz wireless local area network (WLAN) is assumed and the shielding effect when the FSS arrays are placed on the window glass is investigated.

1. INTRODUCTION

Frequency selective surfaces (FSSs) have attracted much attention due to their widespread applications as spatial microwave filters. Their practical applications can be found in electromagnetic-interference (EMI) protection, selective electromagnetic (EM) shielding, and EM absorbers. A typical FSS is composed of planar metallic patterns that are periodically arranged on a dielectric substrate. Various types of FSS have been proposed so far [1–5].

In this study, a ladder-shaped FSS is proposed and it is shown that the proposed FSS has dual-stopband characteristics. Indoor shielding application for 2.4/5.2 GHz wireless local area network (WLAN) is assumed and the shielding effect when the FSS arrays are placed on the window glass is studied.

2. FSS STRUCTURE

The unit-cell geometry of the proposed FSS is shown in Fig. 1. The structure consists of a ladder-shaped metal patch on a dielectric substrate. The metal patch has five parallel slots in it. It has been shown that the equivalent circuits of wire-mesh and square-patch FSSs are expressed by parallel and series LC circuits, respectively [6]. Accordingly, it is thought that the ladder-shaped FSS has dual-stopband characteristics because its equivalent circuit is composed of a combination of parallel and series LC circuit shown in Fig. 2. The unit cell surrounded by symmetric planes is analyzed by using an EM simulator HFSS.

3. RESULTS AND DISCUSSIONS

The frequency characteristics of a transmission coefficient are shown in Fig. 3. The simulation results are compared with the results of the equivalent circuit. The circuit parameters of the equivalent circuit are chosen as follows: $C_1 = 6$ pF, $C_2 = 5$ pF, $C_3 = 8$ pF, $L_1 = 1.4$ nH, $L_2 = 0.1$ nH. The dual stopband characteristics are obtained. It is understood that both results are in good agreement. It is also confirmed that the plural resonances occur at the higher frequency range above 4.8 GHz from the simulation results.

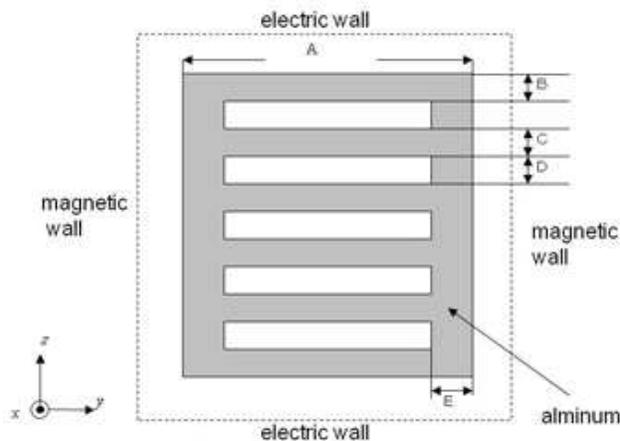


Figure 1: Unit-cell structure.

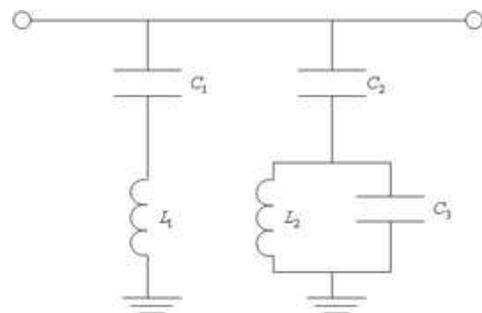


Figure 2: Equivalent circuit.

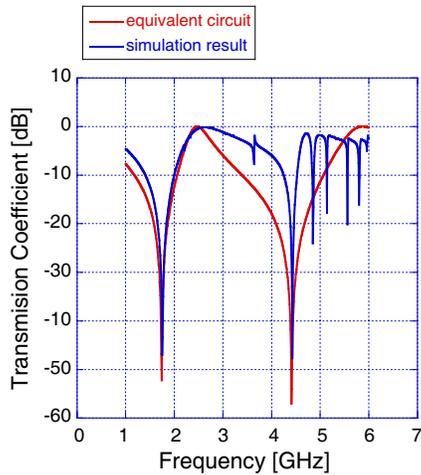


Figure 3: Transmission coefficient.

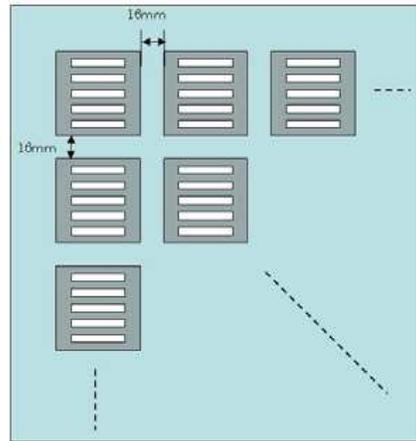


Figure 4: Geometry of FSS arrays.

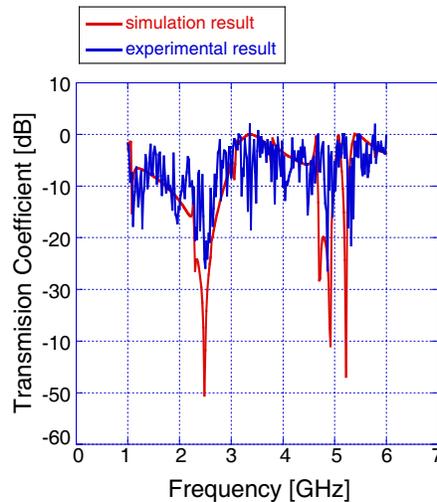


Figure 5: Transmission coefficient.

Finally, indoor shielding application for 2.4/5.2 GHz wireless local area network (WLAN) is assumed and the shielding effect when the FSS arrays are placed on the window glass is studied. The proposed FSS was fabricated by arranging the arrays of the proposed unit cells on both sides of the glass substrate whose relative permittivity is 5.5 as shown in Fig. 4. Aluminum sheet was used for forming metallic patterns. A shield box (WaveLaboratory Co., Ltd.) and a spectrum analyzer N1996A (Agilent Technology) was used for the experiment. The frequency characteristics of a transmission coefficient are shown in Fig. 5. The simulation results are compared with the experimental results. As shown in this figure, the shielding effects in the 2.4 and 5.2 GHz bands were obtained experimentally.

4. CONCLUSION

A ladder-shaped FSS that has dual-stopband characteristics has been proposed. The simulation results of the transmission coefficient has been compared with the results of the equivalent circuit. FSS arrays of the proposed unit cells were fabricated on both sides of the glass substrate. The shielding effects in the 2.4 and 5.2 GHz bands were obtained experimentally.

REFERENCES

1. Bayatpur, F. and K. Sarabandi, "Single-layer high-order miniaturized-element frequency-selective surface," *IEEE Trans. Microw. Theory Tech.*, Vol. 56, No. 4, 774–781, 2008.
2. Campos, A. L. P. S., E. E. C. Oliveira, and P. H. F. Silva, "Miniaturization of frequency

- selective surfaces using fractal koch curves,” *Microw. Opt. Tech. Lett.*, Vol. 51, No. 8, 1983–1986, 2009.
3. Stefanelli, R. and D. Trincherò, “Scattering analysis of frequency selective shields for electromagnetic indoor isolation,” *Microw. Opt. Tech. Lett.*, Vol. 51, No. 11, 2758–2762, 2009.
 4. Hu, X. D., X. L. Zhou, L. S. Wu, L. Zhou, and W. Y. Yin, “A novel dual-band frequency selective surface (FSS),” *Proc. 2009 Asia-Pacific Microwave Conference*, 1227–1230, 2009.
 5. Titaouine, M., N. Raveu, A. Gomes Neto, and H. Gaudrand, “Dual-band and enhanced band FSS characterization using WCIP method,” *Microw. Opt. Tech. Lett.*, Vol. 51, No. 8, 1983–1986, 2009.
 6. Celozzi, S., R. Araneo, and G. Lovat, *Electromagnetic Shielding*, Wiley-IEEE Press, 2008.

Maxwell's Motor Equation and the Mechanical Power

S. L. Vesely¹, A. A. Vesely², and C. A. Dolci³

¹I.T.B.-C.N.R., Italy

²Via L. Anelli 13, Milano, Italy

³Liceo Einstein, Milano, Italy

Abstract— Maxwell's illustration of the Ampère-Maxwell equation as well as his geometrical interpretation of it by means of the field-of-force concept are analyzed. Then a difficulty associated with transferring the same representation to modern electrical machines is clarified, and a possible reinterpretation of the electromagnetic equations alternative to Einstein's is briefly sketched.

1. INTRODUCTION

The quantities \mathbf{E} and \mathbf{B} in the electromagnetic equations are assumed to be generated by electric charges and by currents, perhaps reliant on them, as their sources. The so called “direct problem”, which consists in reckoning vectors \mathbf{E} and \mathbf{B} when the sources are known, is taught to be directly connected with the ancient way of recording the electromagnetic forces acting on a test body. In fact, in the late 1700s, Coulomb conducted the experiments about electric and magnetic actions by means of a torsion balance¹ [1], and interpreted them in accordance with the graphical representations of the static fields of force on a unit mass. That field-of-force notion was not only taken up by electrostatics and magnetostatics, by defining an electric and a magnetic unit mass, but it was also substantially accepted by the pioneers investigating the mutual actions between magnets and conductors. When, at the beginning of 1900, a completely new understanding of the propagating nature of the fields and of radiation-matter interactions broke through, motivating a profound revision of the conceptions of \mathbf{E} and \mathbf{B} (or \mathbf{H}), those new instances led to the modern physical theories of relativity and quantum electrodynamics. Today, classical electromagnetism subsists for its usefulness in technical applications. On the one hand, the wish to improve on the classically defined fields is considered rather conservative and unaware of later developments. On the other hand, in order to participate in the new comprehension trends, an electrical engineer is compelled to adopt the new theoretical frameworks together with the physically and mathematically unsolved problems therein. In our opinion, particularly the need to introduce higher dimensional spaces and the probabilistic interpretation of measurements could hamper practical applications. As a further option we will try and update the interpretation of paragraphs I and III of the IV part of Maxwell's Treatise² [2]. In fact the experiments described there are reminding of the “inverse problem” approach, i.e., of the problem of tracing the received signals back to their sources, which is more commonly encountered today.

In the 19th century thermodynamics seemed general enough to embrace the whole of the electromagnetic interaction mechanisms. The reason is that the conservation principle allows to take all contributions into due account by just adding suitably defined energies to the balance³. Great advances had been made also in analytical mechanics, where Lagrange, and all the more Hamilton, developed the relevant mathematical tools required to generalize the force concept to moving bodies. In spite of those accomplishments, to exhibit more clearly the traits of electromagnetic phenomena, Maxwell choose to stick to Newton's geometric style. Hence, besides appreciating that the experiments don't indeed fit in with Newton's mechanics, one is struck by the use of a different geometry. Those new trends in geometry had come to the fore along with the development of complex analysis. Nowadays they have evolved far beyond the geometric description adopted by Maxwell, and there are many elegant alternative ways to handle his equations. By contrast, modern electric power applications continue to be put on a level with those basic experiments. We will try and show that the experimental setup coming down to the Ampère-Maxwell equation is the one linking the electromagnetic fields to the moving electrified bodies in its common meaning. That link fails as soon as power is transmitted, due to the onset of special conversion mechanisms which cannot be easily represented geometrically. Alternatively, acknowledging that the geometry adopted by

¹Since the apparatus is very light, the twisting of the stretched wire in response to an impulse is swift. If the elastic restoring torque produces a free torsional oscillation with a very low frequency, that twisting action can be compared with a static couple.

²Art. 486–491, 530–532.

³Thermodynamics was supplemented by statistical methods to handle irradiation, conduction and convection phenomena.

Maxwell is different from Newton's, and thus that the usual link to the mechanical representations is without foundation, it becomes possible to scrutinize the connection between the electromagnetic induction signal and the degree of electro-mechanical coupling during power conversion.

2. MAXWELL'S EQUATIONS' SYSTEM

Maxwell's field equations summarize Faraday's experiments and apparently provide an intuitive interpretation for the time-varying electromagnetic quantities involved therein. Yet, electromagnetism had introduced motion into the static pictures, troubling their geometric representation. Neither the mathematical description was univocally settled any more [3], nor there was agreement about what physical properties a fields' mapping embodies⁴. Nowadays, if the system of Maxwell's equations is taken to consist of four coupled vector partial differential equations in the four unknown quantities \mathbf{E} , \mathbf{B} , \mathbf{D} and \mathbf{H} , the geometric representation of the electromagnetic fields is local, unless the equations are integrable⁵. With integration in view, the number of variables is usually reduced, by expressing the flux densities \mathbf{D} and \mathbf{B} by means of the field densities \mathbf{E} and \mathbf{H} , through the linear constitutive relations $\mathbf{D} = \epsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$ ⁶, where ϵ and μ are real constants.

As the interpretation of the relevant quantities was debated already at Maxwell's times, he himself illustrated each of the fundamental equations by means of typical experiments. According to the examples provided, two of the equations are transferred literally from electrostatics and magnetostatics⁷, implying that even the phenomenology associated with motion shall become static whenever the movement ceases. Regardless of the differential equations $\nabla \cdot \mathbf{D} = \rho$ (ρ being defined as the actual free electric charge density) and $\nabla \cdot \mathbf{B} = 0$ satisfying Faraday's request that there be a close by action, it can be shown that finite equations of Coulomb's kind can be recovered, owing to the fact that Gauss' theorem doesn't depend on true boundary conditions. Although the static phenomenology is decidedly not electromagnetic, the corresponding equations are convenient for the simplification of the coupled ones, which are the Ampère-Maxwell law $\nabla \times \mathbf{H} = \mathbf{J} + \partial\mathbf{D}/\partial t$, and Faraday's electromagnetic induction law $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$. Therefore we first give a look at mathematics. In the static limit the electromagnetic equations involve irrotational solutions, which can be expressed as gradients of some scalar functions ψ and φ . Maxwell accepts that the potential function, say, ψ of \mathbf{H} can become multivalued in the time-dependent cases, and defines it on a Riemann surface. For those cases, he adds a term $\partial\mathbf{D}/\partial t$ to Ampère's formula $\nabla \times \mathbf{H} = \mathbf{J}$, bringing the stationary solutions to *his* equations back to those of potential flow analysis and to conformal mappings. Of course if, on account of physical applications, all of the four variables were needed, the solutions could no longer be restricted to surfaces. Anyway, in the complex plane it is easy to relate $\mathbf{E} + i\mathbf{B}$ to equipotential lines and streamlines respectively, and that can be upheld in geometric transformations, according to the right hand sides of the equations: $[\nabla^2\mathbf{E} - \epsilon_0\mu_0\partial^2\mathbf{E}/\partial t^2] = \epsilon_0^{-1}[\nabla\rho + \epsilon_0\mu_0\partial\mathbf{J}/\partial t]$ ⁸ and $[\nabla^2\mathbf{B} - \epsilon_0\mu_0\partial^2\mathbf{B}/\partial t^2] = -\mu_0[\nabla \times \mathbf{J}]$. In fact, since the solutions of the homogeneous equations are harmonic functions, the elements \mathbf{E} and \mathbf{B} can be made harmonic conjugates in every simply connected domain. Let's now come to the physical interpretation. To give a geometric representation, the time dependent equations ought to be exemplified in a strictly linear range, i.e., when the experimental behavior is independent of frequency and intensity⁹. The *electromagnetic induction law* is taken to describe the small current transient elicited in an electric circuit (receiver) by the movement of a magnet (source)¹⁰, leaving

⁴The meaning apparently attributed by Maxwell to the electromagnetic variables is taken from statics' representations of the fields of force. His dynamical picture requires the static frame to be displaced in time according to the convective derivative: $\partial * / \partial t \rightarrow \delta * / \delta \tau - \nabla \times (\mathbf{v} \times *) + \mathbf{v}(\nabla \cdot *)$. As opposed to Maxwell's dynamics, an electromagnetism in the "Newtonian" style represents the solutions of the equations of motion in the potential fields. The electromagnetic potentials are written as $\mathbf{A}(t - r/c)$ and $\varphi(t - r/c)$, and the electric field becomes $\mathbf{E}_{em} = \mathbf{E}_s - \partial\mathbf{A}/\partial t$.

⁵As an alternative, the "Erlanger Program", still unknown in Maxwell's times, permits a global "imaging" interpretation of the electromagnetic field coordinates, which is obtained by transforming the space coordinates into the (finite) field variables. It is different from transformation optics in that the geometry models the imaging properties, i.e., the transforming functions, without entailing (mechanical) space properties.

⁶In the SI units of measure the relations $\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P}$ and $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$ hold, where \mathbf{P} and \mathbf{M} are the electric and magnetic polarizations respectively. Physically the introduction of \mathbf{D} and \mathbf{B} besides \mathbf{E} and \mathbf{H} complies with Faraday's observations about the non-inertness of the interposed media.

⁷They are Faraday's cage and the broken magnet experiments, respectively.

⁸This equation is obtained from $\nabla \times \mathbf{B} = \mu_0\mathbf{J} + \mu_0\epsilon_0\partial\mathbf{E}/\partial t$, making use of one static solution, by the steps: $\nabla \times \partial\mathbf{B}/\partial t = -\nabla \times (\nabla \times \mathbf{E}) = \nabla^2\mathbf{E} - \epsilon_0^{-1}\nabla(\nabla \cdot \mathbf{D}) = \nabla^2\mathbf{E} - \epsilon_0^{-1}\nabla\rho = \mu_0\partial\mathbf{J}/\partial t + c^{-2}\partial^2\mathbf{E}/\partial t^2$. The second one is obtained in a similar way.

⁹In practice the induction phenomena appear to be much stronger in the magnets than in the conductors. Hence magnets also show more impressive deviations from the linear behavior. Nevertheless, in our opinion, neither of both displays the properties of the demodulated electromagnetic signals. Rather the magnetic and conduction properties could be understood as resonance behaviors of the metals near $\nu = 0$.

¹⁰Or by the movement of the circuit with respect to the magnet. Besides, a mechanical movement is not truly necessary to

any feeding in aside. Choosing decibel units, this law gets an independent geometric meaning, whatever the meaning of the induction signal. The *Ampère-Maxwell law* is taken to describe an application of Ørsted’s effect carried out by Faraday. Ørsted had noticed that when a conductor thread (source) is connected to a battery, a magnetized needle (receiver) gets deflected, apart from the way it is moved toward the thread. Faraday exploited that jerky deflection to obtain complete turning cycles of a part of a fed circuit around a magnetic cylinder, or reciprocally. If the distinction between electromagnetic induction and Ørsted’s effect is considered a matter of interchanging the term “receiver” with “source”, then the second equation means that a needle sensor is being used for detection. Although the detection method is very different, indeed, the equation $\nabla \times \mathbf{H} = \partial \mathbf{D} / \partial t$ gets the same interpretation as the induction equation, and the constant \mathbf{J} results from circle integration. If, as currently assumed, $\nabla \times \mathbf{H} = \mathbf{J}$ is taken to be a magnetostatics equation, then closed loops carrying a direct electric current are taken as the outer/inner bounding edges of magnetic shells, and the current circuits substitute for the magnetic dipoles as field sources. All the same, on account of the circulation the space topology differs from Newton’s. For instance, if one current thread pierces the plane, the surface becomes homeomorphic to a torus, and to represent the electromagnetic variables as well as the motion variables on the same diagram, the interchangeability of some circular paths and straight lines has to be allowed for. According to Maxwell there are such inertial circular paths¹¹. However Newton’s gravitational theory does not admit them. Finally the geometric picture of the conservation of angular momentum breaks down, when the motive force operating an engine starts changing, i.e., when there is power delivery.

3. MAXWELL’S ILLUSTRATION OF THE MOTOR EQUATION

To give a Euclidean illustration of a machine functioning is like telling that a micro-motor is similar to a marine propulsion engine. The similarity pertains to both their size and their kinematics. In fact, there is no way to distinguish geometrically between the circling of a big electric engine and that of an electrostatic whirl, though evidently the magnetic energy component drops out for the latter. The crucial point is that, on the one hand no *linear* representation can account for the different behaviors of conductors versus magnetic materials that arise at terrestrial conditions¹². On the other hand, circuitual schemata are useful to analyze the steady electrical behavior of electrical machines, such as transformers, reversible DC motors, synchronous motors, and alternators¹³. Once the analysis is led back to linear circuits, the electromagnetic force gets a symmetric role to an emf, as Einstein remarked, whereas Maxwell in the end had tried to tie Ampère’s law to an unuttered property of magnets.

The device described by Maxwell shows an electromagnetically driven rotation. The machine is actually reversible: the motor is called unipolar or homopolar [4], while the dynamo is called Faraday’s disc, or unipolar inductor, to stress that the current is generated by induction. Currently, neither the motor nor the generator enjoy relevant applications, albeit a quite big generator has been deployed for some electrolytic process [5]. According to Maxwell’s description, a cylindrical bar magnet is inserted on a shaft so as to be capable of revolving about its vertical axis. Attached to it there is one end of a horizontal conducting arm, whose other end dips into a mercury-filled annular trough laying around a ring, which is concentric with the magnet’s section and fastened to the support at about midway between the poles. When the circuit is connected to a battery, as soon as the magnet starts rotating the current around the ring is split into two branches whose path lengths vary cyclically. The direction of the current depends on the polarity of the battery, whereas the sense of rotation of the magnet also depends on whether the upper pole is north or south. Notwithstanding the fact that the split circuit is not exactly like a commutator, once the revolving motion has been started, it is apparently driven by current commutation, as Maxwell notes¹⁴. Since to model the experiment geometrically one has to proceed along a different line of thought, he represents the upper magnetic pole in the magnetic force field that the current generates, and suggests that the magnet exerts a reaction force on the circuit. In fact, if the circuit

elicit induction.

¹¹It is clarified in the Treatise that the battery only feeds the circuit, without the emf being used to maintain the rotation.

¹²Under moistly conditions the electric effects are much smaller than the magnetic ones, but they seem to affect all substances equally. (Ferro-)magnetic properties, on the other side, belong only to peculiar substances.

¹³We point out that in this case the rotational speed matches the electric frequency, thus granting the use of filter analysis to determine the maximum mechanical power transmitted to a load. Although then the electric frequency is understood to be a resonance frequency, the machine usually doesn’t assume a state of resonance.

¹⁴Ørsted’s experiment is delicate. By definition, the battery supplies a direct current. However, to protect it, the needle’s deflection is usually obtained by switching the circuit on and off.

arm is detached from the body of the magnet, and is provided with a sliding contact, it can be made to rotate around the bar magnet. By leaving the field-of-force representation for a kinematical one, the reciprocity can be accounted for by the relativity of motion. However, if the field-of-force is mapped on the same space where the movement is represented, the simultaneous involvement of two reference frames rotating with respect to one another renders the interpretation ambiguous [6].

Einstein faced the problem of the drag of the lines of force by the frame in 1905. He remarked that the magnetostatic force supports an electric current within the circuit of a generator without doing any work, and that therefore the current should be induced therein due to a reaction force [7]. He postulated the invariance of physical laws under coordinate transformations related to his principle of relativity. Furthermore, he required that electric current be identified with convection of electrified particles endowed with a rest mass. Thus the Lorentz force expresses an *equivalence* between the electromagnetic and the mechanical force¹⁵, that depends on the reference frame, and reduces to the electrostatic force in a frame instantaneously at rest relative to an electrically charged particle. Reciprocally a magnetic field reduces to an electrostatic field under relativity transformations [8]. The last statement is at odds with Maxwell's own illustration of the Ampère-Maxwell equation, and with his belief of describing something different from Faraday's induction.

4. IMAGING VERSUS POWER TRANSFER DIAGRAMS

Maxwell's equations system can be considered mathematically coherent as is. Perhaps, his theory is called phenomenological because it illustrates the facts, without implying a unique, general interpretation for the field variables. A major problem is to clarify what is meant by a geometric representation of \mathbf{E} , \mathbf{D} , \mathbf{H} and \mathbf{B} . There are many instances, such as medical or SAR imaging, where the induction phenomenon can be linearized, so as to become altogether similar to TV imaging, with due care. By contrast, the circuit diagrams of electrical machines' performances are rather conceived as graphical means to evaluate energy flow. But, while the circuit's parameters are linearly scalable, the evaluated power refers to fixed values of the electrical and mechanical parameters of the machine. In turn, that means that the *watts* computed using the scalable schematics as a base can by no means represent the absolute values of the converted power, and the *volt-amperes* belong to the resistive and reactive parts of the power factor of the associated electrical circuits.

One way to escape the interpretive problem of the fields consists in considering the variables to describe in every case received signals, and interpreting signals in a different fashion, depending on the process involved in their detection. Actually, while images can often be reconstructed from the recorded electromagnetic input¹⁶, a geometric representation of power machines can only be endorsed at steady state. To show it, let's examine what kind of picture a circuit analysis conveys for the energy flow through a machine. Usually, the transmitted power is not evaluated during the run-up cycle, but rather at full load conditions, when, except for the "dynamical frictions", the motor's torque is exactly balancing the load's torque. At steady conditions, the relation between the rotation frequency of the moving parts and the sinusoidal electric current has hardly anything to do with resonance. In fact, the functioning regime is supposed to be stable, i.e., either the machine is working far away from any resonating condition, or any vibrations, which could have arisen because of couplings with the power supply or the load, have been damped. In that pretty linear regime circuit analysis allows to decompose the electrical signal into its Fourier components, and to evaluate either the transferred power or the efficiency. However, the electro-mechanical couplings responsible for power transfer can arise at the moment when supply conditions or load start changing. For instance, that happens during the time required by an induction motor's slip to stabilize itself. In brushless DC motors the drive electronics controls the speed by sensing the rotor's position, instead. Going back to Maxwell's analysis, we notice that the illustrations of the two electromagnetic equations differ in one respect. Whereas Faraday's induction is an electrical transient, the Faraday's disc is the earliest example of a generator. To describe in every case received signals, we substitute Ørsted's jerky excitation of the magnetic needle for the electromagnetic rotation. By doing so, the improvements to the motional conversion brought about by Faraday become apparent, too: they testify how strongly the motor design contributes to the parametric *mechanical amplification* of Ørsted's effect. Exactly that kind of amplification is never used for

¹⁵This is tantamount to a nonradiation condition.

¹⁶We stress again that, according to the Erlanger Program, the four relations $\mathbf{E} = f_1(x, y, z, t)$, $\mathbf{D} = f_2(x, y, z, t)$, $\mathbf{H} = f_3(x, y, z, t)$, and $\mathbf{B} = f_4(x, y, z, t)$ mean a transform $(x, y, z, t) \rightarrow (\mathbf{E}, \mathbf{D}, \mathbf{H}, \mathbf{B})$ of the space into itself (without necessarily describing focused images).

imaging purposes, even if power amplification is sometimes allowed for in electrical schematics.

Thus far, we have been telling that the mechanical power amplification of Ørsted's effect obtained by means of electrical machines is non-linear, without endeavoring in the reduction of macroscopic mechanisms to microscopic processes. As we ascribe the power conversion capability to a great deal to the machine design, we expect the construction of actuators and MEMS to open up new experimental opportunities, which may also rely on non-linear phenomena. In our opinion the electromechanical performance cannot be assessed by modeling the structure of matter so as to reduce the transduction to linear processes. Still, it is hypothesized by quantum mechanics that the functioning of simple systems is energetically compatible with very few different mechanical structures. It is assumed that functioning is linear at the atomic level, and that coherence, self-organization, or cooperativity takes over at the macroscopic level. That can be considered a valuable enumeration of the possible modelling mechanisms. We owe to quantum mechanics the insight that a peaked signal entrains a transition from a stationary state to another, instead of evidencing a revolution frequency. However, the atomic hypothesis has given rise to an inherently probabilistic theory, which cannot characterize an individual macroscopic motor as such. In addition the aim of that theory at verifying the hypothesized simple functioning on small scale engines fails due to its probabilistic mathematical basis.

5. CONCLUSIONS

Maxwell explains the Ampère-Maxwell equation by means of Faraday's revolving motor, unintentionally suggesting two things: that his description concerns a power generator/motor, and that the geometric representation used to represent dynamics may display electromagnetism as well. In retrospective, his view deserves some revising. A first shortcoming arises because he was unaware of the distinction between a signal and a carrier wave. Today the broadcasted signal is usually obtained by modulating a carrier. It is known that there is no energy intrinsic in the act of communicating, and that the fluctuations of stabilized generators can be held very small. A second shortcoming is due to the general belief that physical quantities require absolute units of measurement. When the chosen geometry is thought of as being the "physical space" rather than just a geometric model alternative to the mechanical one, electromagnetic fields are measured with the same scale used for describing the mechanical processes. However, such stance implies that the electro-mechanical conversion factor is fixed by geometrical considerations, and the possibility of bettering power conversion by monitoring the associated induction and using it as a broadcasted (*radiated*) signal is precluded. We have tried and shown that the usual descriptions of electric power machines apply to steady state functioning only, whereas, in examining power changes, the Ampère-Maxwell equation could be considered by analogy with Faraday's law.

REFERENCES

1. Martínez, A. A., "Replication of Coulomb's torsion balance experiment," *Arch. Hist. Exact Sci.*, Vol. 60, 517–563, 2006.
2. Maxwell, J. C., *A Treatise on Electricity and Magnetism*, Vol. 2, Dover Publications Inc., placeStateNew York, 1954.
3. Silva, C. C. and R. de Andrade Martins, "Polar and axial vectors versus quaternions," *Am. J. Phys.*, Vol. 70, No. 9, 958–963, 2002.
4. Zeleny, J. and L. Page, "The problem of the rotating magnet," *Phys. Rev.*, Vol. 40, 299–306, 1932.
5. Miller, A. I., "Unipolar induction: a case study of the interaction between science and technology," *Ann. Sci.*, Vol. 30, No. 2, 155–189, 1981.
6. Rosen, N. and D. Schieber, "Some remarks on Faraday's law," *Am. J. Phys.*, Vol. 50, No. 11, 974–975, 1982.
7. Becker, R., *Electromagnetic Fields and Interactions*, Dover Publications Inc., placeStateNew York, 1964.
8. Field, J. H., "Classical electromagnetism as a consequence of Coulomb's law, special relativity and Hamilton's principle and its relationship to quantum electrodynamics," *Phys. Scr.*, Vol. 74, 702–717, 2006.

Analysis of Light-beam Scattering from DWDD Disk with Control Layer under Considering Rear Process

Yuya Matsunami and Toshiaki Kitamura

Faculty of Engineering Science, Kansai University, Osaka, Japan

Abstract— Scattering characteristics of a domain wall displacement detection (DWDD) disk were investigated. DWDD is one of the high-density storage technologies of magneto-optical (MO) disks. A DWDD structure with a control layer was assumed and rear process was considered. A control layer is used to suppress ghost signal due to rear process. We studied how the rear process affect the characteristics of the main-polarized and crosspolarized components of the scattered light beam through numerical simulations based on finite-difference time-domain (FDTD) method.

1. INTRODUCTION

Magneto-optical (MO) disk is a rewritable type of optical storage media that is used in combination with magnetic technology. Domain wall displacement detection (DWDD) which utilizes a magnetic domain wall displacement due to a temperature gradient [1, 2] is one of the high-density storage technologies of MO disk system. It provides expansion of very small written bits up to the dimensions compatible with the optical resolution limit.

In this study, we analyze the scattering of blue-laser light from a DWDD disk under considering rear process by means of FDTD method [3]. We assume DWDD structure with a control layer, which is used to suppress ghost signal due to rear process, for the analysis. We investigate how the rear process affect the characteristics of the main-polarized and crosspolarized components of the scattered light beam.

2. DWDD DISK STRUCTURE

Figure 1 shows a DWDD disk with a control layer and its coordinate axes. The structure is assumed to be on a polycarbonate (PC) substrate. We assume that the disk moves toward the $-x$ direction and the upward and downward magnetizations correspond to the recorded and non-recorded states, respectively. The grooves and lands have the same width of 540 nm in the y direction, and the adjacent marks in the lands are assumed to have the same mark patterns as the grooves. The structural parameters are shown in Table 1. In Table 1, ε_0 indicates the permittivity of free space.

Figure 2 shows the recording states for three-layer and four-layer structures. We assume that the recorded and non-recorded marks are alternately allocated.

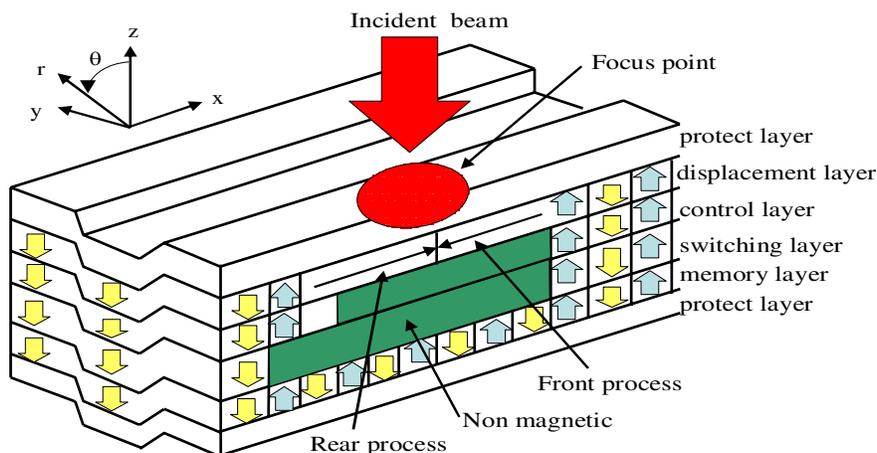


Figure 1: DWDD disk structure.

Table 1: Structural parameters.

| | | Permittivity[F/m] | | Layer Thickness [nm] |
|--------------------|------------|-------------------|-------------------|----------------------|
| | | ϵ | ϵ_{xy} | |
| PC Substrate | | $2.5\epsilon_0$ | 0.0 | Semi-infinite |
| Protect Layer 1 | | $4.0\epsilon_0$ | 0.0 | 35 |
| Displacement Layer | (Non-rec.) | $2.0\epsilon_0$ | $-0.05\epsilon_0$ | d |
| | (Rec.) | $2.0\epsilon_0$ | $0.05\epsilon_0$ | |
| Control Layer | (Non-rec.) | $2.0\epsilon_0$ | $-0.1\epsilon_0$ | 5 |
| | (Rec.) | $2.0\epsilon_0$ | $0.1\epsilon_0$ | |
| | (Non-mag) | $2.0\epsilon_0$ | 0.0 | |
| Switching Layer | (Non-rec.) | $2.0\epsilon_0$ | $-0.05\epsilon_0$ | 15 |
| | (Rec.) | $2.0\epsilon_0$ | $0.05\epsilon_0$ | |
| | (Non-mag) | $2.0\epsilon_0$ | 0.0 | |
| Memory Layer | (Non-rec.) | $2.0\epsilon_0$ | $-0.1\epsilon_0$ | 60 |
| | (Rec.) | $2.0\epsilon_0$ | $0.1\epsilon_0$ | |
| Protect Layer 2 | | $4.0\epsilon_0$ | 0.0 | 55 |

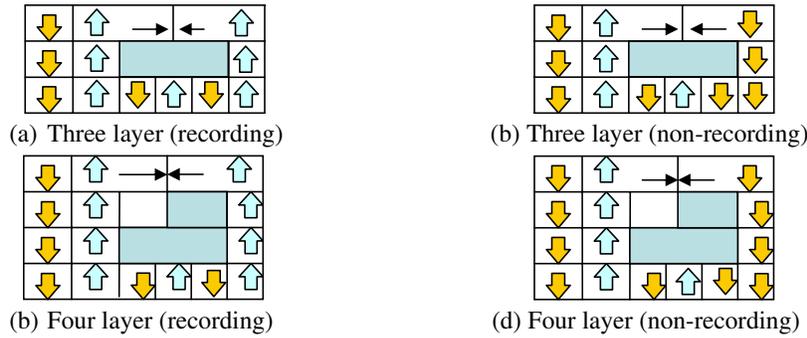


Figure 2: Recording states.

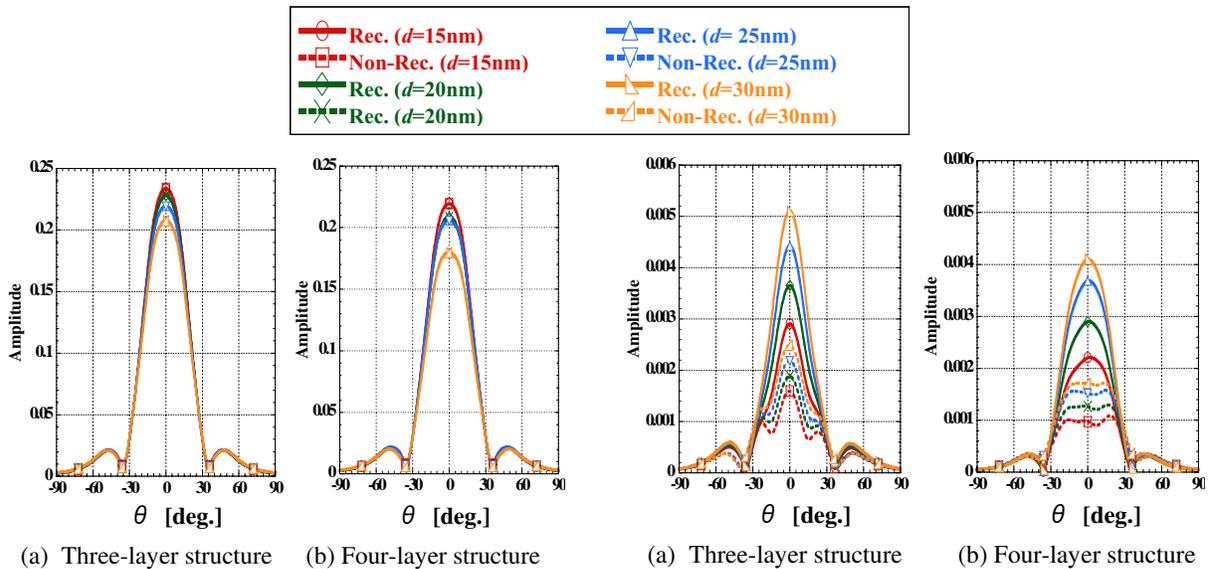


Figure 3: Main-polarized components.

Figure 4: Cross-polarized component.

3. NUMERICAL RESULTS

Figures 3 and 4 show the main-polarized (x -directed) and cross-polarized (y -directed) components, respectively, of far-zone scattered electric fields in y - z plane. Here, the length of the displacement

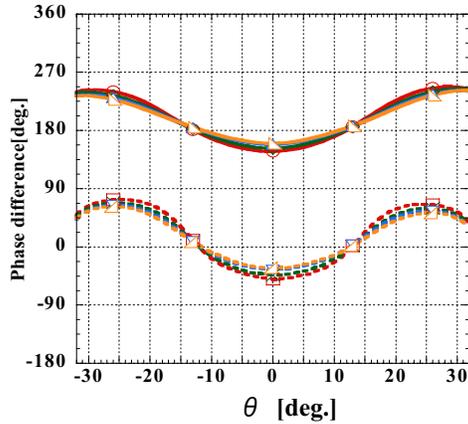


Figure 5: Phase differences (three-layer).

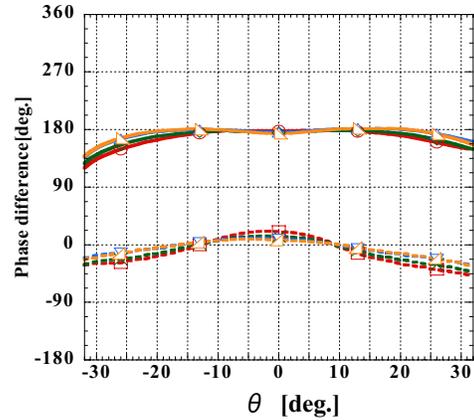


Figure 6: Phase differences (four-layer).

layer (d) is chosen as a parameter. In these figures, the results of (a) three-layer (without a control layer) and (b) four-layer (with a control layer) structures are shown. Here, the scattering patterns are normalized by the maximum field amplitude that is scattered from a perfectly conducting plane. It is understood that the main-polarized components hardly depend on whether the state is recorded or non-recorded. On the other hand, the amplitudes of the cross-polarized components differ depending on the recording states.

The phase differences between the main-polarized and cross-polarized components collected by object lens are shown in Figs. 5 and 6. As shown in Fig. 5, the phase differences of the three-layer structures vary almost sinusoidally with respect to θ for both of recorded and non-recorded states. On the other hand, from Fig. 6, the phase differences of four-layer structures take almost constant values (0 degrees for non-recorded states and 180 degrees for recorded ones).

4. CONCLUSION

We analyzed scattering characteristics of a DWDD disk under considering rear process by means of FDTD method. We have shown that the amplitudes of the cross-polarized components differ depending on the recording states. We have also shown that the the phase differences of four-layer structures take almost constant values.

REFERENCES

1. Kai, S., A. Fukumoto, K. Aratani, S. Yoshimura, K. Tsutsumi, and M. Arai, "Over 15-GB capacity domain wall displacement detection magneto optical recording using a digital versatile disc dimensional optical head," *Jpn. Appl. Phys.*, Vol.39, No.4A, 1757–1761, 2000.
2. Hiroki, T., K. Deguchi, K. Morita, R. Yokoyama, M. Konishi, Y. Miyaoka, and O. Koyama, "Anneal-less domain wall displacement detection of 27 Gbit/in² land/groove recording using a deep groove substrate and a blue laser," *Jpn. J. Appl. Phys.*, Vol. 44, No. 5B, 3568–3573, 2005.
3. Kobayashi, I., T. Kojima, S. Fukai, and Y. He "Numerical analysis of light-beam diffraction from magneto-optical disk medium by FDTD method," *IEICE Trans. Electron*, Vol. E84-C, No. 9, 1189–1195, 2001.

Electromagnetic Wave Propagating in Gyroelectric Slab in the Perpendicular Configuration

Hui Huang^{1,2}, Bo Yi², and Bo Huang³

¹School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China

²State Key Laboratory of Millimeter Waves, Nanjing 210096, China

³School of Electronics Engineering and Computer Science, Peking University, 100871, China

Abstract— This paper present the characteristics of electromagnetic wave propagating in gyroelectric slab with an external magnetic field perpendicular to the interface between gyroelectric medium and a perfect conductor. First, using KDB coordinate system, we decomposed an electromagnetic wave in infinite gyroelectric medium into 2 types and got the dispersion relation respectively. Second, we discuss the reflection from the interface between gyroelectric medium and a perfect conductor for inclined incidence case. The conclusion that the reflected wave has the same ellipticity but the opposite rotate direction to the incident wave is proofed theoretically. Finally, the characteristics of metallic waveguide in Perpendicular Configuration have been discussed, and the guidance condition has been derived. And we found the main mode is zero mode, which is similar to the ordinary wave in infinite gyrotropic medium.

1. INTRODUCTION

Applied a magnetic field on an electron plasma, it becomes gyroelectric medium, an anisotropic medium. The characteristics of electromagnetic waves propagation in gyrotropic plasmas have been theoretically investigated for years. Kushwaha and Halevi have been studied the magnetoplasma modes in Voigt, perpendicular, and Faraday configurations [1–3], Gillies and Hlawiczka have done some researches on gyrotropic waveguide [4–8], and Eroglu, as well as Li, have investigated dyadic Green's functions for gyrotropic medium [9–11]. Furthermore, there are some studies focusing upon the effects of magnetic field on semiconducting plasma slab and negatively refracting surfaces [12, 13]. Moreover, propagation and scattering characteristics in gyrotropic systems [9, 14–18] and surface modes at the interface of a special gyrotropic medium [19] have been investigated extensively.

In this paper, we focus on the characteristics of electromagnetic wave propagating in gyroelectric slab with an external magnetic field perpendicular to the interface between gyroelectric medium and a perfect conductor.

2. DISPERSION RELATIONS

consider the gyroelectric medium with permeability μ_2 and permittivity $\bar{\epsilon}_2$, which is a tensor and takes the following form:

$$\bar{\epsilon}_2 = \begin{bmatrix} \epsilon_{xx} & i\epsilon_g & 0 \\ -i\epsilon_g & \epsilon_{yy} & 0 \\ 0 & 0 & \epsilon_{zz} \end{bmatrix}, \quad (1)$$

where elements are given by

$$\epsilon_{xx} = \epsilon_{yy} = \epsilon_\infty \left(1 - \frac{\omega_p^2}{\omega^2 - \omega_c^2}\right), \quad \epsilon_{zz} = \epsilon_\infty \left(1 - \frac{\omega_p^2}{\omega^2}\right), \quad \epsilon_g = \epsilon_\infty \left[-\frac{\omega_p^2 \omega_c}{\omega(\omega^2 - \omega_c^2)}\right]. \quad (2)$$

Here, $\omega_p = \sqrt{Nq_e^2/m_{eff}\epsilon_\infty}$ and $\bar{\omega}_c = q_e\bar{B}_0/m_{eff}$ are the plasma and cyclotron frequencies respectively, ϵ_∞ is the background permittivity, N is the electron density, m_{eff} is the effective mass, and q_e is the electron charge.

Using KDB coordinate system, we can get the dispersion relations

$$\omega^2 = \frac{\nu}{2} \left[\kappa (k^2 + k_z^2) + \kappa_z k_s^2 \pm \sqrt{(\kappa - \kappa_z)^2 k_s^4 + 4\kappa_g^2 k_z^2 k^2} \right] \quad (3)$$

where k_z is z component of \bar{k} , k_s is its component in xy plane, and

$$\kappa = \frac{\epsilon}{\epsilon^2 - \epsilon_g^2}, \quad \kappa_g = \frac{-\epsilon_g}{\epsilon^2 - \epsilon_g^2}, \quad \kappa_z = \frac{1}{\epsilon_z}.$$

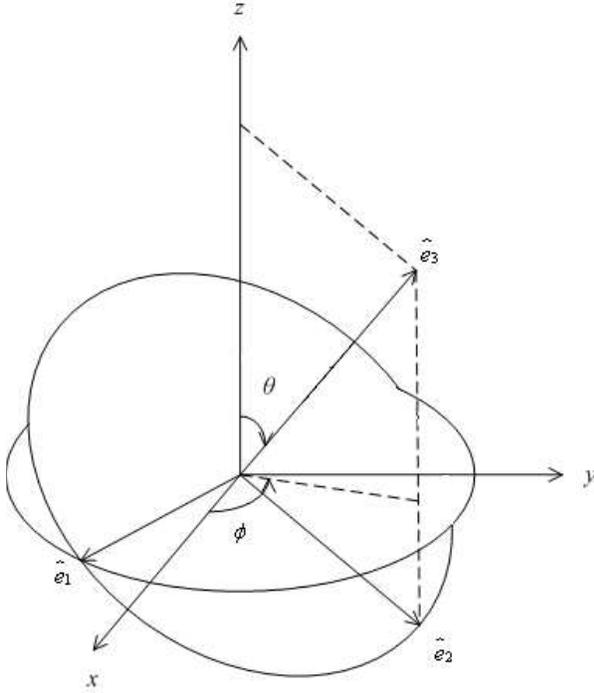


Figure 1: The KDB system.

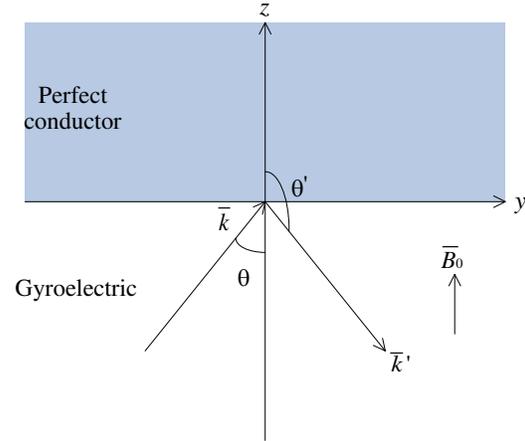


Figure 2: The incidence of elliptical polarization wave to the interface between gyroelectric medium and perfect conductor.

We call the characteristic wave with the plus sign in Eq. (3) Type I wave, while the wave with the minus sign Type II wave.

3. REFLECTION ON THE SURFACE BETWEEN GYROELECTRIC MEDIUM AND PERFECT CONDUCTOR

We consider the configuration in Fig. 2, where a plane wave (suppose it is either Type I or Type II) is incident from an semi-infinite gyroelectric medium into a perfect conductor at an oblique angle θ_i with respect to the normal of the interface.

In KDB system,

$$\bar{D} = \hat{e}_1 D_1 + \hat{e}_2 D_2 \quad (4)$$

We transform it into the Cartesian coordinate system, and in the expression in the matrix form is

$$\bar{D} = (D_1 \quad D_2 \cos \theta \quad D_2 \sin \theta)^T \quad (5)$$

Thus we can get the electric field

$$\bar{E} = \bar{\kappa} \cdot \bar{D} = \begin{pmatrix} \kappa & i\kappa_g & 0 \\ -i\kappa_g & \kappa & 0 \\ 0 & 0 & \kappa_z \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \cos \theta \\ D_2 \sin \theta \end{pmatrix} = \begin{pmatrix} \kappa D_1 + i\kappa_g D_2 \cos \theta \\ -i\kappa_g D_1 + \kappa D_2 \cos \theta \\ \kappa_z D_2 \sin \theta \end{pmatrix} \quad (6)$$

According to the boundary condition of perfect conductor, i.e., continuous tangential component of electric field, we can get that the x and y components of the reflected wave:

$$\begin{aligned} E'_x &= -E_x = -(\kappa D_1 + i\kappa_g D_2 \cos \theta) \\ E'_y &= -E_y = i\kappa_g D_1 - \kappa D_2 \cos \theta \\ D'_x &= \varepsilon E'_x + i\varepsilon_g E'_y = -(\varepsilon \kappa + \varepsilon_g \kappa_g) D_1 - (\varepsilon \kappa_g + \varepsilon_g \kappa) i D_2 \\ D'_y &= -i\varepsilon_g E'_x + \varepsilon E'_y = (\varepsilon \kappa_g + \varepsilon_g \kappa) i D_1 - (\varepsilon \kappa + \varepsilon_g \kappa_g) D_2 \end{aligned} \quad (7)$$

Simplified them in KDB system, we can get

$$\begin{aligned} D'_x &= -D_1 \\ D'_y &= -D_2 \cos \theta \end{aligned} \quad (8)$$

Only on the DB plane can it be electric displacement vector, we can get the electric displacement vector of the reflected wave in kDB coordinate system

$$\bar{D}' = -\hat{e}_1' D_1 + \hat{e}_2' D_2 \quad (9)$$

We can find out from the formula mentioned above that the reflected wave has the same ellipticity but the opposite rotate direction to the incident wave.

Now we will prove the existence of this kind of reflected wave. Notice that the kDB coordinate system has changed compared to the incident wave, that is $\theta' = \pi - \theta$, then,

$$\tan 2\psi' = \frac{2\kappa_g \cos \theta'}{(\kappa - \kappa_z) \sin^2 \theta'} = \frac{2\kappa_g \cos \theta}{(\kappa - \kappa_z) \sin^2 \theta} = -\tan 2\psi \quad (10)$$

so $\psi' = -\psi$, and $\tan \psi' = -\tan \psi$. Which demonstrate that this type of reflected wave does exist, and is the same type as the incident one.

Hlawiczka once mentioned this condition as the supposed condition for the study of gyrotropic slab waveguide. He extended this general condition to gyroelectric medium, supposing that the reflected wave has the same ellipticity and rotate direction to the incident wave, while the calculation above shows that the reflected wave has the same ellipticity but the opposite rotate direction to the incident wave.

4. CHARACTERISTICS OF METALLIC WAVEGUIDE IN PERPENDICULAR CONFIGURATION

Considering the waveguide mode as vector combination of incident and reflected wave, we derived the guidance condition. We suppose that incident and reflected wave vectors are

$$\begin{aligned} \bar{k}_i &= \hat{y}k_y + \hat{z}k_z \\ \bar{k}_r &= \hat{y}k_y - \hat{z}k_z \end{aligned} \quad (11)$$

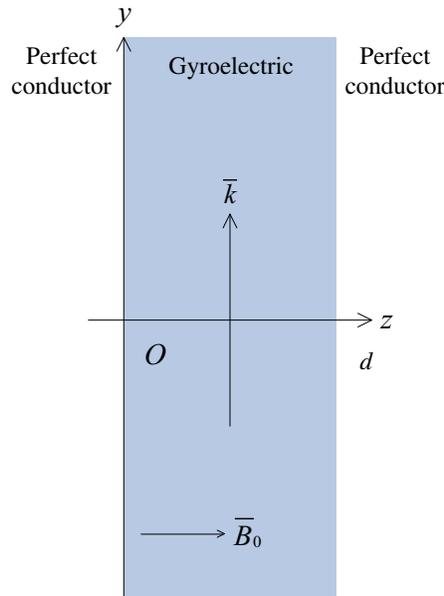


Figure 3: Gyroelectric slab in the perpendicular configuration.

And electric displacement vectors are

$$\begin{aligned}
 \bar{D}_i &= \left(\hat{x}D_1 + \hat{y}\frac{k_z}{k}D_2 - \hat{z}\frac{k_y}{k}D_2 \right) e^{-i(k_y y + k_z z - \omega t)} \\
 \bar{D}_r &= \left(-\hat{x}D_1 - \hat{y}\frac{k_z}{k}D_2 - \hat{z}\frac{k_y}{k}D_2 \right) e^{-i(k_y y - k_z z - \omega t)} \\
 \bar{D} &= -\hat{x}2i \sin(k_z z) D_1 e^{-i(k_y y - \omega t)} - \hat{y}2i \sin(k_z z) \frac{k_z}{k} D_2 e^{-i(k_y y - \omega t)} \\
 &\quad - \hat{z}2 \cos(k_z z) \frac{k_y}{k} D_2 e^{-i(k_y y - \omega t)}
 \end{aligned} \tag{12}$$

According to boundary conditions of perfect conductor, we can get the guidance condition:

$$k_z = \frac{m\pi}{d} \tag{13}$$

where m is integer.

If $m = 0$

$$\bar{D} = -\hat{z}2 \frac{k_y}{k} D_2 e^{-i(k_y y - \omega t)} \tag{14}$$

Electric field only has the z component, but magnetic field intensity has the component of x and y , which equivalents to ordinary wave.

5. CONCLUSION

This paper investigates characteristics of electromagnetic wave propagating in gyroelectric slab with an external magnetic field perpendicular to the interface between gyroelectric medium and a perfect conductor. After deriving the dispersion relation of the medium, we discuss the reflection at the interface between the gyroelectric medium and the perfect conductor. Then we derive the guidance condition of the slab waveguide.

ACKNOWLEDGMENT

This work is sponsored by State Key Laboratory of Millimeter Waves under Contract K201012.

REFERENCES

1. Kushwaha, M. S. and P. Halevi, "Magnetoplasmons in thin films in the Voigt configuration," *Physical Review B*, Vol. 36, 5960–5967, 1987.
2. Kushwaha, M. S. and P. Halevi, "Magnetoplasma modes in thin films in the Faraday configuration," *Physical Review B*, Vol. 35, 3879–3889, 1987.
3. Kushwaha, M. S. and P. Halevi, "Magnetoplasmons in thin films in the perpendicular configuration," *Physical Review B*, Vol. 38, 12428–12435, 1988.
4. Gillies, J. R. and P. Hlawiczka, "TE and TM modes in gyrotropic waveguides," *J. Phys. D: Appl. Phys.*, Vol. 9, 1315, 1976.
5. Gillies, J. R. and P. Hlawiczka, "Elliptically polarized modes in gyrotropic waveguides. II. An alternative treatment of the longitudinally magnetized case," *J. Phys. D: Appl. Phys.*, Vol. 10, 1891, 1977.
6. Hlawiczka, P., "Elliptically polarized modes in gyrotropic waveguides," *J. Phys. D: Appl. Phys.*, Vol. 9, 1957, 1976.
7. Hlawiczka, P., "A gyrotropic waveguide with dielectric boundaries: the longitudinally magnetized case," *J. Phys. D: Appl. Phys.*, Vol. 11, 1157, 1978.
8. Hlawiczka, P., "The gyrotropic waveguide with a normal applied DC field," *J. Phys. D: Appl. Phys.*, Vol. 11, 1941, 1978.
9. Eroglu, A. and J. K. Lee, "Wave propagation and dispersion characteristics for a nonreciprocal electrically gyrotropic medium," *Progress In Electromagnetics Research*, Vol. 62, 237–260, 2006.
10. Li, L.-W., N.-H. Lim, and J. A. Kong, "Cylindrical vector wave function representation of Green's dyadic in gyrotropic bianisotropic media — Abstract," *Journal of Electromagnetic Waves and Applications*, Vol. 17, No. 11, 1589–1591, 2003.

11. Li, L.-W., N.-H. Lim, W.-Y. Yin, and J.-A. Kong, “Eigenfunctional expansion of dyadic Green’s functions in gyrotropic media using cylindrical vector wave functions — Abstract,” *Journal of Electromagnetic Waves and Applications*, Vol. 17, No. 12, 1731–1733, 2003.
12. Ivanov, S. T. and N. I. Nikolaev, “Magnetic-field effect on wave dispersion in a free semiconductor plasma slab,” *J. Phys. D: Appl. Phys.*, Vol. 32, 430, 1999.
13. Boardman, A., N. King, Y. Rapoport, and L. Velasco, “Gyrotropic impact upon negatively refracting surfaces,” *New J. Phys.*, Vol. 7, 1–24, 2005.
14. Zhang, M., L. W. Li, T. S. Yeo, and M. S. Leong, “Scattering by a gyrotropic bianisotropic cylinder of arbitrary cross section: An analysis using generalized multipole technique — Abstract,” *Journal of Electromagnetic Waves and Applications*, Vol. 17, No. 7, 1049–1051, 2003.
15. Yin, W. Y., L. W. Li, and M. S. Leong, “Scattering from multiple bianisotropic cylinders and their modeling of cylindrical objects of arbitrary cross-section — Abstract,” *Journal of Electromagnetic Waves and Applications*, Vol. 14, No. 5, 611–612, 2000.
16. Tan, E. L. and S. Y. Tan, “Cylindrical vector wave function representations of electromagnetic fields in gyrotropic bianisotropic media,” *Journal of Electromagnetic Waves and Applications*, Vol. 13, No. 11, 1461–1476, 1999.
17. Bass, F. and L. Resnick, “Spatial and temporal rotation of the polarization plane of electromagnetic waves reflected from and transmitted through a gyrotropic plate,” *Journal of Electromagnetic Waves and Applications*, Vol. 17, No. 8, 1131–1137, 2003.
18. Censor, D. and M. D. Fox, “Polarimetry in the presence of various external reflection and retrodirection mirroring mechanisms, for chiral and gyrotropic media,” *Journal of Electromagnetic Waves and Applications*, Vol. 11, No. 3, 297–313, 1997.
19. Huang, H., Y. Fan, B.-I. Wu, F. Kong, and J. A. Kong, “Surface modes at the interfaces between isotropic media and uniaxial plasma,” *Progress In Electromagnetics Research*, Vol. 76, 1–14, 2007.

Three Dimensional FDTD Analysis of Near-field Optical Disk

Shingo Iwata and Toshiaki Kitamura

Faculty of Engineering Science, Kansai University, Osaka, Japan

Abstract— Scattering characteristics of a near-field optical disk with a bow-tie shaped metallic nano-aperture were investigated. In the analysis, finite-difference time-domain (FDTD) method into which motion equations of free electrons are installed were used and a three dimensional disk structure was considered. Electromagnetic field distributions of optical near-field around the aperture and far-field scattering patterns from the phase change were analyzed.

1. INTRODUCTION

The recording density of conventional optical recording systems has a limitation due to optical diffraction limit. Recording methods based on near-field optical principles have stimulated much interest in optical storage technologies because they have the potential to overcome the limitation by writing and reading recording marks using a localized optical near-field. Metallic nano-apertures have attracted much attention for the method to obtain near-field light and many types of apertures have been studied [1, 2]. Bow-tie shaped apertures are one of promising method to obtain high emission intensity and a small spot simultaneously [3, 4].

We have proposed a near-field optical disk structure with an acute-edged metallic nano-aperture and analyzed their scattering characteristics by using two-dimensional FDTD method into which motion equations of free electrons are installed to analyze a metallic material [5]. In this study, we expand this method to the three-dimensional analysis. We consider applying a bow-tie shaped nano-aperture to a near-field optical disk that consists of a phase change disk. We analyze electromagnetic field distributions of optical near-field around the aperture. We also analyze far-field scattering patterns from the phase change disk.

2. FDTD FORMULATION

Generally, the permittivity of a metallic material is dispersive and has a negative value in optical frequency. Therefore, in order to deal with a metallic material in the FDTD method, some special treatment is indispensable. In this paper, we introduce the following motion equation into the FDTD method to evaluate conducting current.

$$m \frac{\partial \mathbf{u}}{\partial t} = -e\mathbf{E} - m\nu\mathbf{u}$$

where \mathbf{u} is the electron velocity, \mathbf{E} the electric field, e the elementary electric charge, m the electron mass and ν the collision frequency, respectively. The collision frequency ν is expressed as follows.

$$\nu = \frac{2n_1n_2\omega}{1 - n_1^2 + n_2^2}$$

where ω is the angular frequency of light wave, and n_1 and n_2 are the real and imaginary parts, respectively, of the complex refractive index of metallic medium ($N = n_1 - jn_2$).

When we represent current density by the electron velocity \mathbf{u} and the electron density n_0 , Maxwell's equations are expressed as follows.

$$\begin{aligned} \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} &= \nabla \times \mathbf{H} + n_0 e \mathbf{u} \\ \mu_0 \frac{\partial \mathbf{H}}{\partial t} &= -\nabla \times \mathbf{E} \end{aligned}$$

where \mathbf{H} is the magnetic field, and ε_0 and μ_0 are the electric permittivity and magnetic permeability of free space, respectively. FDTD cell for the analysis are shown in Fig. 1.

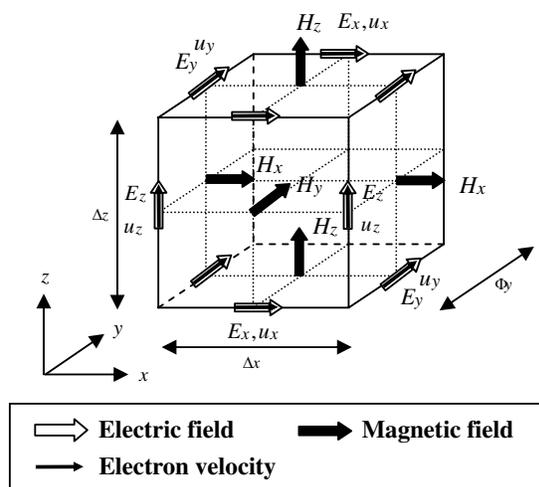


Figure 1: FDTD cell.

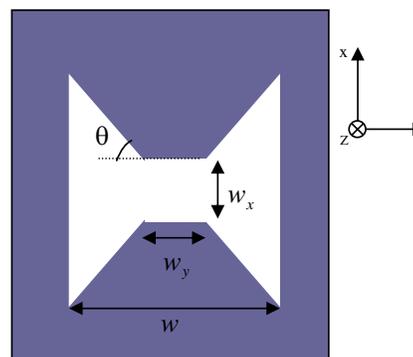
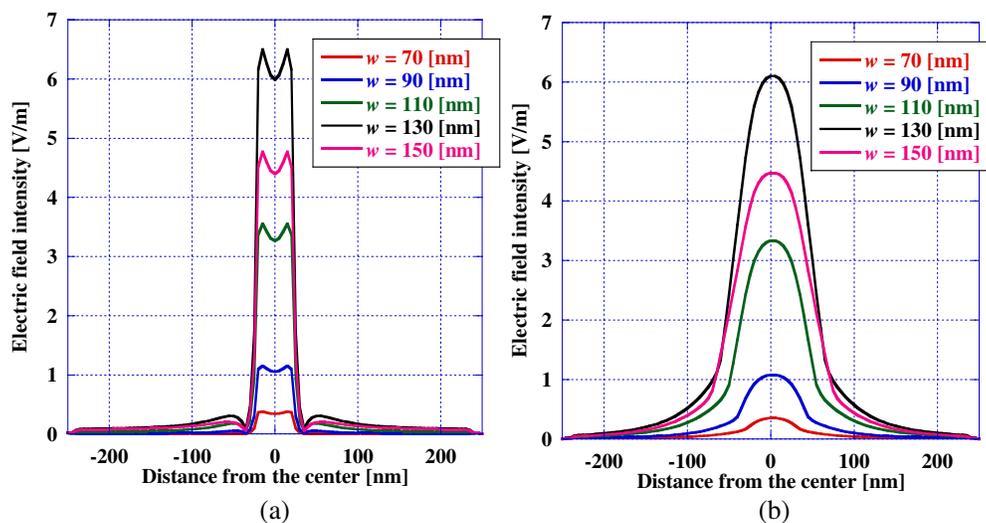


Figure 2: Bow-tie aperture.

Figure 3: Electric field intensity along (a) x and (b) y axes.

3. FIELD INTENSITY OF NEAR-FIELD LIGHT

Figure 2 shows the cross section of a bow-tie shaped nano-aperture. The parameters and coordinate axes are shown in these figures. Here, for the metallic material we assume silver, which has a complex refractive index $N = 0.07 - 4.2j$ and a thickness $t = 100$ nm. A p -polarized plane wave with the wavelength of 650 nm is illuminated vertically to the aperture. The electric field is evaluated at the distance of 10 nm along the z axis from the aperture. Fig. 3 shows the field intensity of near-field light with w as a parameter when $w_x = w_y = 50$ nm, $\theta = 30$ degree. The field intensity becomes maximum when $w = 130$ nm.

4. SCATTERING CHARACTERISTICS

Figure 4(a) shows the structure of a near-field optical disk that consists of a metallic nano-aperture, a phase change disk, two protect films and a substrate. Fig. 4(b) shows the cross section of a bow-tie shaped nano-aperture. Here, we assume silver as a metallic material. The refractive indexes of the optical disk structure are shown in Table 1. The incident light is a p -polarized Gaussian beam with the wavelength of 650 nm and the numerical aperture of an object lens is assumed to be 0.6.

The transmitted far-field patterns in the x - z and y - z planes are shown in Figs. 5(a) and (b), respectively. Here, $t = 100$ nm, $w_x = w_y = 50$ nm, $w = 130$ nm. A cylindrical mark pattern with a diameter (mark size) w_m is assumed. It is understood that the transmitted patterns change according to w_m .

Table 1: Refractive indexes of the optical disk structure.

| | |
|----------------------|--|
| a metallic film (Ag) | 0.070–4.20 <i>j</i> |
| Phase change film | 4.2–1.9 <i>j</i> (amorphous) 4.6–4.2 <i>j</i> (crystalline) |
| Protection film | 2.2 |
| Substrate | 1.5 |

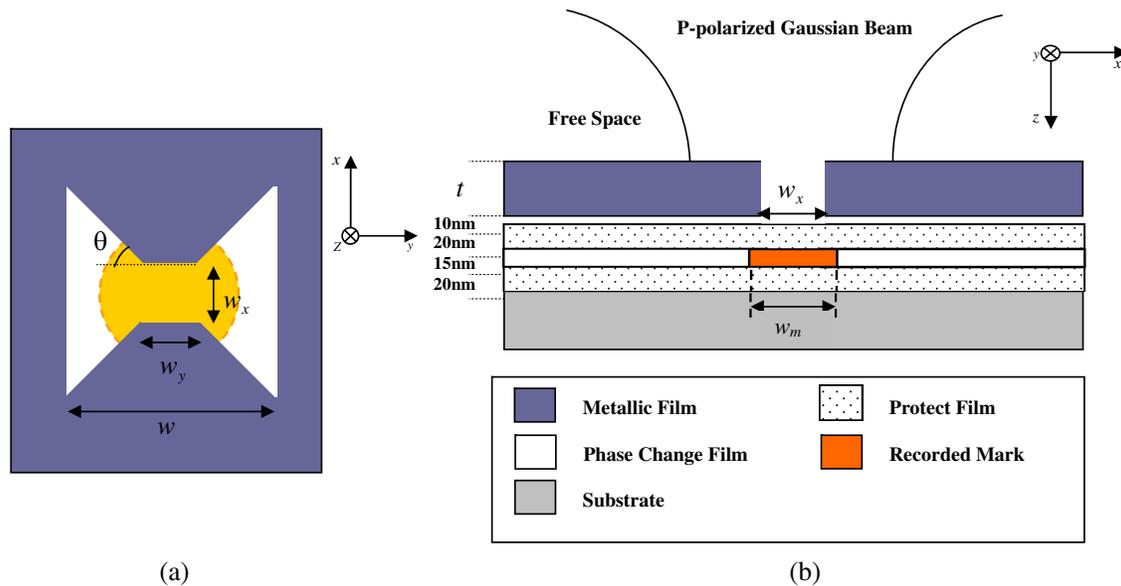


Figure 4: Near-field optical disk.

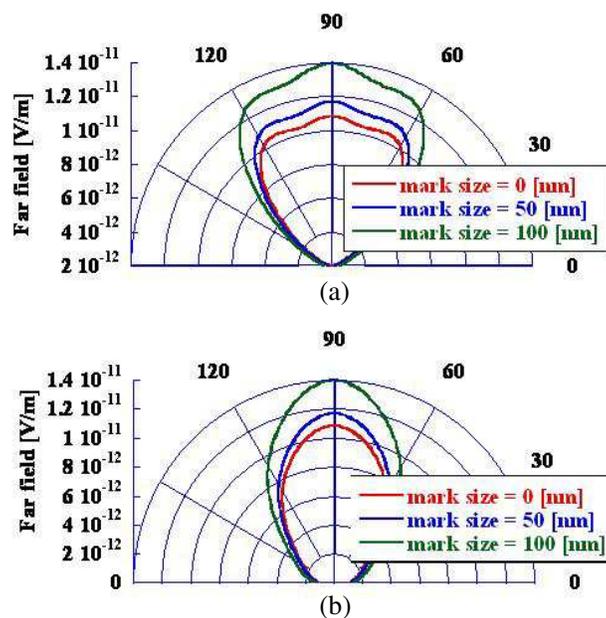


Figure 5: Transmitted far-field.

5. CONCLUSION

Scattering characteristics of a near-field optical disk with a bow-tie shaped metallic nano-aperture were investigated. In the analysis, finite-difference time-domain (FDTD) method into which motion equations of free electrons are installed were used. Electromagnetic field distributions of optical near-field around the aperture and far-field scattering patterns from the phase change were analyzed.

REFERENCES

1. Chen, Y., J. Fang, C. Tien, and H. D. Shien, “Double-corrugated C-shaped aperture for near-field recording,” *Jpn. J. Appl. Phys.*, Vol. 45, No. 2B, 1348–1350, 2006.
2. Eom, G. S., D. Yang, E. Lee, S. Park, Y. Lee, and J. W. Hahn, “Wave propagation characteristics of a figure-eight shaped nanoaperture,” *J. Appl. Phys.*, Vol. 101, 103101, 1–4, 2007.
3. Park, D., H. J. Kim, B. H. O, S. G. Park, E. Lee, and S. G. Lee, “Effect of incident beam width on light transmission enhancement by bow-tie-shaped nano-aperture,” *Jpn. J. Appl. Phys.*, Vol. 46, No. 12, 7991–7994, 2007.
4. Wu, Y. M., L. W. Li, and B. Liu, “Gold bow-tie shaped aperture nanoantenna: Wide band near-field resonance and far-field radiation,” *IEEE Trans. Magn.*, Vol. 46, No. 6, 1918–1921, 2010.
5. Kitamura, T. and S. Iwata, “Analysis of a near-field optical disk with an acute-edged metallic nano-aperture,” *IEICE Trans. Electron.*, Vol. E93-C, No. 9, 1474–1477, 2010.

Mapping Technique of Basic Magnetic Field in MR Tomography

M. Hadinec¹, P. Fiala¹, and K. Bartušek²

¹Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejní 2906/4, Brno 612 00, Czech Republic

²Institute of Scientific Instruments of the ASCR, v.v.i
Královopolsk 147, Brno 612 64, Czech Republic

Abstract— The paper presents a mapping method, which is used to obtain boundary conditions on desired volume. The process of obtaining the magnetic flux density from NMR spectral characteristics is explained. The spherical functions are used for magnetic flux density approximation. The discussion of suitable degree and order for approximation is made. Then maps of magnetic flux density inside measured volume are computed. These maps are compared to measured phase images. Phase images are gained by mapping technique based on MRI methods. The comparison algorithm is made in Matlab and results are discussed.

1. INTRODUCTION

In many modern applications and technologies, the stable magnetic fields of precisely defined shape are necessary. One of these applications is Nuclear Magnetic Resonance (NMR). In order to obtain high-quality images, the high homogeneity of static magnetic field B_0 in the desired imaging volume is required [1]. However, present manufacturing technologies are not able to produce permanent or superconducting magnets. Therefore, a lot of magnetic field corrections methods called shimming was developed in order to improve the magnetic field homogeneity into desired level [2, 3]. In principle, there are two shimming concepts, passive and active corrections. In most of the present NMR applications are used active correction coils, generating shimming gradients of defined shape [4]. For effective shimming, the exact shape of non-homogeneous B_0 field should be known. In present literature, we find a lot of different mapping techniques and instruments, which speed up the whole process [5]. This article describes a method to obtain magnetic flux density maps of basic magnetic field B_0 and reverse approximation technique based on spherical functions expansion, which can be used for selecting the correct shimming gradients. To speed up process of mapping magnetic field in desired volume, the comparison of computed and measured phase images is made, in order to show errors for both of these different techniques. It is obvious, that the obtaining of magnetic flux density maps using GE phase images is much quicker than measuring whole volume in discrete points.

2. APPROXIMATION METHOD

If we will measure a certain number of points on the surface of any volume, which is situated inside basic magnetic field B_0 , these values can be considered as boundary conditions for Laplace's equation solution. Magnetic flux density can be then approximated at any point of measured area. These points can be selected using spherical coordinates $[r, \theta, \varphi]$, so we can define the approximation formula as follows

$$B_a(r, \theta, \varphi) = \sum_{k=0}^{N_K} \cdot \sum_{m=0}^{m=k} \cdot r^k \cdot P_{m,k}(\cos \theta) \cdot [A_{m,k} \cos m \cdot \varphi + B_{m,k} \sin m \cdot \varphi] \quad (1)$$

Coefficients $A_{m,k}$ a $B_{m,k}$ is then possible to find like the minimum value of this formula

$$\Psi = \min \sum_{i=1}^{N_m} (B_{im} - B_{ia})^2 \quad (2)$$

where B_m are measured values of magnetic induction at the desired area (circle, sphere, cylinder) and B_{ia} are approximated values of magnetic induction. This method is known as Least squares method (LSM).

3. MAGNETIC FLUX DENSITY MAPPING METHOD

For discrete mapping of basic magnetic field B_0 is used a small glass NMR probe, which is filled with water. For measuring was used simple one RF-pulse method. For each measured point is we gain frequency spectrum and mean value. This value is shifted against frequency in the centre of measured volume (Figure 1).

If the magnetic flux density B is given by equation

$$B = B_0 + \Delta B \tag{3}$$

where B_0 is homogeneous magnetic field and ΔB is magnetic field inhomogeneity, then we can formulate measured frequency f by equation

$$f = f_0 + \Delta f = f_0 + \gamma' \Delta B \tag{4}$$

where f_0 is frequency in the centre of measured volume and $\gamma' = \gamma/2\pi$ is gyromagnetical ratio of 1H nuclei. We can say, that variable Δf is given by equation

$$\Delta f = \gamma' \Delta B \tag{5}$$

This equation describes the inhomogeneity in measured discrete point.

4. APPROXIMATION RESULTS

Using LSQ in Matlab, we have obtained approximation coefficients $A_{m,k}$ a $B_{m,k}$, which can be used for setting up correction gradients. Using the solution of Laplace's equation, we can compute magnetic field at any point of measured volume. The comparison of measured slice and computed slice is at Figure 2. We can see, that the approximation for degree $Q = 8$ has very good results, as was expected. The big red speckle in the lower part of difference map can be caused by gross error during measuring.

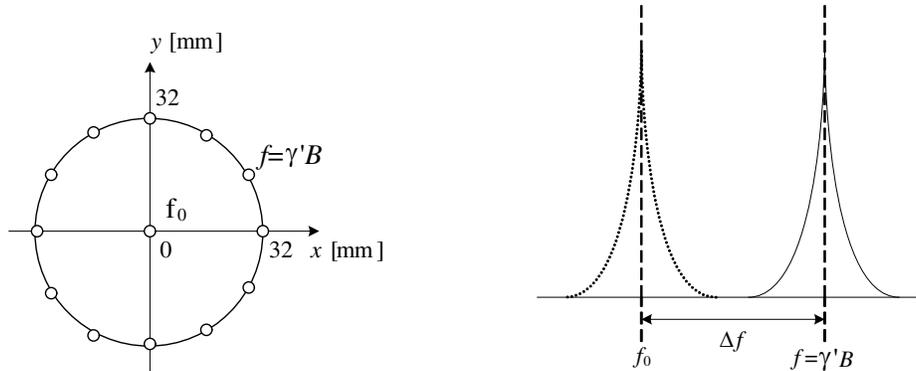


Figure 1: Obtaining of magnetic flux density from measured spectral characteristic.

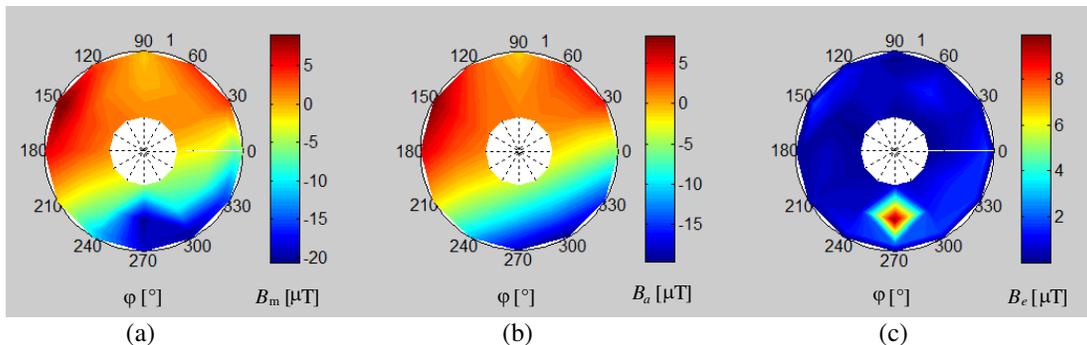


Figure 2: (a) Measured map in discrete points, (b) approximated map, (c) difference map, $Q = 8$.

5. APPROXIMATION DEGREE AND ORDER

It is known, that success of magnetic field shimming depends on quantity and quality of shimming gradients. It is obvious, that also the accuracy of approximation method depends on quantity of spherical functions. The number of spherical functions in magnetic flux density expansion is given by their degree m and order k . The spherical functions degree and order can be described by one variable, approximation degree Q ($Q = k + 1$). We have made a comparison of approximated maps for different approximation orders from $Q = 2$ to $Q = 22$ and computed maximal error between measured and approximated values, named B_{emax} . The result is on Figure 3.

We can see, that the best results are made for approximation degree $Q = 10$ and $Q = 11$. This conclusion is valid only for approximation on surface of measured volume. When we will try to compute magnetic flux density inside the measured volume, the error for $Q = 10$ is too high and maps are very distorted. We can find the reason in analogy with Shannon-Kotelnik theorem for functions sampling and also in number of measured points, which are used for approximation. It is necessary to obtain samples (measured points) at least two times during the period of function of the highest order and degree. The functions are given by expansion (1), it means functions $\cos(m\varphi)$ and $\sin(m\varphi)$. Then is obvious, that we must fulfill equation

$$\Delta w < 0.5t_f \tag{6}$$

where t_f is period of function relating to $A_{m,k}$ a $B_{m,k}$ coefficients, Δw is measuring step of angle φ or shift z . The course of two goniometric functions is presented on Figure 4 and Figure 5.

During measuring of discrete points on the surface of mapped volume we used stepping $\Delta\varphi = 30^\circ$. It is then obvious, that for fulfilling the Equation (6) we should use approximation around $Q = 7$. During our experiments in Matlab we realized, that the matching of measured and approximated maps is also influenced by tolerance limits of LSQ method.

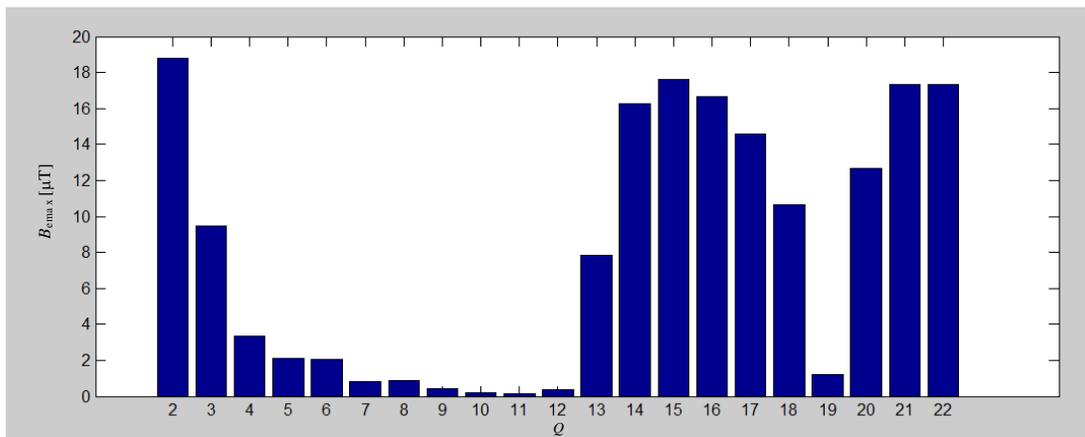


Figure 3: The influence of approximation degree Q on approximation error B_{emax} .

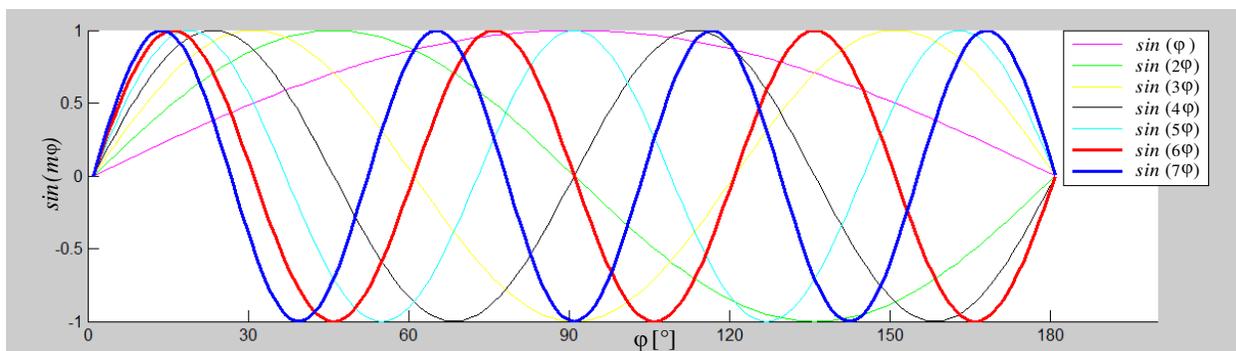


Figure 4: The course of $\sin(m\varphi)$ for parameters $m < 7, n < 7, Q = 8$.

6. PHASE IMAGES AND APPROXIMATED MAPS COMPARISON

Phase images using GE method for different offset frequencies were measured and unwrapped. This enables to obtain maps of magnetic flux density much quicker than measuring in discrete points (Figure 6). The noisy area in the right side of Figure 6(a)) is caused by air bubble in spherical volume sample, which is filled with water.

The GE map is presented as bitmap in Matlab. It means, that there is a matrix of many points representing values in of magnetic flux density in Cartesian coordinates. It is necessary to “sample” this matrix (Cartesian coordinates) and obtain values in points corresponding to polar coordinates at approximated maps. Unfortunately, this transformation is not perfect and some points are not at their optimal positions, as can be seen from Figure 6(b)). However, if we compute differential

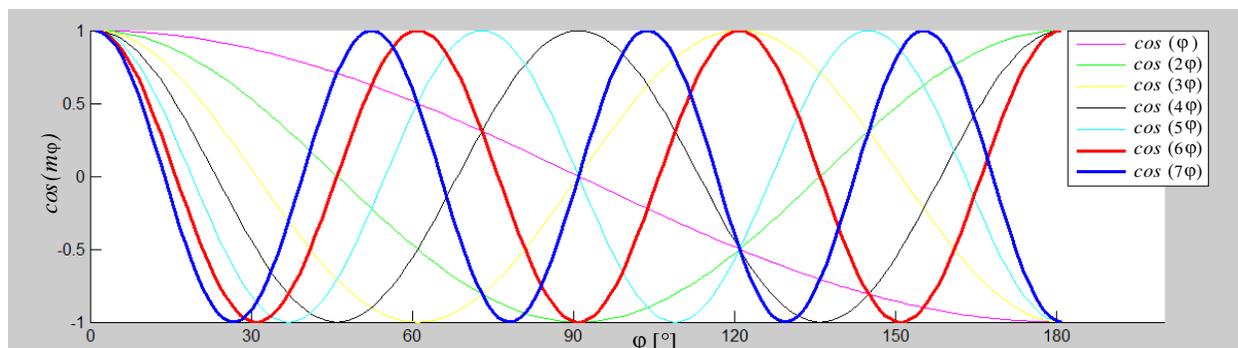


Figure 5: The course of $\cos(m\varphi)$ for parameters $m < 7$, $n < 7$, $Q = 8$.

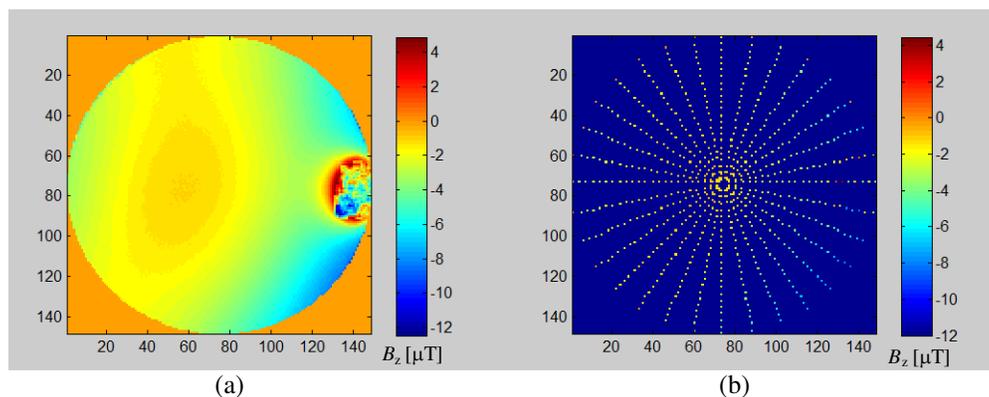


Figure 6: (a) GE map at transversal direction (offset 1000 Hz), (b) sampling points.

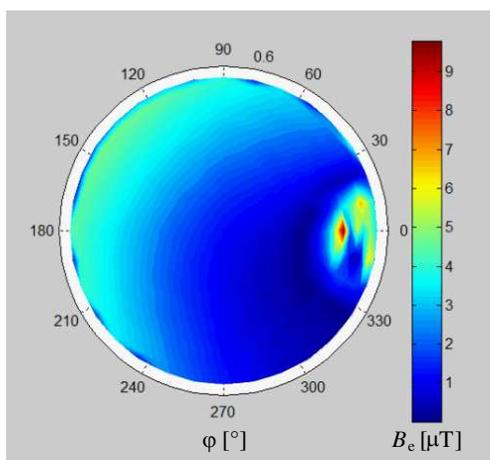


Figure 7: Differential image between approximated and GE measured values.

image, we can see, that the biggest error is around the area of air bubble, which was expected. Other areas of map are quite similar, the maximum error is approximately not more than $4.5 \mu\text{T}$. This error is in the range of used approximation and can be better in the future work.

7. CONCLUSIONS

A mapping and approximation technique used to obtain boundary conditions on desired spherical volume is presented. We made theoretical description of magnetic flux density expansion using spherical harmonics. Approximation accuracy and optimal approximation degree is discussed. The comparison of unwrapped GE phase images and approximated maps is made.

ACKNOWLEDGMENT

The research described in the paper was supported by FRVS (a fund of university development) by research plan No. MSM 0021630513 ELCOM, No. MSM 0021630516 and project of the BUT Grant Agency FEKT-S-10-13.

REFERENCES

1. Haacke, E. M., *Magnetic Resonance Imaging — Physical Principles and Sequence Design*, John Wiley & Sons, 2001, ISBN 0-471-48921-2.
2. Yao, Y. Y. and D. X. Xie, “A hybrid optimal analysis for the design of passive shimming elements in high homogeneous permanent magnets,” *Computation in Electromagnetics*, 2002.
3. Dorri, B. and M. E. Vermilyea, “Passive shimming of MR magnets: algorithm, hardware and results,” *IEEE Transactions on Applied Superconductivity*, Vol. 3, No. 1, March 1993.
4. Gupta, R., M. Anerella, J. Cozzolino, B. Erickson, et al., “Tuning shims for high field quality in superconducting magnets,” *IEEE Transactions on Magnetics*, Vol. 32, No. 4, July 1996.
5. Yan, J., S. K. Huang, Z. L. Hou, Z. X. Liu, L. Zhao, and Z. Zhu, “Magnetic field mapping in the BESIII solenoid,” *IEEE Transactions on Applied Superconductivity*, Vol. 20, No. 3, June 2010.

Modelling of 3D Thin Regions in Magnetostatic NDT Using Overlapping Elements in Dual Formulations

H. Zaidi, L. Santandréa, G. Krebs, and Y. Le Bihan

Laboratoire de Génie Electrique de Paris
(CNRS UMR 8507, Supelec, Univ. Paris-Sud 11, UPMC Paris 6)
11 rue Joliot Curie, 91192 Gif-sur-Yvette, France

Abstract— Ferromagnetic materials can be investigated by observing the reactance of an inductive probe-coil. This paper presents the use of overlapping elements for simplifying the modelling of such a problem in the finite element method.

1. INTRODUCTION

The finite element method (FEM) is widely used for the simulation of non-destructive testing (NDT). Indeed, FEM simulation permits to reduce the experimentation and prototyping costs. In NDT based on inductive probe, the modelling of thin airgap is necessary. The lift-off between the probe and the piece to inspect is usually low (typically lower than 100 mm). The length of the defect is often weak in comparison with the size of the piece to inspect. The skin effect has also to be taken into account in order to have accurate results. Consequently these constraints can lead to very fine meshes (one for each position of the sensor), distorted elements or ill-conditioned matrix. To overcome these numerical problems, a solution can consist in the use of two independent meshes, one for the piece to inspect and the other for the sensor and the surrounding air. The mesh propagation is avoided but a method for connecting the two domains is then required. Several mesh connecting methods can be found [1–3]. Among them, the overlapping finite element method was originally proposed in [4]. The overlapping element method has some attractive aspects. The conditioning of the matrix system to solve is not degraded and the method is relatively straightforward to implement and develop. In a recent contribution [5], the use of overlapping elements with a vector potential formulation was proposed. Using dual formulations (one electric and the other magnetic) exhibits interesting features for mesh refinement and energy calculation [6]. However, the effect of thin regions has not been yet evaluated with dual formulation and overlapping elements.

This paper proposes to evaluate the overlapping elements in a magnetostatic 3D NDT problem using the classical dual formulations based on potentials with nodal and edge Whitney elements. The first and second parts of the paper are dedicated to the description of the used magnetic and electric formulations. In the third part, the overlapping finite element is briefly described. Finally, as an example, a probe-coil characterizing the magnetic permeability of a plate is considered and its inductance computed. Results are then given and discussed.

2. MAGNETIC FORMULATION $\mathbf{t}_0 - \phi$

In the magnetostatic case, the Maxwell's equations become:

$$\operatorname{div}(\mathbf{b}) = 0 \text{ with } \mathbf{b} \cdot \mathbf{n} = 0 \text{ on } \Gamma_e \quad (1)$$

$$\operatorname{curl}(\mathbf{h}) = \mathbf{j} \text{ with } \mathbf{h} \times \mathbf{n} = 0 \text{ on } \Gamma_h \quad (2)$$

with \mathbf{h} the magnetic field, \mathbf{b} the magnetic flux density, \mathbf{j} the current density and Γ_e , Γ_h complementary boundaries that entirely enclose the study domain. The following constitutive law is also considered.

$$\mathbf{b} = \mu \mathbf{h} \quad (3)$$

The excitation current, \mathbf{j} , is introduced by a source vector \mathbf{t}_0 such as:

$$\mathbf{j} = \operatorname{curl}(\mathbf{t}_0) \quad (4)$$

From (2) and (4), a scalar potential ϕ is considered such as:

$$\mathbf{h} = \mathbf{t}_0 - \operatorname{grad}(\phi) \text{ with } \phi \in W_h^0 \{u \in L^2 \text{ and } \operatorname{grad}(u) \in L^2 \text{ and } \phi|_{\Gamma_h} = 0\} \quad (5)$$

The weak variational formulation of (1) leads to.

$$\begin{cases} \text{find } \phi \in W_0^h \text{ such as} \\ \int_{\Omega} \mu \text{grad} \phi \cdot \text{grad} \phi' d\Omega - \int_{\Omega} \mu \mathbf{t}_0 \text{grad} \phi' d\Omega = 0 \quad \forall \phi' \in W_h^0 \end{cases} \quad (6)$$

The scalar potential ϕ is expressed at the nodes of the finite elements and is interpolated using:

$$\phi = \sum_{n=1}^{N_n} \lambda_n \bar{\phi}_n \quad (7)$$

where N_n is the number of nodes of the mesh, λ_n is the shape function of node n . $\bar{\phi}_n$ is the value of ϕ on the n node. Using (7) and the Galerkin method being applied to (6) (The test functions are the shape functions), the following expression to solve is obtained:

$$\sum_{m=1}^{N_n} \left[\int_{\Omega} \mu \text{grad} \lambda_m \cdot \text{grad} \lambda_n d\Omega \right] \bar{\phi}_m - \int_{\Omega} \mu \mathbf{t}_0 \cdot \text{grad} \lambda_n d\Omega = 0 \quad \forall n \in \{1, \dots, N_n\} \quad (8)$$

The expression (8) under matrix form is expressed by (9) where Φ is the vector of unknowns.

$$\begin{cases} R \cdot \Phi = S \text{ with} \\ R_{mm} = \int_{\Omega} \mu \text{grad} \lambda_m \cdot \text{grad} \lambda_m d\Omega \quad \forall n, m \in \{1, \dots, N_n\} \\ S_n = \int_{\Omega} \mu \mathbf{t}_0 \cdot \text{grad} \lambda_n d\Omega = 0 \quad \forall n \in \{1, \dots, N_n\} \end{cases} \quad (9)$$

3. ELECTRIC FORMULATION $\mathbf{a} - \mathbf{t}_0$

Due to expression (1), a magnetic vector potential, noted \mathbf{a} , may be introduced as

$$\mathbf{b} = \text{curl}(\mathbf{a}) \text{ and } \mathbf{a} \times \mathbf{n}|_{\Gamma_e} = 0 \quad (10)$$

Replacing this equality in (2), the following equation is obtained:

$$\text{curl} \left(\frac{1}{\mu} \text{curl}(\mathbf{a}) \right) = \text{curl}(\mathbf{t}_0) \quad \mathbf{a} \in \mathbf{W}_e^1 = \{u \in L^2 \text{ such as } \text{curl}(u) \in L^2 \text{ and } u \times \mathbf{n}|_{\Gamma_e} = 0\} \quad (11)$$

The weak formulation of this equation becomes:

$$\begin{cases} \text{find } \mathbf{a} \in W_1^e \text{ such as} \\ \int_{\Omega} \frac{1}{\mu} \text{curl} \mathbf{a} \cdot \text{curl} \mathbf{a}' d\Omega - \int_{\Omega} \mathbf{a}' \cdot \text{curl}(\mathbf{t}_0) = 0 \quad \forall \mathbf{a}' \in W_1^e \end{cases} \quad (12)$$

The vector potential \mathbf{a} is expressed from its edge circulations and its interpolation inside an element is performed using:

$$\mathbf{a} = \sum_{a=1}^{N_a} \psi_a \bar{a}_a \quad (13)$$

with N_a is the total number of edges in the mesh, ψ_a is the shape function of the edge a and \bar{a}_a is the circulation of \mathbf{a} along the edge a . As previously, the Galerkin method and (13) applied to (12) lead to expression (14).

$$\sum_{m=1}^{N_a} \left[\int_{\Omega} \frac{1}{\mu} \text{curl} \psi_m \cdot \text{curl} \psi_n d\Omega \right] \bar{a}_m - \int_{\Omega} \mathbf{t}_0 \cdot \text{curl} \psi_n d\Omega = 0 \quad \forall n \in \{1, \dots, N_n\} \quad (14)$$

Thereafter, the following matrix system to solve is obtained:

$$\begin{cases} M \cdot A = S \text{ with} \\ M_{mn} = \int_{\Omega} \frac{1}{\mu} \text{curl} \psi_m \cdot \text{curl} \psi_n d\Omega \quad \forall n, m \in \{1, \dots, N_a\} \\ S_n = \int_{\Omega} \mathbf{t}_0 \cdot \text{curl} \psi_n d\Omega = 0 \quad \forall n \in \{1, \dots, N_a\} \end{cases} \quad (15)$$

A is the vector of unknowns.

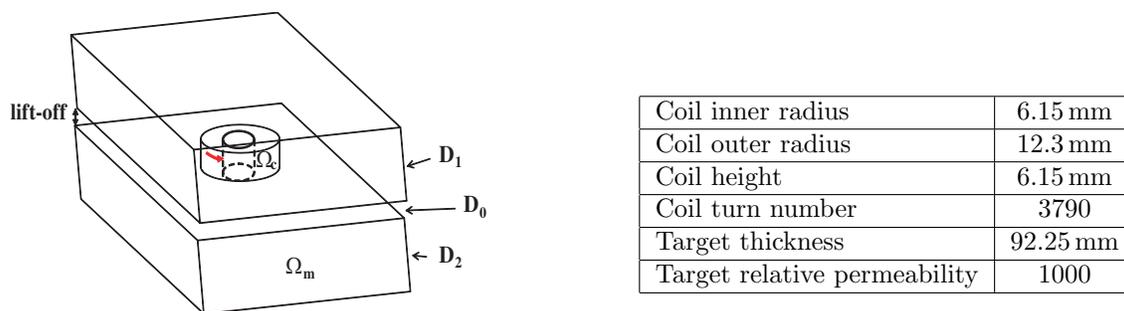


Figure 1: Description of the studied problem.

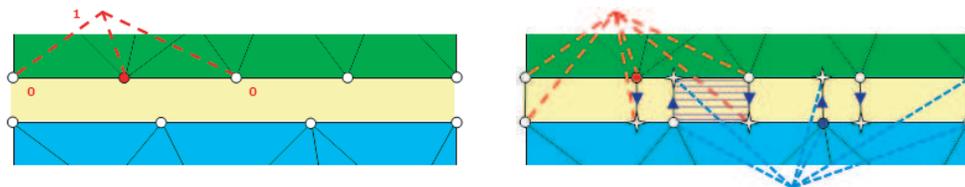


Figure 2: Before and after the node projection for overlapping finite element method.

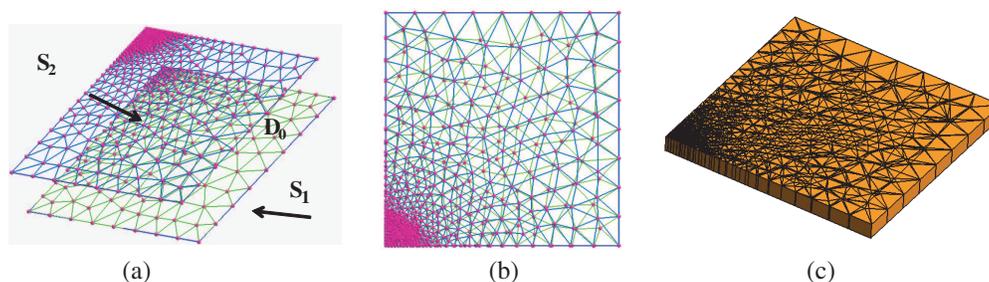


Figure 3: (a) (b) Meshed surfaces to connect 3D and (c) top views and integration volumes.

4. STUDIED PROBLEM

The following figure shows the configuration of the studied geometry. The problem is divided into 3 domains. D_1 is composed of air and includes the coil domain Ω_c where a uniform current density \mathbf{j} is imposed. The second domain D_2 is constituted by the magnetic target Ω_m . The coil and the magnetic target are separated by a non meshed lift-off domain, noted D_0 , where the overlapping element method will be applied. The problem parameters are summarized in the following table.

Only one quarter of the problem will be modelled due to the several planes of symmetry.

5. DESCRIPTION OF THE OVERLAPPING METHOD

Considering two non-conform meshes facing each other, the principle of the approach consists in the projection of the nodes of one boundary on the other and reversely, see Figure 2.

This projection creates overlapping elements. These elements are the support of shape functions and permit to assure the continuity of the interpolated values. The projected nodes (stars in Figure 2) are virtual and are not the support of degrees of freedom. Once the double projection is done, integration areas are created to evaluate the different integral terms of the FEM formulation. In the case of the proposed problem, the meshed surfaces S_1 and S_2 (belonging to D_1 and D_2 , respectively) to connect are given in Figure 3. These surfaces surround the area D_0 .

To calculate the integration volumes, the set of triangles of S_1 and S_2 is considered in the same plane. The intersections between the triangles are then determined. A subdivision of all the intersections gives the integration volumes (prisms). For the studied problem, the prisms in D_0 are given in Figure 3(c). Each integration area (prism) comes from the intersection of two triangles, see Figure 3. For each prism, nodal and edge unknowns are considered. The nodal unknowns are the six nodes of the two triangles which derives the prism. In the vector potential formulation, vertical unknowns are added [5]. The edge unknowns are composed of the six edges of the two triangles plus the three vertical edges of the prism as shown in Figure 4.

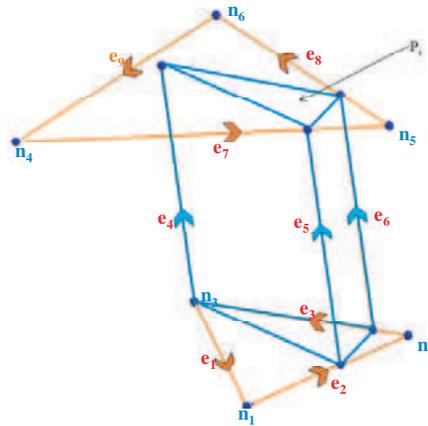


Figure 4: Unknowns associated with a prism.

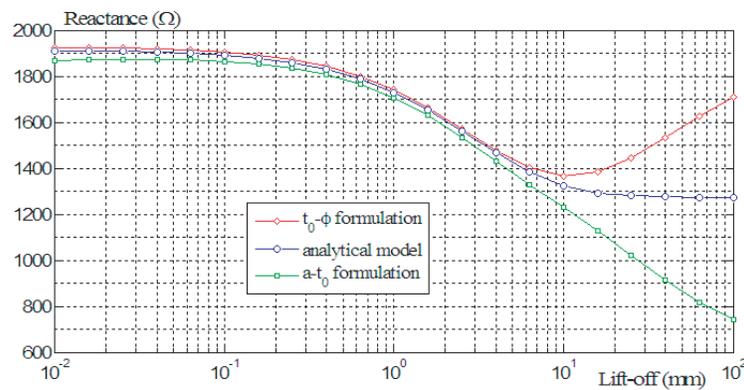


Figure 5: Reactance of the coil function of the lift-off value.

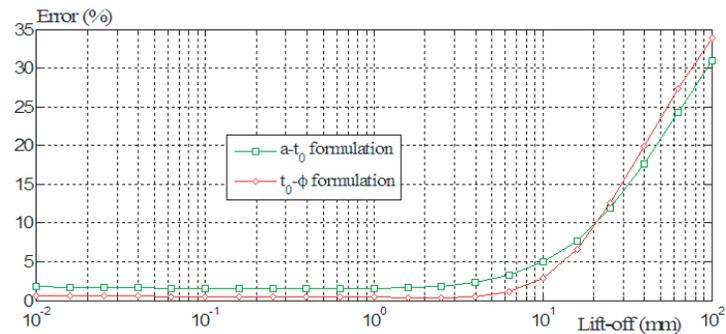


Figure 6: Error between numerical and analytical values of the reactance function of the lift-off value.

6. RESULTS

To compare the dual formulations with overlapping elements, the lift-off between D_1 and D_2 is modified in a range from 0.01 mm to 100 mm. A Dodd and Deeds analytical solution has been implemented and is used as reference solution. The reactance of the coil has been calculated with both formulations, see Figure 5.

The errors between analytical and numerical results are given in Figure 6.

Finite element results and analytical model are in good agreement until a lift-off of 10 mm. This value depends naturally on the size of the coil. Under this limit, the difference between the dual formulations remains almost constant. A mesh refinement around the coil could reduce this error.

7. CONCLUSION

In this paper, the overlapping element method was used in two dual formulations (nodal and edge elements) for a magnetostatic NDT problem. For a lift-off value above the radius of the coil,

the numerical results tend to diverge. Under this limit the finite element results are close from the analytical solution (less than 5% of difference). Furthermore, no oscillation is observed. The so-obtained range of validity of the overlapping elements is then largely sufficient for a practical utilization in modelling of magnetostatic NDT.

ACKNOWLEDGMENT

This work was supported by DIGITEO, Île-de-France Region and CEA LIST.

REFERENCES

1. Trowbridge, C. W. and J. K. Sykulski, "Some key developments in computational electromagnetics and their attribution," *IEEE Trans. on Magn.*, Vol. 42, No. 4, 503–508, April 2006.
2. Racheka, M. and M. Feliachi, "3-D movement simulation techniques using FE methods: Application to eddy current non-destructive testing," *NDT&E International*, 35–42, September 2006.
3. Antunes, O. J., J. P. A. Bastos, N. Sadowski, A. Razek, L. Santandrea, F. Bouillault, and F. Rapetti, "Comparison between nonconforming movements methods," *IEEE Trans. on Magn.*, Vol. 42, No. 4, 2006.
4. Tsukerman, I. A., "Overlapping finite element for problems with movement," *IEEE Trans. on Magn.*, Vol. 28, No. 5, 2247–2249, September 1992.
5. Krebs, G., T. Henneron, S. Clénet, and Y. Le Bihan, "Overlapping finite elements used to connect non-conforming meshes in 3D with a vector potential formulation," *IEEE Conference on Electromagnetic Field Computation (CEFC)*, Chicago, USA, May 2010.
6. Bensetti, M., Y. Choua, L. Santandréa, Y. Le Bihan, and C. Marchand, "Adaptive mesh refinement and probe signal calculation in eddy current NDT by complementary formulations," *IEEE Trans. on Magn.*, Vol. 44, No. 6, 1646–1649, June 2008.

Design and Study of a Permanent Magnet Synchronous Motor for an Electric Compressor

M. Khanchoul¹, G. Krebs¹, C. Marchand¹, F. Alves¹, A. Battelier², and M. Roze³

¹Laboratoire de Génie Electrique de Paris
CNRS UMR 8507, Supelec, Univ. Paris-Sud 11, UPMC Paris 6
11 rue Joliot Curie, 91192 Gif-sur-Yvette, France

²EDC Project, VALEO Compressor, La Verrière 78321, France

³Phenix International, Route de Noiron, Gray 70100, France

Abstract— Air conditioning compressors are important parts of electrical vehicles. These systems permit to cool down the batteries and assure the thermal comfort. In this paper a permanent magnet synchronous motor is proposed to drive such compressors. After analytical and numerical studies, the main features of the proposed structure are given.

1. INTRODUCTION

Nowadays, electrical vehicles (EV) present a technological solution to reduce gas emissions. However the success of EV is facing with some drawbacks, notably the sensitivity of the battery to climatic conditions. In hard conditions, the autonomy of the battery may be reduced within 100 kms instead of 150–200 kms; in addition the thermal comfort of EV is required. Consequently air conditioning compressors are embedded in EV. Mechanical air compressors are now substituted by electrical devices in order to improve efficiency, compactness and flexibility. An example of an electric compressor is given in Figure 1.

The electrical motor in such devices is subject to the following constraints. The necessary electrical power is about 6 kW and the maximal torque and speed are respectively equal to 6 Nm and 10000 rpm. The space required for the electrical motor is limited the reason why active length and outer diameter were reduced. The mass of the motor, must be lesser than 2 kg. Furthermore torque oscillations must not exceed 3% of the average torque to reduce the electromagnetic noise. The supply voltage depends on the battery level (200 to 410 V DC). Due to the restricted space, the cooling gas flows through the motor (stator and airgap).

In this paper, the design and the study of a permanent magnet synchronous motor able to satisfy the previous constraints is proposed. The first part of the paper deals with motor structures. The second part is dedicated to an analytical model (reluctance networks) used to pre-design the motor. In order to have more accurate results, the third part concerns a 2D finite element model. The features of the proposed motor are finally given.

2. PROPOSED ELECTRICAL MOTOR

Different types of electrical motor can be used in air compressor. Synchronous reluctance machine [1], wounded rotor induction motor [2] or synchronous permanent magnet motor with concentrated flux [3] can be found. The main drawback of almost configurations is the torque ripple [1] and some configurations present a low efficiency. Taking into account the previous constraints, a

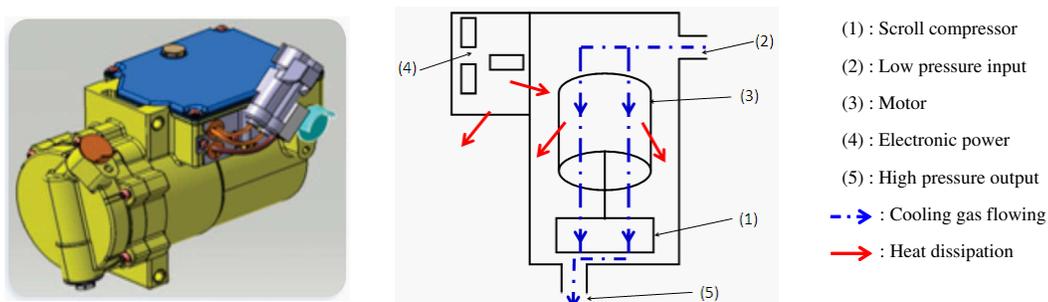


Figure 1: Example and structure of an electric air compressor.

surface-mounted permanent magnet motor with concentrated coils is proposed for driving the compressor, see Figure 2(a). Permanent magnet synchronous motors present attractive features due to the good mass/power ratio and due to lack of excitation windings (no magnetizing currents and lower Joule losses). In addition the proposed structure permits to have very low values for the no-load torque ripple and it shows the constant winding inductance [4]. Due to the different pole numbers of the stator and rotor, this kind of structure can present a higher mass/power ratio than conventional motors [5]. Concentrated coils permit to reduce the motor length, Joule losses [6] and can offer enough space for the gas flow. A number of eight poles in the rotor and six slots in the stator are chosen to limit oscillations and the maximum operating frequency.

In order to design the electromagnetic part of the motor, an analytical model is firstly used. The aim of this study is to check, with reasonable computation times, the choice of the structure in comparison with constraints previously mentioned.

3. ANALYTICAL MODEL

The analytical model is presented under the form of a 2D reluctance network [7]. Only half part of the motor is taken into account due to the magnetic symmetry, see Figure 2(b). In this model, the stator yoke and teeth reluctance are represented by \mathbf{R}_s and \mathbf{R}_t respectively. The flux leakage between the teeth is modelled by \mathbf{R}_L . The reluctance \mathbf{R}_a represents the airgap and magnet thickness and is calculated using Carter factor. The magnetomotive force due to the “ n ” coil is given by \mathbf{f}_n and the one due to the “ p ” magnet is given by \mathbf{f}_{mp} . The magnetomotive forces due to the magnets are function of the rotor displacement (sinusoidal waveform). All the reluctance values are constant. The following procedure has been applied to size the motor. Optimal tooth and magnet openings giving low ripples are used ($\alpha_t = 0.45 \times \lambda_t$ and $\alpha_m = 0.4 \times \lambda_m$). The thicknesses of the stator and rotor yokes have been chosen equal to 6 mm to avoid magnetic saturation. The diameter of the conductors was fixed to 1 mm and the filling factor equal to 0.4. The motor length is fixed (do not exceed 50 mm) and also the airgap (do not exceed 1 mm), thus only two parameters define entirely the geometry: r_a (radius in the middle of the airgap) and l_m (magnet thickness). The variations of these parameters give the maximal electromotive force value (emf) for one phase and torque ranges in Figure 3. The model was built around initial values ($r_a = 19.2$ mm and $l_m = 2.9$ mm).

Considering Figure 3, r_a and l_m equal respectively to 23.9 mm and 3.3 mm can fulfil the con-

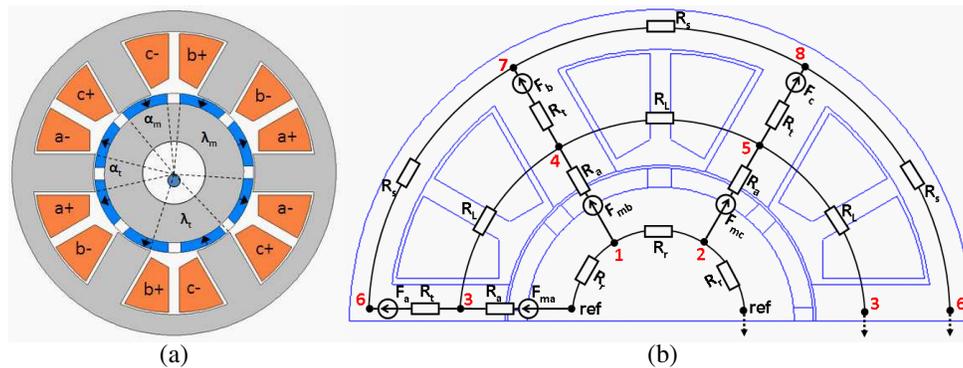


Figure 2: (a) Structure of motor. (b) Analytical.

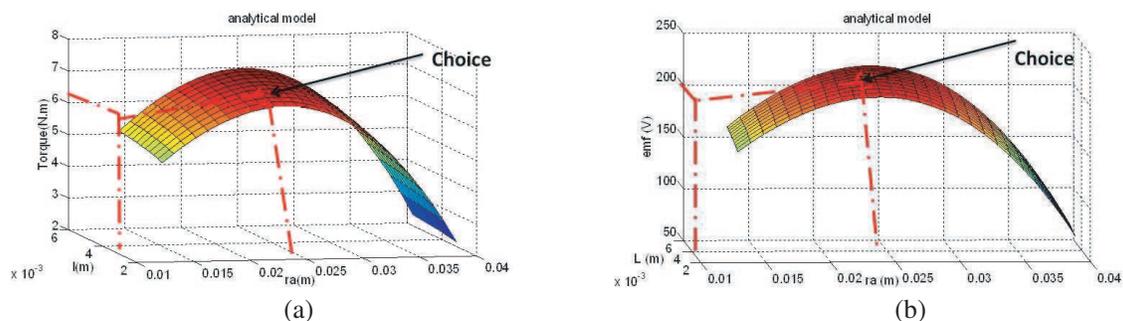


Figure 3: (a) Torque curve. (b) Electromotive force curve with the analytical model.

straints in term of emf and torque values. The number of coil turns is decreased from 108 to 45. To validate the analytical model and obtain more accurate results, a finite element study is proposed afterwards.

4. FINITE ELEMENT STUDY

The model is established with the software Femm [8]. The “weighted stress tensor” method is used to obtain the torque. The movement is simulated either via a complete remeshing at each rotation step or via a method equivalent to the locked step approach (based on the use of periodical conditions). This last permits to work with a constant number of elements. One example of mesh is given in Figure 4(a).

Stator and rotor yokes are composed of laminated steels (thickness of 0.35 mm, length of 41 mm). The steel grade is M330-35A (silicon iron, equivalent to former standard AISI M36). Rare earth magnets (NdFeB) are used with a coercivity of 955 kA/m. The wire diameter and the filling factor are equal to 1.4 mm and 0.4 respectively. The tooth and magnet openings respect the previous constraints on purpose to minimize the no-load torque ripples. The geometry is built again with only the variables r_a and l_m . The average torque and emf curves depending on the variations of theses parameters are represented on Figures 5(a) and (b).

With the same approach, the coil turns and total mass curves are given in Figures 6(a) and (b)

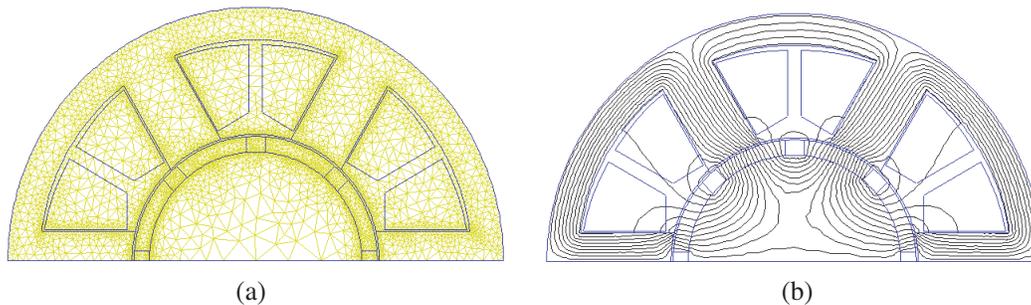


Figure 4: (a) Mesh of the motor and (b) field density at no-load.

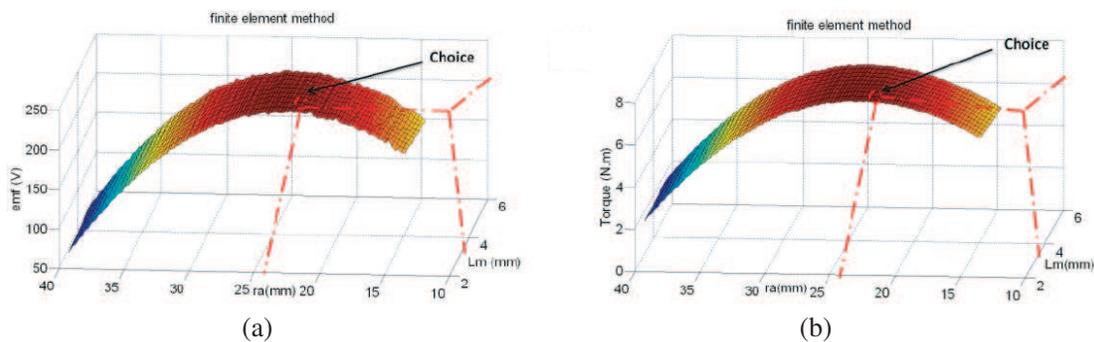


Figure 5: (a) Force curve electromotive. (b) Average torque curve with the finite element model.

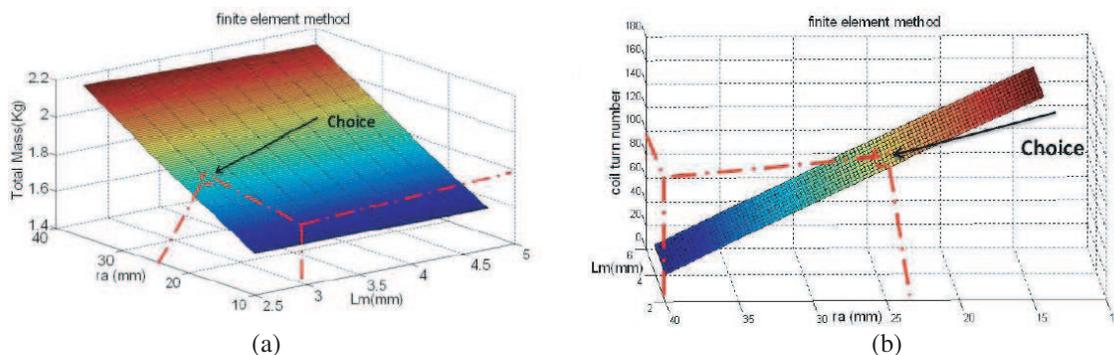


Figure 6: (a) Total mass with the finite element model. (b) Number of coil turns curve.

respectively magnetic linear behaviour has been used here to reduce the computation times.

Results between the analytical and finite element models are in good agreement, the difference is lesser than 10%. The following values have been preferred, $r_a = 24.9$ mm and $l_m = 3$ mm. The average torque is about 6.09 Nm and the maximal value for the emf is 230.36 V between two phases at 6000 rpm. The number of turns is equal to 45 per coil. Furthermore these values of r_a and l_m give a total mass under 2 kg. The modelled motor could be able to satisfy the air compressor constraints, more results are presented afterwards.

5. MOTOR FEATURES

The shapes of the magnets and rotor have been modified to simplify a manufacturing of these pieces, see Figure 7. For a good mechanical strength of the magnets at high speed, a hoop (thickness of 0.15 mm) is added around the rotor. The total airgap is equal to 0.65 mm. The bottom of the slots has a particular shape to place the coils. The model includes a non linear magnetic behaviour and the current density does not exceed 20 A/mm².

The flux linkage of one phase is given in Figure 8(a). A Fourier decomposition of the emf shows that the signal is almost sinusoidal (the harmonics 5, 7 and 9 are lesser than 1%). The maximum induced voltage between two phases is equal to 354 V. The no load torque is presented in Figure 8(b).

The ripples in the load torque, see figure 9(a), are lesser than 3% ($\Delta C = 0.16$ Nm). The motor

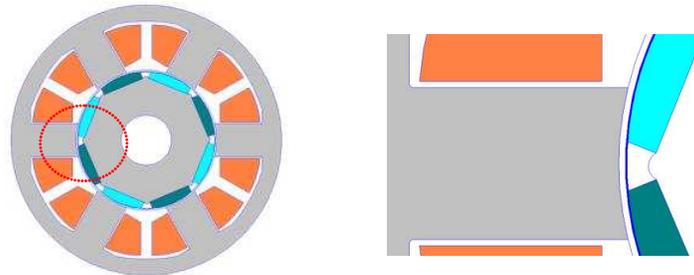


Figure 7: Surface permanent magnet motor for an air compressor.

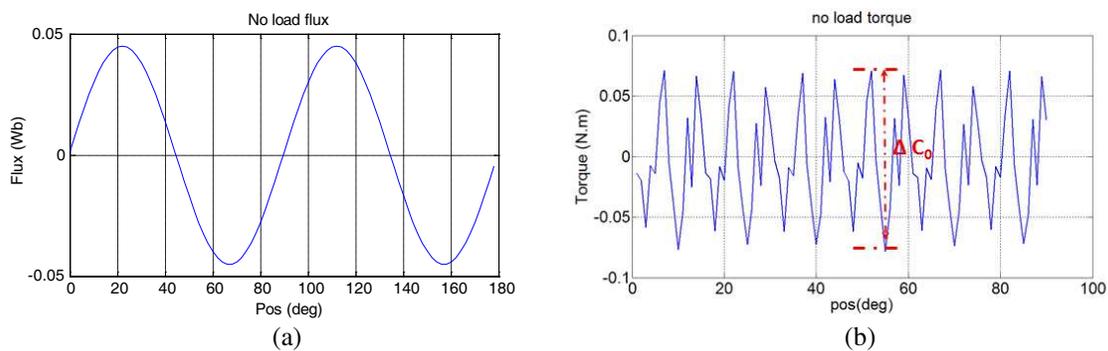


Figure 8: (a) Flux linkage at no-load. (b) No-load torque.

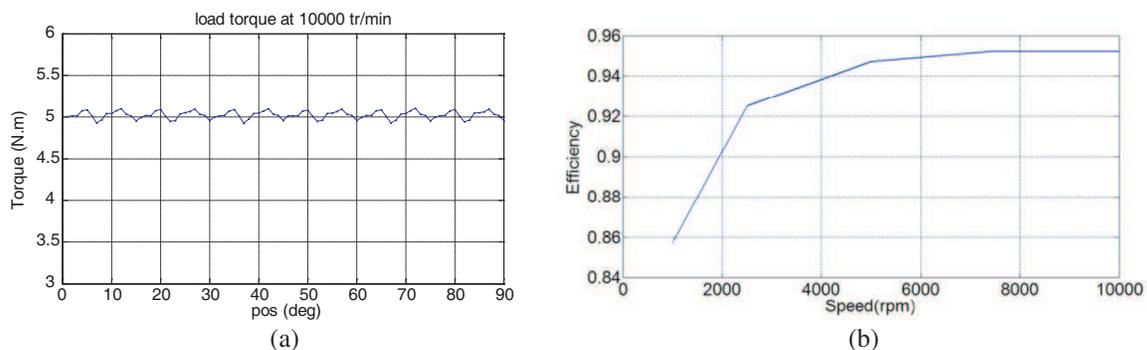


Figure 9: (a) Load torque at 10000 rpm. (b) Efficiency function of the speed.

efficiency (for a constant torque of 6Nm) has been estimated and is represented function of the speed in Figure 9(b). It includes the Joules and mechanical losses (rotor and air friction). The iron losses have been estimated using the finite element model and steel loss data's. The magnet losses are estimated coarsely (20% of the iron losses [9]).

6. CONCLUSION

In this paper, the design of a surface permanent magnet synchronous motor has been proposed. A first sizing has been performed using a simplified but reliable reluctance network. The final sizing of the motor has been achieved with a 2D finite element model. Taking into account the results, the air compressor constraints (in terms of electromagnetic behaviours) could be satisfied with the proposed structure. Further works will deal with multiphysic aspects [10]. At first, the cooling gas flowing through the stator, the heat exchange between the motor and the fluid has to be estimated. Then, a mechanical study will be performed on purpose to evaluate stresses on the rotor hoop and also on the magnets. The vibrations acting on the stator will be also studied.

ACKNOWLEDGMENT

These works are supported by the French program COMPACITE (compact and low consumption electric compressor for electric vehicle air conditioning). The partners are Valeo, Areelis, Leman Industries, Phenix, STMicroelectronics, Bree, Université Paul Sabatier, Lgep, Irseem.

REFERENCES

1. Raminosoa, T., B. Blunier, D. Fodorean, and A. Miraoui, "Design and comparison of high speed switched and synchronous reluctance machines to drive the compressor of an automotive PEM fuel cell," *ICEM Conference*, Vol. 57, 2988–2997, 2008.
2. Kawakami, M. and N. Fujii, "Analytical study of new type of direct drive motor for scroll type of compressor," *ICEM Conference*, 1–6, 2008.
3. Fukuda, Y., P. Patisonticharoen, Y. Kikushi, and H. Nakagawara, "Magnet embedded rotor, electric motor using the same rotor, and compressor using the same rotor," U.S. Patent, No. US2010/0166575 A1, 2010.
4. Zhu, T. Q. and D. Howe, "Electrical machines and drives for electric, hybrid, and fuel cell vehicles," *Proceeding of the IEEE*, Vol. 95, No. 4, 746–765, 2007.
5. Matt, D. and J. F. Llibre, "Performance comparées des machines à aimants et à réluctance variable. Maximisation du couple massique ou volumique," *Journal de Physique III*, 1621–1641, 1995.
6. Magnussen, F. and C. Sadarangani, "Winding factors and Joule losses of permanent magnet machines with concentrated winding," *IEMDC Conference*, Vol. 1, 333–339, 2003.
7. Krebs, G., A. Tounzi, B. Pauwels, D. Willemot, and F. Piriou, "Design and study of a linear actuator using an analytical approach and the 3D FEM," *Compel Journal*, Vol. 26, No. 4, 1005–1016, 2007.
8. [HTTP://WWW.FEMM.INFO/WIKI/HOMEPAGE](http://WWW.FEMM.INFO/WIKI/HOMEPAGE).
9. Benttayeb, A., "Pertes à haute fréquence dans les rotors des machines synchrones à aimants alimentées par onduleur de tension," Ph.D. thesis, Supelec, 2010 (in French).
10. Makni, Z., M. Besbes, and C. Marchand, "Multiphysics design methodology of permanent-magnet synchronous motors," *IEEE Transactions on Vehicular Technology*, Vol. 56, 1524–1530, 2007.

Full-wave Mode Analysis of Asymmetric Coupled Microstrip Structures: Particular Case of Quasi-symmetric Lines

A. Khodja¹, R. Touhami¹, M. C. E. Yagoub², and H. Baudrand³

¹Instrumentation Laboratory, Faculty of Electronics and Informatics
U.S.T.H.B University, Algiers, Algeria

²SITE, University of Ottawa, 800 King Edward, Ottawa, Ontario, K1N 6N5, Canada

³ENSEEIH, 2 rue Charles Camichel 31071, Toulouse Cedex 7, France

Abstract— The aim of this paper consists to highlight in which case asymmetric coupled microstrip structures can be studied by symmetric line model in order to use the quasi-symmetric approach. For this purpose, a rigorous formulation was developed; trial functions of C and π modes were obtained from the quasi-symmetric case, and thus proposed for the asymmetric case. A comparative study of dispersion parameters demonstrated the limits of the quasi-symmetric model validity with respect to asymmetries due to influence of shielding and metal strips widths difference.

1. INTRODUCTION

Asymmetric coupled structures which are used extensively in MMICs can provide additional advantages compared with symmetric ones because of easy matching to external element connections and improvement of characteristic impedance range. In order to reduce radiation effects and improve heat dissipation, the majority of transmission lines are embedded in metallic boxes. The propagation characteristics of asymmetric coupled lines have been computed by using rigorous full-wave techniques [1–3].

In order to efficient full-wave analysing the quasi-symmetric coupled line model, transverse resonance technique combined with the electromagnetism operators formalism is used. This technique consists in more easily formulating the boundary conditions by considering the fictitious propagation in the transverse direction. This paper presents adequate trial functions used in the asymmetric case, and shows how one can define quasi-symmetric lines. The obtained results permit to define the advantages and the limits of the quasi-symmetric model. In this paper the dielectric substrate used to analyze shielded asymmetric planar structures is assumed to be homogeneous, isotropic, and lossless. Its upper face is partially metalized by two parallel uniform zero-thickness and perfectly conducting strips as in Fig. 1.

2. METHODOLOGY

The aim is to model inhomogeneous asymmetric planar structures, in full-wave mode, using the Galerkin's procedure which leads to a homogeneous system of algebraic equations. Propagation constant " β " or effective permittivity " ε_{eff} " are obtained by solving the latter system. The technique assumes that the propagation is in the oy -transverse direction instead of the real oz -longitudinal direction, where oy represents the direction of the fictitious propagation constant. The transverse section of the structure is considered as a set of cascaded transmission lines terminated by short-circuited loads, noted " sc " (see Fig. 2).

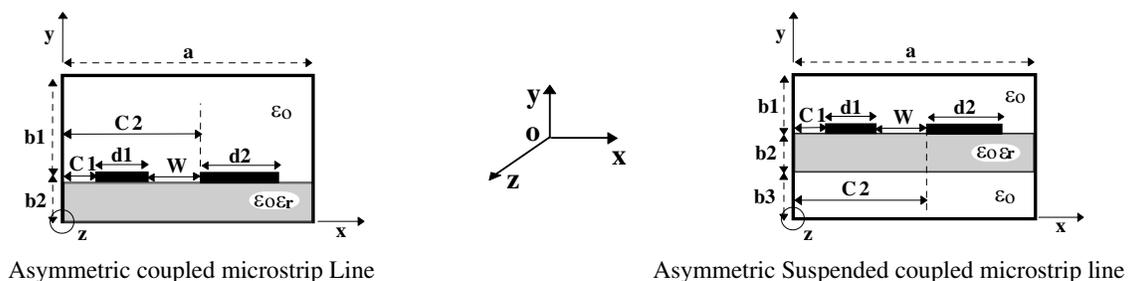


Figure 1: General structure of shielded asymmetric coupled microstrip transmission lines.

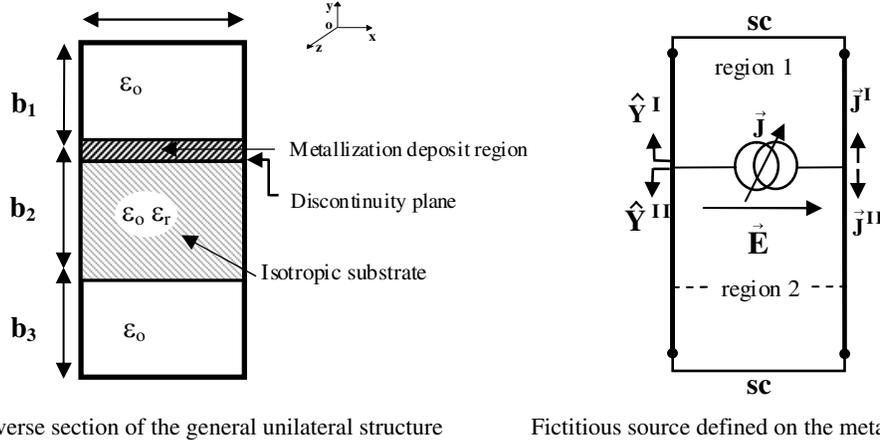


Figure 2: Equivalent network of the cross sectional view of the asymmetric structure with regard to the study domain.

The equivalent circuit of the transverse section of unilateral coupled line with negligible thickness includes the presence of an non-zero fictitious source represented by the transverse surface current density \vec{J} on the conductor strips. The admittance operators \hat{Y}^I and \hat{Y}^{II} , for respectively region 1 and 2, are expressed on the discontinuity plane.

The Kirchoff's law yields to:

$$\vec{J} = \vec{J}^I + \vec{J}^{II} \quad \text{with} \quad \vec{J}^I = \hat{Y}^I \vec{E} \quad \text{and} \quad \vec{J}^{II} = \hat{Y}^{II} \vec{E} \quad (1)$$

Finally:

$$\vec{J} = (\hat{Y}^I + \hat{Y}^{II}) \vec{E} = \hat{Y} \vec{E} \quad (2)$$

With:

$$\vec{E} = \hat{Z} \vec{J} = \vec{0} \quad \text{on the metal} \quad (3)$$

$$\vec{J} = \hat{Y} \vec{E} = \vec{0} \quad \text{on the insulator} \quad (4)$$

Knowing that:

$$\hat{Z} = \hat{Y}^{-1} \quad (5)$$

The matrix representation of the \hat{Z} or \hat{Y} operators should be then determined in order to obtain the dispersion matrix via the Galerkin's technique. This requires the involvement of a set of complete and orthogonal basis functions (noted $\{|f_n\rangle\}_{n=0,\infty}$), and also the calculation at the discontinuity plane of the mode admittances y_n^I and y_n^{II} of region 1 and 2 respectively. The expressions of admittance and impedance operators are given by:

$$\hat{Y} = \sum_{n=0}^{\infty} [|f_n\rangle (y_n^I + y_n^{II}) \langle f_n|]^{(e+h)} = \sum_{n=0}^{\infty} [|f_n\rangle y_n \langle f_n|]^{(e+h)} \quad (6)$$

$$\hat{Z} = \sum_{n=0}^{\infty} [|f_n\rangle \frac{1}{(y_n)} \langle f_n|]^{(e+h)} = \sum_{n=0}^{\infty} [|f_n\rangle z_n \langle f_n|]^{(e+h)} \quad (7)$$

where “e” and “h” states respectively for the TM_n and TE_n modes along the transverse direction (oy).

Here, y_n and z_n , expressed on the discontinuity plane of the considered planar structure, are the mode admittance and mode impedance, respectively.

The expressions of the mode admittances for the TM_n and TE_n modes are given by [4].

Here $|f_n\rangle \langle f_n|$ represents the projection operator on the basis vectors $\{|f_n\rangle\}_{n=0,\infty}$, where the product of vector “bra” ($\langle f_n|$) with vector “cKet” ($|g_n\rangle$) represents inner product of these two vectors.

“ e ” and “ h ” are respectively the TM_n and the TE_n mode along the transverse direction (oy).

The unknown current density is expressed on the base of trial functions taking into account the metallic edge effects. The application of the Galerkin’s procedure requires the projection of the terms of Equation (3) with trial functions, leading to a homogeneous system of algebraic equations which the resolution permit to evaluate the propagation characteristics.

3. QUASI-SYMMETRIC MODEL

The followed approach used in the modeling of the asymmetric lines is based in the general case of the evaluation of the current density on the two conductor strips of the coupled microstrip line in term of trial functions pondered by the (e_{px1}, e_{px2}) and (e_{qz1}, e_{qz2}) coefficients as follows:

$$\begin{cases} J_x = \sum_{p=1}^{kx} e_{px1} \Phi_{px1} + \sum_{p=1}^{kx} e_{px2} \Phi_{px2} \\ J_z = \sum_{q=1}^{kz} e_{qz1} \Phi_{qz1} + \sum_{q=1}^{kz} e_{qz2} \Phi_{qz2} \end{cases} \quad (8)$$

Knowing that (Φ_{px1}, Φ_{px2}) and (Φ_{qz1}, Φ_{qz2}) designate respectively trial functions represented on the 1st and 2nd conductor strip for the widths respectively d_1 and d_2 of the asymmetric coupled line.

k_x and k_z are the number of trial functions per component along the ox and oz directions respectively.

The quasi-symmetric case is determined by the restrictive conditions of the following coefficients:

$$\text{For } C \text{ mode: } e_{px2} = (-1)^{p+1} e_{px1}; \quad e_{qz2} = (-1)^q e_{qz1} \quad (9)$$

$$\text{For } \pi \text{ mode: } e_{px2} = (-1)^p e_{px1}; \quad e_{qz2} = (-1)^{q+1} e_{qz1} \quad (10)$$

These coefficients may be used in the following equations:

$$\text{For } C \text{ mode: } \begin{cases} J_x = \sum_{p=1}^{kx} e_{px1} (\Phi_{px1} + (-1)^{p+1} \Phi_{px2}) \\ J_z = \sum_{q=1}^{kz} e_{qz1} (\Phi_{qz1} + (-1)^q \Phi_{qz2}) \end{cases} \quad (11)$$

$$\text{For } \pi \text{ mode: } \begin{cases} J_x = \sum_{p=1}^{kx} e_{px1} (\Phi_{px1} + (-1)^p \Phi_{px2}) \\ J_z = \sum_{q=1}^{kz} e_{qz1} (\Phi_{qz1} + (-1)^{q+1} \Phi_{qz2}) \end{cases} \quad (12)$$

By taking $e_{px} = e_{px1}$ and $e_{qz} = e_{qz1}$, this permits to have trial functions identical to those used for the symmetric case [5]

$$J_x = \sum_{p=1}^{kx} e_{px} \Phi_{px} \quad \text{and} \quad J_z = \sum_{p=1}^{kx} e_{qz} \Phi_{qz} \quad (13)$$

With:

$$\text{For } C \text{ mode: } \begin{cases} \Phi_{px} = \Phi_{px1} + (-1)^{p+1} \Phi_{px2} \\ \Phi_{qz} = \Phi_{qz1} + (-1)^q \Phi_{qz2} \end{cases} \quad (14)$$

$$\text{For } \pi \text{ mode: } \begin{cases} \Phi_{px} = \Phi_{px1} + (-1)^p \Phi_{px2} \\ \Phi_{qz} = \Phi_{qz1} + (-1)^{q+1} \Phi_{qz2} \end{cases} \quad (15)$$

This approach permits, from modeling asymmetric structures, to analyze quasi-symmetric ones by taking the coefficients expressed on the second half of the considered asymmetric structure in terms of those expressed on the first half of the same structure by using latter Equations (9) and (10).

4. CHOICE OF TRIAL FUNCTIONS

An adequate choice of trial functions is necessary to ensure a reliable solution with a minimum of numerical treatments and processing time. Indeed, a suitable choice of trial functions allows a better configuration of the transverse current density on the metal of the transmission line.

This choice must respect several convergence criteria of this method like Boundary conditions, proportionality conditions and metallic edge effects conditions.

Cosine type trial functions which take into account the metallic edge effects are chosen for the general case deduced from asymmetric model such as:

$$\Phi_{mx}(x) = \begin{cases} \Phi_{mx1}(x) = \sin\left(\frac{m\pi(x-C1)}{d1}\right) & x \in [C1, C1+d1] \\ \Phi_{mx2}(x) = \sin\left(\frac{m\pi(x-C2)}{d2}\right) & x \in [C2, C2+d2] \end{cases}$$

$$\Phi_{mz}(x) = \begin{cases} \Phi_{mz1}(x) = \frac{\cos\left(\frac{(m-1)\pi(x-C1)}{d1}\right)}{\sqrt{\left(\frac{d1}{2}\right)^2 - (x-C1-\frac{d1}{2})^2}} & x \in [C1, C1+d1] \\ \Phi_{mz2}(x) = \frac{\cos\left(\frac{(m-1)\pi(x-C2)}{d2}\right)}{\sqrt{\left(\frac{d2}{2}\right)^2 - (x-C2-\frac{d2}{2})^2}} & x \in [C2, C2+d2] \end{cases}$$

5. NUMERICAL RESULTS

In order to study the dispersive behavior of asymmetric and quasi-symmetric structures with respect to the propagation parameters, we represent the evolution of the dispersion parameters as a function of physical parameters and frequencies. According to Fig. 3, we deduce that when metallic strips which have the same width, move away from the vertical metallic sides of the shielding, the asymmetric structure has the same propagation properties as quasi-symmetric one, moreover when the two strips do not have the same width ($d_2 \neq d_1$), the quasi-symmetric model of C mode is not valid, because the C mode is more affected by the sides of shielding or by conductor strips.

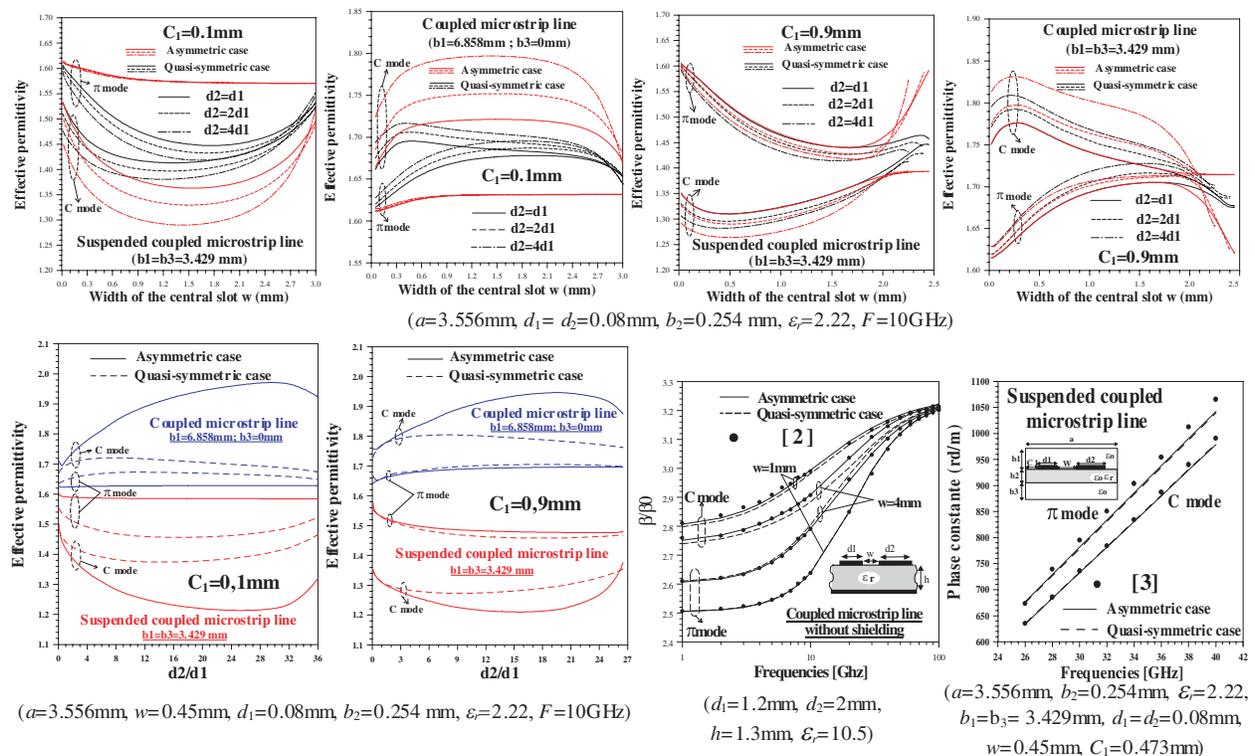


Figure 3: Variation of dispersion parameters versus physical parameters and frequencies.

Nevertheless, according to this figure, a good agreement of our results is shown with those published in the literature [2, 3]. However, when the coupled planar line is without shielding, the asymmetric model is more accurate than the quasi-symmetric case where the maximum average error made by this last case does not exceed 1%.

6. CONCLUSIONS

Through this article, the authors demonstrated that for the π mode and when the metallic strips are relatively far away from the vertical sides of the shielding, the asymmetric line can behave like a quasi-symmetric line. On the other hand, for the C mode and where the two strips do not have the same width ($d_2 \neq d_1$), the dispersive behavior of asymmetric and quasi-symmetric model diverges according to the shielding and slot dimensions.

REFERENCES

1. Srivastava, K. V., S. Awasthi, and A. Biswas, "Dispersion and attenuation characteristics of asymmetric multiconductor lines in suspended substrate structure using full-wave modal analysis," *Microwave and Optical Technology Letters*, Vol. 48, No. 7, 1305–1310, July 2006.
2. Qian, Y. and E. Yamashita, "Characterization of picosecond pulse crosstalk between coupled microstrip lines with arbitrary conductor width," *IEEE Trans. on Microwave Theory and Tech.*, Vol. 41, No. 6/7, 1011–1016, June/July 1993.
3. Achkar, J., O. Picon, V. F. Hanna, and J. Citerne, "Analysis of symmetric and asymmetric coupled suspended striplines and some associated discontinuities," *20th EMC*, Budapest, Hungary, 1990.
4. Khodja, A., "Optimisation des fonctions d'essai dans la modélisation de la ligne à ailettes unilatérale par la méthode de résonance transverse," Master Thesis, USTHB, Algiers, Algeria, April 2000.
5. Khodja, A., M. L. Tounsi, M. C. E. Yagoub, and R. Touhami, "Full-wave mode analysis of coupling effect in microwave planar transmission lines," *The Second IASTED International Conference on Antennas, Radar, and Wave Propagation (ARP2005)*, 49–54, Banff, Alberta, Canada, July 2005.

2D PIM Simulation Based on COMSOL

Xinbo Wang¹, Wanzhao Cui¹, Jingyu Wang^{2,3}, Jingnan Pan², Xiaocheng Zheng²,
Jiangtao Huangfu², and Lixin Ran²

¹National Key Lab of Science and Technology on Space Microwave Technology
Xi'an 710000, China

²Department of Information and Electronic Engineering, Zhejiang University
Hangzhou 310027, China

³Department of Photonics Engineering, Technical University of Denmark
Copenhagen, DK-2800, Denmark

Abstract— Passive intermodulation (PIM) is a problematic type of nonlinear distortion encountered in many communication systems. To analyze the PIM distortion resulting from material nonlinearity, a 2D PIM simulation method based on COMSOL is proposed in this paper. As an example, a rectangular waveguide filled with nonlinear dielectric was taken. Both of the effect of nonlinear coefficients and geometrical discontinuities were taken into consideration. The contribution of this paper is to make the PIM simulation more convenient, especially for those who are not good at analytical methods, and those who are not familiar with computational techniques. This type of analysis will aid in the development of PIM prediction, as well as PIM mitigation in the future.

1. INTRODUCTION

PIM is notoriously known for its detrimental effect on performance of base stations used in the space, military and civil telecommunications. PIM products, resulting from passive nonlinearities, occur in the reception band of antenna and degrade the system signal integrity. There are basically two types of passive nonlinearities, namely the contact nonlinearities and the material nonlinearities. In recent years, more and more efforts have been devoted to the study on the latter one, which refers the bulk materials, such as ferromagnetic materials and carbon fibers, that have nonlinear electrical characteristics [1–4]. Not only the prediction of PIM generation on printed transmission lines has been reported in [1], but also a PIM mitigation method by adding controlled thicknesses of nickel and gold plating to the conductors of a coaxial line, has been proposed in [4].

Although analytical methods have been widely used in these studies, they provided limited discussion for a specific shape of passive devices. If geometric effects are taken into account, the analytical result may not be obtained any longer. In that case, numerical simulation is more attractive. Both of Finite Element Method (FEM) and Finite-Difference Time-Domain (FDTD) are good candidates for such problems. In [5], codes based on FEM were developed to evaluate nonlinear line resistance and inductance of the microstrip lines. While in [6], a modified FDTD method involving current-mode nonlinearity is proposed, and its validity was confirmed through an experiment using the open-ended coaxial method.

However, both of the methods above require a good comprehension of computational electromagnetics, which is a little bit difficult for many engineers in passive device design. In this paper, the well-known commercial software COMSOL Multiphysics is employed for the PIM simulation. Those who are not familiar with computational techniques, but have a solid background in electromagnetics should find it extremely beneficial. The contribution of this paper is to make the PIM simulation more convenient. And this type of analysis should aid in the development of PIM prediction, as well as PIM mitigation.

2. PROBLEM FORMULATION

Maxwell equations are fundamental laws governing the behavior of a electromagnetic field in media. However, they do not contain information about the material properties. To characterize material properties, the so-called constitutive relations should be introduced as a necessary supplement. Constitutive relations in the most general form can be written as

$$\begin{bmatrix} c\bar{D} \\ \bar{H} \end{bmatrix} = \bar{C} \cdot \begin{bmatrix} \bar{E} \\ c\bar{B} \end{bmatrix} \quad (1)$$

where c is the velocity of light, and $\bar{\bar{C}}$ is a constitutive matrix.

According to the functional dependence of $\bar{\bar{C}}$, the media can be classified as nonlinear, if $\bar{\bar{C}}$ is a function of the electromagnetic field strengths [7]. This relation can be very complicated in real world. As an simple example, the relation in [8] is taken in this paper, which satisfies the equations below,

$$\bar{B} = \mu_0 \bar{H} \quad (2)$$

$$\bar{D} = \varepsilon_r \varepsilon_0 \bar{E} + \bar{D}_r \quad (3)$$

where $\bar{D}_r = \chi \varepsilon_0 (E_x^2, E_y^2, E_z^2)^T$, is the nonlinear component of electrical displacement, and χ is the nonlinear coefficients.

Maxwell equations, accomplished with constitutive relations, play a central role in the analysis of nonlinear electromagnetic problems. Under certain circumstances, it can be helpful to formulate the problems in terms of the vector potential \bar{A} and scalar potential Φ . They are given by the equalities

$$\bar{B} = \nabla \times \bar{A} \quad (4)$$

$$\bar{E} = -\nabla \Phi - \frac{\partial \bar{A}}{\partial t} \quad (5)$$

The defining equation for the magnetic vector potential is a direct consequence of the the magnetic Gauss's law, while the electric scalar potential results from Faraday's law.

The potentials would be somehow arbitrary without the Lorenz gauge condition as an additional restriction.

$$\nabla \cdot \bar{A} + \varepsilon_r \varepsilon_0 \mu_0 \frac{\partial \Phi}{\partial t} = 0 \quad (6)$$

Taking all the equations from Eq. (2) to Eq. (6) into account, a wave equation can be derived from Maxwell equations,

$$\varepsilon_r \varepsilon_0 \mu_0 \frac{\partial^2 \bar{A}}{\partial t^2} - \mu_0 \frac{\partial}{\partial t} \bar{D}_r + \nabla \times (\nabla \times \bar{A}) = 0 \quad (7)$$

The main idea of PIM simulaiton is to solve the partial differential equation in Eq. (7), accomplished with proper boundary conditions. In the next section, the well-known commercial FEM software COMSOL Multiphysics is employed to solve it.

3. SIMULATION AND RESULTS

COMSOL Multiphysics is an excellent, state-of-the-art software for the solution of many types of partial differential equations, both stationary and time-dependent, by numerical techniques based on the finite element method. In this section, a rectangular waveguide filled with nonlinear dielectric is analyzed by COMSOL.

Consider a metallic rectangular waveguide having dimensions b along the z axis, and $a > b$ along the y axis. The TE fields, where the electric field is along the z axis only, were investigated. When $a = 100$ mm and $b = 30$ mm, the cutoff frequency of TE_{10} and TE_{01} mode are 1.5 GHz and 5 GHz, respectively.

Although the waveguide is 3D in reality, as shown in Fig. 1(a). It is sufficient to solve an equivalent 2D problem in Fig. 1(b), because the fields have nothing changes along the z axis. Besides, modeling in 2D usually brings two advantages: (a) 2D models are easier to modify and solve; and (b) Given that the necessary approximations are small, the solution will be more accurate in 2D, because the mesh is much denser.

To introduce nonlinear dielectric inside the waveguide, the electromagnetic properties of the sub domain is configured so that it satisfies the nonlinear constitutive relation in Eq. (2) and Eq. (3), as shown in Fig. 2.

The next step is to determine the 4 boundaries in this model. Assuming that the electromagnetic wave propagates in the x direction, and B1 is the input port, while B4 is the output one. The other two boundaries are Perfect Electric Conductor (PEC) for the metallic waveguide. In order to make TE_{10} the only mode in the waveguide, the excitation signal on input port is known analytically to be

$$E_{z0} = \sin \frac{\pi(a-y)}{2a} (\cos 2\pi f_1 t + \cos 2\pi f_2 t) \quad (8)$$

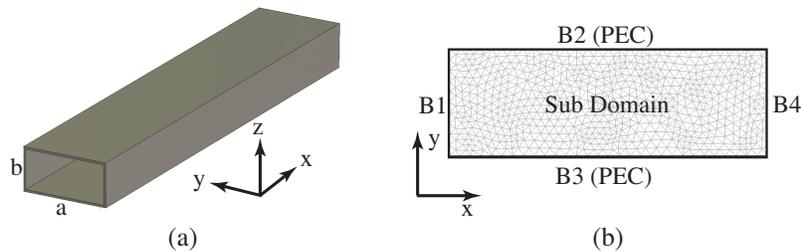


Figure 1: Waveguide model in the simulation: (a) waveguide in 3D; (b) equivalent 2D model.

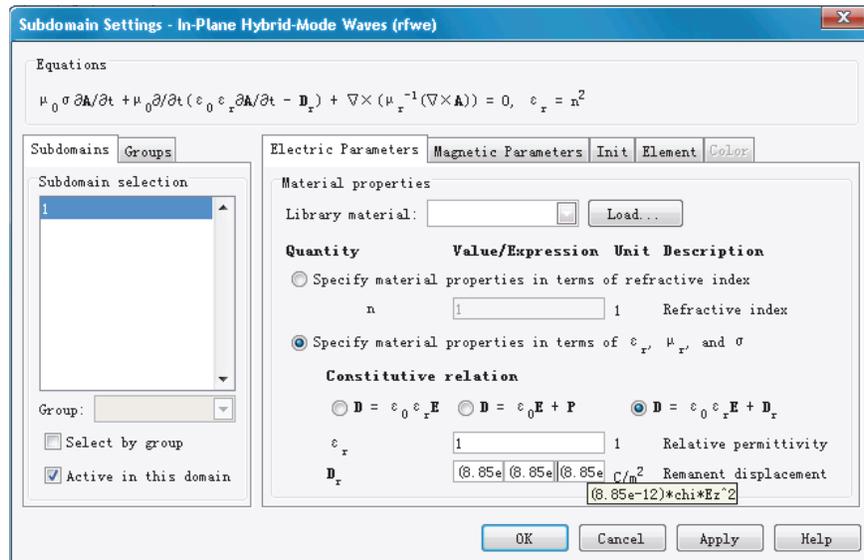


Figure 2: Electromagnetic properties configuration of sub domain.

where $f_1 = 2$ GHz and $f_2 = 3$ GHz, are the frequencies of the 2 input carriers, respectively.

A transient simulation in time domain has been used to evaluate the PIM distortion in the waveguide. In order to achieve enough frequency resolution, the simulation time has been configured to be as long as 2×10^{-8} s. And the time step was 2×10^{-11} s, so that it satisfied the requirements of Nyquist-Shannon sampling theorem. After the field simulation, COMSOL script has been used for post-processing. The basic idea is to monitor the electric field strength at the output first, and the time-domain signal is then converted into the frequency domain by performing a Fast Fourier Transform (FFT).

The spectrums of the output signals with different nonlinear coefficients χ are shown in Fig. 3. The first curve ($\chi = 0$), which means the media inside the waveguide is totally linear, was used for comparison. Because the nonlinearity inside the passive device is believed to be very weak, all the χ used in the simulation were less than 0.1. It is obvious that the 3rd order products PIM3 ($2f_2 - f_1 = 4$ GHz) appeared, if $\chi \neq 0$. And its amplitude increased, as χ got larger.

Further more, a waveguide with specific geometry has been investigated, as shown in Fig. 4(a). This type of waveguide changes the direction of the \vec{H} field components, while leaves the direction of the \vec{E} field unchanged. So it is called an H-bend. Unlike the normal waveguide, the bend will introduce unavoidable losses in signal power because of the unwanted reflections. Fortunately, these reflections can be minimized by keeping the bend smooth with a large enough radius. The H-bend was assumed to be full filled with nonlinear media of $\chi = 0.05$. And a normal waveguide with the same nonlinear media and the same length has also been simulated. Fig. 4(b) is the output spectrums of the these two kinds of waveguides. Given that all the other conditions were the same, the differences in spectrum is believed to result from the geometric effects. To the best of our knowledge, there are limited investigations on PIM analysis of specific geometry in literature at present. Development of the rigorous electromagnetic phenomenology of PIM generation in passive device with geometric effects remains an extremely challenging task. And the method proposed in this paper is a good candidate to solve such problems. Its validity will be confirmed through experiments in the future.

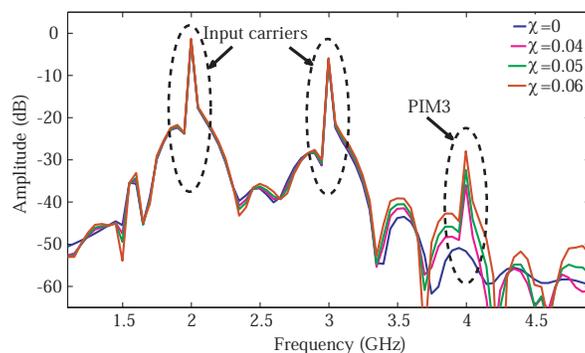


Figure 3: The spectrums of the output signals with different nonlinear coefficients.

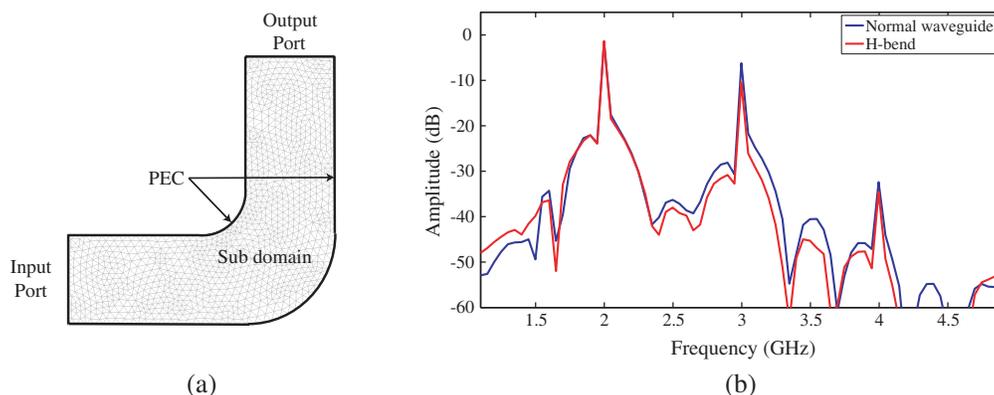


Figure 4: Simulation of waveguide with specific geometry: (a) 2D model of H-bend; (b) output spectrums.

4. CONCLUSION

A PIM simulation method based on commercial software is proposed in this paper. It is helpful in the analysis of PIM distortion resulting from material nonlinearities. Not only the effects of nonlinear coefficients, but also the geometrical discontinuities, have been investigated. This method should aid in the development of PIM prediction, as well as PIM mitigation. The further research will focus on the effective determination methods for nonlinear coefficients. In addition, a system level simulation is also on the schedule.

ACKNOWLEDGMENT

This work is sponsored by the National Key Laboratory Foundation (No. 9140C5304020901), the NSFC (No. 61071063, 60701007), 863 Project (No. 2009AA01Z227), the NCET-07-0750, and the ZJNSF (No. Y1080715).

REFERENCES

1. Zelenchuk, D., A. P. Shitvov, A. G. Schuchinsky, and T. Olsson, "Passive intermodulation on microstrip lines," *Proceedings of the 37th European Microwave Conference*, 396–399, Munich, Germany, October 2007.
2. Allilomes, P. C. and G. A. Kyriacou, "A nonlinear eigenvalue hybrid FEM formulation for two dimensional open waveguiding structures," *PIERS Online*, Vol. 1, No. 5, 620–624, 2005.
3. Yatsyk, V. V., "Effects of the resonant scattering of intensive fields by weakly nonlinear dielectric layer," *PIERS Online*, Vol. 2, No. 5, 524–527, 2006.
4. Henrie, J., A. Christianson, and W. J. Chappell, "Engineered passive nonlinearities for broadband passive intermodulation distortion mitigation," *IEEE Microwave and Wireless Components Letters*, Vol. 19, No. 10, 614–616, 2009.
5. Zhou, S. A. and T. Lewin, "Electromagnetic field simulation of nonlinear microstrip line by FEM," *Proceedings of International Conference on Microwave and Millimeter Wave Technology*, 577–580, Beijing, China, August 1998.

6. Ishibashi, D. and N. Kuga, “Analysis of 3rd-order passive intermodulation generated from metallic materials,” *Proceedings of Asia-Pacific Microwave Conference*, 1–4, Macau, China, December 2008.
7. Kong, J. A., *Electromagnetic Wave Theory*, EMW Publishing, Cambridge, Massachusetts, 2008.
8. Kalinchenko, G. and A. Lerer, “Propagation of electromagnetic waves in non-linear and non-uniform dielectric structures,” *2nd International Conference on Transparent Optical Networks*, 63–66, Gdansk, June 2000.

A Novel Approach for Modeling Diodes into FDTD Method

Hsin-Hsiang Su¹, Chih-Wen Kuo¹, and Toshihide Kitazawa²

¹Department of Electrical Engineering, National Sun Yat-San University, Taiwan

²Department of Electrical and Electronic Engineering, Ritsumeikan University, Kusatsu 525-8777, Japan

Abstract— A novel technique is proposed to incorporate diodes into the finite difference time domain (FDTD) method. The equivalent circuit of the diode is combined with Yee's time-stepping equations to update the fields without using the time-consuming iterative method at each time step. Two types of diodes are investigated by being connected at the end of a microstrip transmission line to demonstrate the validity of this method. Stability of the proposed method is demonstrated numerically and the accuracy is verified by comparison with commercial software ADS. The numerical results agree well with ADS. In these two cases, the numerical results show no sign of instability after a long simulation time (number of time steps = 10^5).

1. INTRODUCTION

The FDTD method can analyze MICs by incorporating the passive and active devices into the simulations [1–4]. In these previous works, for the active devices, the FDTD updating equations for the electric field need to solve a system of nonlinear equations at each time step. This nonlinear system can be solved by an iterative method such as the Newton-Raphson algorithm. However, the simulation time-step (Δt) is often necessarily reduced to ensure the convergence of the scheme. In [3], a time-step adaptive algorithm is proposed to ensure the convergence and increasing computational efficiency of the Newton-Raphson algorithm. In [7, 8], the voltage-current (V - I) relationship of the diode is approximated to a second-order Taylor expansion and incorporated into the FDTD method. This paper proposed a novel technique to efficiently incorporate diodes into the FDTD method by solving a quadratic equation at each time step to update the electric field. Two examples are presented to demonstrate the accuracy of this approach.

2. FORMULATION

According to the lumped element FDTD (LE-FDTD) method [1–3], lumped networks can be expressed as an electric current density term (\mathbf{J}_L) added into Maxwell's curl \mathbf{H} equation. Assuming the lumped network is connected along the z -direction in free space within a single cell-grid in FDTD, the time-dependent Maxwell's curl \mathbf{H} equation can be written as

$$(\nabla \times \mathbf{H})_z = \varepsilon_o \frac{\partial \mathbf{E}_z}{\partial t} + \mathbf{J}_L \quad (1)$$

where the lumped electric current density is $\mathbf{J}_L = \mathbf{I}_L / \Delta x \Delta y$. By multiplying $\Delta x \Delta y$ at both sides, the above equation can be rewritten as

$$I_z = C_z \frac{dV_z}{dt} + I_L \quad (2)$$

where $C_z = \varepsilon_o \Delta x \Delta y / \Delta z$ is the value of the equivalent capacitance of the FDTD grid, $I_z = \Delta x \Delta y (\nabla \times \mathbf{H})_z$ is the total current from the FDTD grid, $V_z = \mathbf{E}_z \Delta z$ is the voltage across the grid (lumped networks), and I_L is the lumped network current.

Suppose that the lumped network is a PN junction diode connected across the grid. The diode is modeled as a shunt combination of a current I_D and a nonlinear capacitance $C(V_D)$. The equivalent circuit of the FDTD grid with the diode is illustrated in Fig. 1, with the device equation as follow:

$$I_z = C_z \frac{dV_z}{dt} + I_D + C(V_D) \frac{dV_D}{dt}, \quad V_z = V_D \quad (3)$$

The current I_D and capacitance $C(V_D)$ are given by [5, 6]

$$I_D(V_d) = I_s (e^{bV_D} - 1), \quad b = q / \eta k T \quad (4)$$

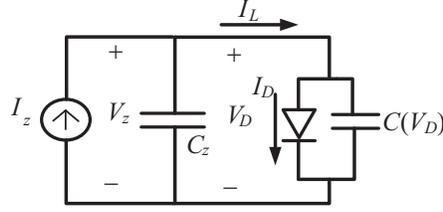


Figure 1: The equivalent circuit of a PN junction diode on the single FDTD grid.

$$C(V_D) = C_j(V_D) + C_d(V_D) = \tau_D b I_s e^{bV_D} + C_{vd},$$

$$C_{vd} = \begin{cases} C_{j0}(1 - V_D/V_j)^{-m} & V_D < (F_C \cdot V_j) \\ (C_{j0}/F_2)(F_3 + mV_D/V_j), & V_D > (F_C \cdot V_j) \end{cases} \quad (5)$$

where I_s is the saturation current, q is the electron charge, k is Boltzmann's constant, T is the temperature of the diode, η is the emission coefficient, τ_D is the sum of transit time for holes and electrons, C_{j0} is depletion capacitance at zero bias, m is the junction grading coefficient, V_j is the built-in junction voltage, F_C is the coefficient which determine when the junction is heavily forward biased, $F_2 = (1 - F_C)^{1+m}$, and $F_3 = 1 - F_C(1 + m)$. Thus the electric field of the grid can be found by solving the non-linear Equations (3)–(5).

The current I_D can be rewritten as

$$I_D = I_{D1} - I_s \quad (6)$$

where $I_{D1} = I_s e^{bV_D}$, and the time derivative of I_{D1} is

$$\frac{dI_{D1}}{dt} = bI_s e^{bV_D} \frac{dV_D}{dt} = bI_{D1} \frac{dV_D}{dt} \quad (7)$$

By inserting both (6) and (7) into (3) and after some manipulations, (3) becomes

$$\left(\frac{C_z + C_{vd}}{b} \right) \frac{dI_{D1}}{dt} + (I_{D1})^2 - I_{D1}(I_z + I_s) + \tau_D I_{D1} \frac{dI_{D1}}{dt} = 0 \quad (8)$$

The finite-difference expression of (8) is

$$\frac{C_z + C_{vd}^n}{b} \frac{I_{D1}^{n+1} - I_{D1}^n}{\Delta t} + \left(\frac{I_{D1}^{n+1} + I_{D1}^n}{2} \right)^2 - \left(\frac{I_{D1}^{n+1} + I_{D1}^n}{2} \right) (I_z + I_s) + \tau_D \frac{I_{D1}^{n+1} + I_{D1}^n}{2} \frac{I_{D1}^{n+1} - I_{D1}^n}{\Delta t} = 0 \quad (9)$$

It is noted that $I_{D1}^{n+1} = I_s e^{bV_D^{n+1}}$, $I_{D1}^n = I_s e^{bV_D^n}$ and I_{D1} is approximated by the arithmetic average of I_{D1}^{n+1} and I_{D1}^n . Furthermore, by setting $x^{n+1/2} = I_{D1}^{n+1} + I_{D1}^n$ and substituting into (9), a quadratic equation is obtained:

$$A(x^{n+1/2})^2 + Bx^{n+1/2} + C = 0 \quad (10)$$

where $A = \frac{\tau_D}{2\Delta t} + \frac{1}{4}$, $B = \frac{C_z + C_{vd}^n}{b\Delta t} - \frac{\tau_D}{\Delta t} I_{D1}^n - \frac{I_z + I_s}{2}$, and $C = -\frac{2(C_z + C_{vd}^n)}{b\Delta t} I_{D1}^n$ with the solution of the quadratic equation as

$$x^{n+1/2} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \quad (11)$$

According to the definition, the current $x^{n+1/2}$ is always non-negative,

$$x^{n+1/2} = I_{D1}^{n+1} + I_{D1}^n = I_s (e^{bV_D^{n+1}} + e^{bV_D^n}) \geq 0, \quad (12)$$

furthermore, the term $B^2 - 4AC$ is always larger than B^2 :

$$B^2 - 4AC = B^2 + 4 \left(\frac{\tau_D}{2\Delta t} + \frac{1}{4} \right) \left(\frac{2(C_z + C_{vd}^n)}{b\Delta t} I_{D1}^n \right) \geq B^2 \geq 0 \quad (13)$$

Thus the correct solution of this quadratic equation is always $x^{n+1/2} = (-B + \sqrt{B^2 - 4AC})/2A$. Finally, substituting $x^{n+1/2}$ into the discrete form of (3), the update equation of voltage V_z is expressed as

$$V_z^{n+1} = V_z^n + \frac{\Delta t}{C_Z + C_{vd}^n + \tau_D b x^{n+1/2}} \left((I_z^{n+1/2} + I_s) - \frac{x^{n+1/2}}{2} \right) \quad (14)$$

and the E_z can be obtained by $E_z = V_z/\Delta z$.

3. EXAMPLE

As shown in Fig. 2, a $50\ \Omega$ microstrip transmission line connected to a diode at the terminal is used to demonstrate the validity of the proposed algorithm. The width and length of the transmission line are 0.8 mm and 25 mm, respectively. For the PCB, the substrate thickness is 0.4 mm and $\epsilon_r=4.3$. The cell size in FDTD are $\Delta x = 0.1$ mm, $\Delta y = 0.4$ mm, $\Delta z = 0.2$ mm and the time step is $\Delta t = 0.99\text{-CFL}$. The excitation is a resistive sinusoidal voltage source with 20 V at 1.2 GHz and the source resistance is $R_s = 50$. A high-level voltage is deliberately selected to test the numerical stability of the proposed algorithm. There are two types of diode used in this structure. The first one is an ideal diode with $I_s = 0.01$ pA and $b = 1/25$ mV and the other is a high speed diode (1N4148) with the parameters listed in Table 1. This structure is also simulated by the commercial circuit simulation package ADS for comparison. The voltage across the diode is recorded and compared with ADS. Fig. 3 shows the voltages on the ideal diode and Fig. 4 shows the voltages on the high speed diode (1N4148). Both results agree very well with those from ADS. In these two cases, the numerical results show no sign of instability after a long simulation time (number of time steps = 10^5).

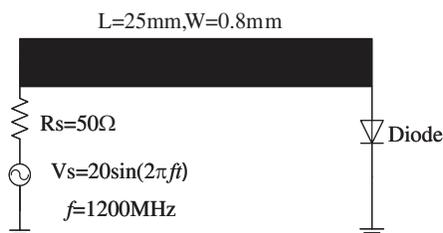


Figure 2: A diode is connected with the end of the transmission line.

Table 1: The parameters of diode 1N4148.

| | | | |
|------------------|-------------|---------------------|-------------|
| $I_s = 2.682$ nA | $M = 0.333$ | $C_{j0} = 4$ pF | $F_C = 0.5$ |
| $n = 1.836$ | $V_j = 0.5$ | $\tau_D = 11.54$ ns | |

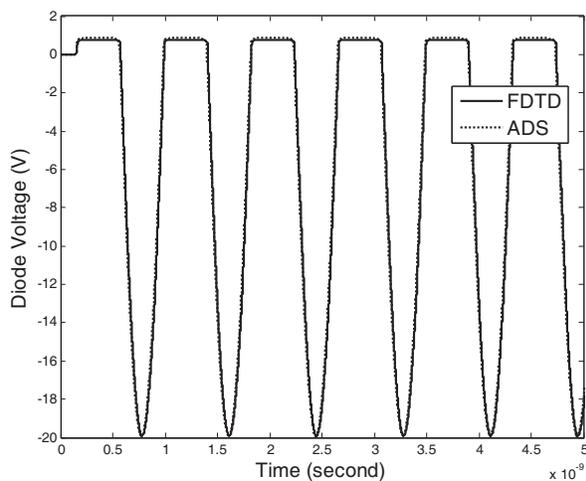


Figure 3: Comparison of voltage across ideal diode of the FDTD with ADS.

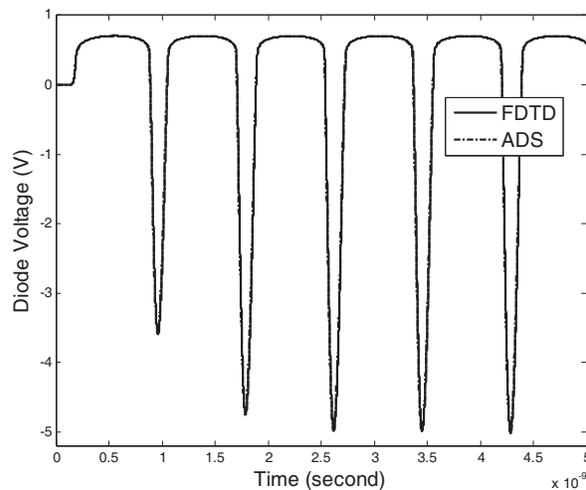


Figure 4: Comparison of voltage across diode (1N4148) of the FDTD with ADS.

4. CONCLUSION

This paper presents a novel accurate and efficient approach for incorporating nonlinear diodes into one FDTD grid. This scheme updates the electric field directly by solving a quadratic equation at each time step. The time-consuming iterative procedure of the Newton-Raphson scheme at each time step is thus avoided.

REFERENCES

1. Sui, W., D. A. Chistensen, and C. H. Durney, "Extending the two dimensional FDTD method to hybrid electromagnetic systems with active and passive lumped elements," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, 724–730, Apr. 1992.
2. Picket-May, M., A. Taflove, and J. Baron, "FDTD modeling of digital signal propagation in 3-D circuits with passive and active loads," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 1514–1532, Aug. 1994.
3. Ciampolini, P., P. Mezzanotte, L. Roselli, and R. Sorrentino, "Accurate and efficient circuit simulation with lumped-element FDTD technique," *IEEE Trans. Microwave Theory Tech.*, Vol. 44, 2207–2215, Dec. 1996.
4. Chen, M., S. T. Chew, and T. Itoh, "Nonlinear analysis of a microwave diode mixer using the extended FDTD," *IEEE MTT-S Int. Symp. Dig.*, 67–70, Denver, CO, June 1997.
5. Massobrio, G. and P. Antognetti, *Semiconductor Device Modeling with SPICE*, 2nd Edition, McGraw-Hill, 1993.
6. Millman, J. and C. C. Halkias, *Integrated Electronics*, McGraw-Hill, 1971.
7. Kung, F. and H.-T. Chuah, "Modeling of diode in FDTD simulation of printed circuit board," *Journal of Electromagnetic Waves and Applications*, Vol. 16, No. 1, 99–110, 2002.
8. Kung, F. and H.-T. Chuah, "A finite difference time domain (FDTD) software for simulation of printed circuit board (PCB) assembly," *Progress In Electromagnetics Research*, Vol. 50, 299–335, 2005.

Magnetic Susceptibility Modelling Using ANSYS

K. Bartusek¹, M. Cap², P. Marcon², and J. Mikulka²

¹Institute of Scientific Instruments, Academy of Sciences of the Czech Republic
Kralovopolska 147, Brno 612 64, Czech Republic

²Department of Theoretical and Experimental Electrical Engineering, Brno University of Technology
Kolejni 2906/4, Brno 612 00, Czech Republic

Abstract— The paper deals with modeling of magnetic induction in close to the paramagnetic materials. Nuclear magnetic resonance (NMR) based method for material susceptibility measuring gives variable results which need to be verified by other reliable method. ANSYS model based on Finite element method (FEM) makes verification of measured results possible. ANSYS is the suitable system for multiphysics modeling and simulation, modeling of anisotropic materials and for possible use of special programming language APDL. Therefore, the use of this program is for the results verification appropriate. Paramagnetic and diamagnetic material samples e.g., aluminum, copper, glass, and water are used. The samples are of cylinder shape, 3 cm in length and 0.5 cm in diameter. The cylinder is placed in an air filled cube with side length 4 cm.

1. INTRODUCTION

Magnetic susceptibility is the physical quantity describing material properties in the external magnetic field [1]. Magnetic susceptibility is defined as ratio between magnetization \mathbf{M} of the material in the magnetic field and the field intensity \mathbf{H} :

$$\mathbf{M} = \chi_m \mathbf{H}. \quad (1)$$

All materials can be classified by value of magnetic susceptibility into three groups:

- diamagnetic materials: $-1 < \chi_m < 0$,
- paramagnetic materials: $0 < \chi_m \ll 1$,
- ferromagnetic materials: $\chi_m \gg 1$.

Several methods are used for magnetic susceptibility measuring such as Faraday's scale, Guoy's scale or inductive method with SQUID magnetometer. For detailed description of these methods see [2]. The MR based method for susceptibility measurement makes also this measurement possible [3–5]. Whereas, this method gives us inaccurate results, other method for results verification is needed. Possible option is use the results of the ANSYS software simulations of the magnetic field close to the different materials for the magnetic susceptibility calculation.

Placing the paramagnetic and diamagnetic materials in homogenous magnetic field cause a distortions, which are corresponding to the magnetic susceptibility χ of the measured material. In this article modeling of this effect is discussed.

2. PROPERTIES AND ASSEMBLY OF THE EXPERIMENT

Space configuration of the sample is shown in Fig. 1. Cubic space ($40 \times 40 \times 40$ mm) filled by an air has magnetic susceptibility $\chi_{m1} = 0$, labeled by number 1 in Fig. 1. In the center of this space is placed the sample of mentioned materials with different magnetic susceptibility χ_m . The samples are cylindrical, 3 cm in length and 0.5 cm in diameter, labeled by number 2 in Fig. 1. There were used the paramagnetic aluminum sample ($\chi_{mAl} = 2.2 \cdot 10^{-5}$), the diamagnetic samples copper ($\chi_{mCu} = -9.2 \cdot 10^{-6}$) and water ($\chi_{mAl} = 2.2 \cdot 10^{-5}$). Section plane is placed to the center of the cubic space, labeled by number 3 in Fig. 1.

The model according to arrangement shown in Fig. 1 was designed in ANSYS environment. The model was designed by using of the finite element method with using of Solid 96 elements [6]. The boundary conditions were chosen according to the formula for the magnetic potential in order to value of a homogenous static magnetic field induction was $B = 4.7$ T.

$$\pm \frac{\phi}{2} = \frac{B \cdot z}{2\mu_0} = \frac{4.7 \text{ T} \cdot 40 \text{ mm}}{2\mu_0}. \quad (2)$$

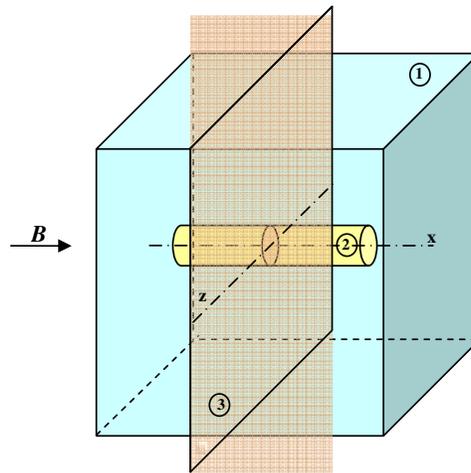


Figure 1: Configuration of the modeled system, cube filled by an air with sample.

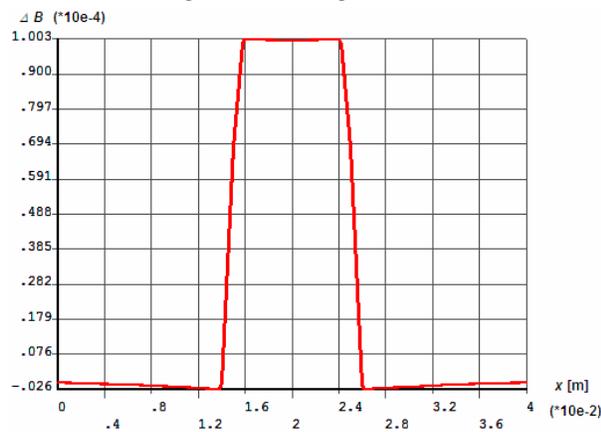


Figure 2: The behaviour of magnetic induction in the ax of the aluminum sample.

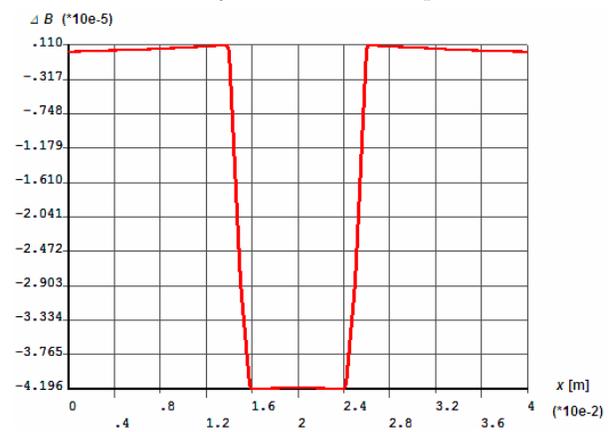


Figure 3: The behaviour of magnetic induction in the ax of the copper sample.

3. NUMERICAL MODELING

ANSYS is the suitable system for multiphysics modeling and simulation, anisotropic materials and for possible use of special programming language APDL. ANSYS numerical analysis is based on Finite element method. For describing of the quasi-stable model by the Finite element method are used the reduced Maxwell's equations:

$$\text{rot}\mathbf{H} = \mathbf{J}, \quad (3)$$

$$\text{div}\mathbf{B} = 0, \quad (4)$$

where \mathbf{H} is the magnetic field intensity vector, \mathbf{B} is the magnetic field induction vector, \mathbf{J} is the current density vector. For case of the irrotational field is the Equation (3) in the form:

$$\text{rot}\mathbf{H} = 0. \quad (5)$$

Material relations are given by equation:

$$\mathbf{B} = \mu_0\mu_r\mathbf{H}. \quad (6)$$

4. RESULTS OF NUMERICAL MODELING

The modeled magnetic induction in the section plane is shown in Fig. 2. Value ΔB is deviation from the static magnetic field $B_0 = 4.7\text{ T}$. As a sample was used the aluminum sample and because of the paramagnetic properties is the magnetic field in the sample space magnified.

Computed ΔB curve is used in the magnetic susceptibility equation:

$$\chi_m = \frac{Bs - Bm}{B_0}. \quad (7)$$

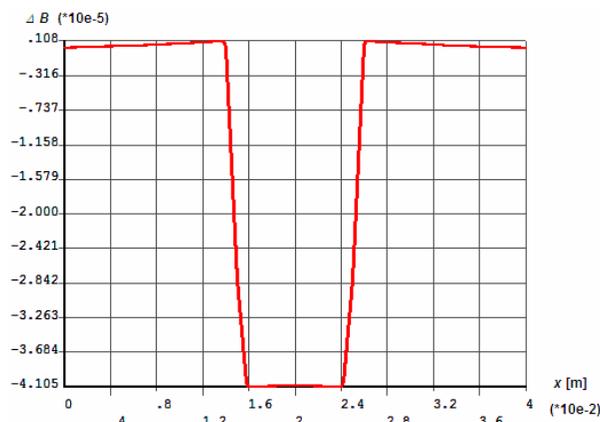


Figure 4: The behaviour of magnetic induction in the ax of the water sample.

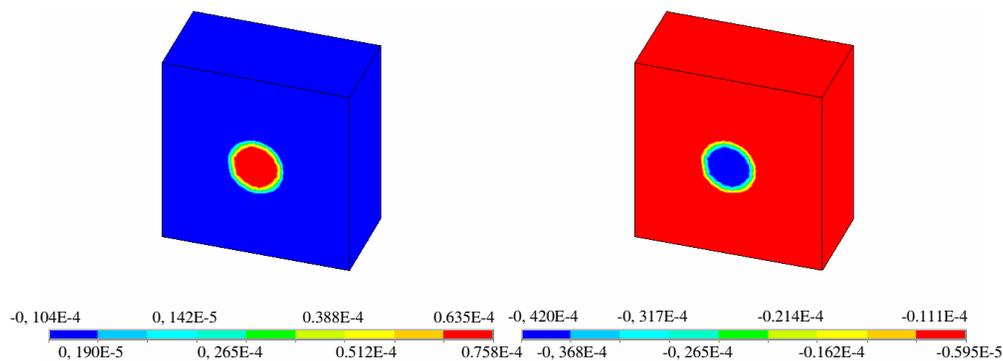


Figure 5: Slice through the center of the modeled system.

Table 1: Errors between the actual value of the magnetic susceptibility χ and the calculated value of $\Delta\chi$.

| Samples | χ [ppm] | $\Delta\chi$ [ppm] |
|----------|--------------|--------------------|
| Copper | -9.2 | 0.04 |
| Water | -9.0 | 0.04 |
| Aluminum | 2.2 | -0.10 |

Value B_s and B_m represent the stable value of B in space of the sample and the minimal B value just next to the material sample. By using ΔB values from Fig. 2 and Equation (7) the magnetic susceptibility of aluminum has been calculated to: $\chi_{cAl} = 2.19 \cdot 10^{-5}$. When we substitute the value of magnetic flux from Fig. 3 to Equation (7), we obtain the calculated value of susceptibility of copper: $\chi_{cCu} = 9.16 \cdot 10^{-6}$. With values of the magnetic flux from Fig. 4 we obtain the calculated susceptibility for water: $\chi_{cH_2O} = 8.96 \cdot 10^{-6}$.

Results on section planes in the ANSYS models are shown in Fig. 5. On the left the red color in the space of the aluminum sample shows the local increment of the magnetic field and the blue color the static magnetic field B_0 . On the right is the local decrement (blue area) of the magnetic field in the place of the diamagnetic material sample (copper, water). The red area corresponds to static magnetic field.

5. CONCLUSIONS

The ANSYS environment was used for modeling of interactions of the magnetic field with paramagnetic and diamagnetic samples. Proposed Equation (7) gives values of the magnetic susceptibility with sufficient accuracy. Higher level of the accuracy is possible with increasing of the element size. In Table 1 is shown an overview of the actual values of magnetic susceptibility χ of measured materials and absolute errors $\Delta\chi$ calculated from the model of individual samples.

It is possible to verify the experimental results of the magnetic susceptibility measuring for different material samples by modeling of the magnetic. This is important for example in MR methods.

ACKNOWLEDGMENT

The research described in the paper was financially supported by research plan No. MSM 0021630513, No. MSM 0021630516, grant of Czech ministry of industry and trade No. FR-TI1/001, Czech Science Foundation (102/09/0314) and project of the BUT Grant Agency FEKT-S-10-13.

REFERENCES

1. Dedek, L. and J. Dedkova, *Elektromagnetismus*, Vol. 2, 232, Vutium, Brno, 2000, ISBN 80-214-1548-7.
2. Steinbauer, M., *Mereni Magneticke Susceptibility Technikami Tomografie Magneticke Resonance*, VUT v Brne, FEKT, Brno, 2006.
3. Vladingerbroek, M. T. and J. A. Den Boer, *Magnetic Resonance Imaging*, Springer-Verlag, Heidelberg, Germany, 1999, ISBN 3-540-64877-1.
4. Bartusek, K., Z. Dokoupil, and E. Gescheidtova, “Mapping of magnetic field around small coils using the magnetic resonance method,” *Measurement Science and Technology*, Vol. 18, 2223–2230, 2007.
5. Bartusek, K., Z. Dokoupil, and E. Gescheidtova, “Magnetic field mapping around metal implants using an asymmetric spin-echo MRI sequence,” *Measurement Science and Technology*, Vol. 17, No. 12, 3293, 2006.
6. Fiala, P., “Finite element method analysis of a magnetic field inside a microwave pulsed generator,” *2nd European Symposium on Non-lethal Weapons*, Ettlingen, SRN, May 13–15, 2003.

Simulation of Defects in Photonic Band Gap Structures

Laurent Oyhenart and Valérie Vignéras

Laboratoire IMS, UMR CNRS 5218, 16 Avenue Pey-Berland, Pessac 33607, France

Abstract— The modelling of defects in photonic crystals (PC) by traditional numerical methods is difficult because it requires much memory and computing time. An analytical method of multiple-scattering solves this typical problem for spherical inclusions. Only this method does the study of lattices made up with a great number of spheres. This method finds the minimal number of spheres to obtain equivalence between finite and infinite photonic crystals. From ordered to disordered systems, we have used this method to study the effects of random variations of the radii and positions of the spheres in the lattice.

1. INTRODUCTION

The scattering and diffraction of electromagnetic waves are closely-dependent phenomena. A different approach is required, according to the scales of the wavelengths studied. When dimensions of the target are large compared to the wavelength, a geometrical or semi-geometrical approach is adequate to describe the problem. If this is not the case, computational methods are required, but may be difficult to implement, depending on the number of particles to be analyzed.

Many analytical and numerical methods provide easy solutions for studying the scattering of a single particle by an electromagnetic wave. For canonical particles, such as spheres, cylinders, ellipsoids etc., the Helmholtz wave equation is separable. The problem is simplified, requiring only a multipole expansion of the electromagnetic fields on an appropriate basis. Particles with more complex geometry may be analyzed either by special methods, such as discrete dipole approximation [1] or extended boundary condition [2], or general methods, such as finite difference [3] or finite element [4]. The computing time of these two categories of methods are very different, in favour of the special methods.

General methods are useful for analyzing several particles, provided that the wavelength is large compared to particle size. The problem becomes more and more difficult to solve as the wavelength decreases. The situation is simpler if the particles are positioned in the shape of an infinite periodic lattice from the side, as only one period is calculated. In the case of finite periodic structures, it is a challenge to find general methods that do not use too much memory and computation time [5].

A code of multiple-scattering has been developed at the laboratory to avoid these limitations [6]. Our code calculates the scattered fields in heterogeneous structures made up of inclusions scattered in a matrix. In the case studied here, inclusions have a spherical shape. A recall of the broad outline of the theory of multiple-scattering and a study of the simple cubic photonic crystal [7] will enable us to understand the influence of the defects in these photonic crystals (PC).

2. MULTIPLE-SCATTERING METHOD

The multiple-scattering method (MST) is separated in two parts. The first part allows the calculation of the scattered wave by one sphere by using the Mie scattering theory. The incident and scattered fields are developed in a spherical harmonics basis. Those are given thanks to the geometry of the object and its optical properties. We write the transfer matrix between the incident field and the scattered field.

The second part of the method is an iterative algorithm which turns the scattered field calculation from one sphere to N spheres. For the first order, we calculate the diffraction of the incident field for each sphere. For the second order, the scattered field of first order for one sphere becomes the incident field for the $N - 1$ other spheres. With this new incident field, the scattered field is calculated as at first order and so on for higher orders. This iterative process stops when a criteria of residual is satisfied. The total scattered field is the contribution of all spheres and all orders used.

The nature and the size of these N spheres can be different. Moreover, spheres may be put in a random way, without conditions of symmetries in the spheres positions. These two last remarks show all the interest of this method in calculation of PC with defects and random structures. The method was developed for cylindrical and spherical objects, but can be extended to other shapes.

3. PRELIMINARY STUDY OF PHOTONIC CRYSTAL WITHOUT DEFECT

Initially, we study a photonic crystal without defect. We chose a simple cubic lattice of dielectric spheres of permittivity $\epsilon_r = 16$. They are placed in a cubic lattice of 10 layers. This number of layers will be kept constant on this study. The spheres radius is standardized to the lattice constant a . In our example, the spheres radius is equal to $r = 0.2a$. This standardization is also used for the frequency.

Figure 1(a) shows the standing wave created by the PC if the standardized frequency is set to $af/c = 0.46$.

The wave is sufficiently plane to be able to calculate the transmission of the structure. Transmission is calculated from the flow of energy crossing a surface in front of the PC. Figure 1(b) represents the transmission curves of the dielectric PC. The transmission of the finite PC ($5 \times 5 = 25$ spheres by layer and $7 \times 7 = 49$ spheres by layer) is modelled by the MST and the infinite PC is modelled by the FDTD method. If we increase the number of spheres per layer i.e., the surface of the PC, the finite PC will be equivalent to the infinite PC.

4. ARBITRARY VARIATIONS OF LAYERS POSITIONS

In our code of multiple-scattering, no condition of symmetry is required. We can choose the diameter and the positions of all spheres. This characteristic is very important for the defects study. The defects in the PCs can be of two types: random variations on the whole of the lattice PC and the arbitrary variations localized on a site of the PC. The PC selected is the previous dielectric one. We present here a sample of the possibilities of variations of layers. This PC can be cut out in layers according to three directions of space. We moved the first five layers of the PC. Shift is expressed as a percentage of the lattice constant. Same shift is applied to the five layers; they are linked. We are interested here only in the normal incidence. We will study two different cases: variations according to the axis of polarization and variations according to the axis of propagation.

On the Figure 2, the band gaps go up gradually when the lattice is shifted. There is almost a linear law between the shift of the lattice and the increase. In fact, it is not astonishing, if the geometry of the problem is looked at.

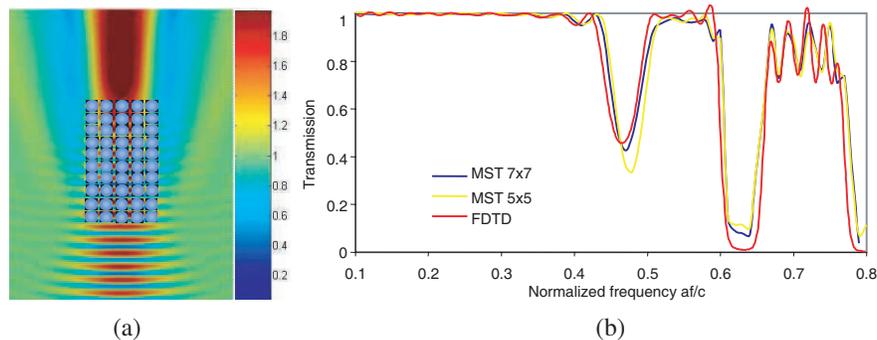


Figure 1: Calculation for an FSS composed of dielectric spheres. (a) Visualization of the amplitude of the total electric field reflected and transmitted by the FSS. (b) Influence of PC size on transmission plot.

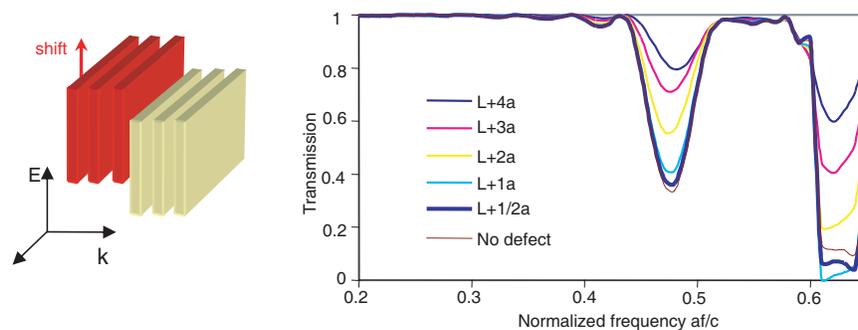


Figure 2: Shift of layers perpendicularly to the propagation axis: effect on the transmittance.

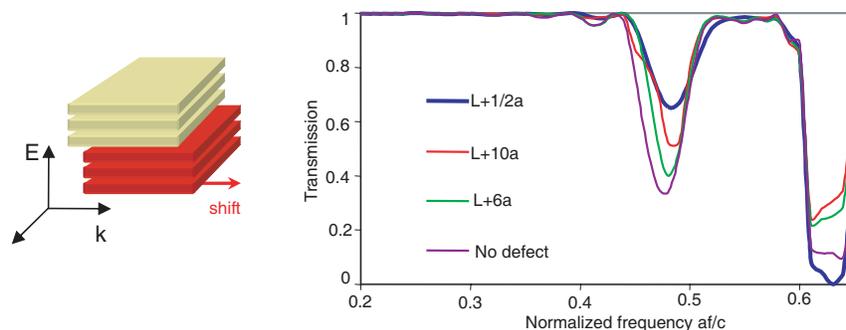


Figure 3: Shift of the layers in parallel to the propagation axis: effect on the transmittance.

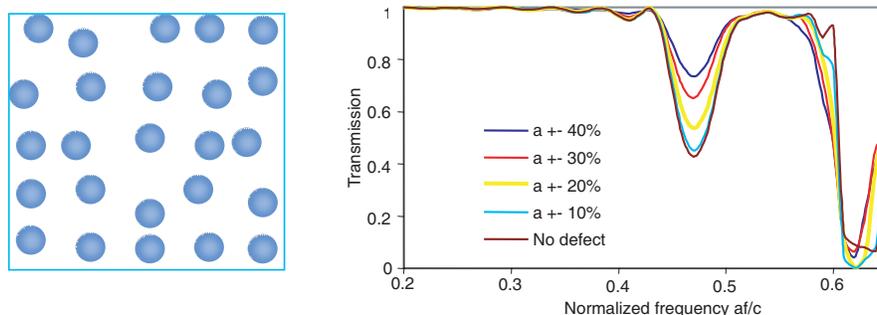


Figure 4: Effect of the random variation of spheres positions on the transmittance as a function of the frequency.

On the Figure 3, the band gap varies slightly for shifts equal to an integer value of lattice constant. In fact, the core of the structure does not change. When the shift is not proportional to an integer value of the lattice constant, the first band gap goes up very strongly. the PC is disorganized much more than the PC with the previous shifts.

For shifts of half lattice constant, the PC is much more sensitive in the case of Figure 3 than of the Figure 2. It's not surprising that the PC is more sensitive to shifts according to the axis of propagation than the axis of polarization. We find what we had already noted in experiments. The PC is very sensitive to the variations thicknesses of the layers according to the axis of propagation. That is due to the origin of the band gaps.

5. RANDOM VARIATIONS OF RADII AND SPHERES POSITIONS

We studied the influence of random variations of radii and spheres positions. We keep the same structure dielectric PC as previously. On Figure 4, we operate a random variation on the spheres positions around the initial value.

We keep constant the average and the standard deviation of spheres positions. The first forbidden band goes up gradually by increasing the percentage of random variation of the lattice constant. This band gap has a traditional behaviour. The lattice is disorganized as the random variation increases.

On the other hand the second band almost does not vary. The first two bands have a radically different behaviour. By modifying the lattice gradually, the first forbidden band will become complete whereas it is not the case for the second. The second band does not result from constructive or destructive interferences between all the layers. If only one layer is kept, this band remains. This band has behaviour of frequency selective surface.

In the second time, we make a random variation of spheres radii in Figure 5. Here also, the average and the standard deviation of spheres radii are constant. A random variation of spheres radii has a weak influence until $\pm 10\%$. Beyond this percentage, the two bands disappear rather quickly. There is almost a percentage threshold where the lattice does not behave any more out of photonic crystals.

The cubic structure is more sensitive to the variations of spheres sizes than the variations of positions. For small random variations preserving the average parameters of the initial lattice, the physical characteristics of the PC remain the same ones. The simple cubic structure is robust to position defects.

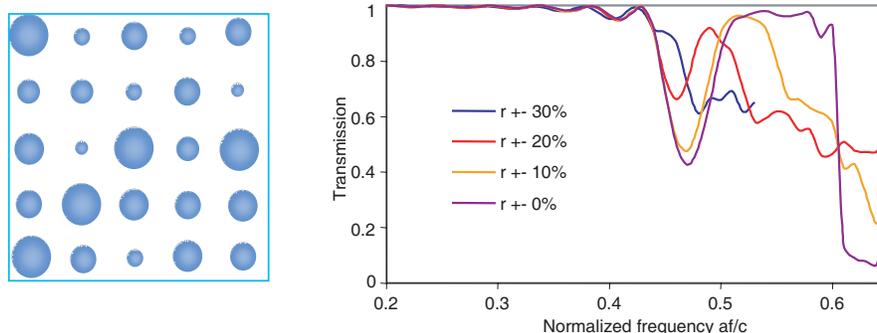


Figure 5: Effect of random variation of sphere radii on the transmittance as a function of the frequency.

6. CONCLUSIONS

We have developed a new code of multiple-scattering allowing the study of the photonic crystals having a great number of spheres. The optimisation of this code allows calculating easily several thousand spheres with radii near the wavelength in the photonic band gap structure, without any symmetry consideration. The reserved memory and the computing time much weaker are compared with other usual methods. This code converges whatever the lattice or the materials are. The next step consists in extending the code to non-spherical inclusions by T -matrix. That will enable us to study the random effects of a very large variety of photonic band gap structures.

REFERENCES

1. Purcell, E. M. and C. R. Pennypacker, "Scattering and absorption of light by nonspherical dielectric grains," *Astrophys. J.*, Vol. 186, 705, 1973.
2. Waterman, P. C., "Matrix formulation of electromagnetic scattering," *Proceedings of the IEEE*, Vol. 53, 805–812, 1965.
3. Taflove, A. and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method*, Artech House, Boston, 2005.
4. Jin, J., *The Finite Element Method in Electromagnetics*, Wiley & Sons, New York, 2002.
5. Mittra, R., "A look at some challenging problems in computational electromagnetics," *IEEE Ant. Prop. Mag.*, Vol. 46, No. 5, 18–32, 2004.
6. Oyhenart, L. and V. Vignéras, "Study of finite periodic structures using the generalized mie theory," *Eur. Phys. J. Appl. Phys.*, 95–100, 2007.
7. Joannopoulos, J. D., R. D. Meade, and J. N. Winn, *Photonic Crystals*, Princeton University Press, 2008.

Microstrip Ultra-Wide-Band Filter

Abdel-Fattah Sheta and Ibrahim Elshafiey

Electrical Engineering Department, College of Engineering
King Saud University, P. O. Box 800, Riyadh 11421, Saudi Arabia

Abstract— A new UWB filter is proposed. The filter structure is based on identical lines and short circuited stubs. The merit of the proposed structure lies in the ability to design a simple structure with arbitrary bandwidth. The proposed filter can be designed for a fractional bandwidth from 40% to more than 100%. The concept is validated through the design of a bandpass filter to cover the entire band of the UWB standard and another filter is designed to cover few groups of the UWB standard (from 6 to 10 GHz). Simulations have been performed using IE3D software package.

1. INTRODUCTION

Since the Federal Communications Committee (FCC) authorized the unlicensed use of the ultra-wideband (UWB) frequency spectrum for short-range and high-speed wireless communication in 2002 [1], remarkable research interests in the analysis and design of various microwave devices have been investigated [2]. UWB bandpass filters (BPFs) with low insertion loss, high out-of-band rejection, and a flat group delay performance within that band, are highly needed to meet the required UWB frequency mask (3.1 to 10.6 GHz). Several prototype UWB filters based on different techniques have been studied. Coupled microstrip resonators are proposed for 128% fractional bandwidth by cascading two stages fork resonator [3]. Another microstrip coupled resonators is designed by cascading two interdigital hairpin resonators [4]. The filter in this case designed to suppress wireless local area network in the band 5.7 GHz to 5.8 GHz. Another method based on circuit model for an optimum distributed high pass filter (HPF) [5] has been implemented in various techniques [6, 7]. The filter structure in its simplest form consists of cascaded short circuit stubs separated by identical connecting lines. The electrical length of the connecting lines is twice that of the length of the short circuited stubs at the lower cutoff frequency. The characteristic impedances of the short-circuited stubs and the connecting lines are calculated based on the HPF model [5].

In this paper, a filter consists of short circuited stubs separated by identical connecting lines is investigated. Based on this arrangement, UWB filter with fractional bandwidth from about 40% to 115% can be designed with one or more short circuited stubs. The stubs electrical length is half of the connecting lines electrical length. The structure is sensitive to the relative electrical lengths.

2. FILTER STRUCTURE

As shown in Fig. 1, the filter structure under considerations consists of four stubs of electrical length θ_s and characteristic impedances Z_{o1} , Z_{o2} , Z_{o3} , and Z_{o4} separated by connecting lines of electrical length θ_c and characteristic impedance Z_{oc} .

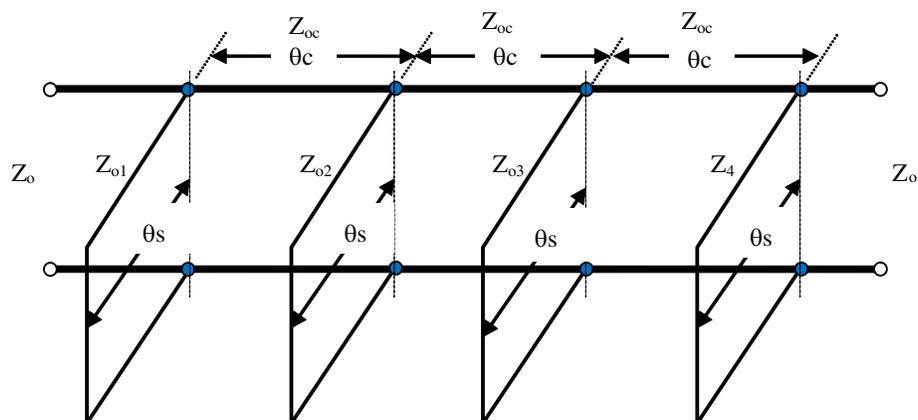


Figure 1: Circuit diagram of the proposed UWB filter. $Z_{o1} = Z_{o4}$ and $Z_{o2} = Z_{o3}$.

The stubs electrical lengths are half of the connecting lines electrical length, such that $\theta_c = 2\theta_s$. The stubs length is one quarter wavelength at the filter center frequency. In this design the middle stubs characteristic impedances are equal. The end stubs characteristic impedances are also equal, such that $Z_{o1} = Z_{o4} = Z_1$ and $Z_{o2} = Z_{o3} = Z_2$. It is observed also that the characteristic impedance of the middle lines are half of the end lines, i.e., $Z_1 = 2Z_2$. The filter bandwidth depends on the lines impedances Z_{oc} and Z_1 . Design curves that relate the filter fractional bandwidth to lines characteristic impedances are shown in Fig. 2. These parameters are obtained based on return loss better than 16 dB.

It is thus clear that, the filter bandwidth can be controlled by the selection of the characteristics impedances of the stubs as well as the connecting lines. The filter in this case can be designed to cover either the entire ultra-wide band (3.1 to 10.3 GHz) or some groups of the sub-bands of the entire this band. UWB-filter with fractional bandwidth from 40% to 120% can be achieved.

3. DEIGN CASES

The design concept has been verified through the design of two UWB filters. The first is designed to cover the entire band from 3.1 to 10.3 GHz. The characteristic impedance of the connecting lines is $60\ \Omega$. The characteristic impedance of the two stubs located at the edge is $80\ \Omega$, while the characteristic impedance of stubs in the middle is $40\ \Omega$. The filter is designed at 6.85 GHz filter center frequency with approximately 110% fractional bandwidth. From Fig. 1, the connecting lines impedance is $Z_c = 58\ \Omega$ and the stubs impedances are $Z_1 = 94\ \Omega$, and thus $Z_2 = 47\ \Omega$. The filter has been simulated using the IE3D simulator and the simulation results are shown in Fig. 3. Dielectric material with 4.5 dielectric constant and 0.78 mm thickness is used in this analysis. The filter layout is shown in Fig. 4. Very good performance is observed with return loss better than 12 dB in the entire frequency band. The mid band insertion loss is less than 0.5 dB. Another narrow band filter has been designed with the aid of the design curves presented in Fig. 2. The filter is designed for a fractional bandwidth of 50% centered at 8 GHz. The connecting line characteristics impedance is $102\ \Omega$, and the stubs impedance is $Z_1 = 20\ \Omega$ and so, $Z_2 = 10\ \Omega$. Since the middle lone stubs have low impedances, two parallel lines with $20\ \Omega$ characteristic impedance are used instead. The filter is analyzed using a dielectric substrate with 6.15 dielectric constant and 0.635 mm thickness. The simulation results are shown in Fig. 5, and the layout is shown in Fig. 6. The midband insertion loss is about 1 dB, and the return loss is better than 10 dB in the entire band.

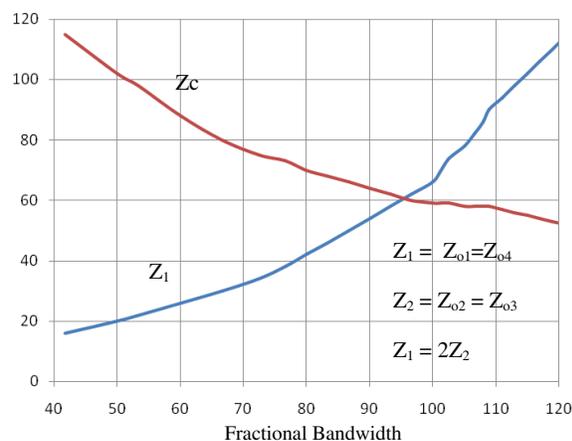


Figure 2: Stubs and connecting lines characteristic impedances against filter fractional bandwidth.



Figure 3: Simulation results of the proposed UWB filter.

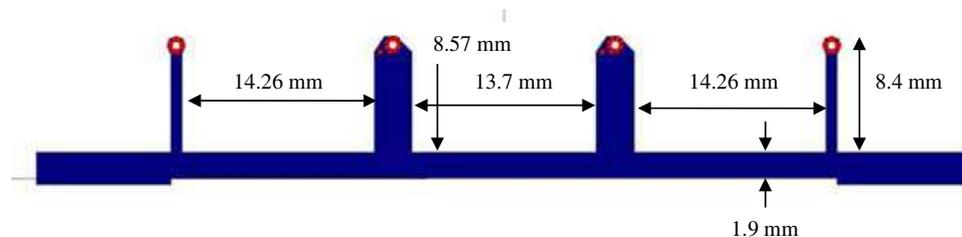


Figure 4: Layout of the UWB filter designed on FR4 material of 4.5 dielectric constant and 0.78 mm thickness.



Figure 5: Simulation results of the proposed UWB filter designed to cover the band from 6 GHz to 10 GHz.

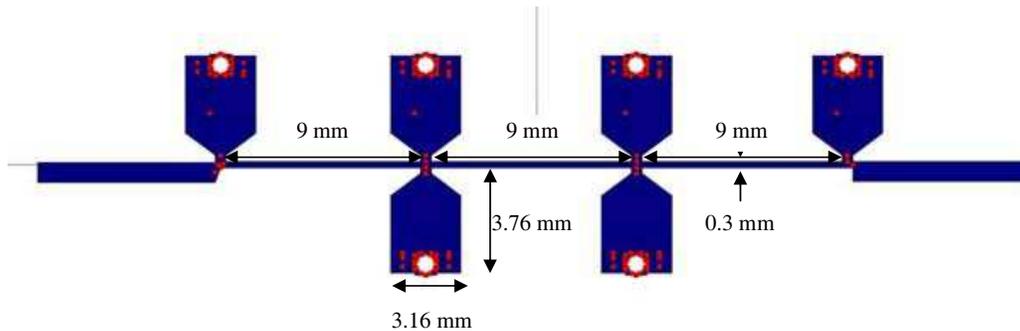


Figure 6: Layout of the UWB filter designed on Duroid material of 6.15 dielectric constant and 0.635 mm thickness.

4. CONCLUSIONS

A new UWB filter structure is introduced. The bandwidth can be controlled using only two lines impedances. Fractional bandwidth from 40% to 120% can be achieved using the proposed structure. Two filters have been designed to verify the design concept. Design curves prove to be helpful for arbitrary fractional bandwidth requirements.

ACKNOWLEDGMENT

This research is funded by *The National Plan for Science & Technology*, Kingdom of Saudi Arabia, under project No. 08-ELE262-2.

REFERENCES

1. "Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems," ET-Docket 98-153, First note and Order, Federal Communications Commission, Feb. 14, 2002.
2. Aiello, G. R. and G. D. Rogerson, "Ultra-wideband wireless systems," *IEEE Microw. Mag.*, Vol. 4, No. 2, 36–47, Jun. 2003.
3. Chen, H. and Y. Zhang, "A novel and compact UWB bandpass filter using microstrip fork-form resonators," *Progress In Electromagnetics Research*, Vol. 77, 273–280, 2007.
4. Wei, F., L. Chen, X.-W. Shi, X. H. Wang, and Q. Huang, "Compact UWB bandpass filter with notched band," *Progress In Electromagnetics Research C*, Vol. 4, 121–128, 2008.
5. Hong, J. S. and M. J. Lancaster, *Microstrip Filters for RF/Microwave Applications*, Wiley, New York, 2001.
6. Razalli, M. S., A. Ismail, M. A. Mahdi, and M. N. Hamidon, "Ultra-wide band microwave filter utilizing quarter-wavelength short-circuited stubs," *Microwave and Optical Technology Letters*, Vol. 50, No. 11, 2981–2983, Nov. 2008.
7. Shaman, H. and J. S. Hong, "A novel Ultra-Wideband (UWB) Bandpass Filter (BPF) with pairs of transmission zeroes," *IEEE Microw. Wireless Compon. Lett.*, Vol. 17, No. 2, 121–123, Feb. 2007.

Study on Electromagnetic Properties of Reinforced Concrete Construction Wall

A. Pruksanubal

King Mongkut's University of Technology North Bangkok, Thailand

Abstract— This paper presents the electromagnetic properties of reinforced concrete construction wall. The shielding effectiveness of the construction wall has been investigated. The setup of an example of reinforced concrete construction wall with metal grid structure (grid size of 10×10 cm.) is shown. The results from the calculations with Schelkunoff's Model for metal grid structure, the 3D MoM simulations and measurements are compared and discussed in the frequency range of 30 MHz–1 GHz. The comparison results show that the calculation results are agreed with the measurement ones at the frequency below resonance frequency, which is caused by the cabinet. The metal grid structure offers the SE of around 20–40 dB at the frequency below 200 MHz and 10–20 dB below 400 MHz.

1. INTRODUCTION

Nowadays, everyday life is concerned with IT devices such as PDAs, mobile phones, laptops, etc., both directly and indirectly. These devices can be used to send or receive the information to or from the large distance. By this reason, our devices can also be disturbed during working in this electromagnetic environment. The elimination or attenuation of the disturbances can be done by shielding the sensitive area. An excellent design of a room or building structure can provide an excellent shielding effectiveness (SE). The application of materials, which offer better shielding properties, can increase SE or reduce the disturbing signals.

In scope of this work, the electromagnetic properties of around 20 construction materials, such as Gypsum boards, Aluminum blinds, etc., are studied. The study procedures start with setting up the measurement method and determining the test materials. The measurement setup is discussed. Then the simulations with Method of Moments (Mom) are modeled in 3D by copying the measurement setup. For verification of the measurement results, the analytical calculation and MoM simulation are applied. The properties of the materials are determined by using the technical data from the literatures or from reasonable assumptions.

2. MEASUREMENT SETUP

A measurement setup for SE is arranged by following CENELEC EN 50147-1, which refers to the measurement of SE in the frequency range of 75 MHz–10 GHz. The measurement setup consists of signal generator, amplifier, EMI receiver, sending antenna, receiving antenna and shielding cabinet, at which the test materials are integrated to. Since the cabinet is a full anechoic one, it can be sure that the external field won't interfere the measuring system. Fig. 1 shows the measurement setup for SE.

The shielding cabinet with the inner dimension of $a \times b \times h = 3 \text{ m} \times 3.5 \text{ m} \times 2.4 \text{ m}$ must be considered to be a cavity resonator, which standing waves can occur at some exact resonance

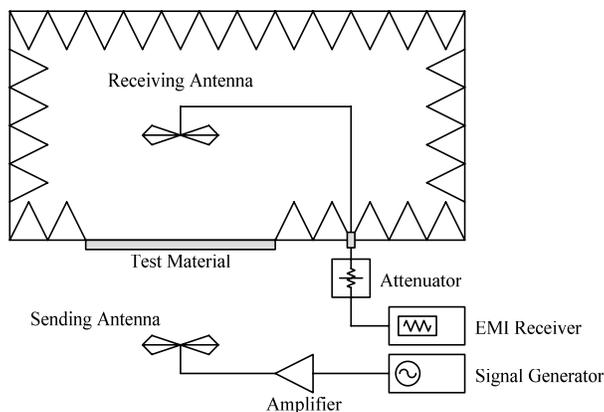


Figure 1: Measurement setup for shielding effectiveness following CENELEC EN 50147-1.

frequencies. Fig. 2 shows the structure of shielding cabinet, which is used for simulation. According to the standing wave within the cabinet, the measurement results can be distorted. The resonance frequency can be calculated by:

$$f_{res} = \frac{c_0}{2} \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{k}{h}\right)^2} \quad (1)$$

where m, n, k are integers, which not more than one of them can be zero at the same time.

The test materials are placed at the door of the cabinet, which are large enough, so that only the incident field through the test material has been measured and the field from other parts of the cabinet can be neglected. The door of the cabinet has the dimension of $w \times h = 1.7 \text{ m} \times 1.85 \text{ m}$. The first mode of resonance (H_{01} -resonance) is

$$f_{res} = \frac{c_0}{1} \sqrt{\left(\frac{m}{w}\right)^2 + \left(\frac{n}{h}\right)^2} = \frac{3 \times 10^8}{2} \sqrt{0 + \left(\frac{1}{1.85}\right)^2} = 81.08 \text{ MHz} \quad (2)$$

which is called cut-off frequency. The fields at lower frequency are strongly attenuated through cabinet wall.

3. SHIELDING EFFECTIVENESS OF THE CONSTRUCTION WALL

Under investigation of several materials, the reinforced concrete construction wall obviously shows its electromagnetic properties, which can attenuate the electromagnetic field at some frequency ranges. The construction wall is made of concrete with metal grid structure inside. The ingredients of the concrete are cement, water and stone. The structure of the construction wall is shown in Fig. 3.

The SE of the horizontal polarized field is measured in the frequency range of 30 MHz–1 GHz and the measurement results are shown in Fig. 5 comparing with the simulation results described in next chapter. With the 10 cm. spacing of the grid structure, the cut-off frequency becomes

$$f_{cut-off} = \frac{3 \times 10^8}{2} \sqrt{0 + \left(\frac{1}{0.1}\right)^2} = 1.5 \text{ GHz} \quad (3)$$

which is higher than the considered frequency range.

4. SIMULATION MODEL

Because of the complex geometry of the measurement setup, a numerical simulation as Method of Moment (MoM) is applied in 3D. In order to limit the simulation matrix size, all metal planes and grids are considered to be without thickness and contain the ideal conductivity. The absorber materials are neglected. A horizontal polarization plane wave with the amplitude $|E|$ of 1 V/m is generated and set to be an incident field. The origin of Cartesian coordinate system, at which the

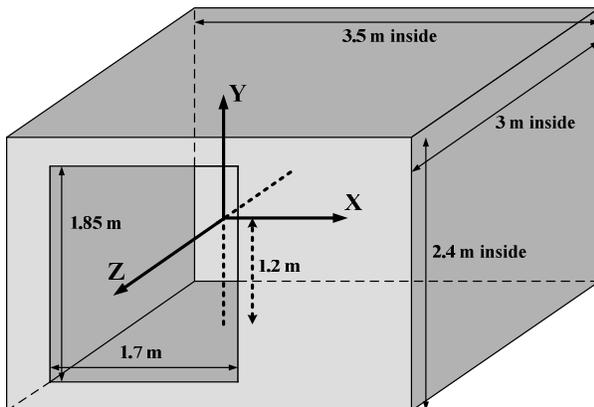


Figure 2: Shielding cabinet considered to be a cavity resonator.

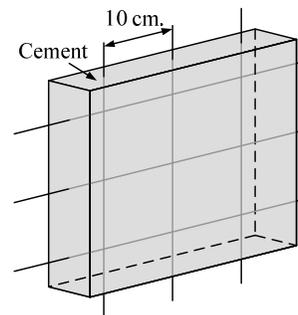


Figure 3: The reinforced concrete construction wall as a test material.

field intensity E_x is observed, is inside the cabinet, 1 m distance from the door and 1.2 m above the floor (Fig. 2).

The model has been simulated, when the door has been opened, in order to obtain the distribution of field intensity E_x at origin as shown in Fig. 4. The resonances occur at several frequencies, especially at 240 MHz, the field intensity is increased almost 7 times of the incident wave.

The next step is to study the electromagnetic properties of the reinforced concrete construction wall by putting the metal grid structure in the model at the door position and simulating in the same frequency range of 30 MHz–1 GHz. The simulation results of the field intensity distributions are transformed to be the SE by:

$$SE = 20 \cdot \log \left| \frac{E_{withoutShield}}{E_{withShield}} \right| \quad (4)$$

and are shown in Fig. 5.

For metal grid structure, its SE can be calculated by an analytical method of Schelkunoff as:

$$SE = 20 \cdot \log \left(\frac{\lambda}{2w} \right) \text{ dB} \quad (5)$$

where w is the grid spacing and $\lambda/100 < w < \lambda/2$.

Figure 5 shows the comparison of SE of the reinforced concrete construction wall detected by three methods, MoM simulation, Schelkunoff method and measurements.

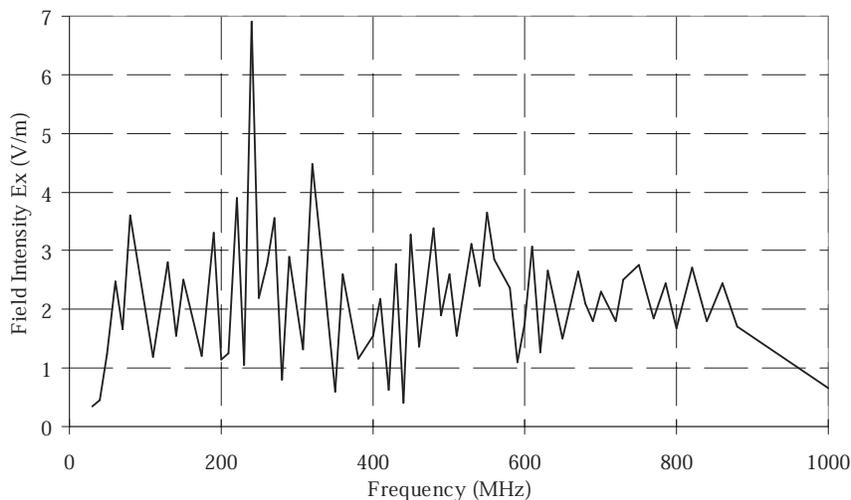


Figure 4: Field distribution at observation point.

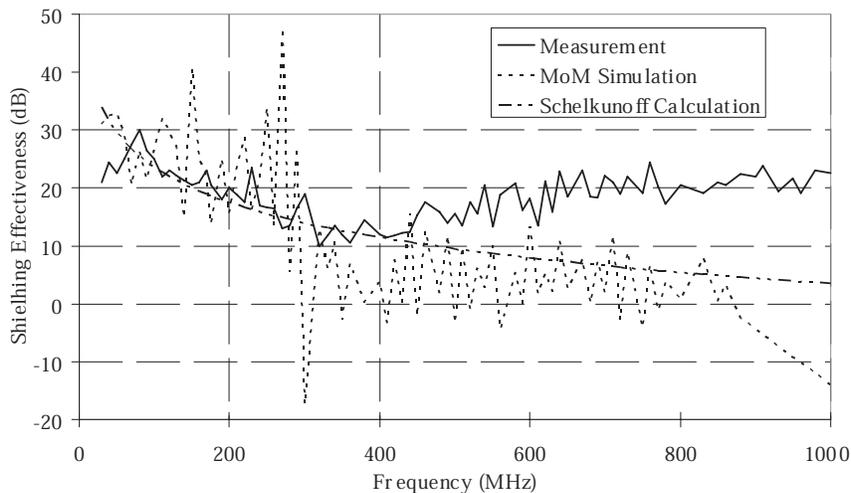


Figure 5: Comparison of SE of the reinforced concrete construction wall detected by three methods.

According to Fig. 5, the SE falls from 30 dB to 0 dB in the measuring frequency range. The comparisons of both analytical and numerical results with the measurement one obviously show that all results are in the same trend until 350 MHz with up and down distribution because of the resonance effects.

5. CONCLUSION

In this work, the investigations of the electromagnetic properties of building materials have been discussed. The reinforced concrete construction wall with metal grid structure is chosen to be a test example. A measurement setup for shielding effectiveness is arranged by following CENELEC EN 50147-1. The field distribution is strongly effected by cavity resonance, however the absorbers can reduce these resonance effects. The results from the calculations with Schelkunoff's Model for metal grid structure, the 3D MoM simulations and measurements are compared and discussed in the frequency range of 30 MHz–1 GHz. The comparisons obviously show that all results are in the same trend until 350 MHz with up and down distribution because of the resonance effects.

REFERENCES

1. CENELEC EN 50147-1 Anechoic Chambers Part 1: Shield Attenuation Measurement, 1996.
2. IEEE Standard Method for Measuring the Effectiveness of Electromagnetic Shielding Enclosures, 1997.
3. White, D. R. J., *A Handbook Series on Electromagnetic Interference and Compatibility, Vol. 3, Electromagnetic Shielding*, Interference Control Technologies, Inc., Virginia, USA, 1988.
4. White, D. R. J., *Electromagnetic Shielding and Performance*, Don White Consultants, Inc., Maryland, USA, 1975.
5. Clayton, R. P., *Introduction to Electromagnetic Compatibility*, John Wiley & Sons, Inc., New York, 1992.

Effect of Friction Layer Creep Deformation on Dynamic Behavior of Traveling Wave Rotary Ultrasonic Motor

Chao Chen, Huafeng Li, Fan Li, and Kang Yang

Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 210016, China

Abstract—The contact mechanism of traveling wave rotary ultrasonic motor (TRUM) provides us with numerous advantages. Usually polymer friction material is bonded on the contact surface of the TRUM rotor, in order to obtain excellent mechanical characteristics, such as high efficiency, long life and so on. Unfortunately, the creep happens for the friction layer made of polymer material, due to the large pressure at the contact interface between the stator and rotor. In this paper the creep of the friction layer is taken into account at the contact interface, especially its effect on the startup characteristics of TRUM in long-term storage. Based on the contact model of TRUM the creep behavior of the friction layer is investigated at the contact interface between the stator and rotor. With the consideration of the friction layer creep, we tried to explain how the distributing contact forces applied on the interface affect the speed response characteristics of TRUM, as is shown in our experiments.

1. FOREWORD

As a novel class of actuators, traveling wave type rotary piezoelectric ultrasonic motor (TRUM) shows an increasing importance in the field of electromechanical sensors, actuators and electro-optic modulators, and attained attention from many investigators. TRUM has a high ratio of torque to weight, low speed output, in particular also compact size and fast response, so that it gains an advantage over the conventional motor in some applications such as auto focus camera lenses, precise positioning and robot actuation [1–3]. TRUM makes use of contact mechanism so that vibration movement of the stator can be converted into movement of the rotor. Therefore, the dynamics of the contact interface between the stator and rotor greatly affect the final output characteristics of TRUM, especially the important response characteristics. Usually polymer friction material is applied at the contact surface of the rotor, in order to obtain extra excellent mechanical characteristics. Creep behavior of the polymer material friction layer occurs in case of the long-term storage, due to the large pressure at the contact interface between the stator and rotor. In this case, TRUM has the poor response characteristics. Some experiments show the motors cannot even start up with superfluous creep deformation. Although in the recent years the modeling and design techniques have been developed through using equivalent circuits, analytical approaches and finite element method [3–6], the effect of creep deformation on the dynamics of TRUM have not been investigated.

In this paper, we presented an investigation into the contact mechanism with the consideration of the creep of the friction layer on rotor. Based on the mathematical model of TRUM, The distributing contact forces at the contact interface between the stator and rotor are analyzed. Then the contact problem is numerically modeled at length to calculate the response characteristics of TRUM. With the creep deformation of the friction layer considered, we tried to explain how the normal preload force applied on the rotor affect the startup characteristics of TRUM, as is shown in our experiments.

2. CONTACT MECHANISM WITH CONSIDERATION OF FRICTION LAYER CREEP BEHAVIOR

Generally TRUM works based essentially on the sequence of two energy conversion levels. One is the electromechanical energy conversion in stator: ultrasonic vibration produced by piezoelectric ceramics. The other is the mechanical energy conversion: rotary movement of rotor derived by friction, which occurs at contact interface, which is illustrated in Figure 1.

Due to the electromechanical coupling involved in piezoceramic material, two orthogonal flexural modes with the same frequency and number of nodal diameters can be excited in stator, if two phase voltages are respectively induced across two groups of piezoceramic half-ring. Two standing wave are simultaneously excited with a temporal phase shift of 90 degrees, and a traveling wave comes into being in stator. In this case a rotor pressed against the stator can be propelled along in the reverse direction from the propagating wave.

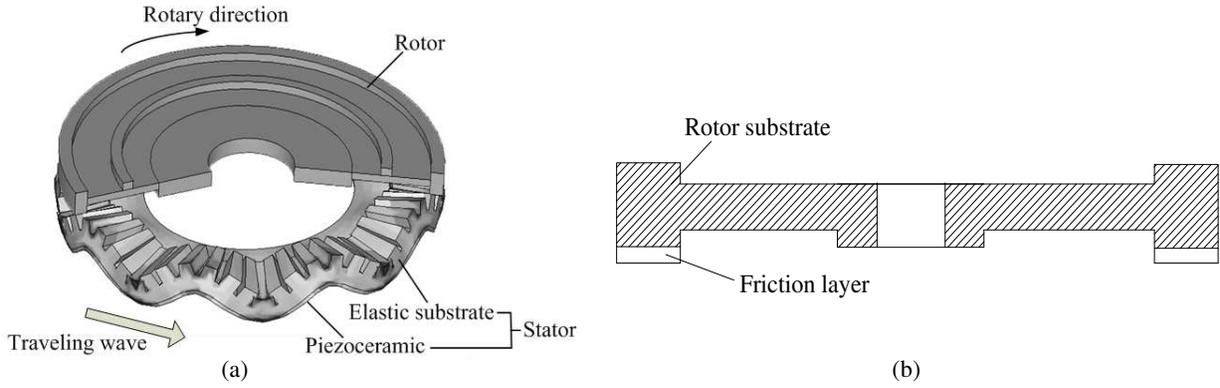


Figure 1: TRUM based on contact mechanism. (a) Rotor propelled by stator with a traveling wave induced. (b) Rotor.

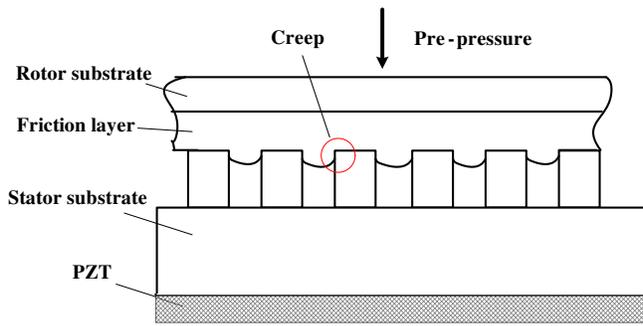


Figure 2: Creep behavior of friction layer.

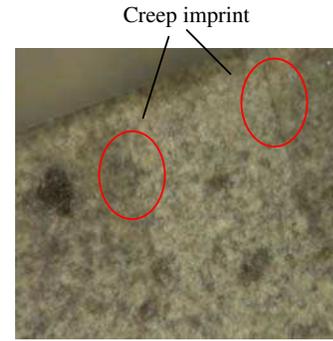


Figure 3: Creep imprint on the friction layer through microscope.

It is well known that polymer material tends toward some permanent strain named by creep behavior. Moreover creep occurs for the polymer contact layer on the rotor of TRUM in case of long-term storage sometimes. Figure 2 shows the friction layer with creep behavior. It is found that the friction layer made of polymer material flows into gaps between the teeth slowly during the long-term storage of TRUM. By microscope system VHX-600 of KEYENCE, the creep imprint can be observed after some time storage.

3. CONTACT MODEL CONSIDERING FRICTION LAYER CREEP BEHAVIOR

Usually a thin contact layer, which is made of rather soft material (in comparison to phosphor bronze and duralumin), is bonded to the rotor in the contact region and can be modeled as a linear spring [7]. Figure 6 shows the interface forces applied to the stator. Except for tangential force, $\mathbf{f}_\tau(\mathbf{e})$, and normal pressure applied on the top surface of teeth, $\mathbf{f}_n(\mathbf{e})$, another force component on the gaps of teeth, $\mathbf{f}_{\tau c}(\mathbf{e})$, exists, where e is used for the number of any tooth.

Then the distributing force vector, which acts on the stator, includes two components in the axial and tangential direction, which can be written as

$$\bar{f}_n = f_n + f_{\tau c} \quad (1a)$$

$$f_\tau = \mu_d f_n \quad (1b)$$

where f_n denotes the normal force on the surface on top of teeth, $f_{\tau c}$ is the extra axial force due to the creep behavior of friction layer. It is noteworthy that only f_n exists without the creep behavior. The tangential value of the contact pressure distribution at the interface, f_τ , is defined on the basis of Coulomb law, and μ_d is the Coulomb's friction coefficient.

According to [8–10] the governing equation of the TRUM stator with two degrees of freedom can be derived based on Hamilton's principles as follows

$$\mathbf{M}\ddot{\mathbf{q}} + \mathbf{C}\dot{\mathbf{q}} + \mathbf{K}\mathbf{q} = \Delta\nu + \bar{\mathbf{F}}_C \quad (2)$$

where \mathbf{M} , \mathbf{K} are the mass and stiffness matrices, and $\bar{\mathbf{F}}_C$, which derives from Eq. (1), is the generalized forcing vector at the interface. The equation above takes the damp matrix \mathbf{C}_s , into

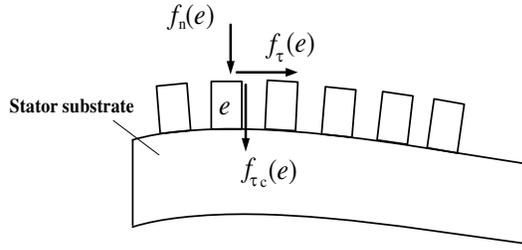


Figure 4: Distributing forces applied on stator.

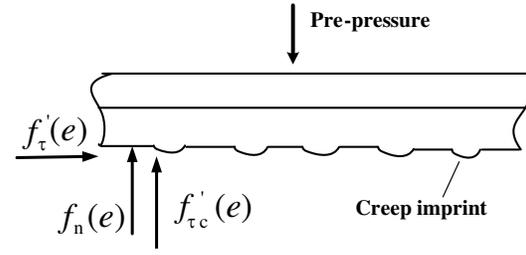


Figure 5: Distributing forces applied on rotor.

account. The matrix ν is the applied voltages vector and Δ is composed of the electromechanical coupling matrix. The more detailed derivation and symbol definition can be obtained in [8].

The dynamic equations of the TRUM rotor read

$$M_R \ddot{z} + C_R \dot{z} = F_{con} - F_{app} \quad (3a)$$

$$J_R \ddot{\beta} + C_R^J \dot{\beta} = T_{con} - T_{load} \quad (3b)$$

where M_R is the mass of rotor and C_R , the corresponding damping in the z -direction. The F_{app} is applied axial preload. The J_R is rotor inertia, T_{load} output torque, C_R^J the spinning damping. The axial force, F_{con} , and the driving torque, T_{con} , from the contact interface read respectively

$$F_{con} = \iint (f'_n + f'\tau_c) dS \quad (4a)$$

$$T_{con} = \iint \gamma f'_\tau r dS \quad (4b)$$

The sign function, γ , is defined by the velocities of the contact points, which are given in [8]. Without the creep of the friction layer, there would not be the extra axial force, $f'_{\tau c}$, as shown in Figure 5. So the creep behavior of the friction layer will have effect on the performance of TRUM. The explicit model representing the complete motor is obtained by Eqs. (2) and (3), and the dynamic response of TRUM can be achieved based on numerical method.

4. SIMULATIONS AND EXPERIMENTS

In these simulations of TRUM-60, which means the ultrasonic motor with the diameter of 60 millimeter, the input value for F_{in} is 200 N, and the applied voltages is $260V_{pp}$. Figure 6 shows the calculated speed response of TRUM-60.

In Figure 6, the percentage denotes the ratio of the creep to thickness of the friction layer. With the creep ratio of over 20%, the ultrasonic motor with the diameter of 60 mm can not operate. Figure 7 shows the speed response of TRUM-60 with the creep ratio of 20%, which is experimentally

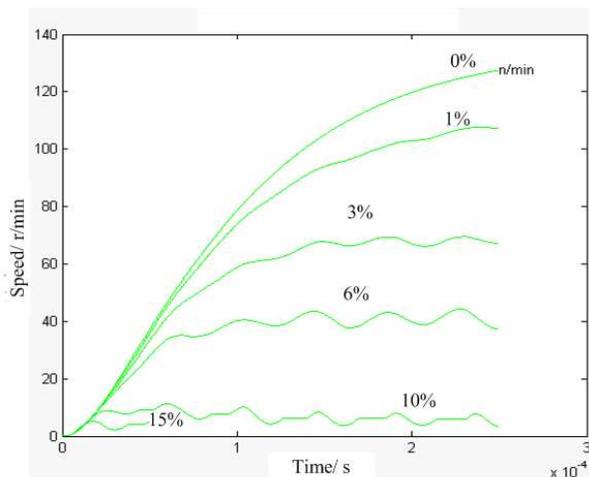


Figure 6: Speed response of TRUM-60 with various creep ration.

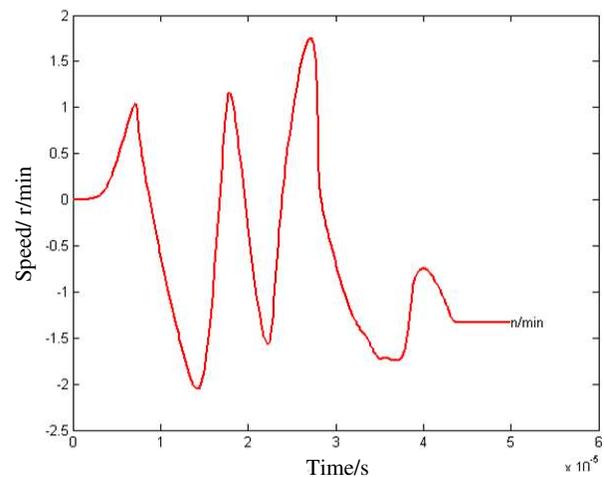


Figure 7: Abnormal operation of TRUM-60.

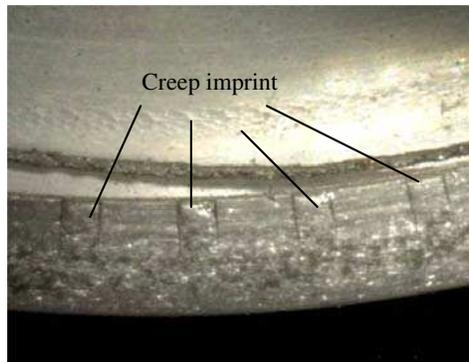


Figure 8: Creep ratio over 20% on the friction layer.

validated too. We conducted an accelerated testing for the creep of polymer friction layer, in order to observe the dynamic response of TRUM-60. More pressure was applied at the contact interface between the stator and rotor with high temperature. Several months later, the pressure and temperature returned to the normal level. Figure 8 illustrates the permanent deformation on the polymer friction layer of the rotor, which exceed 20%. In this case it was found the motor can hardly operate with the normal excitation conditions.

5. CONCLUSION

In this paper, the creep behavior of polymer friction layer and its effect on the contact mechanism are investigated, in order to evaluate the response characteristics of TRUM in long-term storage. The contact interface model is presented with the consideration of polymer material creep in the friction layer, and the contact problem is numerically analyzed at length to calculate startup characteristics of TRUM, as is shown in our experiments. It is found that ultrasonic motors can't operate if the creep ratio is over 20%. So the special design for the material of friction layer should be noticed to obtain the good speed characteristics of TRUM.

ACKNOWLEDGMENT

This project is supported by National Natural Science Foundation of China (10604032, 50975135) and Specialized Research Fund for the Nanjing University of Aeronautics and Astronautics (No. NS-2010001).

REFERENCES

1. Sashida, T., "Trial construction and operation of an ultrasonic vibration — driven motor," *Japan Applied Physics*, Vol. 51, No. 6, 713–720, 1982.
2. Sashida, T. and T. Kenjo, *An Introduction to Ultrasonic motors*, 4–10, Oxford, London, UK, 1993.
3. Zhao, C. S., *Ultrasonic Motors Technologies and Applications*, 125–129, Springer Press, 2009.
4. Nesbitt, W. H. and J. M. Andrew, "Modeling of a piezoelectric rotary ultrasonic motor," *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.*, Vol. 42, No. 2, 210–224, 1995.
5. Frangi, A., A. Corigliano, M. Binci, and P. Faure, "Finite element modelling of a rotating piezoelectric ultrasonic motor," *Ultrasonics*, Vol. 43, No. 9, 747–755, 2005.
6. Zhao, C. S., "Ultrasonic motor techniques for 21st century," *Engineering Science*, Vol. 4, No. 2, 86–91, 2002 (in Chinese).
7. Wallaschek, J., "Contact mechanics of piezoelectric ultrasonic motors," *Smart Mater. Struc.*, Vol. 7, 369–381, 1998.
8. Chen, C. and C. S. Zhao, "A novel model of ultrasonic motors with effect of radial friction in contact mechanism," *Journal of Electroceramics*, Vol. 20, No. 3–4, 293–300, 2008.
9. Friend, J. R. and D. S. Stutts, "The dynamics of an annular piezoelectric motor stator," *Journal of Sound and Vibration*, Vol. 204, No. 3, 421–437, 1997.
10. Young, H. K. and K. H. Sung, "Analysis of a disk-type stator for the piezoelectric ultrasonic motor using impedance matrix," *Journal of Sound and Vibration*, Vol. 263, No. 3, 643–663, 2003.

Doorway State Mechanism with Electromagnetic Waves in the Optical Regime

Celia Sánchez-Pérez¹, Karen Volke-Sepúlveda², and Jorge Flores²

¹Centro de Ciencias Aplicadas y Desarrollo Tecnológico, Universidad Nacional Autónoma de México
Apdo. Postal 70-186, 04510 México D. F., México

²Instituto de Física, Universidad Nacional Autónoma de México
Apdo. Postal 20-364, 01000 México D. F., México

Abstract— The doorway state mechanism appears whenever two oscillating systems — one of them having a low density of states, called the “doorway states”, and the other one having a much larger density of states — couple together. The output of the coupled system is an important enhancement of the oscillation amplitude, and therefore of the energy, for the frequencies of the high-density states around the frequencies of the doorway states. We experimentally analyze the case of two resonators fabricated by means of two thin layers of metal oxides that satisfy the condition of high optical contrast. We show, both numerically and experimentally, that the doorway state mechanism takes place in this system in the optical regime.

The “doorway state” mechanism has been observed in many different systems. It was introduced several decades ago in nuclear physics when dipole giant resonances and isobaric analogue states were analyzed [1–3]. In more recent times, this mechanism has attracted attention since doorway states have been found in atoms and molecules [4], quantum dots [5], fullerenes [6, 7] and, furthermore, in classical wave-like systems: microwave resonators [8] and in seismic response of sedimentary valleys [9]. This mechanism constitutes, therefore, a unifying concept in physics, covering a large range of scales, from femtometers to tens of kilometers. However, it has not been discussed in the optical regime, which is the purpose of the present contribution.

What is the doorway state mechanism? Consider two oscillating systems, one of them having a low density of “simple” states, and the other one a sea of “complicated” states coupled together. Each simple or “distinct” state is coupled to the entrance channel and the complicated states are more weakly coupled to the continuum or not at all. The distinct states are called doorway states. They are not eigenmodes of the complete system and the strength width phenomenon takes place: The strength of the doorway state is distributed among the eigenstates of the complete system. In the simplest case this distribution follows a Breit-Wigner form, a Lorentzian, as a function of energy or frequency.

In order to experimentally demonstrate the doorway state mechanism in the optical regime, we implemented a system of two resonators, which are based on the successive deposition of thin films on a glass substrate by a classical sputtering technique. The resonators consist of metal oxide layers: Al₂O₃ (aluminum oxide) and WO₃ (tungsten oxide), referred to as R1 and R2, respectively, whose refractive indices differ from one another in more than 40%. An additional, partially reflective, metallic aluminum layer was deposited on the glass before the first resonator, with the aim of isolating the system from the substrate and also to discard external reflections; we will refer to this film as the buffer layer (Fig. 1(a)). The widths of R1 and R2 were chosen so as to satisfy the condition $\Lambda_2 \gg \Lambda_1$, where Λ_j ($j = 1, 2$) represents the optical path inside the corresponding resonator. In this case, the doorway states are associated with R1.

We also analyzed the case when the two resonators are separated from one another by a second metallic aluminum film (Fig. 1(b)), here called the coupling layer, whose purpose is to increase the finesse of R1, and thus to improve the coupling between R1 and R2. Both layers of metallic aluminum are partially reflective and also absorptive. Hereafter, the two systems illustrated in Figs. 1(a) and 1(b) will be referred to as A and B, respectively.

Hence, our system consists of N layers of different materials, each one of width d_j and refractive index n_j , which can be, in general, a complex function of the wavelength (or frequency): $n_j(\lambda) = \text{Re}\{n_j(\lambda)\} + i\text{Im}\{n_j(\lambda)\}$, with $j = 1, \dots, N$. For system A: $N = 3$, while for B: $N = 4$. The glass substrate is not considered as a layer but just as the incidence medium with refractive index n_0 , since it is much wider than the thin films. The theoretical modelling of systems A and B was done by determining the characteristic matrix of this stratified medium [10] as a function of the wavelength, from which the spectral transmittance can be obtained. We consider the case of normal

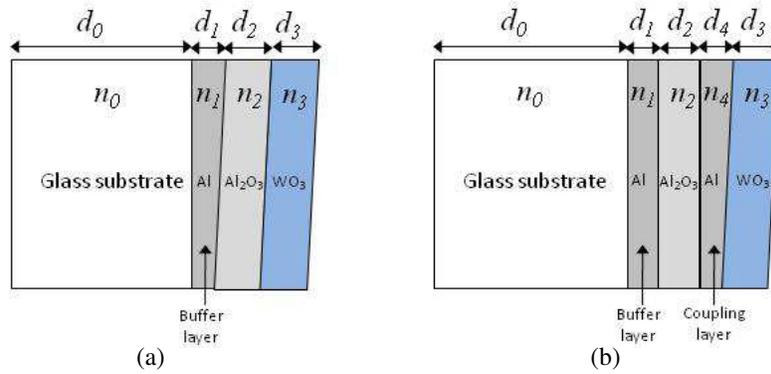


Figure 1: Schematic of the thin layer configurations for systems (a) A and (b) B described in the text.

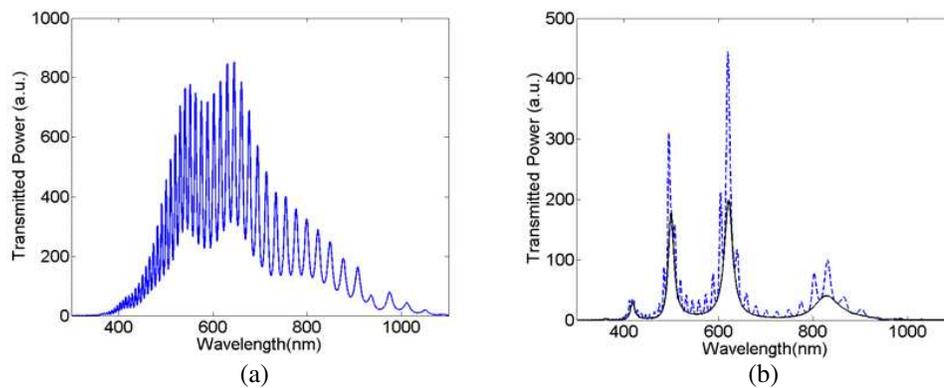


Figure 2: Theoretical response of systems (a) A and (b) B. The solid line in (b) corresponds to the transmittance of the doorway states.

incidence. The data for the dispersion of the refractive index $n_j(\lambda)$ for Al_2O_3 , WO_3 and Al were obtained from references [11, 12].

Figures 2(a) and (b) show the results of the theoretical calculations for the transmittance of systems A and B (dashed line in (b)), respectively. It can be seen, from the fact that there is an enhancement in the response around some wavelengths (the distinct states), that in both cases there exists a coupling between the two resonators. However, the inclusion of the coupling layer (system B), dramatically improves the effect of the doorway states in the coupling. In Fig. 2(b), we have included, for comparison, the response of the distinct states alone (solid line), in order to appreciate that the transmission is indeed increased for the coupled system.

The aluminum oxide layer (R1), of thickness $d_2 \sim 500$ nm, was grown using an aluminum oxide AlO target (Lesker 99.99% purity) and an oxygen reactive deposition process. The tungsten oxide layer (R2), of thickness $d_3 \sim 6000$ nm, was deposited using a tungsten oxide WO target (Lesker 99.99% purity) and also an oxygen reactive deposition process. As shown schematically in Fig. 1(a), the thickness of the buffer layer in system A gradually decreases. It was designed in this way in order to experimentally determine the optimum width for obtaining a good isolation while keeping a relative high transmission. On the other hand, in system B (Fig. 1(b)), the width of the buffer layer was fixed, but now the width of the coupling layer was varied, in this case with the aim of determining the thickness leading to an optimum balance for the best coupling between the resonators and the highest possible transmittance. Due to the fact that system B uses two metallic layers, the transmittance decreases with respect to that obtained for system A. For this reason, the thickness of the buffer layer in system B was reduced to a value of $d_1 \sim 40$ nm. In both cases, the signals must be large enough as to be measured with a conventional photodetector.

The experimental setup we used to evaluate the doorway state mechanism is schematically shown in Fig. 3. White light from a halogen source (DH2000 from Ocean Optics) is coupled to a pigtailed optical fibre and directed onto the sample plane at normal incidence. The transmitted beam is collected by another, 600 nm diameter, optical fibre, which is in turn coupled to a miniature

spectrophotometer (HR2000 from Ocean Optics). We use a collimator for the input and output optical fibers in order to avoid the divergence of the beam. Transmission spectra were measured in the VIS-NIR spectral range. In addition, a reflectance measurement system was also implemented, which consists of a laser (wavelength of 545 nm) impinging at an oblique angle onto the sample plane, on the same incidence point of the white light, and collected by a Si photodetector. The measured reflected power allows us to identify the zones of the sample having a given reflectivity, where each value corresponds to a different thickness of the aluminum layers.

In the experiments, we measured the transmission spectra for each resonator separately, as a reference, and then for the coupled system. For this purpose, our samples were prepared so as to leave a “witness” of each of the individual resonators, but it is worth mentioning that the Al buffer layer was present in all cases. Results for the individual resonators are presented in Fig. 4: (a) corresponds to R1, the doorway resonator, and (b) corresponds to R2, the sea of complicated states. As can be seen from Fig. 4(a), there are three eigenmodes in the analyzed spectral range, at wavelengths 508 nm, 607 nm and 757 nm. For the resonator R2, there are 17 eigenmodes in the region between 500 nm and 750 nm, shown in the inset.

We show the measured spectral response for the coupled resonators in Fig. 5. It will be seen that the doorway state mechanism occurs for the two systems producing a spectral selectivity at the wavelengths of the eigenmodes of the low-density resonator. In Fig. 5(a), we plotted the spectral transmission for different widths of the buffer layer. In this, the higher transmission corresponds to the lower thickness and it decreases as the metallic layer becomes thicker ($d_{1,1} < d_{1,2} < d_{1,3} < d_{1,4} < d_{1,5}$). The spectral selectivity obtained for this case is quite low as we expected from the theoretical results (Fig. 2(a)).

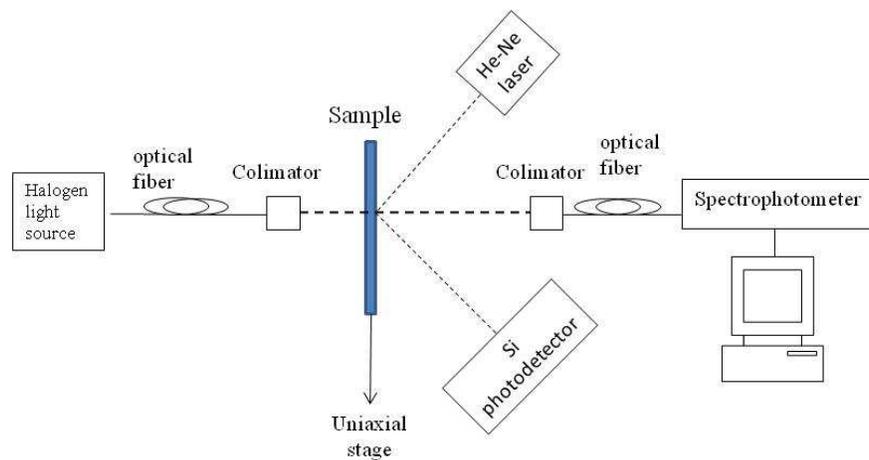


Figure 3: Experimental setup.

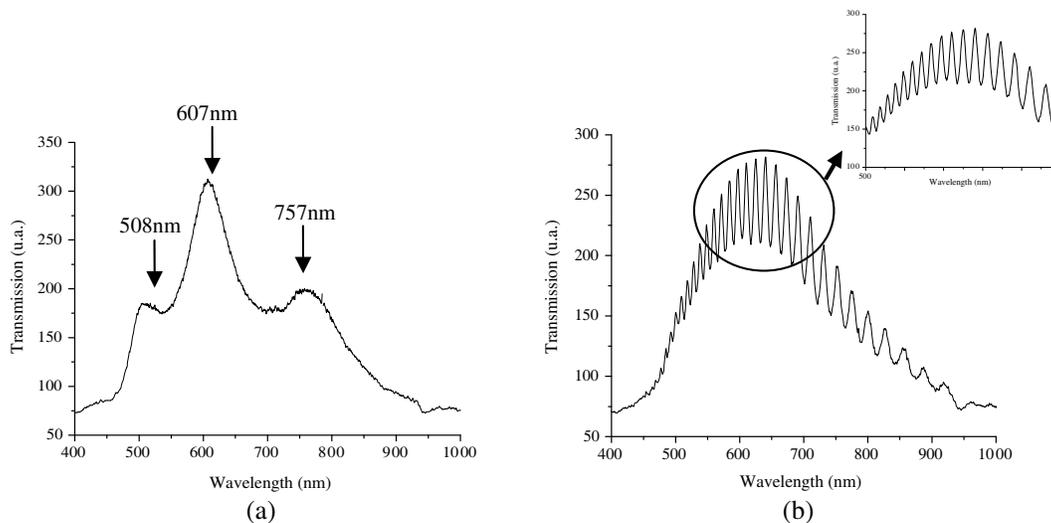


Figure 4: Spectral response for (a) R1 doorway resonator and (b) R2 high density eigenmodes resonator.

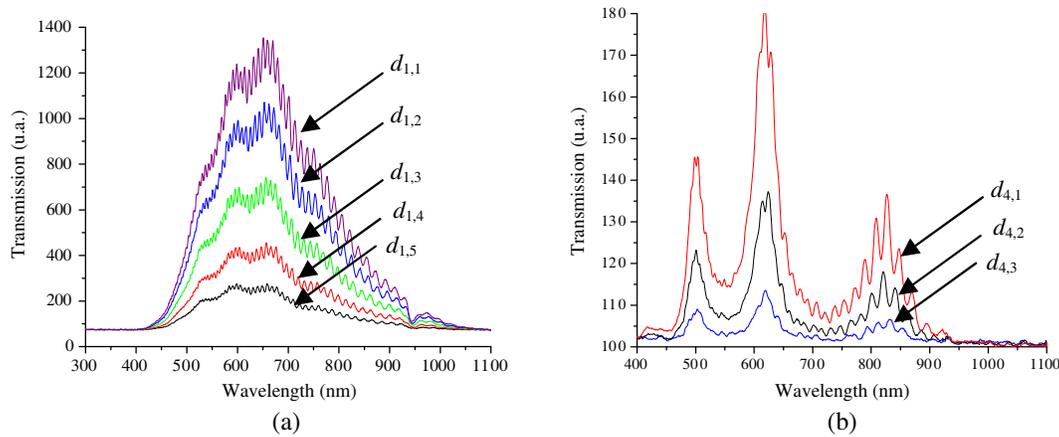


Figure 5: Spectral response for: (a) the two resonators system A for different values of d_1 and (b) two coupled resonators system B for different values of d_4 .

In Fig. 5(b) the spectral response of the coupled system is shown for different widths of the coupling layer ($d_{4,1} < d_{4,2} < d_{4,3}$). Regarding the spectral response when the coupling layer is added to the system, one can appreciate a better spectral selectivity, but one should be aware that the transmission of the signal decreases due to the two Al metallic layers that produce a non-negligible attenuation.

In summary, we have shown, both numerically and experimentally, that doorway states can be present when two layers of different optical paths are coupled. The observed spreading width phenomenon leads to a better spectral selectivity.

ACKNOWLEDGMENT

The authors would like to thank Alejandro Esparza-García from CCADET, UNAM, for its technical contribution and the deposition of the thin films.

REFERENCES

1. Goldhaber, M. and E. Teller, "On nuclear dipole vibrations," *Phys. Rev.*, Vol. 74, 1046, 1948.
2. Brown, G. E., *Unified Theory of Nuclear Models*, North Holland, Amsterdam, 1964.
3. Harney, H. L., A. Richter, and H. A. Weidenmueller, "Breaking of isospin symmetry in compound-nucleus reactions," *Rev. Mod. Phys.*, Vol. 58, 607, 1986.
4. Kawata, I., H. Kono, Y. Fujimura, and A. D. Bandrauk, "Intense-laser-field-enhanced ionization of two-electron molecules: Role of ionic states as doorway states," *Phys. Rev. A*, Vol. 62, 031401(R), 2000.
5. Baksmaty, L. O., C. Yannouleas, and U. Landman, "Nonuniversal transmission phase lapses through a quantum dot: An exact diagonalization of the many-body transport problem," *Phys. Rev. Lett.*, Vol. 101, 136803, 2008.
6. Laarmann, T., et al., "Control of giant breathing motion in C60 with temporally shaped laser pulses," *Phys. Rev. Lett.*, Vol. 98, 058302, 2007.
7. Hertel, I. V., et al., "Fragmentation and ionization dynamics of C60 in elliptically polarized femtosecond laser fields," *Phys. Rev. Lett.*, Vol. 102, 023003, 2009.
8. Aberg, S., T. Guhr, M. Miski-Oglu, and A. Richter, "Superscars in billiards: A model for doorway states in quantum spectra," *Phys. Rev. Lett.*, Vol. 100, 204101, 2008.
9. Flores, J., et al., "Novel doorways and resonances in large-scale classical systems," submitted.
10. Born, M. and E. Wolf, *Principles of Optics*, 7th (expanded) Edition, Cambridge University Press, UK, 1999.
11. Smith, D. Y., E. Shiles, and M. Inokuti, *Handbook of Optical Constants of Solids*, Palik, Ed., Academic Press, 1985.
12. Tropic, W. J. and M. E. Thomas, *Handbook of Optical Constants of Solids III*, Palik, Ed., Academic Press, 1998.

A Hemi-directional Antenna Array Concept for Automotive Radar

S. A. Askeland, T. Cella, and J. Hjelmstad

Norwegian University of Science and Technology (NTNU), Norway

Abstract— New advances in mm-wave technologies are making phased array antenna systems feasible for automotive radar systems. In this paper, we introduce a new automotive radar system capable of wide angle traffic monitoring. A narrow radar beam that is scanned over a wide field of view makes it possible to detect imminent collisions from vehicles approaching our vehicle head on or slantingly, and to monitor traffic from the sides when our vehicle has stopped at a crossing road. The system will ultimately use a multidomain approach (angular, frequency, temporal) to achieve target detection and classification. In order to obtain both wide angle coverage and narrow beamwidth we employ a phased array antenna consisting of three 15 elements planar arrays mounted in three sectors. Each sector has 15 patch antenna elements. The sectors are placed at a 120° angle with each other. When scanning out to angles larger than $\pm 60^\circ$ from broadside only the elements in the respective side sectors are used. When the scan angle is $< \pm 60^\circ$ the elements in the front sector are used together with the elements in either the left or right sector (depending on which side the array is scanned to). For broadside direction, all elements are used. This means that the elements on the side sectors are scanned $\pm 60^\circ$ relative to their broadside directions. By using all elements the array looks electronically larger, and hence a narrower beamwidth can be achieved. To evaluate the performance of the array antenna radiation pattern simulations were run for 0° , 45° and 90° scan angle. The single element pattern and the pattern for the linear arrays in each sector were simulated in CST Microwave studio using a transient time domain solver, while the simulations for angles including elements from multiple sectors were simulated in Matlab Phased Array Design Toolbox by importing the single element pattern from CST and using conformal array theory. The simulation results show that the 3 dB-beamwidths and the side lobe levels for the individual beams are sufficient to obtain good target detection and classification over a more than 180° coverage when used with the multidomain approach. Hence, imminent collisions from head-on, front-side or straight from the sides at intersections can be detected and avoided.

1. INTRODUCTION

With the introduction of adaptive cruise control, radar sensor systems started to enter the automotive industry. As advances are made in the design, manufacture and cost of millimeter wave components, new automotive applications are emerging. Pre-crash systems using the adaptive cruise control radar to monitor the traffic in front of the vehicle are already on the market (e.g., [1]). These systems give the driver a warning in critical situations at long range, and activate the preventive protective systems in critical short range situations.

In addition safety systems that can detect probable front-side collisions are now emerging the market. Toyota [2] have announced a technology that uses diagonally projected, front-side millimeter-wave radar to more broadly detect vehicles moving into the vehicle's path (especially in intersections where there is a clear range of view) or a vehicle that has crossed the centerline. A detection scenario for such a system is shown in Figure 1(a).

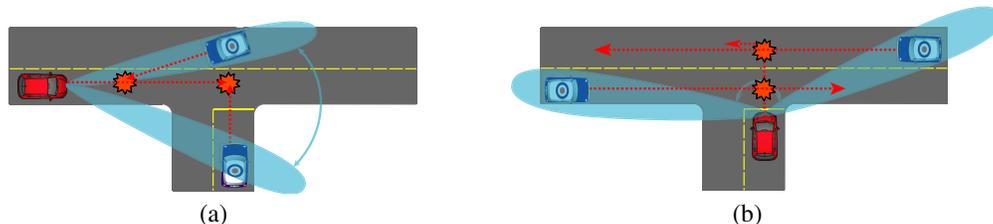


Figure 1: Traffic scenarios.

New advances in RF-MEMS phase shifter design [3] are making phased array antennas at 77 GHz feasible for automotive security systems. In this paper, we introduce such a system that has the capability of both monitoring the traffic straight ahead of the vehicle, and the ability to detect

probable front-side collisions. In addition the system is able to look $\pm 90^\circ$ out to the sides. This provides a new safety feature in scenarios like the one shown in Figure 1(b).

The vehicle carrying the radar sensor has stopped at a crossing road waiting for a gap in traffic in order to turn left. By scanning the radar beam interchangeably from left to right, the position and velocity of lateral moving vehicles can be estimated. Hence, the system can give the driver a signal of when the time gap before the next vehicle is crossing is big enough to avoid dangerous situations.

For close ranges ($< \sim 20$ m) the azimuth resolution is high enough to separate targets in cross-range when scanning out to $\pm 90^\circ$. For longer ranges a separation in azimuth is not possible without increasing the number of antenna elements and thus increasing the cost and complexity of the system. However, for the traffic scenario in Figure 1(b) it is enough to know if there is any vehicles approaching from the sides, their velocities and ranges. Vehicles moving towards the radar have a positive Doppler shift. Further, we know that targets in the radar main beam move on the intersecting road as long as the radar beam doesn't spread out so wide that it scans other nearby roads. At 100 m a 10° beam spreads out to 17.3 m in azimuth. One lane is around 3.5 m wide. Hence, if there are no other roads closer than 5 m to the sides of the intersecting road, the system is able to give a proper warning about intersecting traffic.

2. DESIGN AND SIMULATIONS

A rectangular patch antenna is designed as the basic elements for the array. The substrate material is Rogers 5880. The thickness of the substrate is $127 \mu\text{m}$ and dielectric constant ϵ_r of the substrate is 2.2. A 2D view of the patch antenna is shown in Figure 2(a). The length and width of the patch antenna is 1.31 and 1.54 mm respectively. Other parameters of the antenna are shown in Figure 2(a).

CST [4] microwave studio was used to simulate and optimize the antenna. In order to match the antenna to 50 Ohm, the feed point is recessed 0.35 mm in the antenna. The S_{11} of the patch antenna is shown in Figure 2(b). From Figure 2(b) it can be easily seen that the resonant frequency of the patch antenna is 77 GHz. The S_{11} of the antenna is better than -30 dB at the resonant frequency.

The 2D view of the array configuration is shown in Figure 3. A 1 D array is used for the desired radiation pattern. Each side has 15 elements and they are placed in 120° angle with each other.

In order to reduce the sidelobe level (SLL) to around -20 dB a Taylor window was used for

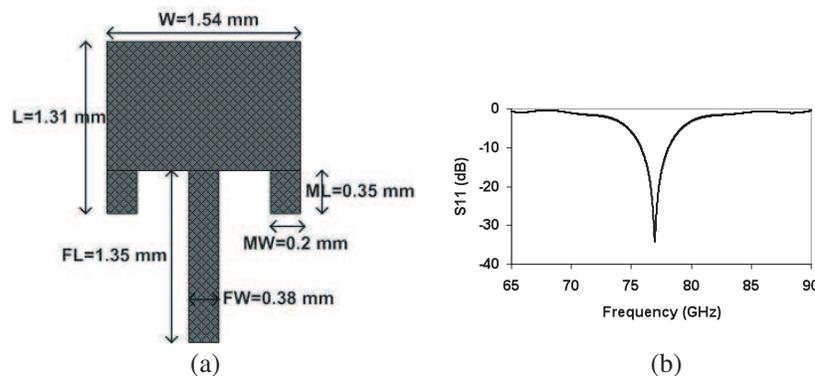


Figure 2: (a) The 2D view of the patch antenna. (b) The simulated S_{11} .

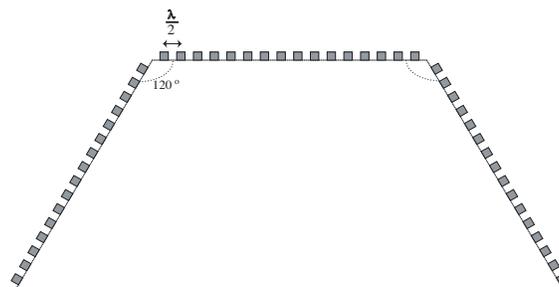


Figure 3: Conformal array configuration.

tapering. The Taylor window SLL was set to -23 dB in order to take into account the pattern degradation when the array is focused off broadside.

To evaluate the performance of the array, the radiation pattern for one element and for one sector was simulated in CST. The more complex configurations using elements in multiple sectors at the same time were simulated in Matlab by importing the radiation pattern for one element from CST and then using the Array Design Toolbox [5]. The effect of coupling was evaluated by running the same simulations in CST and Matlab. Figure 4(a) shows the simulated directivity for one sector with 0° scan angle applied. While Figure 4(b) shows the directivity with -90° scan angle applied. As can be seen the coupling mainly affects the level of the secondary sidelobes, and hence the Matlab simulations for the other configurations are considered to be accurate enough for our purpose.

When scanning out to angles larger than $\pm 60^\circ$ only the elements in the respective side sectors are used. The interelement phase difference α can be found as

$$\alpha = kd \sin \phi_0 \quad (1)$$

where, k is the wavenumber, d is the element spacing ($\lambda/2$) and ϕ_0 is the scan angle for each sector measured away from each sector's broadside.

To show the array performance when scanning out to the sides, a simulation with -90° scan angle was run in CST. A 2D view of the directivity is shown in Figure 4(b). α here equals 90° .

The 3 dB beamwidth for this scan-angle was 10° , while the SLL was -17.5 dB, this is considered to be good enough for this application.

When the scan angle is $< \pm 60^\circ$ the elements in the front sector are used together with the

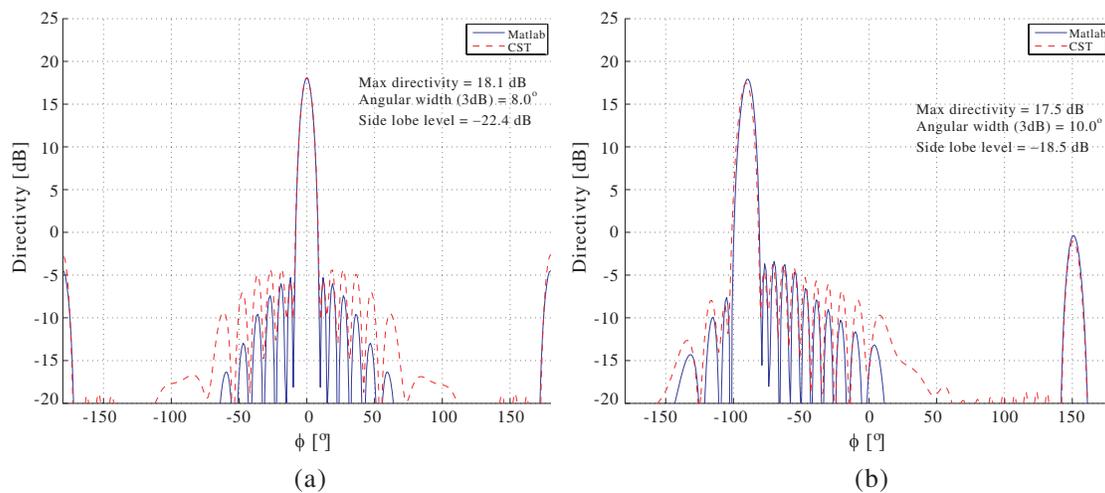


Figure 4: (a) Directivity for one sector with 0° scan angle. (b) Directivity for -90° scan angle.

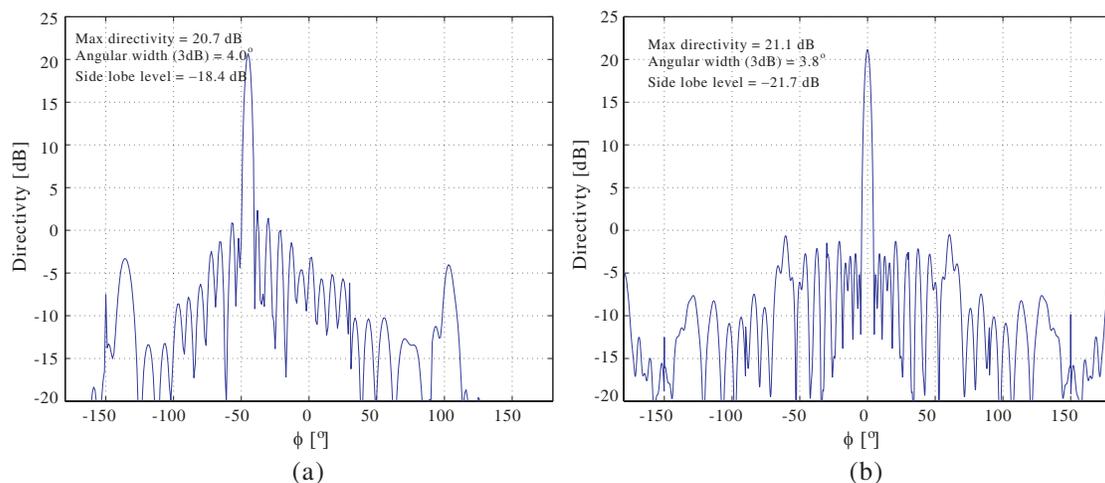


Figure 5: (a) Directivity for 45° scan angle. (b) Directivity for 0° scan angle.

elements in either the left or right sector (depending on which side the array is scanned to). As an example, a 2D view of the simulated directivity for -45° scan angle is shown in Figure 5(a). α equals 46.6° for the left sector (15° scan angle) and α equals -127.3° for the left sector (-45° scan angle).

From the figure it can be seen that the 3 dB beamwidth is 4° and the SLL is -18.4 dB. This is considered to be good enough to detect and track other vehicles that are at a collision course with the vehicle carrying the radar sensor.

For broadside direction, all elements are used. This means that the elements on the side sectors are scanned $\pm 60^\circ$ relative to their broadside directions. By using all elements the array looks electronically larger, and hence a narrower beamwidth can be achieved. A 2D view of the simulated directivity for 0° scan angle is shown in Figure 5(b). α equals $\pm 155.9^\circ$ for the side sectors ($\pm 60^\circ$ scan angle) while it is 0 for the front sector (broadside scan angle).

The 3 dB beamwidth for broadside direction is 3.8° and the SLL is -21.1 dB. Which is considered good enough to detect and track vehicles approaching the radar sensor in the same lane, the radar could also be used in adaptive cruise control applications.

3. CONCLUSION

In this paper, the antenna configuration for a new automotive radar safety system capable of wide angle detection was presented. The system uses a phased array antenna consisting of three 15 planar arrays mounted in three sectors. Each sector has 15 patch antenna elements. The sectors are placed at a 120° angle with each other.

The system capabilities for three different scenarios were analyzed; detection of onward traffic in the same lane as the vehicle carrying the radar sensor (broadside), detection of imminent front-side collisions and estimation of the time gap available when turning left at a road crossing. Hence, the radiation patterns for 0° , 45° and 90° view angle were simulated using CST and Matlab with the Array Design Toolbox.

The simulation results show that the 3 dB beamwidths and side lobe levels are good enough to provide proper detections for the three scenarios.

ACKNOWLEDGMENT

The authors like to thank Shimul Chandra Saha for his valuable contribution. The authors would also like to thank Norwegian Research Council for their financial support.

REFERENCES

1. Audi pre-safe, http://www.audi.ru/audi/master_public/en2/new_cars/a8/equipment/assistance_systems/audi_pre_sense.html, Accessed August 16, 2010.
2. Toyota news releases, <http://www.toyota.co.jp/en/news/09/0226.html>, Accessed September 1, 2010.
3. Stemme, G., N. Somjit, and J. Oberhammer, "Binary-coded 4.25-bit W-band monocrystalline-silicon MEMS multi-stage dielectric-block phase shifters," *IEEE/MTT-S Transactions on Microwave Theory and Technology*, Vol. 57, 2834–2840, 2009.
4. CST STUDIO SUITE 2010 TM, www.cst.com.
5. Tucker, N., "Phased array design toolbox V2.3 for matlab," Technical Report, 2009.

A Dual Polarization Bow-tie Slot Antenna for Broadband Communications

Chang-Ju Wu, I-Fong Chen, and Chia-Mei Peng

Graduate Institute of Computer and Communication, Jinwen University of Science and Technology
23154 No. 99, An-Chung Road, Hsin-Tien City, Taipei, Taiwan, R.O.C.

Abstract— A dual polarization bow-tie slot antenna for broadband communications is proposed in this paper. This antenna consists of two bow-tie slot antennas which is perpendicular to each other and with coplanar waveguide (CPW)-fed. The proposed antenna has a very simple antenna structure and wide impedance bandwidth ($\sim 400\%$ for $|S_{11}|$ and $|S_{22}| \geq 10$ dB) which can cover the 1.5 ~ 6 GHz frequency band for Global Positioning System (GPS, 1575 MHz) and dual ISM band (2.4 GHz and 5.8 GHz) applications. Good isolation between the two input ports ($|S_{21}| \geq 20$ dB) is also achieved at the operating band. The radiation pattern and efficiency of the proposed antenna are also measured, and radiation pattern data are compared with simulation results.

1. INTRODUCTION

With rapid progress in wireless communication systems, the demand to enhance the information accessibility and wideband utility has become major importance in wireless technology. An efficient way to increase the capability is the employment of polarization diversity, and thus the dual-polarized antennas have gained more and more popularities. In [1], a tri-polarization antenna was proposed, but isolation between some ports were not sufficient and were hence unacceptable in high-performance applications. Several papers have been published [2, 3] to improve isolation in similar antenna applications. However, the antennas of the bandwidth are still not wide enough for modern wireless communication systems. To have wider bandwidth and simple planar antenna configuration, bow-tie dipole and slot antennas are good candidates [4, 5]. To enhance the bandwidth of CPW-fed bow-tie slot antennas, some techniques have been proposed, including the use of a tapered metal stub to achieve impedance matching, the use of inductive coupling, and the adjustment of slot flare angle to enhance bandwidth.

To meet the specification of wide bandwidth, simplicity, and high isolation, a dual-polarization CPW-fed bow-tie slot antenna is proposed in this paper. The proposed design is brought by the embedded slot with a pair of stubs and combined two antennas at the scattered CPW feeding on left and right both sides, the horizontal and vertical polarization of slot antenna are excited by the 0° and 90° directional rotation of different CPW feed line. Measured performance of the developed antenna includes -10 dB reflection coefficient, with the bandwidths of 4.7 GHz (461%) and 4.74 GHz (476%) for the dual polarization. The isolation between two ports in the WLAN band is better than -23.1 dB (2.4 GHz) and -37.1 dB (5.8 GHz).

2. ANTENNA STRUCTURE AND DESIGN

Figure 1 shows the configuration of the proposed antenna. The overall dimensions of the antenna are 190×100 mm². The antenna is made of FR4 ($\epsilon_r = 4.4$, $\tan \delta = 0.02$), its thickness is 1.6 mm. A 44×38 mm² the embedded triangle slot with a pair of stubs, it combined two antennas and scattered feeding on left and right both sides. The CPW-fed line is designed to be 50Ω ($L_3 = 3.2$ mm), gap spacing ($S_1 = 0.2$ mm) and taper to the CPW of signal strip length should be 0.175λ ($W_2 = 35$ mm) that parameter is determined by the required band of the lowest frequency. Etched in the left side, the antenna serves as the vertical polarization radiation. The CPW is fed in port 1. In the right side, antenna rotates 90° to cover the horizontal polarization radiation. The CPW is fed in port 2.

Antenna dimensions of $L(= 0.5\lambda)$ and $W(= 0.4\lambda)$ are determined by the lowest frequency of desired band, $\theta = 40^\circ$ and $S_3 = 0.5$ mm are proposed for impedance matching. To ensure isolation between two antennas distance, $D(= 10$ mm) should be 0.05λ at the desired band of the lowest frequency. The final values of each parameter are listed in Table 1.

It is well known that a pair of stubs embedded triangle slot can support wonderful reflection coefficient and utilize the CPW-fed to achieve a much wider impedance bandwidth. Owing to the slot antenna has a directional radiation, it is also suitable to be designed into a dual-polarized antenna for horizontal and vertical polarization, as shown in Fig. 2 and Fig. 3, respectively. To

Table 1: Detailed dimensions of the proposed antenna.

| Parameter | L | L_1 | L_2 | L_3 | L_s | S_1 | S_2 |
|------------|-------|-------|-------|-------|-------|-------|------------|
| Value (mm) | 190 | 100 | 91.2 | 3.2 | 43.8 | 0.2 | 1.2 |
| Parameter | S_3 | W | W_1 | W_2 | W_s | D | θ |
| Value (mm) | 0.5 | 100 | 80 | 35 | 38.3 | 10 | 40° |

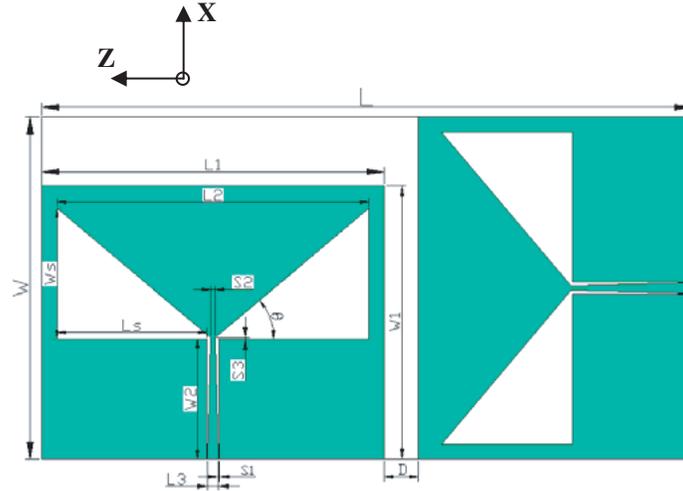


Figure 1: Geometry and dimensions of the proposed antenna.

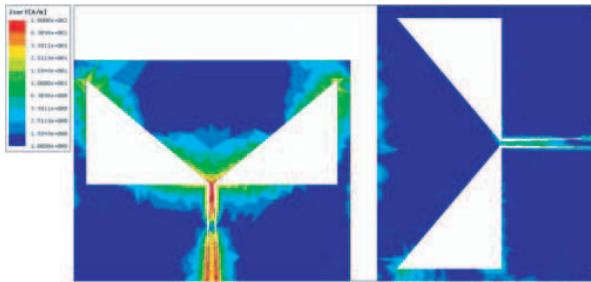


Figure 2: Electric field magnitude in slot and CPW: feeding in port-1 (2.4 GHz).

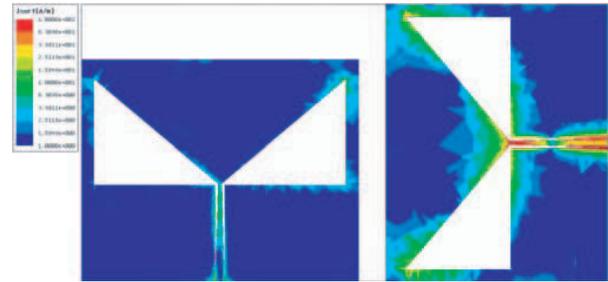


Figure 3: Electric field magnitude in slot and CPW: feeding in port-2 (2.4 GHz).

excite the two polarizations simultaneously, two antennas feeding by the perpendicular to each other must be used. As shown in Fig. 2 that feeding from port 1 ($50\ \Omega$ terminal load in port 2), a vertical polarization is excited in the left CPW structure. On the contrary, when feeding from port 2 (terminate $50\ \Omega$ load in port 1), a horizontal polarization is excited in the right CPW structure, as shown in Fig. 3. The antenna is an attractive candidate for multiple band antenna applications that require polarization diversity and also for some of the MIMO systems applications.

3. MEASUREMENT RESULTS

To validate the design, the S -parameters of the proposed antenna are simulated by using Ansoft High Frequency Structure Simulator (HFSS). The antenna has also been fabricated and measured. The left and right views of the antenna prototype are shown in Fig. 4.

Figure 5 shows the measured S -parameter of the proposed antenna (solid lines) compared to the simulated ones (dash lines). As shown in the Fig. 5, the frequencies of the dual polarization are from $1.5 \sim 6$ GHz. The bandwidths of -10 dB reflection coefficient are 4.7 GHz ($1.3\text{--}6$ GHz, 461%) and 4.74 GHz ($1.26\text{--}6$ GHz, 476%) for port 1 (vertical polarization) and port 2 (horizontal polarization), respectively. Throughout the WLAN frequency band ($2.4\text{--}2.484$ GHz) and ($5.15\text{--}5.825$ GHz), the

worst cases of reflection coefficient value for ports 1 and 2 are -13.4 dB and -18.1 dB. The isolation between two ports in the required band is lower than -20 dB.

When feeding from port 1 and 2, the radiation patterns of the propose antenna are shown in Figs. 6 to 11. For port 1, the vertical polarization is the dominant polarization. The 3-dB beamwidths are 80° and 60° in E -plane (yz plane) and H -plane (xy plane) at the 1575 MHz, the same as 60° in E - and H - at the 2400 MHz, 40° and 20° in E - and H - at the 5800 MHz, the cross polarization levels are lower than 5.51 and 0.43 dB in E -plane and H -plane at the 2400 MHz, respectively. For port 2, the horizontal polarization case, the 3-dB beamwidths are 100° and 120° in E -plane (yz plane) and H -plane (xy plane) at the 1575 MHz, 90° and 60° in E - and H - at the 2400 MHz, 50° and 40° in E - and H - at the 5800 MHz, the cross polarization levels are lower

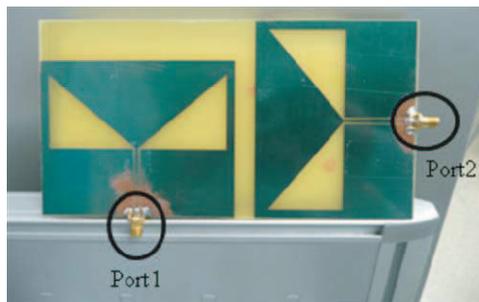


Figure 4: Photograph of the proposed antenna.

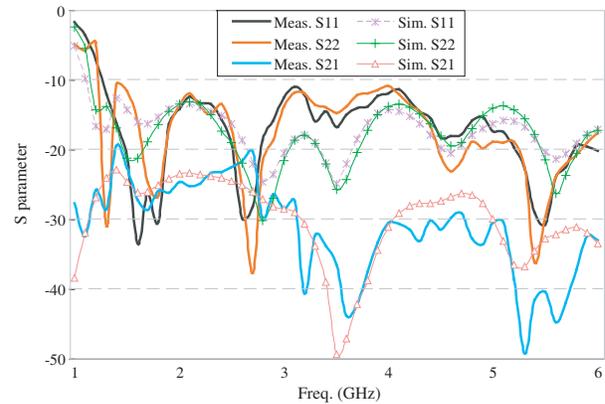


Figure 5: Measured and simulated S -parameter of proposed antenna.

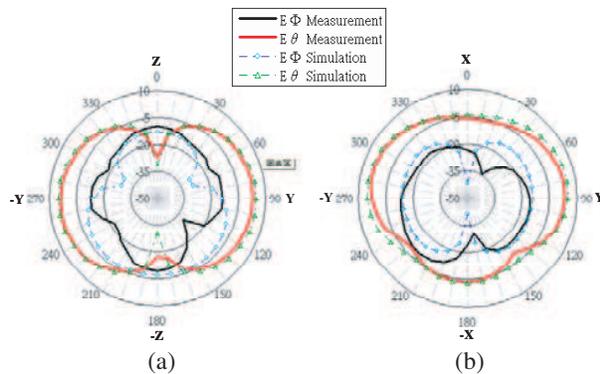


Figure 6: Measured and simulated radiation patterns when feeding from port 1 at 1.57 GHz: (a) yz plane, (b) xy plane.

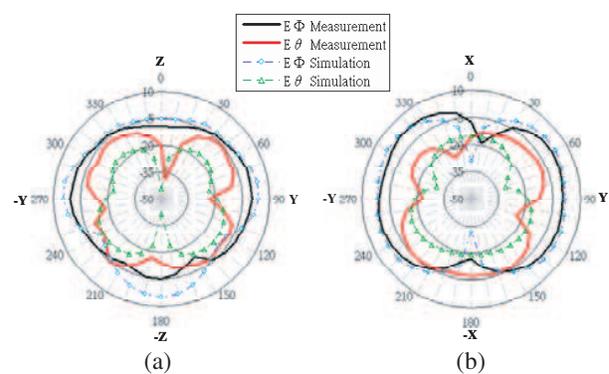


Figure 7: Measured and simulated radiation patterns when feeding from port 2 at 1.57 GHz: (a) yz plane, (b) xy plane.

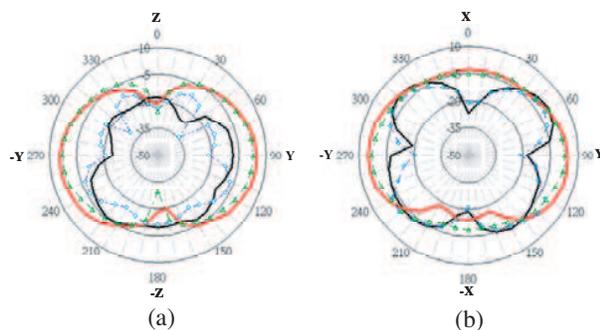


Figure 8: Measured and simulated radiation patterns when feeding from port 1 at 2.4 GHz: (a) yz plane, (b) xy plane.

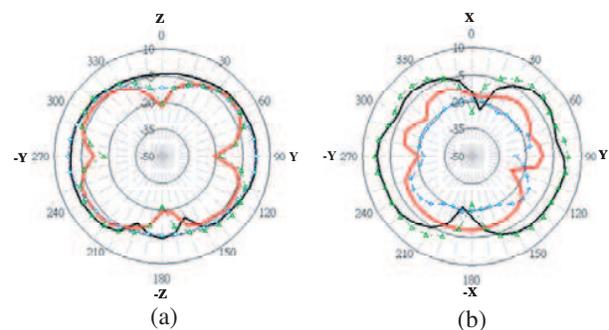


Figure 9: Measured and simulated radiation patterns when feeding from port 2 at 2.4 GHz: (a) yz plane, (b) xy plane.

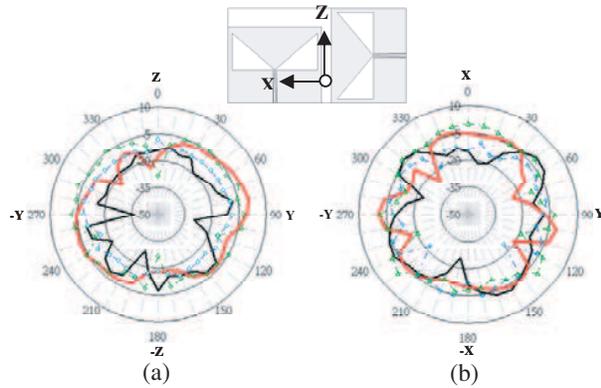


Figure 10: Measured and simulated radiation patterns when feeding from port 1 at 5.8 GHz: (a) yz plane, (b) xy plane.

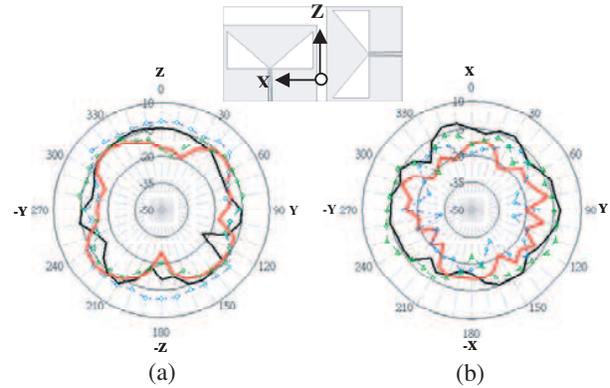


Figure 11: Measured and simulated radiation patterns when feeding from port 2 at 5.8 GHz: (a) yz plane, (b) xy plane.

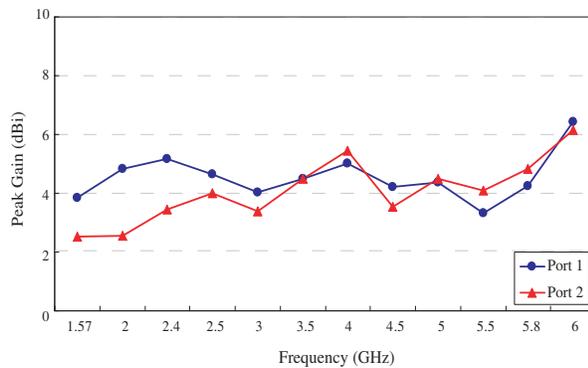


Figure 12: Measured radiation efficiency of the proposed antenna.

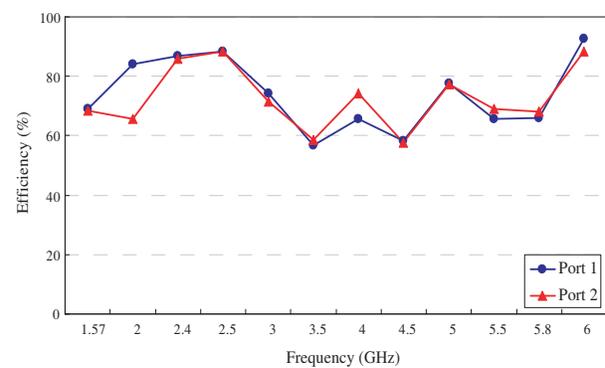


Figure 13: Measured gain of the proposed antenna.

than 0.17 and 9.08 dB in E -plane and H -plane at the 2400 MHz, respectively. From the radiation patterns, we can observe that the port 1 and port 2 of the yz plane and xy plane are almost corresponding to the maximum of each other at the three measured frequencies. According to the results, it can be proved that the two bow-tie slot antennas are placed 90 degrees of difference can achieve good isolation and dual polarization.

The radiation efficiency and gain of the proposed antenna are also measured. The results are shown in Figs. 12 and 13, respectively. In the WLAN band of 2.4 and 5.8 GHz, the efficiency are 86.7% and 65.9% for ports 1 and 85.9% and 68.1% for ports 2; the peak gain are 5.16 and 3.45 dBi for port 1 and 4.25 and 4.82 dBi for port 2.

4. CONCLUSION

A Dual Polarization coplanar waveguide (CPW)-fed bow-tie slot antenna for broadband communications has been proposed and tested. The antenna structure is brought by the embedded slot with a pair of stubs and combined two antennas at the scattered CPW feeding on both left and right sides, it can be obtained a much wider impedance bandwidth, dual-polarization and high isolation. Measured performance of the developed antenna includes -10 dB reflection coefficient, with the bandwidths of 4.7 GHz (461%) and 4.74 GHz (476%) for the two polarizations. The isolation between two ports in the required band is lower than -20 dB. From simulated and measured radiation patterns, the signals received by the two ports are proved to be uncorrelated. The radiation efficiency is better than 56% and 57%, and the peak gain is better than 2.5 and 3.3 dBi. Although this antenna was designed for frequency 1.5 GHz \sim 6 GHz applications, this design concept can be extended to other frequency bands of interest.

REFERENCES

1. Zhong, H., Z. Zhang, W. Chen, Z. Feng, and M. F. Iskander, "A tripolarization antenna fed by proximity coupling and probe," *IEEE Antennas Wireless Propag. Lett.*, Vol. 8, 465–467, 2009.

2. Lee, C.-H., S.-Y. Chen, and P. Hsu, “Isosceles triangular slot antenna for broadband dual polarization applications,” *IEEE Trans. Antennas Propag.*, Vol. 57, No. 10, 3347–3351, Oct. 2009.
3. Li, Y., Z. Zhang, W. Chen, Z. Feng, and M. F. Iskander, “A dual-polarization slot antenna using a compact CPW feeding structure,” *IEEE Antennas Wireless Propag. Lett.*, Vol. 9, 191–194, 2010.
4. Chen, I.-F., C.-M. Peng, C.-C. Hung, and H.-C. Chen, “Printed modified bow-tie dipole antenna for multi-band applications,” *IEICE Trans. on Communications*, Vol. E92-B, No. 4, 1404–1405. Apr. 2009.
5. Huang, C. Y. and D. Y. Lin, “CPW-fed bow-tie slot antenna for ultrawideband communications,” *Electron. Lett.*, Vol. 42, No. 19, 1073–1074, 2006.

Broadband Fractal Circular-monopole Antenna

Wen-Yi Tsai, I-Fong Chen, Chia-Mei Peng, Pei-Cheng Hu,
Hsu-Hung Tung, and Hsuan-Chi Lin

Department of Electronic Engineering, Institute of Computer and Communication Engineering
Jinwen University of Science and Technology, No. 99, An-Chung Rd., Hsien Tien, Taipei, Taiwan, R.O.C.

Abstract— In this paper, a novel broadband Fractal circular-monopole antenna is presented. This antenna consists of printed circular iteration with two iterating level and ground-plane with radius 25 mm, making it easy to making it easy to combine directional, high gain and wide bandwidth. A prototype is designed to operate at 1.5 GHz–5.86 GHz, the measured 10 dB bandwidth is nearly 1 : 2 at the center frequency of 3.775 GHz. Experimental results are shown to verify the validity of theoretical work. Fractal monopole antenna is formed with hollow of circle, which featuring of multi-frequency bands and wide bands.

1. INTRODUCTION

The Sierpinski triangle, also called the Sierpinski gasket or the Sierpinski Sieve, is a fractal named after the Polish mathematician Waclaw Sierpinski who described it in year 1916 [1]. Fractal antenna engineering is the field, which utilizes fractal geometries for antenna design. It has become one of the growing field of antenna engineering due to its particular characteristics and advantages over conventional antenna design. An algorithm of the Sierpinski triangle as follows: Any triangle in a plane and canonical Sierpinski triangle uses an equilateral triangle with a base parallel to the horizontal axis (first image). Reduce the size of the triangle to 1/2 height and 1/2 width, make three copies, and position the three shrunken triangles so that each triangle touches the other two triangles at a corner. Note the emergence of the central hole — because the three shrunken triangles between them can cover only 3/4 of the area of the original. (Central Holes are an important feature and rule of Sierpinski’s triangle). Repeat step 2 with each of the smaller triangles (image 3 and so on). In Fig. 1 was shown iterate level.

For the typical Sierpinski gasket configuration, the patch dimensions from 88.9 mm × 88.9 mm to 800 mm ground plane, a triangle generator with a scale fractal ratio of 1/2 is slotted in the triangle patch to form the monopole antenna [2]. For the modified Sierpinski gasket with 88.9 mm × 88.9 mm patch and 25 cm × 25 cm ground plane, a circular generator with a non-constant fractal ratio depending on the Descartes circle theorem is slotted in the triangle patch to form the monopole antenna [3]. Most of the fractal geometries have the characteristics mentioned as above, with various periods that form of fractal structure dimension and self-similarity. Utilization the Sierpinski gasket or carpet geometries characteristic of fractals can be improved and used to design both type of monopole and dipole antennas [4–8]. It has become one of the growing fields of antenna engineering due to its advantages over conventional antenna design. The results suggest a method for perturbing the Sierpinski structure in such way as to control the position of multiple bands where is necessary.

Modified Sierpinski category fractal of circular monopole antenna are plotted and inward iteration in Fig. 2. The antenna overall height of h_n is 54 mm. The Sierpinski fractal has the geometrical

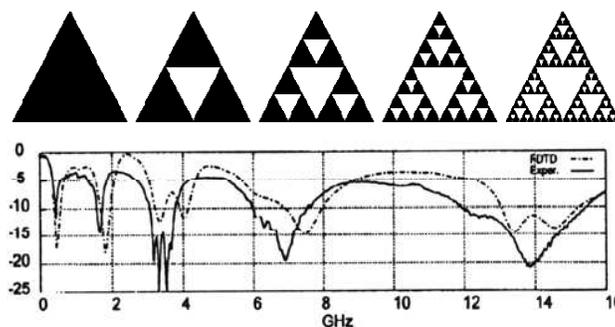


Figure 1: Sierpinski triangle iteration, and iterate level by 0~5.

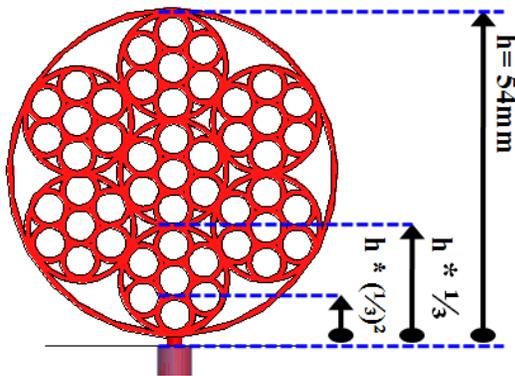


Figure 2: Circular-shaped Sierpinski geometric antennas.

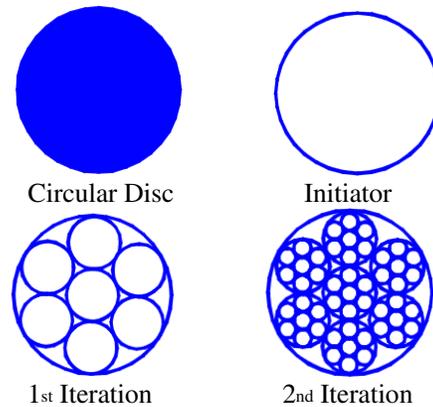


Figure 3: Circular-shaped Sierpinski geometric antennas, disc, initiator and iterations.

scale factor,

$$\delta = \frac{h_{n+1}}{h_n} = 0.33$$

where h represents the height of sub-generation and n is a natural number represents the number of iterations proposed. The following sections will describe the antenna design, implementation, and the performances of the proposed antenna.

2. BASIC GUIDELINES

The planar fractal monopole antenna is based on the Sierpinski carpet concept and to modified, the Sierpinski fractal has generated two iterations, use the decomposition algorithm by circular, and compare these with integrators and initiator type. Firstly, base on generally circular disc to generate hollow of circular, this initiator has a circumference of $D\pi$ and 1 mm trace width. In the decomposition algorithm, a geometrical ratio of circle is taken and joining the midpoints of union of the circles central, reduce the circle to $1/3$ diameter and canonical Sierpinski carpet algorithm to fill the circle with twice mathematics equals $1/9$ diameter by Sec. iteration, is shown in Fig. 3 the iterate level. The geometrical structure and dimensions of the proposed monopole type antenna is printed on FR4 substrate and is simulated by using FEM based electromagnetic simulator, Ansoft HFSS.

The Sierpinski gasket is deterministic fractal which generalizes the objects into two dimensions. A circle in plane will be divided into seven smaller congruent circle of fractal self similar geometrically. The process continuing as long as the limitation of the subdivided is not too small.

However, using Sierpinski gasket equation and the fractal counting Dimension $D \cong 1.77$, the formula defined by:

N_n : Number of self similar objects covers the original object.

L_n : Reduce its ratio for the length

$$\begin{aligned} Db &= \lim_{n \rightarrow \infty} \left(\frac{\ln N_n}{\ln L_n} \right) = \lim_{n \rightarrow \infty} \frac{\log 7^n}{\log 3^n} \\ &= \lim_{n \rightarrow \infty} \frac{n \cdot \log(7)}{n \cdot \log(3)} \approx 1.77 \end{aligned}$$

Relevant Circular-shaped Sierpinski Fractal antenna information of proposed are summarized include outside circle of the initiator with 27 mm radius, the radius are 9 mm of 1st iteration follow initiator, and the last or 2nd iteration radius are 3 mm. the dimensions of ground plane radius is 25 mm, were shown in Fig. 4.

The antenna has been fabricated on 0.8 mm thick FR4 substrate, mounted on a circle ground plane 25 mm and fed down through the ground with a SMA connector structure, the measured three-axis position define as shown in Fig. 5.

We've designed the circular-shaped of monopole antenna based on the Sierpinski gasket geometry are mode in the antenna in order to improve the width-band and radiation pattern, the circular

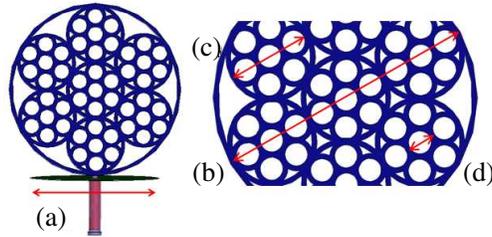


Figure 4: Geometry of the proposed circular-shaped fractal antenna. Dimensions of ground plane radius (a): 25 mm; PCB material: FR-4 with $\epsilon_r = 4$ used; Circular of diameter (b): 54 mm; Circular of diameter (c): 18 mm; Circular of diameter (d): 6 mm.

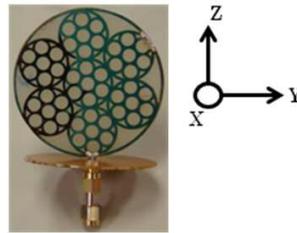


Figure 5: Fabricated Sierpinski monopole antenna.

fractal was printed on FR4 and mounted over a radius of 25 mm circle ground plane. The structure was fed through a 1.5 mm diameter, 50 Ω coaxial probe with a SMA connector on the bottom side of the circle plane. The geometrical structure and dimensions of the proposed monopole circular-shaped has constructed through three iterations in this particular case and circular disc, Sierpinski gasket and carpet is known as a shelf similar fractals.

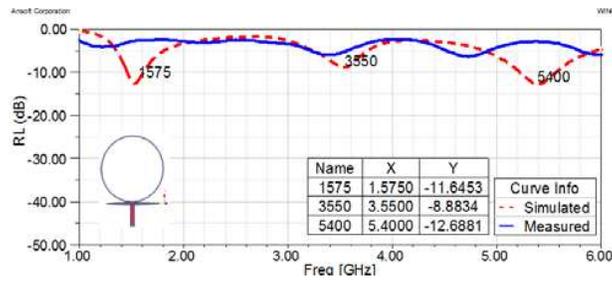
Antennas are designed and simulated by using Ansoft HFSS and a prototype is fabricated, the antenna return loss and VSWR is measured, using the Agilent 8753ES Network Analyzer to measure and simulate it responses in 1-6GHz frequency range is demonstrated in Fig. 6. The result we had from simulated and measured test is similar while we observed. It is known from Fig. 6, that to perform width-band characteristics and the second iteration antenna structure covering the frequency range from 1.5 GHz to 6 GHz is best among them. A actual return loss and VSWR is better than -10 dB and < 2 in 1.5 GHz to 6 GHz in 2nd iteration mode.

3. HELPFUL HINTS

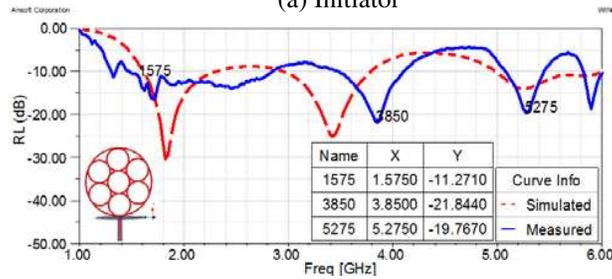
The antenna is built in a monopole configuration over an 25 mm radius of circle ground plane, and antenna structured height 54 mm with FR4 substrate (thickness = 0.8 mm, $\epsilon_r = 4.0$). A fifty-ohm coaxial probe directly feeds the driven SMA connector. The resonances the circular-shaped of fractal monopole antenna including 1st and 2nd type of return loss, are shown in Fig. 7(a). In 1st of iteration return loss, the bandwidth and resonances are slightly matching impedance. As the 2nd of iteration, the resonances are very remarkable increased and rich multi-band responses occur in the S_{11} spectrum. The measured gain of the antenna at the 1st and 2nd iterations band is given in Fig. 7(b). The table shows the fractal antenna peak gain on 1st and 2nd iterations frequencies resonance point, respectively. The measured antenna gain against frequency and offers a peak gain of 1st and 2nd iterations, it was apparent peak gain that the 2nd iteration more than before models on the same scope.

This 2nd of iteration was appeared multi- and width range from 1.5 GHz~6 GHz for all ISM band requirement applications. The radiation pattern of iterations antenna were measured inside a 3D anechoic chamber.

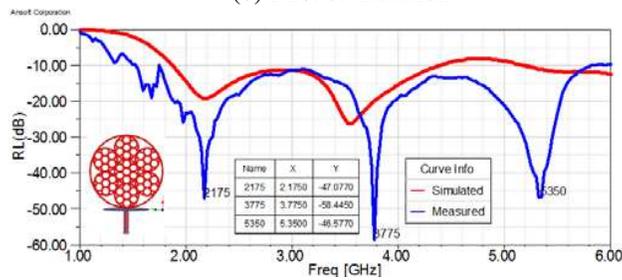
The main cuts measured radiation patterns of the antenna at 1st and 2nd iterations are shown in Fig. 8. All the popular bands measured of antenna radiation patterns appear at the same position that shows the radiation patterns of iterations antennas in the X/Y/Z planes, respectively. The patterns are observed to be nearly Omni-directional in the H-plane, and resonance frequencies span are illustrated in Fig. 8.



(a) Initiator

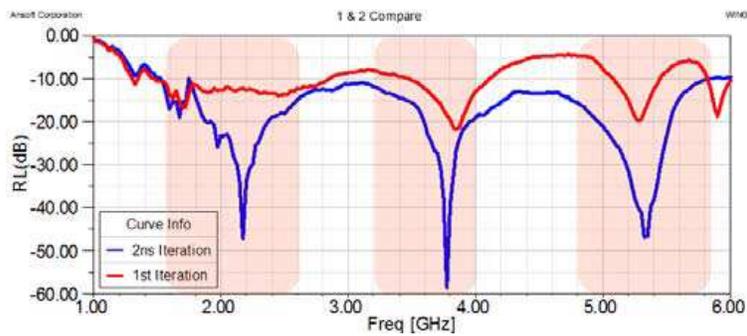


(b) First of iteration



(c) Second of iteration

Figure 6: Simulated and measured of return loss in 1 ~ 6 GHz range and As defines simulated (Red color) and measured (Blue color).



(a)

| Type of Iteration | Resonant Frequency (GHz) | Return Loss(dB) | Gain (dBi) |
|-------------------|--------------------------|-----------------|------------|
| 1st Iteration | 2.175 | -14.03 | 3.8 |
| | 3.85 | -21.84 | 6.1 |
| | 5.275 | -19.40 | 4.7 |
| 2nd Iteration | 2.175 | -33.13 | 4.5 |
| | 3.775 | -27.28 | 6.3 |
| | 5.35 | -41.44 | 6.8 |

(b)

Figure 7: (a) Comparing 1st and 2nd iterations with Return loss and gain. (b) Comparing 1st and 2nd iterations with Return loss and gain.

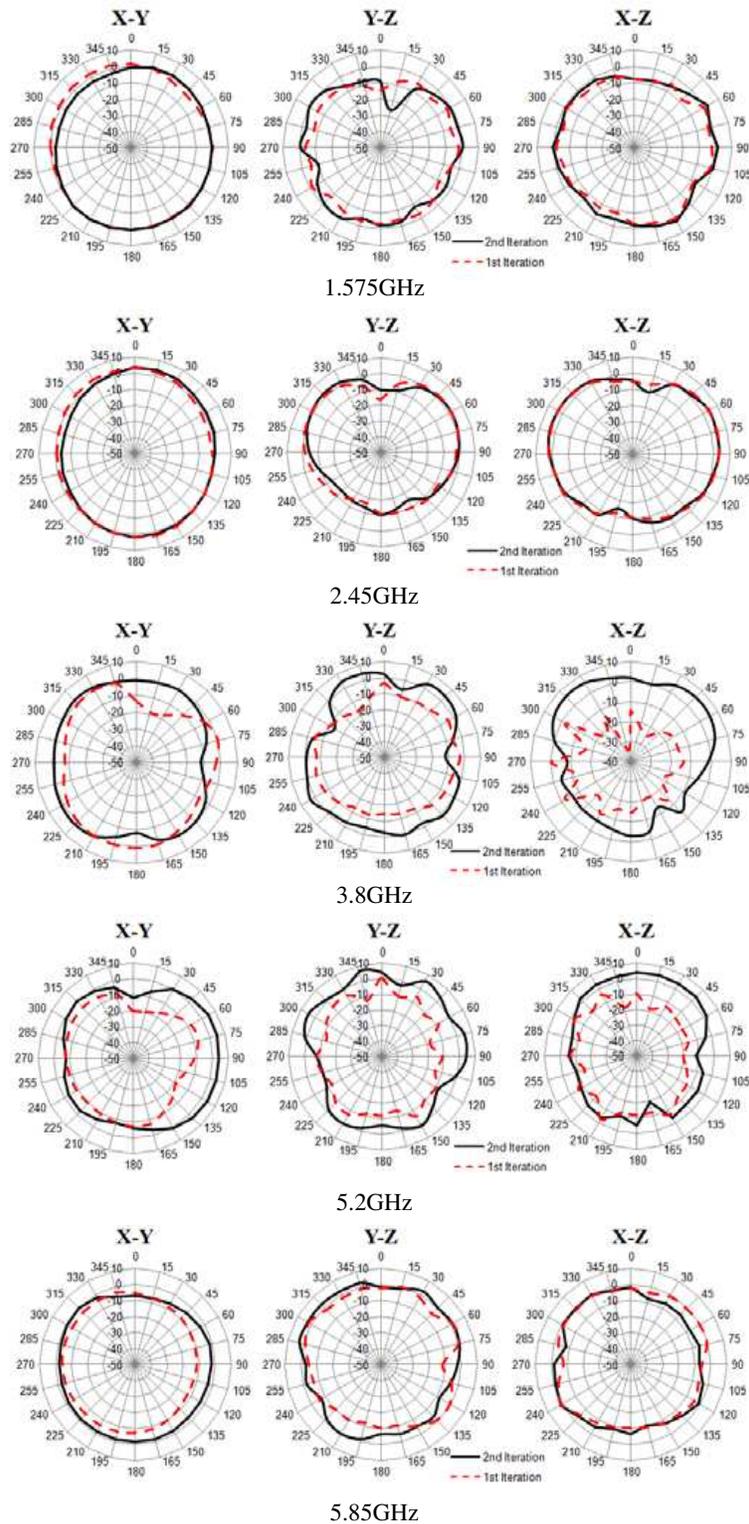


Figure 8: Measured peak gain antenna gain for the proposed antenna.

4. CONCLUSION

Both experimental and numerical results on the Sierpinski antenna have been presented. In an actual example, as the Sierpinski fractal geometry into circular structure, use HFSS simulate software to test and analysis, it shows the result is match to fractal characteristic.

All of them describe a multiband behavior of fractal antenna. This behavior is consistent from the input return loss and gain; moreover radiation patterns planes of view. The same scale factor

existing among similar structures in the fractal circular-shape. It can be summarized that the self-similarity properties of the fractal structure are translated into its electromagnetic behavior. The current density distributions have a similar and vary in complicated among bands as well. Such distributions allows flexibility in matching multi- and width band operations in which a larger frequency required, such as FemtoCell and UMTS base station application including LTE, UMTS, GPS L1, WIFI, and WiMax. The circular monopole type is based on fractal structure and refer to the Sierpinski gasket self-similarity algorithm, a prototype of the design is successfully implemented with close agreement between measurement and simulation. The fractal geometry and overall size can be effectively utilized ID or Logo surface for integrating with other components in IT products.

ACKNOWLEDGMENT

The authors are appreciated the reviewers for their valuable comments, which is significantly help to this paper, and their colleague, Dr. I-Fong Chen and Dr. Chia-Mei Peng for his assistance and advice during the revision of this paper.

REFERENCES

1. Sierpiński, W., “Sur une courbe cantorienne qui contient une image biunivoque et continue de toute courbe donnée,” *C. R. hebd. Séanc Acad. Sci.*, Vol. 162, 629–632, Paris, 1916.
2. Puente-Baliarda, C., J. Romeu, R. Pous, and A. Cardama, “On the behavior of the Sierpinski multiband fractal antenna,” *IEEE Trans. Antenna Propag.*, Vol. 46, 517–524, 1998.
3. Liu, J.-C., C.-Y. Wu, D.-C. Chang, and C.-Y. Liu, “Relationship between Sierpinski gasket and Apollonian packing monopole antennas,” *Electronics Letters*, 847–848, 2006.
4. Ding, M., R. Jin, J. Geng, Q. Wu, and W. Wang, “Design of a CPW-fed ultra wideband crown circular fractal antenna,” *IEEE Antennas and Propagation Society International Symposium*, 2049–2052, 2006.
5. Hee, T. W., P. S. Hall, and K. Y. Liew, “Wideband stacked Sierpinski carpet dipole antenna,” *IEEE Antennas and Propagation Society International Symposium*, 242–245, 2003.
6. Vsetula, P. and Z. Raida, “Sierpinski conical monopole antennas,” *15th International Conference on Microwave Techniques (COMITE)*, 55–58, 2010.
7. Yao, N. and X.-W. Shi, “Analysis of the multiband behavior on Sierpinski carpet fractal antennas,” *Microwave Conference Proceedings 2005, APMC 2005, Asia-Pacific Conference Proceedings*, IEEE Conferences, 2005.
8. Chen, I.-F. and C. M. Peng, “Microstrip-fed dual U-shaped printed monopole antenna for dual band wireless communication application,” *IET Journals*, 955–956, 2003.
9. 3rd Antennas version.
10. Chen, I.-F., *Antennas Design Hand Book*, The Computer and Communication System of Jinwen University of Science & Technology.

A Monopole Antenna with CPW-fed for Digital Video Broadcasting Applications

Mau-Phon Houn¹, Yu-Jen Chou¹, Ding-Bing Lin², and I-Tseng Tang³

¹Institute of Microelectronics, National Cheng Kung University, Tainan, Taiwan

²Institute of Computer and Communication Engineering, National Taipei University of Technology
Taipei, Taiwan

³Department of Greenergy, National University of Tainan, Tainan, Taiwan

Abstract— This paper is to design the miniaturized Digital video broadcasting (DVB) antenna. The miniaturized consideration must be paid attention in DVB antenna design, because many consumer products and communication facilities are integration with the DVB. The monopole antenna is designed in the paper. The antenna characteristics including return loss and radiation patterns were analyzed and discussed. In this thesis, a monopole antenna conformed standard (470 ~ 862 MHz) is proposed. The obtained results show that the impedance bandwidth, determined by -6 dB return loss. The radiation E -plane patterns are displayed with traditional 8-shaped patterns.

1. INTRODUCTION

Recently, the trend of personal wireless communications is growing thin, light, and relatively miniature. Consequently, the area of antennas should be ease integration. Printed coplanar waveguide (CPW) fed antenna have attracted much attention due to its ease of integration with monolithic microwave integrated circuits (MMIC). Numerous designs of antenna with CPW-fed have been reported, including the use of a center-fed slot antenna [1], or sleeve monopole antenna [2], a Bow-tie antenna [3], a loop monopole antenna [4], a bent folded monopole antenna [5], and dual-frequency monopole antenna [6].

Moreover, monopole antennas have found widespread applications in wireless mobile communication systems. In the paper, the proposed a novel design of uniplanar monopole antenna fed by a CPW line. The proposed antenna is specifically easy in manufacturing, because it is single dielectric and single metal layer. The content of this paper is focused on the analysis and investigation based on the DVB band of a novel DVB antenna. We apply many solutions for those considerations. We proposed a novel notion of printed monopole antenna for DVB (470 ~ 862 GHz) applications. The radiation E -plane patterns are displayed with traditional 8-shaped patterns.

2. ANTENNA DESIGN

Antenna structure and design: Fig. 1 shows the geometry and configuration of the proposed antenna. The proposed antenna is to fabricate on a low-cast FR-4 substrate with dielectric constant $\epsilon_r = 4.4$, loss tangent = 0.02, and thickness $h = 1.6$ mm. The antenna for overall size is $50 \times 50 \times 1.6$ mm. The proposed planar coplanar waveguide fed antenna is easy to be integrated for low manufacturing cost. The CPW-fed line is fixed at 4.6 mm and the distance of the gap between the line and the symmetric ground planes is fixed at 0.4 mm to achieve 50Ω , characteristics impedance. The optimized dimensions for the proposed antenna are as follows: $L = 50$ mm, $W = 50$ mm, $L_1 = 39$ mm, $W_1 = 1$ mm, $L_2 = 10$ mm, $W_2 = 10$ mm, $L_3 = 1$ mm, $W_3 = 3$ mm, $L_4 = 31$ mm, $W_4 = 45$ mm, $L_5 = 10$ mm, $W_5 = 36.4$ mm.

Generally speaking, the design notion of the resonant frequency is the adjusting the length of the monopole antenna to be about quarter-wavelength at the desired resonant frequency making the input impedance singular. The resonant frequency can be postulated as:

$$f = \frac{C}{4L\sqrt{\epsilon_{re}}} \quad (1)$$

where L is the total length of the monopole antenna, ϵ_{re} is the effective dielectric constant, and C is the speed of the light. We can take (1) into account in obtaining the total length of the monopole antenna at very beginning of the design and then adjust the geometry for the final design.

3. SIMULATION RESULTS AND DISCUSSION

Figure 2 shows that the stub length L_4 mainly influences the impedance at higher frequencies. The simulated return loss against frequency for various length $L_4 = 29$, 31 and 33 mm, where the optimal bandwidth is obtained when the length is about 33 mm. The input impedance is well

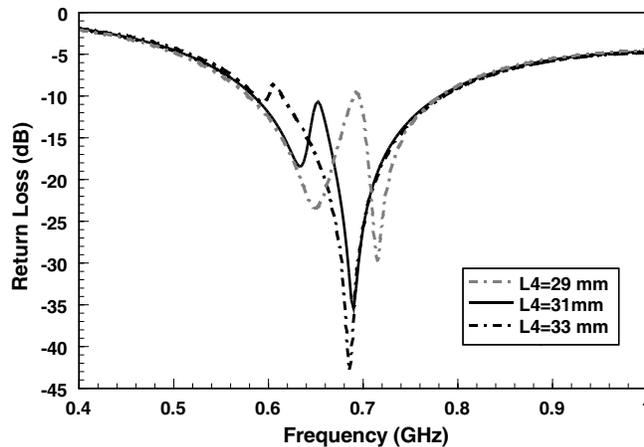
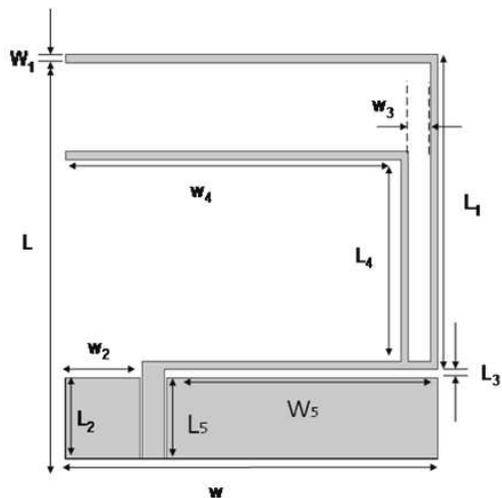


Figure 1: Geometry and configuration of the proposed antenna.

Figure 2: Effects of the stub length L_4 on the return loss.

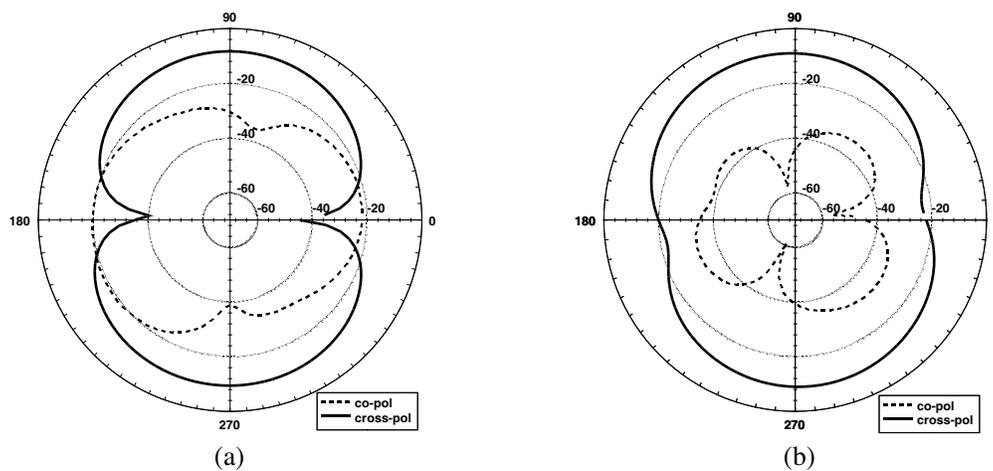


Figure 3: The first resonant mode at 620 MHz. (a) E -plane (yz -plane), and (b) H -plane (xy -plane).

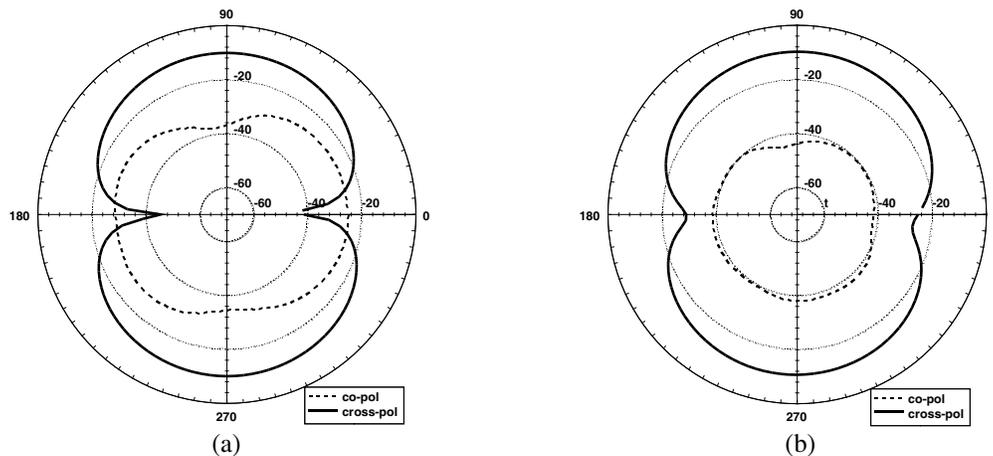


Figure 4: The second resonant mode at 680 MHz. (a) E -plane (yz -plane), and (b) H -plane (xy -plane).

matched as the 6-dB return loss bandwidth covers the entire DVB band (470 ~ 862 MHz). The radiation patterns of the antennas were also simulation at 620 and 680 MHz are shown in Figs. 3(a), (b) and Figs. 4(a), (b). From the radiation pattern of this mode is like a small monopole oriented in the z -axis leading to a bidirectional pattern in the E -plane (yz -plane) and H -plane (xy -plane)

4. CONCLUSION

This paper is to design the miniaturized Digital video broadcasting (DVB) antenna. Simple designs of CPW-fed monopole antennas for achieving wideband operation have been successfully design and proposed. For a fixed resonant frequency, adjust series length to a monopole antenna for achieving wideband application. This technique is also used in combination with other wideband techniques to further decrease the size of the antenna structure. Finally, the proposed antennas have a simple geometry and successfully implemented.

REFERENCES

1. Chiou, J.-Y., J.-Y. Sze, and K.-L. Wong, "A broad-band CPW-fed strip-loaded square slot antenna," *IEEE Trans. Antenna Propagat.*, Vol. 51, No. 4, April 2003.
2. Chen, H.-D. and H.-M. Chen, "Planar CPW-fed sleeve monopole antenna for ultra-wideband operation," *IEE Proc. Microw. Antenna Propagat.*, Vol. 152, No. 6, December 2005.
3. Durgun, A. C., M. S. Reese, C. A. Balanis, C. R. Birtcher, D. R. Allee, and S. Venugopal, "Flexible bow-tie antennas," *IEEE Antennas and Propagation Society International Symposium*, 2010.
4. Song, C. T. P., P. S. Hall, H. Ghafouri-Shiraz, and D. Wake, "Multi-circular loop monopole antenna," *Electron. Lett.*, Vol. 36, 391–393, 2000.
5. Lee, E., P. S. Hall, and P. Gardner, "Dual band folded monopole/loop antenna for terrestrial communication system," *Electron. Lett.*, Vol. 36, 1990–1991, 2000.
6. Chen, H.-D. and H.-T. Chen, "A CPW-fed dual-frequency monopole antenna," *IEEE Trans. Antenna Propagat.*, Vol. 52, No. 4, April 2004.

Reflection Characteristics of Microstrip Base on Finite Element Method

Qi Liu¹, Hui Huang^{1,2}, and Xin Wang¹

¹School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China

²State Key Laboratory of Millimeter Waves, Nanjing 210096, China

Abstract— A detailed study on the reflection characteristics of microstrip line with right angle bend cut by 30° on the Printed Circuit Board (PCB) based on Finite Element Method (FEM) is presented. We discuss two cases, respectively. Firstly, keeping the microstrip width and permittivity as constant, we focus on reflection characteristics of microstrip with different thickness of the dielectric material. Secondly, maintaining the microstrip width and dielectric material thickness, we calculate reflection characteristics of microstrip with different permittivity. Finally, we calculate the reflection parameters of a special case above with different cut size. It is found that there exists an optimum cut size that minimizes reflection in high frequency band.

1. INTRODUCTION

With the development of technology, the operation frequency of the circuits increases rapidly, and causes negative effects. Microstrip, as one of the main transmission lines for the components interconnecting on PCB, is more and more important for us to study [1, 2]. And the microstrip corner structure is an important element which affects the transmission characteristics. At the microstrip corner, accumulation effect of charge, which makes the capacitance bigger, leads to the discontinuous characteristic impedance, more reflection and less transmission [3].

References [4] and [5] analyzed the different structures of microstrip line with right angle bend based on the experimental and numerical methods. The conclusion is that the transmission characteristics of the right angle bend cut by 45° is better than that with no cut. However, they didn't discuss the bend cut by other angles. In this paper, we calculate the reflection parameters of the bend cut by 15°, 30°, 45°, 60° and the bend with no cut based on the Finite Element Method. It is found that the 30° cut angle has better reflection characteristics than the others during specific frequency.

2. THEORY AND METHODS

Usually we hope that the microstrip works in the quasi-TEM mode of the weak dispersion to avoid the signal distortion. The increasing of operation frequency may cause high-order modes. The cutoff frequency depends on the section size of the microstrip and substrate materials in high-order modes [6]. When the shortest wave length is given, the microstrip width w , dielectric material thickness h and the permittivity of the medium ϵ_r should meet the condition $w < \frac{\lambda_{\min}}{2\sqrt{\epsilon_r}}$; $h < \frac{\lambda_{\min}}{2\sqrt{\epsilon_r}}$; $h < \frac{\lambda_{\min}}{4\sqrt{\epsilon_r-1}}$, which ensures the microstrip transport in main quasi-TEM mode. However, the condition that the phase velocity of surface wave mode is the same as that of the quasi-TEM mode will result in a strong coupling between surface wave mode and the quasi-TEM mode. So the operation frequency should be less than f_{TE} and f_{TM} to avoid the coupling. While if the operation frequency is too large, f_{TE} and f_{TM} should be improved by changing the parameters of the design.

In our case, the maximum working frequency is 20 GHz. According to the analysis above, we have $w < 2.37154$ mm, $h < 1.25$ mm. The microstrip structure is shown in Figure 1, where θ is the cut angle, x is the length of the cutting edge, and d is the maximum value of x . We define m as x/d , and the amount of cut size is the quantity m . where $m = 0$ indicates a bend with no cut and $m = 1$ indicates a cut large enough to barely separate the two sections of microstrip from the bend.

3. SIMULATION AND ANALYSIS

Figure 2 illustrates that during the frequency band from 11 GHz to 20 GHz, the reflective parameter S_{11} reduces a lot when the cut angle $\theta = 30^\circ$. It means that transmission characteristics of 30° cut angle are much better than the others in this high frequency band. So it is meaningful to study the reflection characteristics of 30° cut angle.

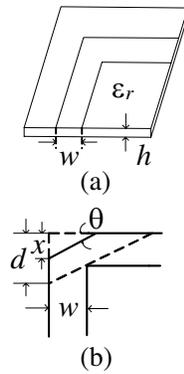


Figure 1: (a) Microstrip geometry. (b) Top view and detail of the bend.

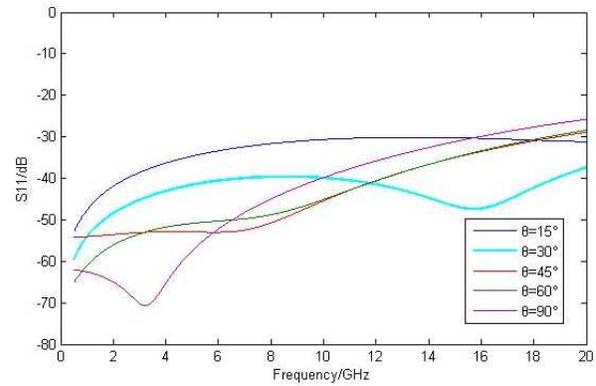


Figure 2: S_{11} versus frequency for case: $w = 0.1$ mm, $h = 0.2$ mm, $\epsilon_r = 10$, $x = 0.1$ mm.

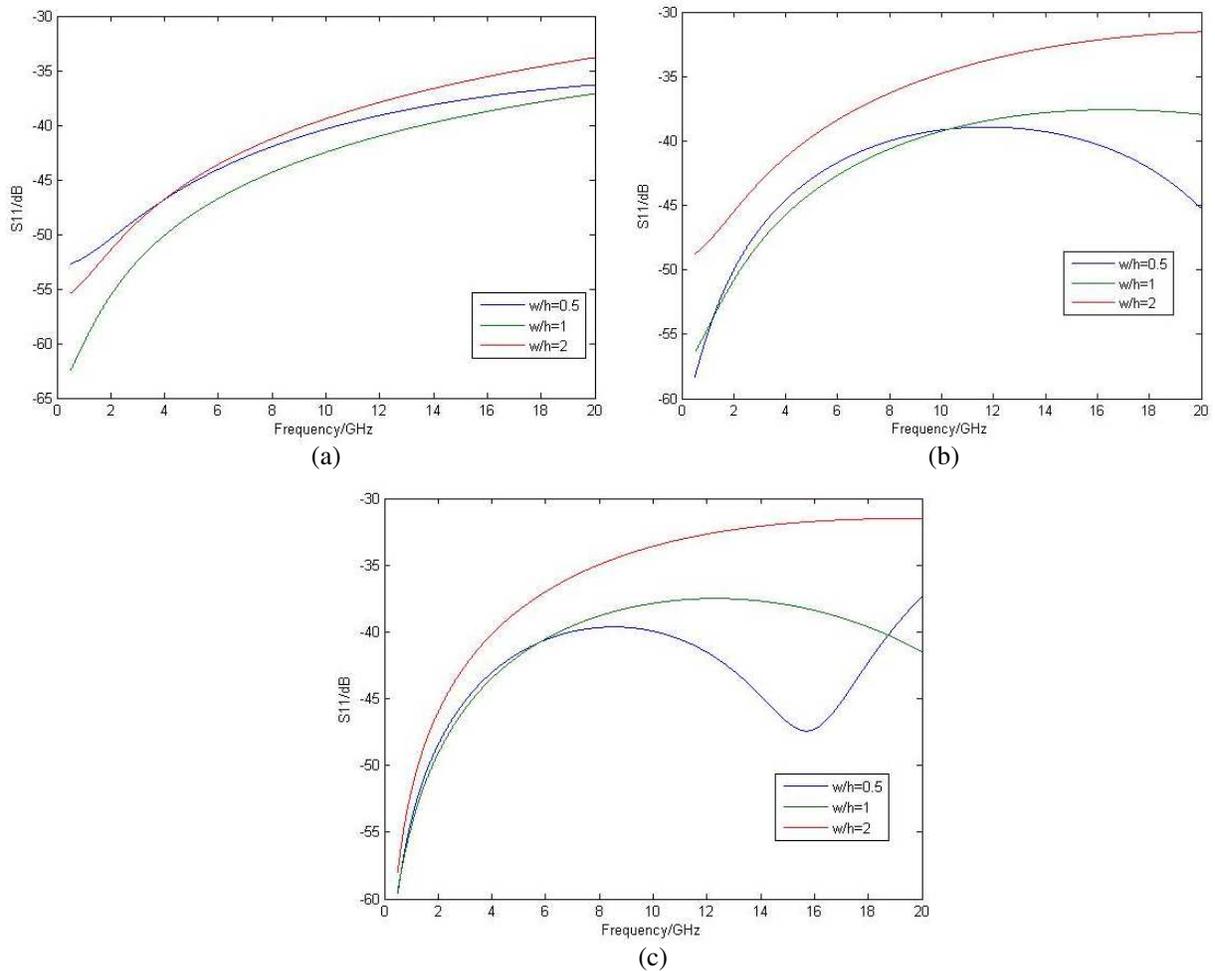


Figure 3: S_{11} versus width-to-height ratio w/h for $w = 0.1$ mm and $x = 0.1$ mm. (a) $\epsilon_r = 2$. (b) $\epsilon_r = 5.5$. (c) $\epsilon_r = 10$.

3.1. Reflection Characteristics of 30° with Different Width-to-height Ratio w/h

We show our simulation results in Figure 3. The parameters we used are $w = 0.1$ mm and $x = 0.1$ mm. It reveals that when the frequency is lower than 10 GHz, value of S_{11} with $w/h = 1$ is less than those with other width-to-height ratio. However, the value of w/h affects the reflection characteristics slightly in low frequency band. It is found that a change from $w/h = 1$ to $w/h = 0.5$ results in a sharp decrease of S_{11} in high frequency band, especially when $\epsilon_r = 10$. In theory,

characteristic impedance of uniform microstrip

$$Z_0 = (v_p C_0)^{-1}, v_p = c \cdot (\epsilon_e)^{-1/2} \quad (1)$$

where C_0 is capacitance of transmission line in per unit length, v_p is the phase velocity, c is the speed of light, and ϵ_e , affected by the geometry and material, is the equivalent permittivity of microstrip. The signal that transmits along the microstrip with characteristic impedance as a constant is not reflected. For the microstrip with right-angle bend, larger the width of the bend line is, the more accumulation effects of charge result. Thus the capacitance is bigger, and the characteristic impedance is smaller. That is to say, bend cut by 30° will reduce the bend area and weaken the accumulation effects of charge, which can decrease the discontinuity of characteristic impedance. However, the characteristic impedance is not affected by only one variable. In the frequency band we discuss, S_{11} changes irregularly versus width-to-height ratio w/h , and it decreases rapidly only in high frequency band.

3.2. Reflection Characteristics of 30° with Different ϵ_r

We keep the w and x as above, and do the simulation with different permittivity ϵ_r . The results are shown in Figure 4. It shows that the transmission characteristic with $\epsilon_r = 2$ is better than the others when the frequency is lower than 10GHz . However, S_{11} has a down trend when permittivity becomes larger and w/h becomes smaller in high frequency band. The result that the minimum value of S_{11} appears when $\epsilon_r = 10$ and $w/h = 0.5$ is consistent with the conclusion of Figure 3.

3.3. Discussing the Optimum Cut Size of 30°

As we know, the capacitance C can be calculated as

$$C = \frac{\epsilon_r S}{h}, \quad (2)$$

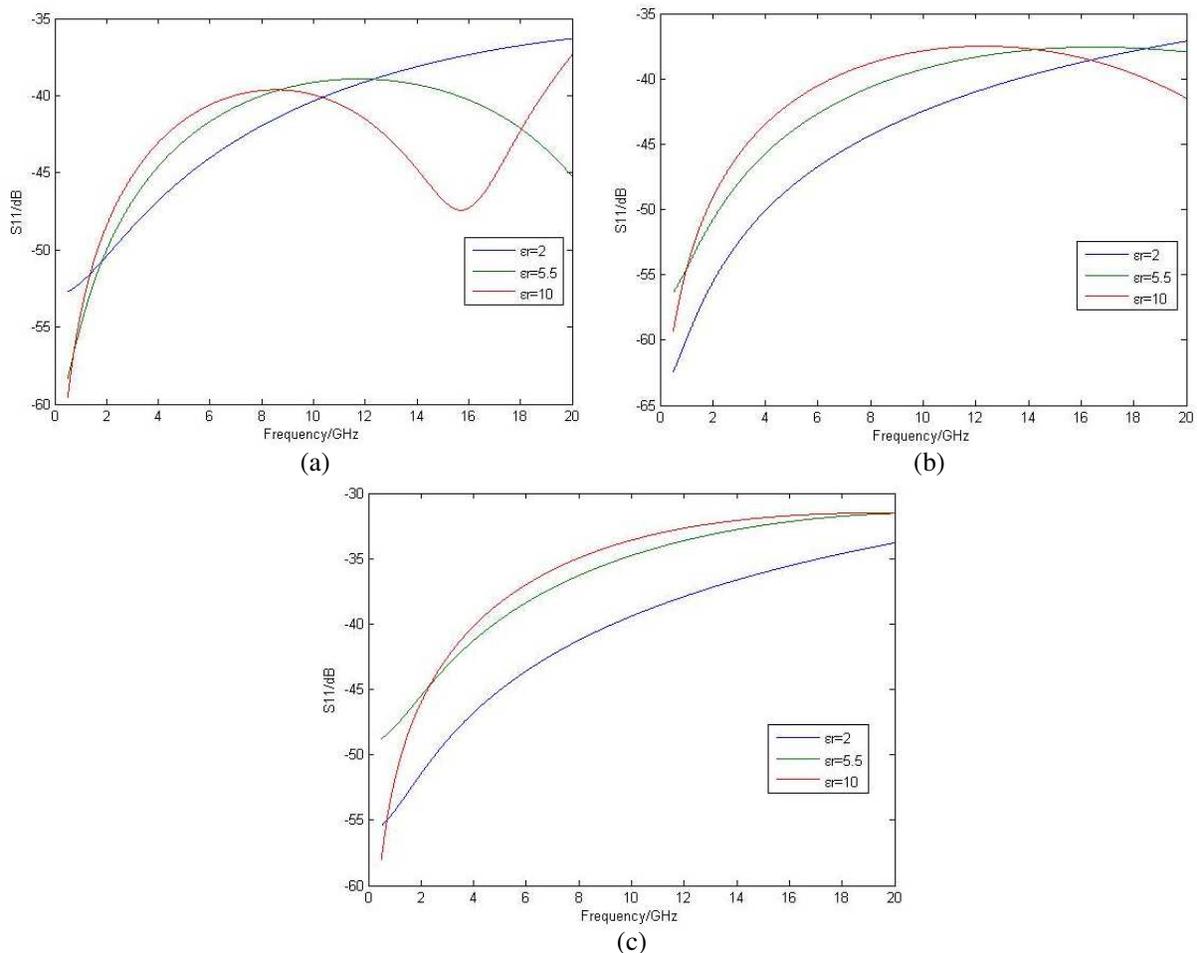


Figure 4: S_{11} versus ϵ_r for $w = 0.1\text{ mm}$ and $x = 0.1\text{ mm}$. (a) $w/h = 0.5$. (b) $w/h = 1$. (c) $w/h = 2$.

where ε_r is permittivity of the dielectric material, s is equivalent area of the discontinuity microstrip, h is the thickness of the dielectric material. It can be drawn from formulation (2) that the value of C is affected only by s when h and ε_r are constant. So cut length x in Figure 1 is the key element which affect the characteristic impedance. In Figure 3 and Figure 4, bend cut by 30° with $w/h = 0.5$ and $\varepsilon_r = 10$ has better transmission characteristic in high frequency band. Based on this case, we will discuss different cut length x to obtain the lowest S_{11} .

From Figures 5(a) and (b) we can see that the minimum value of S_{11} in different frequency band depends on the value of cut length x . The minimum value of S_{11} in high frequency band we focus on appears when $x = 0.10$ mm. Comparing the minimum value of S_{11} in low frequency band in Figure 5(a) with the curve of S_{11} with bend cut by 45° in Figure 2, the conclusion is that the reflection characteristics of bend cut by 30° is not worse than that with bend cut by 45° . This shows that under the condition we discuss, the optimum cut length x that meets the transmission performance can be found both in high and low frequencies respectively.

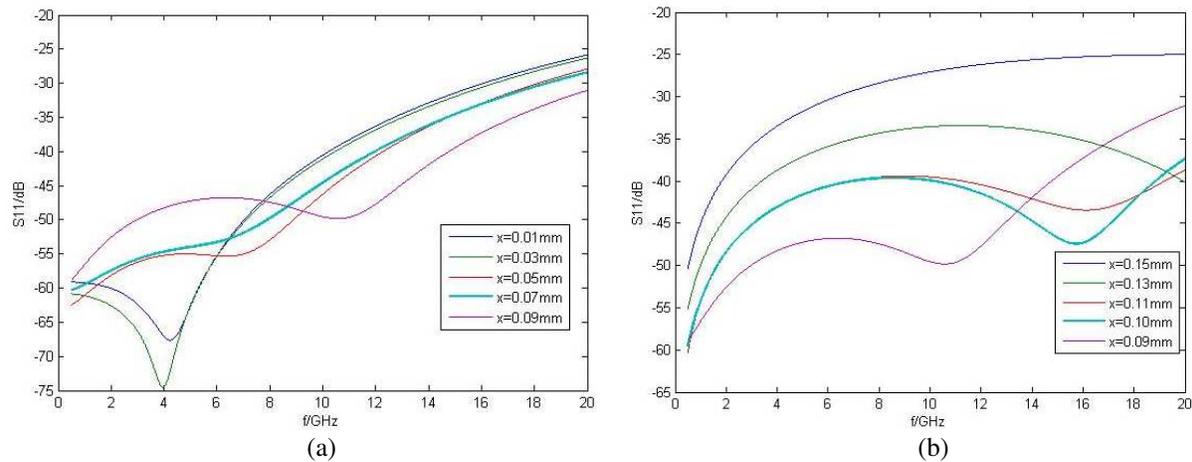


Figure 5: S_{11} versus x for $w/h = 0.5$ and $\varepsilon_r = 10$. (a) x is no more than 0.09 mm. (b) x is no less than 0.09 mm.

4. CONCLUSION

Reflection characteristics of bend cut by 30° are largely affected by w/h and ε_r in high frequency band. The optimum cut size changes with frequency band. Thus, it is helpful to choose the suitable microstrip structure to obtain the optimum transmission characteristics before designing a given high frequency circuit.

ACKNOWLEDGMENT

This work is sponsored by State Key Laboratory of Millimeter Waves under Contract K201012.

REFERENCES

1. Montrose, M. I., *EMC and the PCB-Design, Theory and Layout Made Simple*, IEEE, Piscataway, 1998
2. Bogatin, E., *Signal Integrity-simplified*, Prentice Hall, Upper Saddle River, 2004
3. Wang, K. and J. Wang, "Study on the transmission and reflection properties of the microstrip line bend," *Journal of Microwaves*, Vol. 22, No. 3, 33–35, 2006.
4. Douvilie, R. J. P. and D. S. James, "Experiment study of symmetric microstrip bends and their compensation," *IEEE Trans. on Microwave Theory Tech.*, Vol. 26, No. 3, 175–182, 1978.
5. Lu, H., J. An, Y. Zhao, et al., "Frequency domain analysis for right angle corners," *Journal of Xi'an Jiaotong University*, Vol. 12, No. 41, 1451–1454, 2007.
6. Fan, S., J. Li, et al., *Microwave Technique, Microwave Circuit and Antenna*, 71–72, China Machine Press, Bei Jing, 2008.

Circularly Polarized Rectangular Microstrip Antenna Using Ring Slots on the Ground Plane

Jeung-Keun Park, Dang-Oh Kim, and Che-Young Kim
 School of Electronics Engineering, Kyungpook National University
 Sankyuk-dong, Buk-gu, Daegu 702-701, South Korea

Abstract— In this paper, a single-feed circularly polarized rectangular microstrip antenna with a slotted ground plane is proposed. To obtain a circularly polarized operation, two ring slots are etched on the ground plane. These slots are optimally placed below the diagonal corner of radiating rectangular patch. Additionally, with the optimum position of the ring slots, either of right-handed CP (RHCP) or left-handed CP (LHCP) operation is worked out. By embedding the two ring slots on the ground plane and generating electric field from the patch, a magnetic resonance can be established in the operating bandwidth of the proposed antenna. Using this phenomenon, two orthogonal nearly degenerate resonant modes for circular polarization (CP) radiation can be excited. In order to demonstrate this phenomenon, surface current distributions on the ground plane for proposed antenna and reference antenna are simulated by using the simulation software SEMCAD X and HFSS [1, 2]. The structure and size of the proposed antenna are the same as those of the reference antenna without ring slots on the ground plane. The proposed antenna shows more improvements on the directivity, gain, and radiation efficiency than the reference antenna. Comparisons between the simulated and measured results are given, and in particular the role of two ring slots is fully explained during the descriptive illustrations.

1. INTRODUCTION

Many studies using diverse types of slits in the microstrip patch antenna for CP operation have been performed [3, 4]. A single feed microstrip patch antennas using various types of slots have been investigated to produce CP operation [5–8]. In this paper, to radiate CP operation, two ring slots etched on the ground plane is proposed, which are located below the diagonal corner of radiating rectangular patch. By means of optimizing the length and the position of the ring slots, the radiation of the antenna is to be a CP. In order to prove this phenomenon, surface current distributions on the ground plane for proposed antenna and reference antenna are simulated and discussed. The size and structure of the proposed antenna are the same as those of the reference antenna without ring slots on ground plane. The objectives are to compare and investigate the performances such as polarization, return loss, peak gain and radiation efficiency; when ring slots are optimally placed. Details of the antennas design are described and the simulated and measurement results of the prototype antennas are presented.

2. ANTENNA DESIGN

Figure 1(a) shows the geometry of proposed single-feed circularly polarized rectangular microstrip antenna. This is printed on a microwave substrate (Rogers 4003) of thickness t and relative permittivity 3.55. The size of the nearly square rectangular radiating patch of length P_L and width P_W are also chosen. The ground plane is chosen to be square and has dimensions $L \times W$. A quarter wavelength transformer is used for impedance matching between the patch and the feeding port. In Figure 1(a), ring slots having equal length of R and width R_W are etched on square ground plane. These slots are placed below the diagonal corner of rectangular radiating patch on the ground plane. By placing ring slots, left-hand CP operation can be obtained. Right-hand CP operation is also obtained by ring slots at the exactly opposite location. P is the distance from the corner of rectangular patch to ring slot. All of the structure parameter of the proposed antenna is introduced in Table 1.

The geometry of the reference antenna is presented in Figure 1(b). The basic dimension of this antenna is the same as those of the proposed antenna, except that the ring slots are embedded on the ground plane. Comparisons between the simulated and measurement results are given, and in particular the role of two ring slots is fully explained during the descriptive illustrations in Chapter 3.

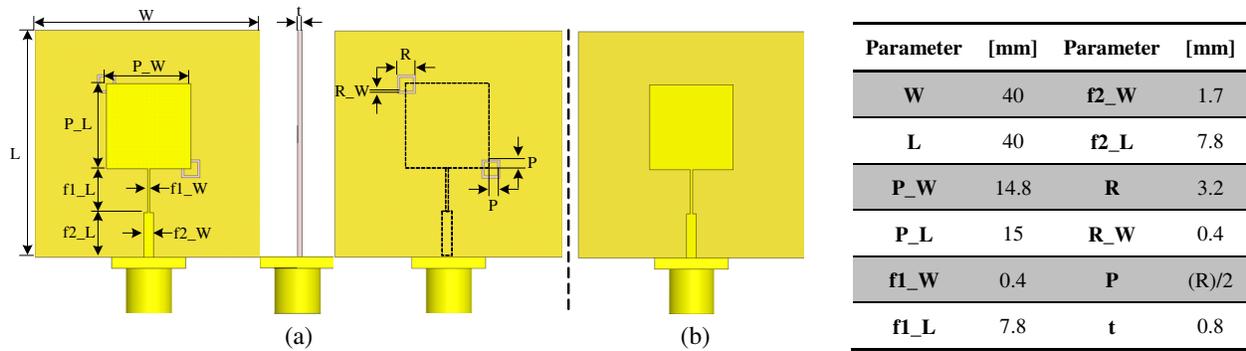


Figure 1: Geometry of the proposed antenna (a) and reference antenna (b).

Table 1: The size of the proposed antenna.

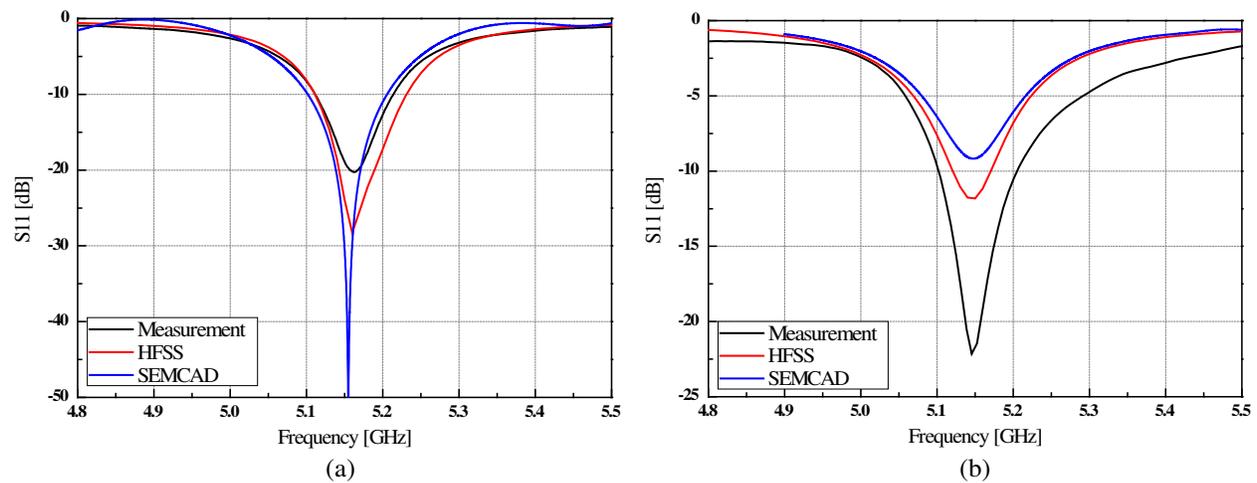


Figure 2: Measured and simulated return loss diagram of the proposed antenna (a) and reference antenna (b).

Table 2: Detail of measurement and simulation results.

| Type | Fc [GHz] | AR[dB] | Gain[dB] | Dir.[dBi] | Eff.[%] |
|-------------------------------|----------|-----------------|----------|-----------|---------|
| Pro. (Simulation_HFSS) | 5.16 | 0.6@5.18[GHz] | 6.174 | 7.177 | 79.39 |
| Pro. (Measurement) | 5.15 | 2.72@5.15[GHz] | 6.03 | 6.7 | 90 |
| Ref. (Simulation_HFSS) | 5.14 | 46@5.14[GHz] | 5.796 | 6.909 | 77.38 |
| Ref. (Measurement) | 5.15 | 18.81@5.15[GHz] | 5.02 | 6.97 | 72 |

3. EXPERIMENTAL RESULTS AND DISCUSSION

Measurement and simulation results for the proposed single-feed circularly polarized rectangular microstrip antenna with a slotted ground plane and reference antenna are obtained. The return loss diagram of the proposed antenna and reference antenna as depicted in Figures 2(a) and 2(b) respectively demonstrated a measured 10 dB impedance bandwidth of 100 MHz centered at 5.15 GHz. By these figures, good impedance matching is achieved.

The axial ratio diagram of the proposed antenna is shown in Figure 3. The result shows the CP operating bandwidth of 50 MHz centered at 5.15 GHz. The corresponding simulated and measured data are listed in Table 2 for comparison. It is of note that the simulated results in general agree with the measured data. The measured radiation efficiency of the proposed antenna is found to be 90%, which is around 18% higher than those of the reference antenna. This characteristic may be the reason for the enhanced gain. The peak gain for the proposed antenna measured at 6.03 dBic is around 1 dB higher than the reference antenna. Furthermore, the simulated directivity of the proposed antenna increased.

The simulated vector and magnitude of surface current distributions on the ground plane are shown in Figure 4 and Figure 5. Careful examination of the electromagnetic fields in the configuration of proposed antenna between radiating patch and the etched ring slots on ground plane suggests that the two ring slots resonate. Hence, optimizing the length and the position of the ring slots, they can be made to resonate like effective magnetic loops. This phenomenon is shown in Figure 4. By this means, two orthogonal nearly degenerate resonant modes for CP radiation can be excited. Hence, it is clearly seen that the proposed antenna radiated LHCP and the reference antenna operated linear polarization as shown in Figure 5.

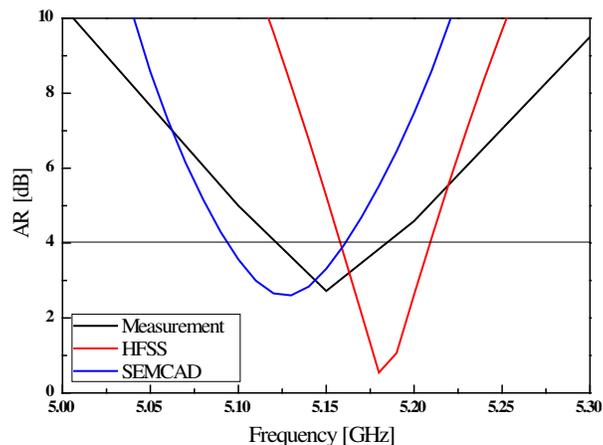


Figure 3: Measured and simulated axial ratio diagram of the proposed antenna.

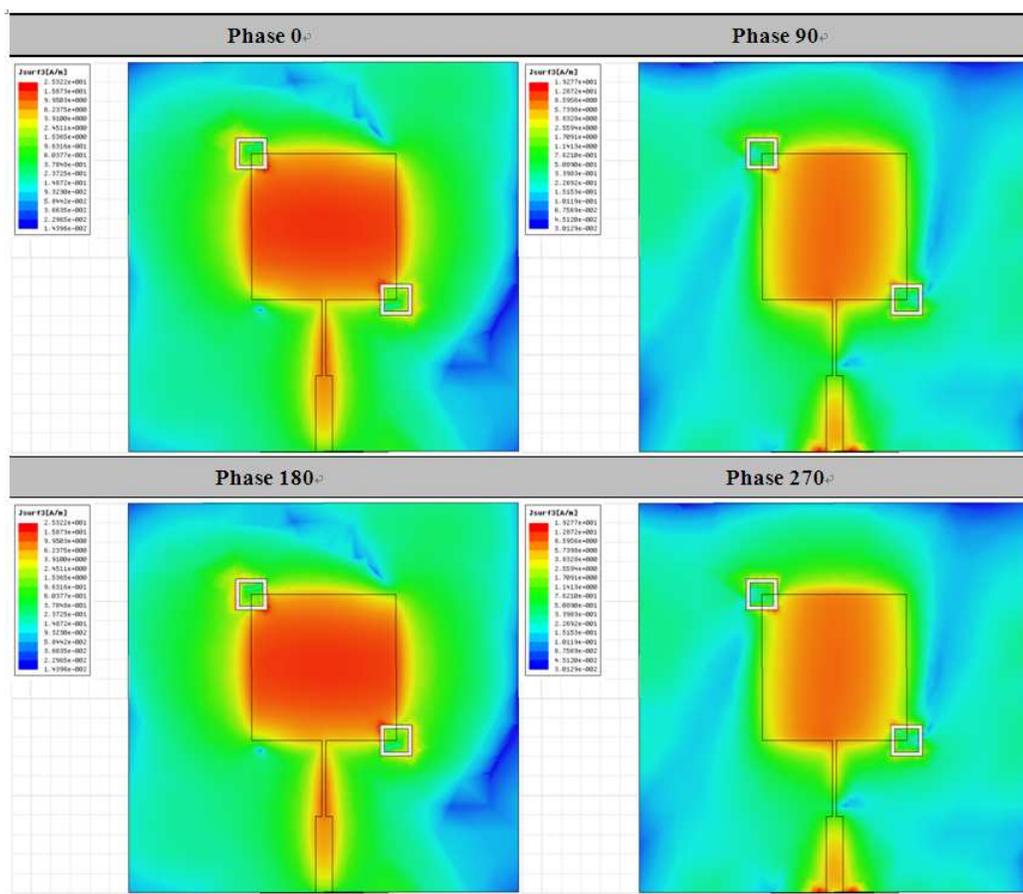


Figure 4: Simulated magnitude of surface current distributions on the ground plane at 5.15 GHz.

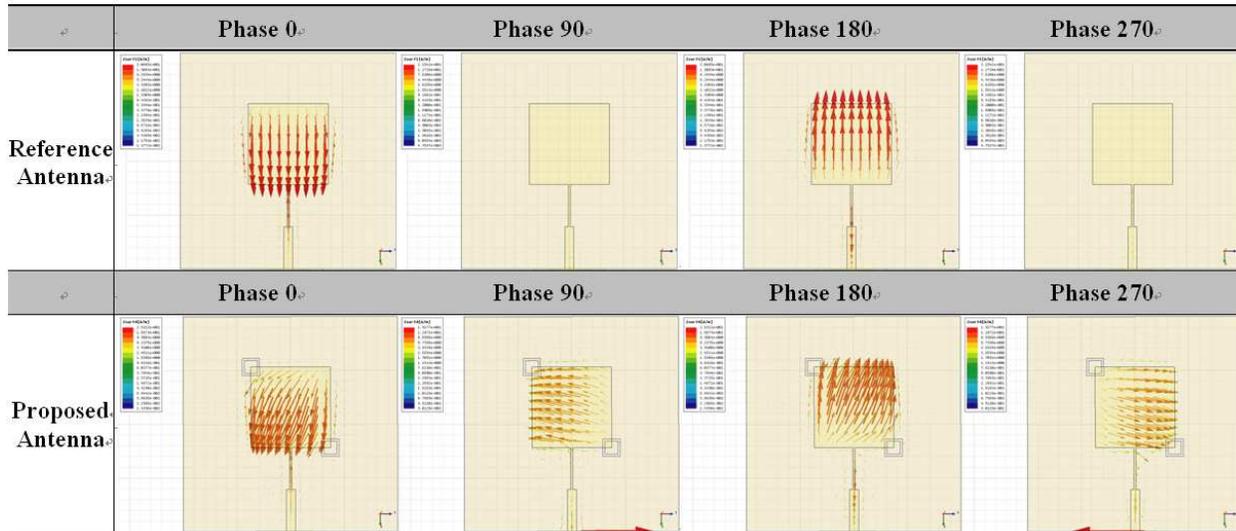


Figure 5: Simulated vector of surface current distributions on the ground plane at 5.14 GHz.

4. CONCLUSIONS

A single-feed circularly polarized rectangular microstrip antenna with a slotted ground plane is proposed and experimentally studied. Further, this paper described the antenna structures and provided the simulated and measurement results of the prototype antennas. By comparing results, the proposed antenna observes more increase on the radiation efficiency and gain than the reference antenna. With the proposed ring slot arrangement, the CP radiation for the proposed CP design can easily be obtained. Also, the role of two ring slots is explained by the simulated surface current distributions on the ground plane.

In the future, studies will investigate for the reduction of the size. Furthermore, the proposed antenna will be applied to the base station of the wireless LAN system.

REFERENCES

1. www.speag.com.
2. www.ansoft.com.
3. Sim, C. Y. D. and T. Y. Han, "GPS antenna design with slotted ground plane," *Microwave and Optical Technology Letters*, Vol. 50, No. 3, 818–821, 2008.
4. Chen, H. D., "Compact circularly polarized microstrip antenna with slotted ground plane," *Electronics Letters*, Vol. 38, No. 13, 616–617, 2002.
5. Chen, W. S., C. C. Huang, and K. L. Wong, "Microstrip line fed printed shorted ring slot antennas for circular polarization," *Microwave and Optical Technology Letters*, Vol. 31, No. 2, 137–140, 2001.
6. Baik, J. W., S. J. Kim, and Y. S. Kim, "Circularly polarized microstrip antenna using asymmetric ring sector slots embedded on the ground plane," *Microwave and Optical Technology Letters*, Vol. 49, No. 3, 602–604, 2007.
7. Kuo, J. S. and G. B. Hsieh, "Gain enhancement of a circularly polarized equilateral triangular microstrip antenna with a slotted ground plane," *IEEE Transactions on Antennas and Propagation*, Vol. 51, No. 7, 1652–1656, 2003.
8. Kumar, C. K. and S. N. Sinha "A patch loaded ring slot antenna for wide band circular polarization," *Journal of Electromagnetic Waves and Applications*, Vol. 23, No. 17–18, 2409–2419, 2009.

Resonance of Rectangular Microstrip Patch over Ground Plane with Rectangular Aperture in the Presence of High-permittivity Dielectric Layer below the Aperture

S. Benkouda and T. Fortaki

Electronics Department, University of Batna, Algeria

Abstract— In this communication, Galerkin’s method in the Fourier transform domain is applied to the determination of the resonant frequencies of rectangular microstrip patch over ground plane with rectangular aperture when there is a high-permittivity dielectric layer below the aperture. The numerical calculations show that the results obtained for the resonant frequencies of the patches with apertures in contact with a high-permittivity dielectric layer not only substantially differ from the results without apertures, but they also appreciably differ from the results obtained with apertures in contact with air. Unlike microstrip patch resonators without apertures, the resonant frequency does not vary monotonically with increasing substrate thickness.

1. INTRODUCTION

In recent years, microstrip patches have aroused great interest in both theoretical research and engineering applications due to their low profile, light weight, conformal structure, and ease in fabrication and integration with solid state devices. Because of the inherent narrow bandwidth of microstrip patch resonators, many attempts have been made to improve their bandwidth characteristics. In general, there are two efficient approaches to broaden the bandwidth of the microstrip resonators, i.e., by increasing the substrate thickness and by combining several patches with similar resonance frequencies on the same or different layers of a microstrip structure [1]. Since the bandwidth of microstrip patch resonators around their operating resonant frequencies is very narrow, it is important to develop accurate algorithms for the prediction of those resonant frequencies [2, 3].

When a microstrip patch resonator acts as an antenna, it can be fed by a microstrip line located below the ground plane of the antenna through an aperture in the ground plane. This feeding configuration has been found very advantageous for several reasons [4]. For instance, it makes it possible to use a high dielectric-constant substrate for the feeding network and a low dielectric-constant substrate for the antenna element, which yields optimal performance for both the feeding network and antenna element. Also, the radiation arising from the feeding network cannot interfere with the main radiation pattern generated by the antenna since the ground plane separates the two radiation mechanisms. In addition, the presence of apertures in the ground planes of microstrip patches adds new design parameters that can be used as a way to tune their resonant frequencies without modifying the patch itself. A rectangular aperture introduces two physical parameters—its length and width. Since the presence of apertures in the ground planes of microstrip patches affects the resonant properties of microstrip-patch antennas and microstrip-patch circuit components, the development of an accurate theoretical method for predicting the aperture effect is of considerable interest.

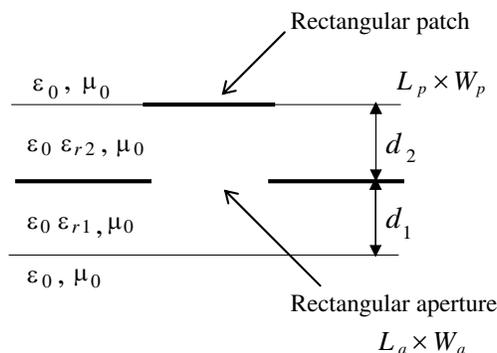


Figure 1: Geometrical structure of a rectangular microstrip patch over a ground plane with rectangular aperture. The aperture is in contact with a high-permittivity dielectric layer.

In this paper, the authors analyze how the resonant frequencies of rectangular microstrip patch resonators with rectangular apertures in the ground planes are affected by the presence of a high-permittivity dielectric layer below the apertures. To the best of our knowledge, this subject has not been reported in the open literature; the only published results on full-wave analysis of rectangular microstrip-patch resonators over ground planes with rectangular apertures refer to apertures in contact with air [2]. The effect of substrate thickness on the resonant frequency of rectangular microstrip patch over ground plane with rectangular aperture is also investigated. We show that the effect of the substrate thickness on the operating frequency of rectangular microstrip patch resonator with rectangular aperture in the ground plane is absolutely different from that known in the case of rectangular microstrip resonator without aperture in the ground plane.

2. OUTLINE OF THE NUMERICAL PROCEDURE

The problem to be solved is illustrated in Fig. 1. We have a rectangular patch of length L_P along the x direction and width W_P along the y direction over a ground plane with a rectangular aperture of length L_a and width W_a . Both the center of the patch and the center of the aperture have the coordinate value $(x, y) = (0, 0)$. Also, the metallic patch and the ground plane are assumed to be perfect electric conductors of neglecting thickness. The aperture is embedded in a two-layered substrate. The ambient medium is air with constitutive parameters ε_0 and μ_0 . All fields and currents are time harmonic with the $e^{i\omega t}$ time dependence suppressed. Let $\mathbf{j}(\mathbf{k}_s)$ ($\mathbf{j}_0(\mathbf{k}_s)$) be the vector Fourier transform of the current $\mathbf{J}(x, y)$ ($\mathbf{J}_0(x, y)$) on the patch (ground plane with a rectangular aperture). Also, let $\mathbf{e}(\mathbf{k}_s, d_1 + d_2)$ ($\mathbf{e}(\mathbf{k}_s, d_1)$) be the vector Fourier transform of the transverse electric field $\mathbf{E}(x, y, d_1 + d_2)$ ($\mathbf{E}(x, y, d_1)$) at the plane of the patch (ground plane with a rectangular aperture). Following a mathematical reasoning similar to that shown in [2, Eqs. (2)–(22)], we obtain a relation among $\mathbf{j}(\mathbf{k}_s)$, $\mathbf{j}_0(\mathbf{k}_s)$, $\mathbf{e}(\mathbf{k}_s, d_1 + d_2)$ and $\mathbf{e}(\mathbf{k}_s, d_1)$ given by

$$\mathbf{e}(\mathbf{k}_s, d_1 + d_2) = \bar{\mathbf{G}}(\mathbf{k}_s) \cdot \mathbf{j}(\mathbf{k}_s) + \bar{\mathbf{\Omega}}(\mathbf{k}_s) \cdot \mathbf{e}(\mathbf{k}_s, d_1) \quad (1)$$

$$\mathbf{j}_0(\mathbf{k}_s) = \bar{\mathbf{\Lambda}}(\mathbf{k}_s) \cdot \mathbf{j}(\mathbf{k}_s) + \bar{\mathbf{Y}}(\mathbf{k}_s) \cdot \mathbf{e}(\mathbf{k}_s, d_1) \quad (2)$$

The four diagonal matrices $\bar{\mathbf{G}}(\mathbf{k}_s)$, $\bar{\mathbf{\Omega}}(\mathbf{k}_s)$, $\bar{\mathbf{\Lambda}}(\mathbf{k}_s)$ and $\bar{\mathbf{Y}}(\mathbf{k}_s)$ stand for a set of dyadic Green's functions.

Now that we have the necessary Green's functions, the well-known Galerkin's procedure of the moment method can be easily applied to Eqs. (1) and (2) to obtain the resonant frequencies of the resonant modes of the structure shown in Fig. 1. Concerning the choice of basis functions, the unknown patch current is expanded in terms of a set of transverse magnetic modes of a rectangular cavity with magnetic side walls and electric top and bottom walls, and whose Fourier transforms are obtained in closed form. Also, the same basis functions are used for approximating the magnetic current density on the aperture in accordance with the concept of complementary electromagnetic structures [2].

3. NUMERICAL RESULTS AND DISCUSSION

In Fig. 2, the numerical results obtained for the resonant frequencies of rectangular microstrip patches over ground planes without apertures are compared with the results obtained for the same microstrip patches over ground planes with apertures in two different situations. In the first situation, the air occupies the space below the apertures and, in the second situation, there is a dielectric layer of high permittivity below the apertures. This latter situation has an interest for antenna applications since when microstrip patch antennas are fed by microstrip lines through apertures, the feeding microstrip lines may be printed on high-permittivity dielectric substrates [4]. It can be noticed that the results obtained for the resonant frequencies of the patches with apertures in contact with a high-permittivity dielectric layer not only substantially differ from the results without apertures, but they also appreciably differ from the results obtained with apertures in contact with air. Thus, the differences between the results obtained for the resonant frequencies when $L_a \times W_a = 0.5W_P \times 0.5W_P$ and $\varepsilon_{r1} = 1$ and those obtained when $L_a \times W_a = 0.5W_P \times 0.5W_P$ and $\varepsilon_{r1} = 10$ reach 4.31% for a thickness $d_1 = 1.2$ mm and 7.89% for a thickness $d_1 = 2$ mm. This clearly indicates that the placement of a dielectric layer of high permittivity below the ground plane of a rectangular microstrip patch over ground plane with aperture may considerably change the resonant frequencies of the patch.

In Fig. 3, the influence of the substrate thickness on the resonant frequency of rectangular microstrip patch over ground plane with rectangular aperture is studied. The results of Fig. 3 indicate

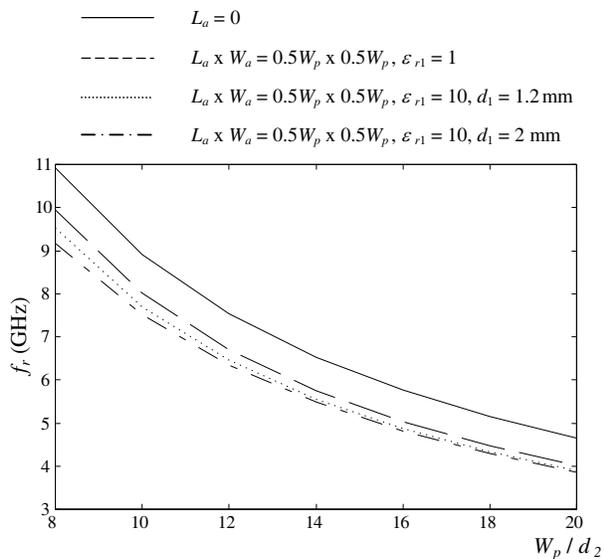


Figure 2: Resonant frequencies of rectangular microstrip patches over ground planes with and without rectangular apertures. Below the aperture there is a dielectric layer of high permittivity; $L_p = 1.5W_p$, $\epsilon_{r2} = 2.35$, $d_1 = 1$ mm.

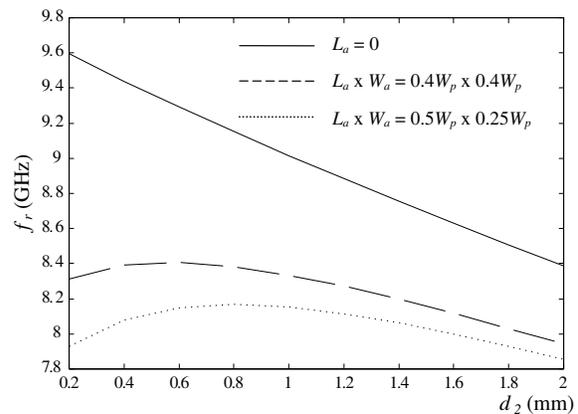


Figure 3: Resonant frequency of rectangular microstrip patch over ground planes with and without rectangular apertures versus the substrate thickness; $L_p = 1.5$ cm, $W_p = 1$ cm, $\epsilon_{r2} = 2.35$, $\epsilon_{r1} = 1$.

that the effect of the substrate thickness on the operating frequency of rectangular microstrip patch resonator with rectangular aperture in the ground plane is absolutely different from that known in the case of rectangular microstrip patch resonator without aperture in the ground plane. In effect, for the case of rectangular patch over ground plane with rectangular aperture, the increase in the substrate thickness causes initially an increase in the resonance frequency. This increase is owed to the reduction in the effective size of the rectangular aperture.

4. CONCLUSIONS

The Galerkin's method in the Fourier transform domain has been used for the numerical calculation of the resonant frequencies of rectangular microstrip patches over ground planes with rectangular apertures in the case in which there is a high-permittivity dielectric layer below the apertures. The numerical calculations have shown that the results obtained for the resonant frequencies of the patches with apertures in contact with a high-permittivity dielectric layer not only substantially differ from the results without apertures, but they also appreciably differ from the results obtained with apertures in contact with air. The numerical results obtained have also indicated that the effect of the substrate thickness on the operating frequency of rectangular microstrip patch resonator with rectangular aperture in the ground plane is absolutely different from that known in the case of rectangular microstrip patch resonator without aperture in the ground plane.

REFERENCES

1. Fortaki, T., L. Djouane, F. Chebara, and A. benghalia, "On the dual-frequency behavior of stacked microstrip patches," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 7, 310–313, 2008.
2. Fortaki, T. and A. benghalia, "Rigorous full-wave analysis of rectangular microstrip patches over ground planes with rectangular apertures in multilayered substrates that contain isotropic and uniaxial anisotropic materials," *Microwave Opt. Technol. Lett.*, Vol. 41, No. 6, 496–500, 2004.
3. Fortaki, T., L. Djouane, F. Chebara, and A. benghalia, "Radiation of a rectangular microstrip patch antenna covered with a dielectric layer," *Int. J. Electron.*, Vol. 95, No. 9, 989–998, 2008.
4. Losada, V., R. R. Boix, and M. Horno, "Resonant modes of circular microstrip patches over ground planes with circular apertures in multilayered substrates containing anisotropic and ferrite materials," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 1756–1762, 2000.

Detailed Modified Iwasawa Decomposition of Ray Transformation Matrix and Its Applications

T. Alieva

Facultad de Ciencias Físicas, Universidad Complutense de Madrid
Avda. Complutense, s/n, Madrid E-28040, Spain

Abstract— The formalism of ray transformation matrices allows establishing a close relation between the geometric and wave paraxial optics and providing their description in phase space. The modified Iwasawa decomposition of ray transformation matrix is crucial for classification of phase space transformations. Moreover further concretization of its components significantly simplifies system design, beam synthesis and characterization. The applications of the detailed ray transformation matrix decomposition are demonstrated on several examples: design of optical set up for phase space tomography, generation of stable, spiral and auto-reciprocal beams, characterization of optical signals.

1. INTRODUCTION

Ray transformation matrix formalism is widely used for the analysis and synthesis of optical systems considered in paraxial approximation. We remind that a real symplectic ray transformation 4×4 matrix \mathbf{T} relates the position \mathbf{r}_i and direction \mathbf{q}_i of an incoming ray to the position \mathbf{r}_o and direction \mathbf{q}_o of the outgoing ray,

$$\begin{bmatrix} \mathbf{r}_o \\ \mathbf{q}_o \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{r}_i \\ \mathbf{q}_i \end{bmatrix} = \mathbf{T} \begin{bmatrix} \mathbf{r}_i \\ \mathbf{q}_i \end{bmatrix}, \quad (1)$$

where $\mathbf{r} = (x, y)^t$ and $\mathbf{q} = (q_x, q_y)^t$. Here the normalized dimensionless variables are used and therefore \mathbf{q} can also be interpreted as spatial frequency or momentum. The ray transformation matrix parametrizes the kernel of the canonical integral linear transform which describes the evolution of the complex field amplitude during beam propagation through the system. Moreover the same matrix describes the corresponding affine transformation of the Wigner distribution in phase space, defined by the position and spatial frequency coordinates.

The Iwasawa decomposition of the proper normalized symplectic matrix adapted for the analysis of the first-order optical systems [1], called also as a modified Iwasawa decomposition, consists in its presentation as a product of three

$$\mathbf{T} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{G} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ -\mathbf{Y} & \mathbf{X} \end{bmatrix} = \mathbf{T}_L \mathbf{T}_S \mathbf{T}_O, \quad (2)$$

with \mathbf{I} denoting the identity matrix, in which the first matrix represents a lens action described by the symmetric matrix

$$\mathbf{G} = -(\mathbf{C}\mathbf{A}^t + \mathbf{D}\mathbf{B}^t)(\mathbf{A}\mathbf{A}^t + \mathbf{B}\mathbf{B}^t)^{-1} = \mathbf{G}^t. \quad (3)$$

The second matrix corresponds to a scaler described by the positive definite symmetric matrix

$$\mathbf{S} = (\mathbf{A}\mathbf{A}^t + \mathbf{B}\mathbf{B}^t)^{1/2} = \mathbf{S}^t \quad (4)$$

and the third is an orthosymplectic [2] (i.e., both orthogonal and symplectic) matrix, which can be shortly represented by the unitary matrix

$$\mathbf{U} = \mathbf{X} + i\mathbf{Y} = (\mathbf{A}\mathbf{A}^t + \mathbf{B}\mathbf{B}^t)^{-1/2}(\mathbf{A} + i\mathbf{B}). \quad (5)$$

Since the ray transformation matrix \mathbf{T} is symplectic and therefore

$$\mathbf{A}\mathbf{B}^t = \mathbf{B}\mathbf{A}^t, \quad \mathbf{C}\mathbf{D}^t = \mathbf{D}\mathbf{C}^t, \quad \mathbf{A}\mathbf{D}^t - \mathbf{B}\mathbf{C}^t = \mathbf{I}, \quad (6)$$

it has only ten free parameters. We call the system associated with \mathbf{T} *separable* if the block matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , and \mathbf{D} and \mathbf{G} , \mathbf{S} , \mathbf{X} , and \mathbf{Y} correspondingly are diagonal that reduces up to six degrees of freedom. The rotational symmetric or isotropic system (IS) corresponding to scalar block matrices is described by three parameters.

The ten parameters of the ray transformation matrix can be better interpreted if we use the detailed modified Iwasawa decomposition. Note that the first two matrices in Eq. (2) are defined by three parameters, while the last one is described by four. Let us start from the orthogonal matrix which describes the phase space rotations. The phase space rotators include signal rotator, separable fractional FT and gyrator among others. The signal rotator can be expressed by the unitary matrix,

$$\mathbf{U}_r(\alpha) = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix} \quad (7)$$

associated with the clockwise rotation in the (x, y) and (q_x, q_y) planes at angle α . The separable fractional FT described by the unitary matrix

$$\mathbf{U}_f(\gamma_x, \gamma_y) = \begin{bmatrix} \exp(i\gamma_x) & 0 \\ 0 & \exp(i\gamma_y) \end{bmatrix}, \quad (8)$$

corresponds to rotations in the (x, q_x) and (y, q_y) planes through the angles γ_x and γ_y , respectively. The gyrator transform, associated with

$$\mathbf{U}_g(\vartheta) = \begin{bmatrix} \cos \vartheta & i \sin \vartheta \\ i \sin \vartheta & \cos \vartheta \end{bmatrix}, \quad (9)$$

produces twisting, i.e., rotations in the mixed (x, q_y) and (y, q_x) planes of the phase space.

It has been shown [3] that any phase space rotator can be represented as a separable fractional FT (two parameters) embedded into two signal rotators (two parameters) $\mathbf{U} = \mathbf{U}_r(\beta) \mathbf{U}_f(\gamma_x, \gamma_y) \mathbf{U}_r(\alpha)$. By the same manner the astigmatic lens transform can be written as a separable lens (two parameters) embedded into direct-inverse rotators (one parameter). Similarly the astigmatic magnifier can be interpreted as a separable magnifier (two parameters) embedded into direct-inverse rotators (one parameter). This detailed decomposition is useful for example for system analysis and in particular allows describing the canonical integral transform associated with the ray transformation matrix with singular but non zero block matrix \mathbf{B} [3]. Nevertheless as we will see further there are other possibilities for detailed matrix decompositions.

2. SYSTEM DESIGN FOR PHASE SPACE TOMOGRAPHY

The tomographic reconstruction of the WD from its projections is an attractive method for the characterization of coherent and partially coherent signals [4]. It is based on the rotation of the Wigner distribution under the fractional Fourier transform and the fact that the WD projections coincide with the fractional Fourier power spectra which can be measured as intensity distributions. Since a system performing the separable fractional FT is constructed from the basic elements such as lenses, mirrors and free space intervals, the ray transformation matrix associated with $\mathbf{U}_f(\gamma_x, \gamma_y)$ has to be written as a product of ones corresponding to these elements. The main requirements for the system are the following: minimum number of elements, the same scaling factor for all WD projections and the fast change of the transformation parameters. Using the above mentioned decomposition of the fractional FT matrix it has been shown [5, 6] that the optimal system consists from three generalized lenses separated by the same distance between them. The first and the last lenses are identical. For the measurements of the WD projections the last lens is substituted by the CCD camera for intensity distribution registration. The generalized lenses can be implemented by the spatial light modulators that allows the fast change of their focal distances and therefore the transformation parameters. Alternatively the pair of the crossed glass cylindrical lenses can be used for every generalized lens performance. In the last case only the fractional transformations for the antisymmetric angles $\gamma_x = -\gamma_y$ are obtained. From these power spectra the WD for signals separable in the Cartesian coordinates can be reconstructed [7].

This example demonstrates that based on the modified Iwasawa decomposition the appropriate transformation for the signal characterization can be found, while the further representation of the corresponding ray transformation matrix as a product of basic elements allows constructing proper optical system.

3. BEAM GENERATION

The modified Iwasawa decomposition also can be used for the generation of optical beams with desired properties. Usually the beams are designed for the propagation through the isotropic

systems corresponding to the scalar block matrices **A**, **B**, **C**, and **D**. In this case the orthogonal matrix \mathbf{T}_O corresponds to the symmetric fractional FT ($\gamma_x = \gamma_y$). Therefore the construction of the stable beam, which does not change its intensity distribution apart from a scaling, is reduced to the search of the eigenfunctions for the symmetric fractional FT. It is known that the Hermite-Gaussian (HG) mode

$$HG_{m,n}(\mathbf{r}) = \sqrt{\frac{2}{\pi}} \frac{H_m(\sqrt{2}x)H_n(\sqrt{2}y)}{i^{-n}\sqrt{2^m m!}\sqrt{2^n n!}} \exp(-r^2), \quad (10)$$

where H_m represents the Hermite polynomial of order m , is an eigenfunction for the symmetric fractional FT at angle γ with the eigenvalue $\exp(-i(m+n+1)\gamma)$. It is easy to see that a linear combination of the HG modes with the same sum of the indices $m+n = \text{const}$ form a stable beam. Some stable beams of interest are the Laguerre-Gaussian (LG), Hermite-Laguerre-Gaussian (HLG) [8, 9], Ince-Gaussian [10] ones.

A linear combination of the HG modes with equal eigenvalues only for concrete angle values of the fractional FT corresponds to an auto-reciprocal beam, whose intensity distribution is repeated (apart from the scaling) at certain distances during the propagation [11].

The LG modes which are eigenfunctions for both, the symmetric fractional FT and the signal rotator, are the proper basis for the construction of spiral beams [12–14]. In this case the intensity distribution of the beam, again apart from the scaling, rotates during its propagation. In particular any combination of two LG modes forms a spiral beam. Analogously, if instead of the LG modes the HLG orthonormal basis is used the beam, which rotates at another phase space plane during the propagation through the IS, can be constructed [14].

As we can learned from these examples the application of modified Iwasawa decomposition allows often reducing the task of beam design to the solution of the eigenvalue problem. It also explain that origin of the Gouy phase accumulation is the phase space rotation of the beam [15]. Thus the dynamic phase is accumulated in IS, related to the symmetric fractional FT, while the geometric phase accumulation requires a system with astigmatic elements, similar to the antisymmetric fractional FT, as well as mode asymmetry, $m \neq n$.

4. SIGNAL CHARACTERIZATION

The express characterization of the paraxial beams is usually related to the analysis of the second-order moments of its WD [16, 17] ordinated in form of matrix 4×4 . They provide the information about the beam width, divergence, the longitudinal projection of the carried orbital angular momentum (OAM), etc.

It has been underlined [18] that knowledge of two eigenvalues of the moment matrix allows dividing two-dimensional signals into two classes: isotropic (equal eigenvalues) and anisotropic (different eigenvalues). Nevertheless since a signal has ten second-order moments more specific classification can be done [19]. Thus applying the modified Iwasawa decomposition to the ray transformation matrix which the inverse to one which brings the moment matrix to the diagonal form we are able to find the quadratic phase of the beam and the proper scaling. Moreover the further decomposition of the orthogonal matrix into rotator and gyrotator defines the principal axis of the signal and vortex part of its OAM. Following this procedure we are able to associate the signal with a certain point on the sphere, called Poincaré sphere, similar to the one applied for Gaussian mode representation [20–22] and characterization of the polarization. The latitude of this point describes the vorticity of the signal, while its longitude corresponds to the orientation of the beam's principal axes. This method of signals normalization is useful in particular for their comparison.

5. CONCLUSION

We conclude that the detailed modified Iwasawa decomposition of ray transformation matrix into a product of three: lens, magnifier and phase space rotator with different kinds of further cascade representation of every of them is a powerful tool for the analysis and synthesis of paraxial optical systems and beams.

ACKNOWLEDGMENT

The financial support from the Spanish Ministry of Science and Innovation under project TEC2008-04105 and from the Universidad Complutense de Madrid (Proyecto de Innovación y Mejora de la Calidad Docente, N 178, 2010) are acknowledged.

REFERENCES

1. Simon, R. and N. Mukunda, "Iwasawa decomposition in first-order optics: Universal treatment of shape-invariant propagation for coherent and partially coherent beams," *J. Opt. Soc. Am. A*, Vol. 15, 2146–2155, 1998.
2. Simon, R. and K. B. Wolf, "Structure of the set of paraxial optical systems," *J. Opt. Soc. Am. A*, Vol. 17, 342–355, 2000.
3. Alieva, T. and M. J. Bastiaans, "Alternative representation of the linear canonical integral transform," *Opt. Lett.*, Vol. 30, 3302–3304, 2005.
4. Raymer, M. G., M. Beck, and D. F. McAlister, "Complex wave-field reconstruction using phase-space tomography," *Phys. Rev. Lett.*, Vol. 72, 1137–1140, 1994.
5. Rodrigo, J. A., T. Alieva, and M. L. Calvo, "Optical system design for orthosymplectic transformations in phase space," *J. Opt. Soc. Am. A*, Vol. 23, 2494–2500, 2006.
6. Rodrigo, J. A., T. Alieva, and M. L. Calvo, "Programmable two-dimensional optical fractional Fourier processor," *Opt. Express*, Vol. 17, 4976–4983, 2009.
7. Cámara, A., T. Alieva, J. A. Rodrigo, and M. L. Calvo, "Phase space tomography reconstruction of the Wigner distribution for optical beams separable in Cartesian coordinates," *J. Opt. Soc. Am. A*, Vol. 26, 1301–1306, 2009.
8. Abramochkin, E. and V. Volostnikov, "Beam transformations and nontransformed beams," *Opt. Comm.*, Vol. 83, 123–135, 1991.
9. Allen, L., M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, "Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes," *Phys. Rev. A*, Vol. 45, 8185–8189, 1992.
10. Bandres, M. A. and J. C. Gutierrez, "Ince Gaussian beams," *Opt. Lett.*, Vol. 29, No. 2, 144–146, 1986.
11. Alieva, T. and A. M. Barbé, "Self-fractional Fourier images," *J. Mod. Optics*, Vol. 46, 83–99, 1999.
12. Schechner, Y. Y., R. Piestun, and J. Shamir, "Wave propagation with rotating intensity distribution," *Phys. Rev. E*, Vol. 54, No. 1, 50–53, 1996.
13. Abramochkin, E. and V. Volostnikov, "Spiral light beams," *Phys. Usp.*, Vol. 47, 1177–1203, 2004.
14. Alieva, T., E. Abramochkin, A. Asenjo-Garcia, and E. Razueva, "Rotating beams in isotropic optical system," *Opt. Express*, Vol. 18, 3568–3573, 2010.
15. Alieva, T. and M. J. Bastiaans, "Dynamic and geometric phase accumulation by Gaussian-type modes in first-order optical systems," *Opt. Lett.*, Vol. 33, 1659–1661, 2008.
16. Bastiaans, M. J., "Second-order moments of the Wigner distribution function in first-order optical systems," *Optik*, Vol. 88, 163–168, 1991.
17. Serna, J., R. Martínez-Herrero, and P. M. Mejías, "Parametric characterization of general partially coherent beams propagating through ABCD optical systems," *J. Opt. Soc. Am. A*, Vol. 8, 1094–1098, 1991.
18. Sundar, K., N. Mukunda, and R. Simon, "Coherent mode decomposition of general anisotropic Gaussian Schell model beams," *J. Opt. Soc. Am. A*, Vol. 12, 560–569, 1995.
19. Bastiaans, M. J. and T. Alieva, "Signal representation on the angular Poincaré sphere, based on second-order moments," *J. Opt. Soc. Am. A*, Vol. 27, 918–927, 2010.
20. Padgett, M. J. and J. Courtial, "Poincaré sphere equivalent for light beams containing orbital angular momentum," *Opt. Lett.*, Vol. 24, 430–432, 1999.
21. Calvo, G. F., "Wigner representation and geometric transformations of optical orbital angular momentum spatial modes," *Opt. Lett.*, Vol. 30, 1207–1209, 2005.
22. Alieva, T. and M. J. Bastiaans, "Phase-space rotations and orbital Stokes parameters," *Opt. Lett.*, Vol. 34, 410–412, 2009.

Radon-Wigner Display Implemented by Spatial Light Modulators

A. Cámara¹, J. A. Rodrigo², T. Alieva¹, and M. L. Calvo¹

¹Facultad de Ciencias Físicas, Universidad Complutense de Madrid
Avda. Complutense, s/n, Madrid E-28040, Spain

²Imaging and Vision Department, Instituto de Optica (CSIC)
Serrano 121, Madrid E-28006, Spain

Abstract— We present an optical system formed by spatial light modulators that allows measuring the Radon-Wigner transform of 1D signals in the range $[\pi/2, 3\pi/2]$. From it, the Wigner distribution (WD) of the signal can be reconstructed. The experimental feasibility of the setup is demonstrated on the examples of a slit and a windowed chirp signals. The quality of the WD reconstructions can be improved by increasing the number of the measured WD projections.

1. INTRODUCTION

The Radon-Wigner Transform (RWT) is an important tool for analysis and synthesis of signals [1, 2]. It is defined for the 1D case as the set of projections of the Wigner distribution (WD) of the signal, or equivalently, as the set of squared modulus fractional Fourier transforms (FRFTs) for parameter α in a π -interval [3]. Therefore, the WD of the signal can be obtained from the RWT applying a tomographic reconstruction algorithm [4]. Note that from the WD the phase of coherent signals, or the mutual intensity in case of partially coherent ones, can be recovered.

A Radon-Wigner Display (RWD) is an optical system that generates the RWT of a 1D signal as the intensity distribution in a certain transverse plane. There are several successful implementations of RWDs that involve the use of combinations of Fresnel zone plates or analogic varifocal glass lenses [5–7]. We propose the use of two reflection spatial light modulators (SLMs) working in the phase-only modulation regime for the construction of a flexible RWD that provides a scale invariant RWT for the range $[\pi/2, 3\pi/2]$ at the output plane. The comparison of the WD reconstructed from experimental data with one numerically simulated is presented for two examples: a slit and a windowed chirp signals.

Let us define a scalar one-dimensional quasi-monochromatic completely coherent beam by its complex amplitude $f(x)$. Assuming the paraxial approximation, its RWT is equivalent to the set of its fractional power spectra (the squared moduli of the FRFT) for angle α covering a π -interval:

$$P^\alpha [f(x)](u) = \frac{1}{s |\sin \alpha|} \left| \int dx f(x) \exp \left\{ \frac{i\pi}{s} [(x^2 + u^2) \cot \alpha - 2xu \csc \alpha] \right\} \right|^2, \quad (1)$$

where s is a parameter that depends on the optical implementation of the transformation and has dimension of squared length. It also corresponds to the projection of the WD along a line that spans the angle α with the x coordinate axis [3]:

$$P^\alpha [f(x)](u) = \int dq W(x \cos \alpha - sq \sin \alpha, q \cos \alpha + s^{-1}x \sin \alpha), \quad (2)$$

being $W(x, q)$ the WD of the signal:

$$W(x, q) = \int dx' f(x + x'/2) f^*(x - x'/2) \exp(-i2\pi x'q). \quad (3)$$

An optical implementation of the 2D FRFT for changeable angles (α_x, α_y) in the range $[\pi/2, 3\pi/2]$ using spatial light modulators have been proposed recently [8]. For the purpose of measuring the fractional power spectra, the setup consists of two generalized lenses implemented by SLMs separated by a fixed optical path length, d . Each generalized lens is formed by two orthogonally crossed cylindrical lenses whose power change depending on the transformation angle. The fractional power spectra of the input signal are measured by a CCD camera separated by the optical path length d from the second generalized lens.

In the case of 1D signals, the system can be modified for one-shot registration of the RWT. Firstly, we remove the cylindrical lens modulation of the wavefront in the y coordinate of each

generalized lens. To compensate the diffraction caused by the propagation between the SLMs and the CCD we place two cylindrical convergent lenses, L , of power $4/d$ at a distance $d/2$ from each SLM with its phase modulation along the y coordinate. Then, we have to replace the cylindrical lenses actuating in the x coordinate with varifocal lenses so that the fractional power spectra of different order are obtained along the y coordinate at the output plane of the setup (see Fig. 1). As the lenses are generated by the SLMs, we consider each row of pixels as a different channel i , where $i = 0, 1, \dots, N-1$ and N is the number of channels, and in each one we implement the phase modulation given by the functions:

$$t_i^{(1)}(x, y_i) = \exp \left[-\frac{i\pi}{s} x^2 \left(2 - \cot \frac{\alpha_i}{2} \right) \right], \quad (4)$$

$$t_i^{(2)}(x, y_i) = \exp \left[-\frac{i\pi}{s} x^2 4 (1 - \sin \alpha_i) \right], \quad (5)$$

where λ is the light wavelength, s is again an optical system parameter, and α_i is the transformation angle for the i th-channel. In Fig. 2 the phase of the functions $t^{(1)}$ and $t^{(2)}$ are displayed. At the output plane, the CCD camera captures the intensity distribution associated with the RWT of the 1D signal.

We demonstrate the capabilities of this system by the reconstruction of the WD from the measured RWT of a slit and a windowed chirp signal. The slit signal of width w is defined by

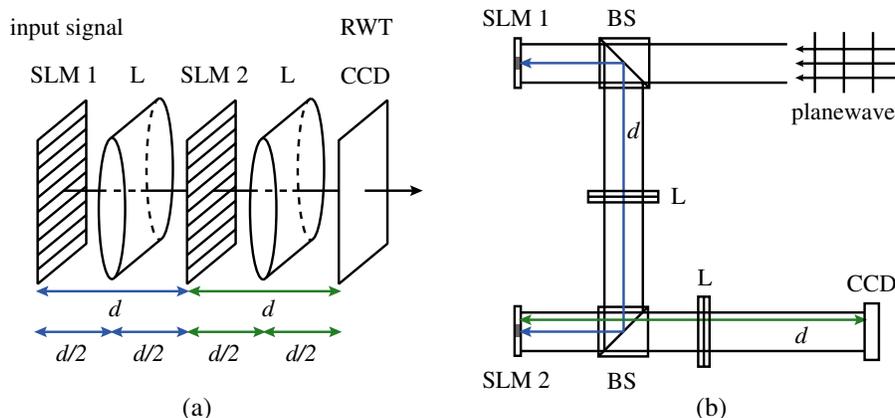


Figure 1: (a) The proposed optical setup and (b) the actual disposition of the elements of the system.

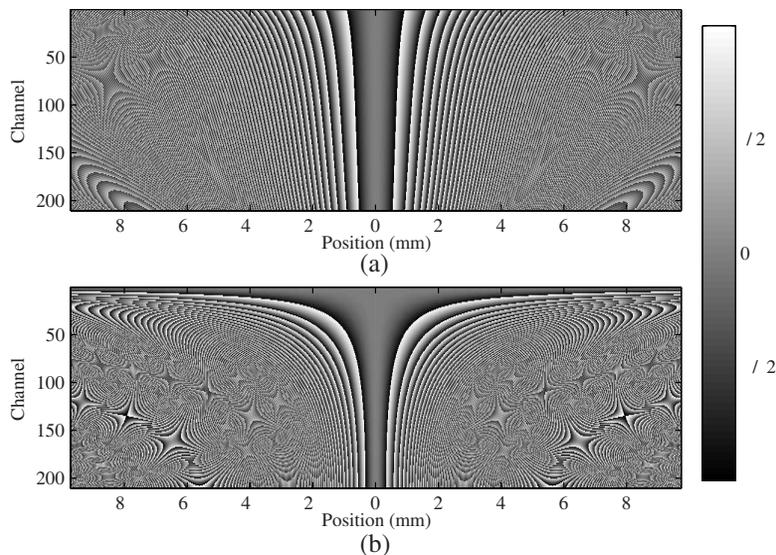


Figure 2: Phase modulation implemented (a) in the first SLM, $t^{(1)}$, and (b) in the second SLM, $t^{(2)}$, both cases for $N = 210$ channels.

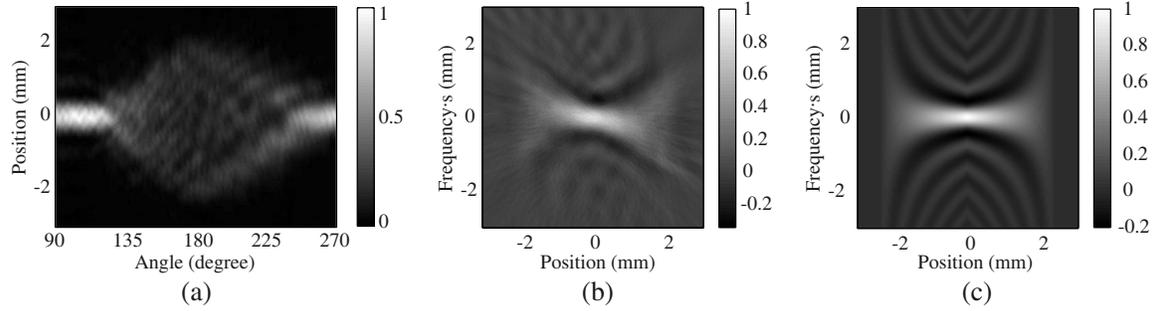


Figure 3: (a) The experimental RWT for $\alpha \in [\pi/2, 3\pi/2]$, (b) the WD reconstructed from the RWT, and (c) the theoretical WD of a slit signal of width $w = 1.1$ mm.

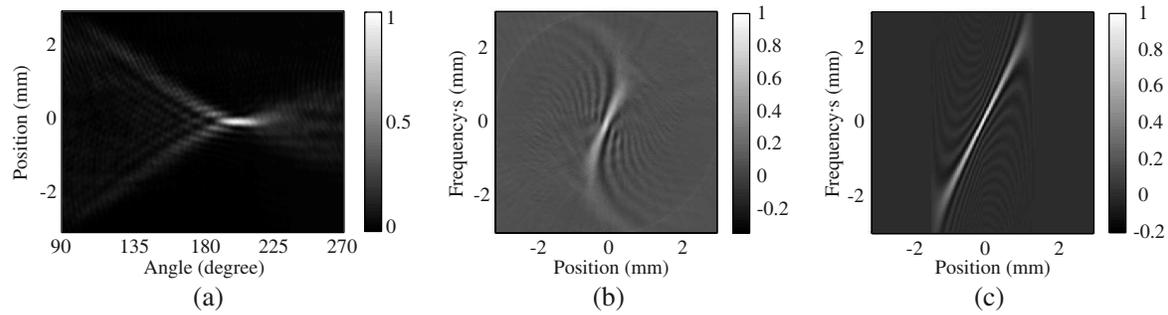


Figure 4: (a) The experimental RWT for $\alpha \in [\pi/2, 3\pi/2]$, (b) the WD reconstructed from the RWT, and (c) the theoretical WD of a windowed chirp signal of width $w = 1.4$ mm and power $p = 0.24 \text{ m}^{-1}$.

$\text{rect}(x/w)$, while the windowed chirp is defined by $\text{rect}(x/w) \exp(-\frac{i\pi}{\lambda} px^2)$, where p is the power of the chirp and w is the width of its window. Both signals can be generated illuminating the system with a plane wave and introducing an additional phase modulation in the first SLM.

In our experiment we use a Nd:YAG laser source of $\lambda = 532$ nm that is filtered and collimated. Each reflection SLMs is composed by an array of 1024×768 pixels of pitch $s_{\text{SLM}} = 19 \mu\text{m}$. The effective distance between SLMs is $d = 0.8$ m, and this defines the optical parameter $s = 2\lambda d = 0.85 \text{ mm}^2$. At the output plane, at $d = 0.8$ m from the second SLM, there is a 1600×1200 px CCD camera with pixel pitch $s_{\text{CCD}} = 4.65 \mu\text{m}$.

The RWT, the experimental WD, and the theoretical WD of the slit signal of $w = 1.1$ mm and the windowed chirp of $w = 1.4$ mm and $p = 0.24 \text{ m}^{-1}$ are displayed in Figs. 3 and 4, respectively. Each RWT is composed of $N = 210$ WD projections in the range $[\pi/2, 3\pi/2]$. The upper bound of the number of projections is determined by the pixel pitch relation between the SLM and the CCD camera and their number of pixels. For instance, our experimental setup can measure at most $1600s_{\text{CCD}}/s_{\text{SLM}} \approx 390$ projections. The experimental WD has been reconstructed from the measured RWT applying the inverse Radon transform algorithm.

The comparison with the theoretical WD demonstrates the feasibility of the proposed system. However, the experimental reconstructions are only accurate around the center of the phase-plane, while the quality decreases as we depart from it. This is caused by the combination of various factors like the limited dynamic range of the SLMs, the possible crosstalk between channels, or the insufficient number of WD projections. Simulations demonstrate that the number of projections of the RWT is crucial for the reconstruction quality. Note that the application of SLMs for the codification of lenses permits an increase of N , as several shots can be used to cover the needed angular region.

2. CONCLUSION

In conclusion, we have presented a RWD using two SLMs that allows measuring the RWT for any subrange of angles included in the interval $[\pi/2, 3\pi/2]$. From the RWT, the WD of the signal can be reconstructed. The experimental feasibility is demonstrated for the case of the slit and the windowed chirp signals. The precision of the reconstruction can be increased taking several

measurements to cover the needed range of angles of the RWT. The proposed setup can be used for phase retrieval, analysis of cylindrical lenses and their combinations, signal characterization, etc.

ACKNOWLEDGMENT

The financial support from the Spanish Ministry of Science and Innovation under project TEC2008-04105 and from the Universidad Complutense de Madrid (Proyecto de Innovación y Mejora de la Calidad Docente, N 178, 2010) are acknowledged. J. A. Rodrigo gratefully thanks “Juan de la Cierva” grant and A. Cámara acknowledges the financial support from the Consejería de Educación de la Comunidad de Madrid and the European Social Fund.

REFERENCES

1. Kay, S. and G. Boudreaux-Bartels, “On the optimality of the Wigner distribution for detection,” *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP’85*, Vol. 10, 1017–1020, April 1985.
2. Wood, J. and D. Barry, “Linear signal synthesis using the Radon-Wigner transform,” *IEEE Trans. Sig. Proc.*, Vol. 42, 2105–2111, 1994.
3. Lohmann, A. W., “Image rotation, Wigner rotation, and the fractional Fourier transform,” *J. Opt. Soc. Am. A*, Vol. 10, 2181–2186, 1993.
4. Raymer, M. G., M. Beck, and D. F. McAlister, “Complex wave-field reconstruction using phase-space tomography,” *Phys. Rev. Lett.*, Vol. 72, 1137–1140, 1994.
5. Mendlovic, R., R. G. Dorsch, A. W. Lohmann, Z. Zalevsky, and C. Ferreira, “Optical illustration of a varied fractional Fourier-transform order and the Radon-Wigner display,” *Appl. Opt.*, Vol. 35, 3925–3929, 1996.
6. Granieri, S., W. D. Furlan, G. Saavedra, and P. Andres, “Radon-Wigner display: A compact optical implementation with a single varifocal lens,” *Appl. Opt.*, Vol. 36, 8363–8369, 1997.
7. Zhang, Y., B.-Y. Gu, B.-Z. Dong, and G.-Z. Yang, “Optical implementations of the Radon-Wigner display for one-dimensional signals,” *Opt. Lett.*, Vol. 23, 1126–1128, 1998.
8. Rodrigo, J. A., T. Alieva, and M. L. Calvo, “Programmable two-dimensional optical fractional Fourier processor,” *Opt. Express*, Vol. 17, 4976–4983, 2009.

Orbital Angular Moment Density of Beam Given as a Superposition of Hermite-Laguerre-Gauss Functions

A. M. Caravaca Aguirre and T. Alieva
Universidad Complutense de Madrid, Madrid, Spain

Abstract— Orbital angular moment (OAM) is an important global parameter of beam which can be found from its second order moments. Nevertheless often, for example for the description of the interaction of light with microparticles, the knowledge of the OAM density is required. Hermite-Laguerre-Gauss (HLG) modes which form a complete orthonormal basis are frequently used for beam design. In this contribution we derive the expression for the OAM density of beam given as a superposition of the HLG modes. The result is useful for the generation and characterization of stable, spiral, rotating beams and their application for microparticle manipulation.

1. INTRODUCTION

The light-matter interaction is based on the transfer of energy, linear and angular momenta. Therefore industrial, medical, and research applications of lasers often require a beam with not only a specific intensity but also phase distribution in some plane, since the carried angular momentum depends on both terms. Frequently for beam synthesis or analysis its presentation as a linear composition of the orthonormal Gaussian modes whose characteristics are well established is used [1–4]. Thus for example the combination of Hermite-Gaussian (HG) modes with the constant sum of index $m + n = \text{const}$ forms a stable beam which propagates through the isotropic (or rotationally symmetric) system (IS) without change of the form of its complex field amplitude apart from some scaling and additional quadratic phase factor. Depending on the coefficients used in the composition one can obtain for example Laguerre-Gaussian (LG), Hermite-Laguerre-Gaussian (HLG) [1] or Ince-Gaussian modes [2]. Other rules for mode indices allow to generate the autoreciprocal [3] and rotating [4, 5] beams. The complex field amplitude of spiral beams [4, 6] which belong to the last family rotates (again, up to the scaling and quadratic phase factors) during the propagation through the ISs.

The diverse applications of the beams require the knowledge of their properties. Orbital angular momentum (OAM) is one of the important global characteristics of the beam, which can be found from its second order moments, see for example [7]. Thus the helical LG modes carried the OAM are used as optical spanners [11]. Nevertheless often, for example for the description of the interaction of light with microparticles, the OAM density has to be known. In this contribution we derive the expression for the OAM density of beam given as a superposition of the HLG modes. Note that in this paper we use the monochromatic scalar paraxial approximation.

2. HERMITE-LAGUERRE-GAUSS MODES

Since for beam design different orthonormal sets of modes might be used we consider here a variety of them called orthosymplectic ones [8]. Such set of modes is associated to the unitary matrix $\mathbf{U} = \mathbf{X} + i\mathbf{Y}$ with components $U_{11}, U_{12}, U_{21}, U_{22}$, which is related to the normalized orthosymplectic ray transformation matrix $\mathbf{T} = \begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ -\mathbf{Y} & \mathbf{X} \end{bmatrix}$ and describes the rotation in phase space. A mode of this set can be expressed by

$$f_{m,n}^{\mathbf{U}}(x, y) = \frac{(-1)^{m+n} \exp[(x^2 + y^2)]}{2^{m+n-1/2} \sqrt{\pi m! n!}} \left(U_{11} \frac{\partial}{\partial x} + U_{12} \frac{\partial}{\partial y} \right)^m \left(U_{21} \frac{\partial}{\partial x} + U_{22} \frac{\partial}{\partial y} \right)^n \exp[-2(x^2 + y^2)] \quad (1)$$

and for each \mathbf{U} matrix $\{f_{m,n}^{\mathbf{U}}\}$ forms a complete orthonormal function set. For the orthosymplectic modes, $f_{m,n}^{\mathbf{U}}(x, y) = f_{m,n}$, the following derivative relations [9]

$$\begin{aligned} x \frac{\partial f_{m,n}}{\partial y} &= x 2\sqrt{\pi} [U_{21}^* \sqrt{m} f_{m-1,n} + U_{22}^* \sqrt{n} f_{m,n-1}] - 2\pi xy f_{m,n} \\ y \frac{\partial f_{m,n}}{\partial x} &= y 2\sqrt{\pi} [U_{11}^* \sqrt{m} f_{m-1,n} + U_{12}^* \sqrt{n} f_{m,n-1}] - 2\pi xy f_{m,n} \end{aligned} \quad (2)$$

and the recurrence relations [9]

$$\begin{aligned}
 2\sqrt{\pi} x f_{m-1,n} &= U_{11}\sqrt{m} f_{m,n} + U_{12}\sqrt{n+1} f_{m-1,n+1} + U_{11}^*\sqrt{m-1} f_{m-2,n} + U_{12}^*\sqrt{n} f_{m-1,n-1} \\
 2\sqrt{\pi} y f_{m-1,n} &= U_{21}\sqrt{m} f_{m,n} + U_{22}\sqrt{n+1} f_{m-1,n+1} + U_{21}^*\sqrt{m-1} f_{m-2,n} + U_{22}^*\sqrt{n} f_{m-1,n-1} \\
 2\sqrt{\pi} x f_{m,n-1} &= U_{11}\sqrt{m+1} f_{m+1,n-1} + U_{12}\sqrt{n} f_{m,n} + U_{11}^*\sqrt{m} f_{m-1,n-1} + U_{12}^*\sqrt{n-1} f_{m,n-2} \\
 2\sqrt{\pi} y f_{m,n-1} &= U_{21}\sqrt{m+1} f_{m+1,n-1} + U_{22}\sqrt{n} f_{m,n} + U_{21}^*\sqrt{m} f_{m-1,n-1} + U_{22}^*\sqrt{n-1} f_{m,n-2}
 \end{aligned} \quad (3)$$

are hold.

If $\mathbf{U} = \mathbf{U}_g = \begin{pmatrix} \cos \alpha & i \sin \alpha \\ i \sin \alpha & \cos \alpha \end{pmatrix}$, and thus $U_{11} = U_{21} = \cos \alpha$, $U_{12} = U_{22} = i \sin \alpha$ the Eq. (1) reduces to HG modes for $\alpha = 0$, and to LG modes if $\alpha = \pm\pi/4$

$$f_{m,n}^{\mathbf{I}}(\mathbf{r}) = HG_{m,n}(\mathbf{r}) = \sqrt{\frac{2}{\pi}} \frac{H_m(\sqrt{2}x)H_n(\sqrt{2}y)}{i^{-n}\sqrt{2^m m!}\sqrt{2^n n!}} \exp(-r^2), \quad (4)$$

$$\begin{aligned}
 f_{m,n}^{\mathbf{U}_g(\pm\pi/4)}(\mathbf{r}) &= \mathcal{L}_{m,n}^{\mp}(\mathbf{r}) = \frac{(\pm i)^n (-1)^{\min(m,n)}}{\sqrt{\pi/2}} \sqrt{\frac{\min(m,n)!}{\max(m,n)!}} \left[\sqrt{2}(x \pm y) \right]^{|m-n|} \\
 &\times L_{\min(m,n)}^{|m-n|}(2\mathbf{r}^2) \exp(-\mathbf{r}^2), \quad (5)
 \end{aligned}$$

where H_m represents the Hermite polynomial of order m , and $L_{\min(m,n)}^{|m-n|}$ stands for the generalized Laguerre polynomial with radial $p = \min(m,n)$ and topological $l = |m-n|$ indices. The topological charge defined by $\pm l$ describes the phase change (in 2π units) of the associate field along a circular trajectory around the optical axis (geometrical mode center), where the $-$ sign corresponds to the counterclockwise direction and is determined by $\text{sign}(m-n)$. For other angles α we obtain intermediate modes referred as Hermite-Laguerre-Gaussian ones.

3. ORBITAL ANGULAR MOMENTUM DENSITY

In this section, we obtain an expression for the density of the OAM projection on the direction of the beam propagation (further mentioned as OAM density) for a beam expressed as a linear combination of orthonormal modes. The OAM projection in the propagation direction [7] is given by

$$L_z = C \int_{-\infty}^{+\infty} \text{Im} \left\{ E(x,y)^* \left[x \frac{\partial E(x,y)}{\partial y} - y \frac{\partial E(x,y)}{\partial x} \right] \right\} dx dy, \quad (6)$$

where $E(x,y)$ is the complex amplitude of electric field. The proportionality constant $C = 1/\omega I_0$ contains the beam intensity, $I_0 = \int |E(x,y)|^2 dx dy$, and the radiation frequency ω . Correspondingly the density of the OAM is written as

$$M(x,y) = \text{Im} \left\{ E(x,y)^* \left[x \frac{\partial E(x,y)}{\partial y} - y \frac{\partial E(x,y)}{\partial x} \right] \right\}. \quad (7)$$

The complex field amplitude of the electric field can be represented as a superposition of the orthosymplectic modes $f_{m,n}$ of the same set (\mathbf{U}) as $E(x,y) = \sum_{m,n} a_{m,n} f_{m,n}(x,y)$. Using this decomposition and the derivative and recurrence relations (2)–(3) after some simplifications we derive the expression for the OAM density

$$\begin{aligned}
 M_E(x,y) &= \sum_{m,n} |a_{m,n}|^2 M_{f_{m,n}} + \sum_{m,n,m',n'} 2 \text{Re} (a_{m,n}^* a_{m',n'} f_{m,n}^* f_{m',n'}) \text{Im} [U_{11} U_{21}^* (m+m') \\
 &+ U_{12} U_{22}^* (n+n')] + (U_{11} U_{22}^* - U_{21} U_{12}) \left[\text{Im} (a_{m,n}^* a_{m',n'} f_{m,n}^* f_{m'+1,n'-1}) \sqrt{n'(m'+1)} \right. \\
 &+ \text{Im} (a_{m,n} a_{m',n'}^* f_{m+1,n-1} f_{m',n}') \sqrt{n(m+1)} \left. \right] + (U_{12} U_{21}^* - U_{22} U_{11}) \\
 &\times \left[\text{Im} (a_{m,n}^* a_{m',n'} f_{m,n}^* f_{m'-1,n'+1}) \sqrt{m'(n'+1)} + \text{Im} (a_{m,n} a_{m',n'}^* f_{m-1,n+1} f_{m',n}') \sqrt{m(n+1)} \right] \quad (8)
 \end{aligned}$$

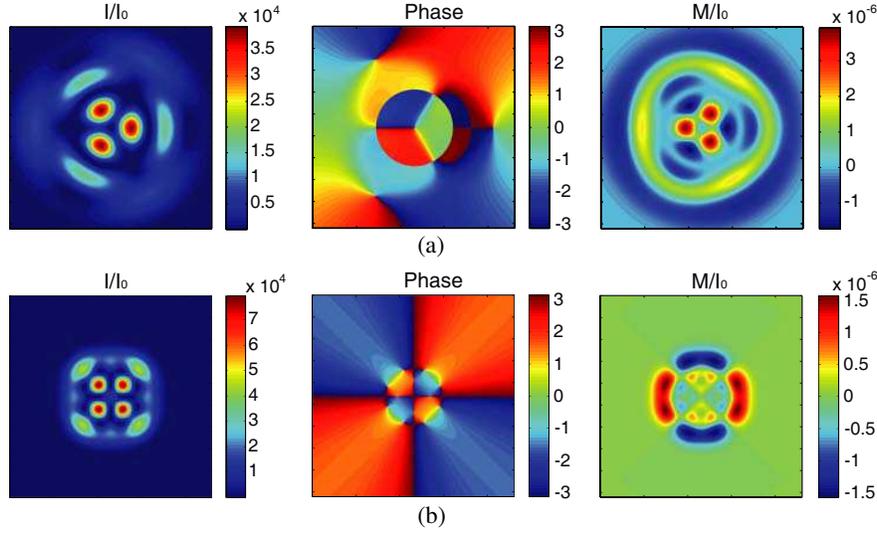


Figure 1: Intensity, phase and OAM density of the beams (a) $\mathcal{L}_{2,4} + \sqrt{2}\mathcal{L}_{2,1}$ and (b) $f_{2,2}^{\alpha=\pi/8}$ that have total OAM equals to 0, while OAM density is not 0.

where $M_{f_{m,n}}$ correspond to the OAM density of one mode $f_{m,n}$ and the second summation runs over every couple of index m, n, m', n' , except $m = m'$ and $n = n'$. Let us first analyse the OAM density of one orthosymplectic mode

$$M_{f_{m,n}}(x, y) = 2|f_{m,n}|^2 [\text{Im}(U_{11}U_{21}^*)m + \text{Im}(U_{12}U_{22}^*)n] + \text{Im} \left\{ f_{m,n}^* f_{m-1,n+1} \sqrt{m(n+1)} \right. \\ \left. \times [U_{12}U_{21}^* - U_{22}U_{11}^*] + f_{m,n}^* f_{m+1,n-1} \sqrt{n(m+1)} [U_{11}U_{22}^* - U_{21}U_{12}^*] \right\}. \quad (9)$$

Particularizing the formula for modes associated with $\mathbf{U}_g(\alpha)$, $f_{m,n}^{\mathbf{U}_g(\alpha)} \equiv f_{m,n}^\alpha$, we obtain

$$M_{f_{m,n}^\alpha} = \sin 2\alpha |f_{m,n}^\alpha|^2 (n - m) \\ + \cos 2\alpha \left[\text{Im}(f_{m,n}^{\alpha*} f_{m+1,n-1}^\alpha) \sqrt{n(m+1)} - \text{Im}(f_{m,n}^{\alpha*} f_{m-1,n+1}^\alpha) \sqrt{m(n+1)} \right]. \quad (10)$$

It is easy to see from this expression that only for the LG modes ($\alpha = \mp\pi/4$) the OAM density is proportional to the intensity distribution of the beam $M_{\mathcal{L}_{m,n}(\mathbf{r})} = \pm |\mathcal{L}_{m,n}(\mathbf{r})|^2 (m - n)$. The OAM of the mode can be obtained by the integration over x and y that leads to the known result $L_z = \pm (m - n)$ [10]. Moreover, we observe that some modes with zero global OAM have non zero OAM density. In particular it happens for the mode $f_{m,n}^\alpha$ with equal indices, $m = n$, and $\alpha \neq k\pi/4$ with an integer k or for the sum of LG modes with appropriate weights and topological charges of different signs. For example in Fig. 1, the intensity, phase and OAM density distributions of beams with zero global OAM are shown. Note also that the maxima of the intensity and the OAM density distributions do not coincide.

Let us now analyze the expression for OAM density for a linear combination of two HLG modes

$$M_{af_{m,n}^\alpha + bf_{m',n'}^\alpha} = \text{Im} \left\{ (af_{m,n}^\alpha + bf_{m',n'}^\alpha)^* \left[x \frac{\partial (af_{m,n}^\alpha + bf_{m',n'}^\alpha)}{\partial y} - y \frac{\partial (af_{m,n}^\alpha + bf_{m',n'}^\alpha)}{\partial x} \right] \right\} \\ = \text{Im} \left\{ |a|^2 f_{m,n}^{\alpha*} \left[x \frac{\partial f_{m,n}^\alpha}{\partial y} - y \frac{\partial f_{m,n}^\alpha}{\partial x} \right] \right\} + \text{Im} \left\{ |b|^2 f_{m',n'}^{\alpha*} \left[x \frac{\partial f_{m',n'}^\alpha}{\partial y} - y \frac{\partial f_{m',n'}^\alpha}{\partial x} \right] \right\} \\ + \text{Im} \left\{ a^* b f_{m,n}^{\alpha*} \left[x \frac{\partial f_{m',n'}^\alpha}{\partial y} - y \frac{\partial f_{m',n'}^\alpha}{\partial x} \right] + ab^* f_{m',n'}^{\alpha*} \left[x \frac{\partial f_{m,n}^\alpha}{\partial y} - y \frac{\partial f_{m,n}^\alpha}{\partial x} \right] \right\} \\ = |a|^2 M_{f_{m,n}^\alpha} + |b|^2 M_{f_{m',n'}^\alpha} + M_{int}^{mnm'n'}(\alpha), \quad (11)$$

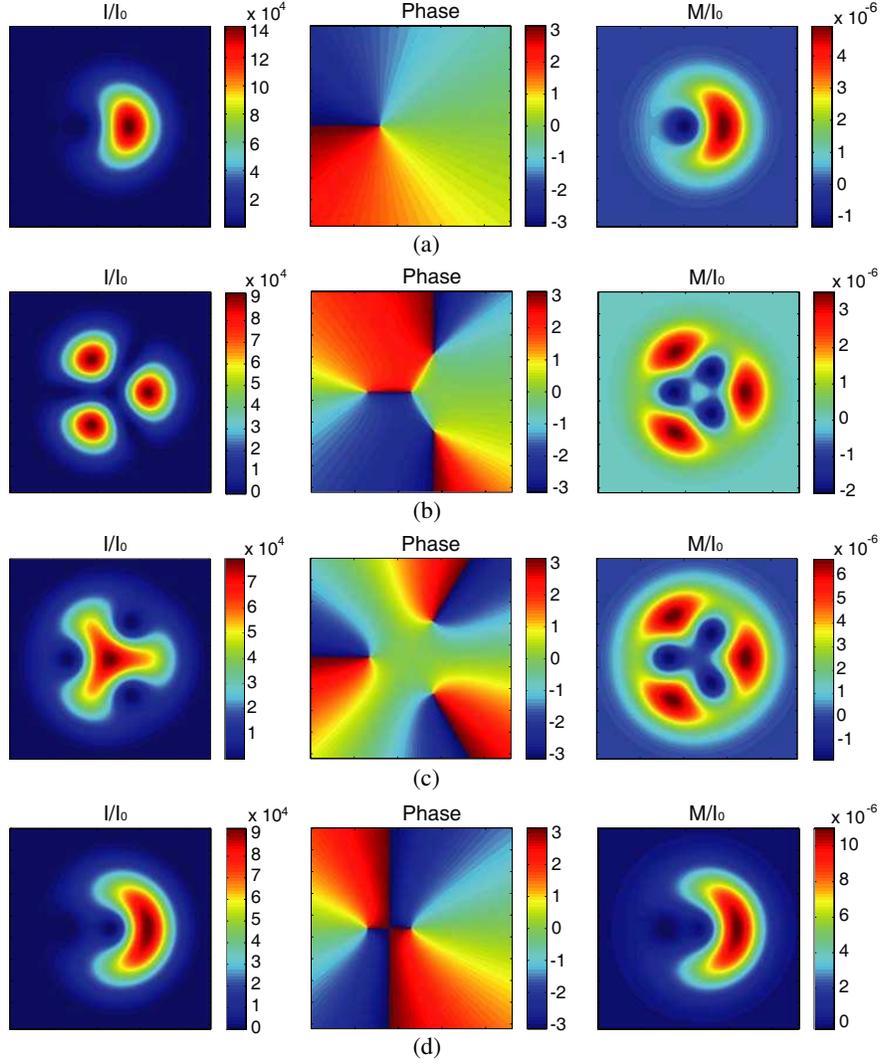


Figure 2: Intensity, phase and OAM density of the combination of the LG modes: (a) $\mathcal{L}_{0,0} + \mathcal{L}_{1,0}$, (b) $\mathcal{L}_{2,0} + \mathcal{L}_{0,1}$, (c) $\mathcal{L}_{0,0} + \mathcal{L}_{3,0}$ and (d) $\mathcal{L}_{1,0} + \mathcal{L}_{2,0}$ modes.

where $M_{f_{m,n}^\alpha}$ is the density of OAM of one mode, and $M_{int}^{mnm'n'}(\alpha)$ is

$$\begin{aligned}
 M_{int}^{mnm'n'}(\alpha) = & \operatorname{Re}(a^* b f_{m,n}^{\alpha*} f_{m',n'}^\alpha) \sin 2\alpha [(n' - m') + (n - m)] \\
 & + \cos 2\alpha \left[\operatorname{Im}(a^* b f_{m,n}^{\alpha*} f_{m'+1,n'-1}^\alpha) \sqrt{n'(m'+1)} + \operatorname{Im}(a b^* f_{m',n'}^{\alpha*} f_{m+1,n-1}^\alpha) \sqrt{n(m+1)} \right. \\
 & \left. - \operatorname{Im}(a^* b f_{m,n}^{\alpha*} f_{m'-1,n'+1}^\alpha) \sqrt{m'(n'+1)} - \operatorname{Im}(a b^* f_{m',n'}^{\alpha*} f_{m-1,n+1}^\alpha) \sqrt{m(n+1)} \right]. \quad (12)
 \end{aligned}$$

The total OAM can be obtained integrating Eq. (11) and using the orthonormality of the HLG modes. We observe that first term (proportional to the $\sin 2\alpha$ in Eq. (12)) has no input to the L_z , while the second one has it only if the mode indices differ by one, for example $m = m' + 1, n = n' - 1$. Thus for the combination $a f_{5,1}^\alpha + b f_{3,3}^\alpha$ the total OAM is defined as $L_z = |a|^2 L_{z,f_{5,1}^\alpha} + |b|^2 L_{z,f_{3,3}^\alpha}$.

The generalization of the Eq. (11) for any field, written as a linear combination of the HLG modes $E(\mathbf{r}) = \sum_{m,n} a_{m,n} f_{m,n}^\alpha$ yields to

$$\begin{aligned}
 M_E = & \sum_{m,n} |a_{m,n}|^2 M_{f_{m,n}^\alpha} + \sum_{m,n,m',n'} \left\{ \sin 2\alpha \operatorname{Re}(a_{m,n}^* a_{m',n'} f_{m,n}^{\alpha*} f_{m',n'}^\alpha) [(n' - m') + (n - m)] + \cos 2\alpha \right. \\
 & \times \left[\operatorname{Im}(a_{m,n}^* a_{m',n'} f_{m,n}^{\alpha*} f_{m'+1,n'-1}^\alpha) \sqrt{n'(m'+1)} + \operatorname{Im}(a_{m,n} a_{m',n'}^* f_{m+1,n-1}^\alpha f_{m',n'}^{\alpha*}) \sqrt{n(m+1)} \right. \\
 & \left. \left. - \operatorname{Im}(a_{m,n}^* a_{m',n'} f_{m,n}^{\alpha*} f_{m'-1,n'+1}^\alpha) \sqrt{m'(n'+1)} - \operatorname{Im}(a_{m,n} a_{m',n'}^* f_{m-1,n+1}^\alpha f_{m',n'}^{\alpha*}) \sqrt{m(n+1)} \right] \right\}, \quad (13)
 \end{aligned}$$

where the second summation runs over every couple of index m, n, m', n' , except $m = m'$ and $n = n'$. Eq. (10) and Eq. (13) allows us to analyze the OAM density of stable, rotating and autoreciprocal beams, written in terms of HLG modes used as a natural basis. It can be observed that the OAM density in general is not proportional to the intensity distribution that is given by

$$I = \left| \sum_{m,n} a_{m,n} \cdot f_{m,n}^\alpha(\mathbf{r}) \right|^2 = \sum_{m,n} |a_{m,n}|^2 |f_{m,n}^\alpha|^2 + \sum_{m,n,m',n'} a_{m,n}^* a_{m',n'} f_{m,n}^{\alpha*} f_{m',n'}^\alpha. \quad (14)$$

Particularizing the expression (13) for a lineal combination of LG modes, $E(\mathbf{r}) = \sum_{m,n} a_{m,n} \mathcal{L}_{m,n}(\mathbf{r})$, we obtain

$$M_E = \sum_{m,n,m',n'} |a_{m,n}|^2 |\mathcal{L}_{m,n}|^2 (m-n) + \text{Re}(a_{m,n}^* a_{m',n'} \mathcal{L}_{m,n}^* \mathcal{L}_{m',n'}) [(m'-n') + (m-n)] \quad (15)$$

The maximum and minimum of intensity don't have to match with the ones of the OAM density, as we can see in Fig. 2. A clear example is Fig. 2(c), where the central maximum of intensity corresponds to a zero of the OAM density. It is important to stand out, that the total OAM is given by $\sum_{m,n} |a_{m,n}|^2 (m-n)$. Beams with the same total OAM might have rather different OAM density distributions. In particular, we observe that the paires of spiral beams $\mathcal{L}_{0,0} + \mathcal{L}_{1,0}$ and $\mathcal{L}_{2,0} + \mathcal{L}_{0,1}$, $\mathcal{L}_{0,0} + \mathcal{L}_{3,0}$ and $\mathcal{L}_{1,0} + \mathcal{L}_{2,0}$ displayed in Fig. 2 have the same total OAM, 1/2 and 3/2 respectively, despite their intensity distributions and OAM densities are completely different.

4. CONCLUSION

We have derived the expression for the OAM density of a beam given as a linear superposition of the HLG modes. It has been demonstrated that there are a variety of beams with zero OAM but nonzero OAM density. The maxima of the intensity distribution in general do not coincide with the maxima of the OAM density (except of helical LG beams). These facts have to be taken into account for design of the beams for different applications.

ACKNOWLEDGMENT

The financial support from the Spanish Ministry of Science and Innovation under project TEC2008-04105, from the Universidad Complutense de Madrid (Proyecto de Innovación y Mejora de la Calidad Docente, N 178, 2010) and from Obra Social Ibercaja (Beca Ibercaja de Investigación 2010) are acknowledged.

REFERENCES

1. Abramochkin, E. and V. Volostnikov, "Beam transformations and nontransformed beams," *Opt. Comm.*, Vol. 83, 123–135, 1991.
2. Bandres, M. A. and J. C. Gutierrez, "Ince Gaussian beams," *Opt. Lett.*, Vol. 29, No. 2, 144–146, 1986.
3. Alieva, T. and A. M. Barbé, "Self-fractional Fourier images," *J. Mod. Optics*, Vol. 46, 83–99, 1999.
4. Schechner, Y. Y., R. Piestun, and J. Shamir, "Wave propagation with rotating intensity distribution," *Phys. Rev. E.*, Vol. 54, No. 1, 50–53, 1996.
5. Alieva, T., E. Abramochkin, A. Asenjo-Garcia, and E. Razueva, "Rotating beams in isotropic optical system," *Opt. Express*, Vol. 18, 3568–3573, 2010.
6. Abramochkin, E. and V. Volostnikov, "Spiral light beams," *Phys. Usp.*, Vol. 47, No. 12, 1177–1203, 2004.
7. Serna, J. and J. M. Movilla, "Orbital angular momentum of partially coherent beams," *Opt. Lett.*, Vol. 26, 405–407, 2001.
8. Alieva, T. and M. J. Bastiaans, "Orthonormal mode sets for the two-dimensional fractional Fourier transformation," *Opt. Lett.*, Vol. 32, No. 10, 1226–1228, 2007.
9. Alieva T. and M. J. Bastiaans, "Mode mapping in paraxial lossless optics," *Opt. Lett.*, Vol. 30, 1461–1463, 2005.
10. Allen, L., M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, "Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes," *Phys. Rev. A*, Vol. 45, 8185–8189, 1992.
11. Simpson, N. B., K. Dholakia, L. Allen, and M. J. Padgett, "Mechanical equivalence of spin and orbital angular momentum of light: An optical spanner," *Opt. Lett.*, Vol. 22, No. 1, 52–54, 1997.

Light Propagation in Tapered Optical Fibers: Spatial Light Confinement and Generation of Plasmonic Waves

A. Hartung, F. Wirth, and H. Bartelt

Institute of Photonic Technology, Albert-Einstein-Str. 9, Jena 07743, Germany

Abstract— Light propagation in optical waveguides is fundamentally different from free space propagation since light spreading by diffraction is avoided. The light is confined to modes which stay constant in shape along perfect waveguides and which can propagate with low attenuation for long propagation lengths. In the case of rotationally symmetric fiber waveguides the fundamental mode exists, in principle, for any core diameter, i.e., also for small core diameters even below the scale of the optical wavelength. However, the shape of the modes is changing greatly in this range of core diameters and coupling to non-guiding modes becomes possible. Under realistic assumptions concerning the homogeneity of the waveguide, a cut-off condition for the guided wave propagation can be derived theoretically and observed experimentally. Additional reflective coatings (metallic layers) on the outside of the fiber guiding structure would support better confinement of the modes, but such metallic components forming part of a waveguide entail stronger attenuation. Therefore the possibilities of further light confinement are limited. At the metallic-dielectric interface plasmonic waves (plasmonic surface polaritons) can be generated under specific conditions, which can be observed as wavelength-dependent attenuation peaks in the light spectrum transmitted. Such plasmonic wave generation is well-known from planar metallic-dielectric interfaces. Specific conditions apply to the circular fiber waveguide geometry. The strength of such attenuation peaks is strongly dependent not only on the refractive index conditions but also on the geometrical parameters of the waveguide. The characteristics of propagating modes and their coupling under the conditions of small-scale fiber diameters and in combination with additional metallic layers will be discussed theoretically and supported by experimental results.

1. INTRODUCTION

Tapered optical fibers with diameters on the sub wavelength scale have been studied extensively in recent years concerning their unique properties compared to standard optical fibers with tens or hundreds of micrometers in diameter. They have for instance been utilized for sensing [1] and in nonlinear applications [2] as well as proposed as atom trapping and guiding schemes [3].

There has been a controversy about strongly increasing losses observed when the fiber diameter is reduced below a specific value. According to [4], these losses are due to the scattering of growing evanescent fields by surface contaminants and could be prevented by clean surfaces. Sumetsky traced these losses back to input and output losses, which therefore occur in the tapered regions with explicitly varying diameter [5]. We show that these losses mainly occur in the smallest diameter region, even if they are perfectly clean and high demands concerning homogeneity are satisfied.

If confined guiding of light on smaller structures is desired, Surface Plasmon Polariton (SPP) waves at metallic dielectric interfaces have to be used [6, 7]. SPPs are electromagnetic waves propagating along the interface of a material with a positive and one with a negative dielectric constant (i.e., a metal). They are strongly confined to the interface and can be excited under specific conditions. In case of metal-coated fiber tapers with a diameter of a few microns, excitation can be realized via the coupling of the guided fiber modes with the SPP modes [8, 9]. In contrast to infinite plane interfaces, the mode field of nano scaled plasmonic fiber modes is inherently confined in two dimensions due to the circular symmetry. We show that cylindrical interfaces guide a mode at arbitrarily small diameters with increasing confinement. Unfortunately, ohmic losses in the metal cause high propagation losses of the modes making such fibers feasible only for short-ranged applications.

2. MODE CONFINEMENT IN DIELECTRIC OPTICAL FIBERS

Confinement of light is limited in dielectric optical fibers. The mode field diameter (MFD) can be shrunk to some extent by reducing the diameter of the fiber core. In the case of a pure silica fiber in air, the evolution of the MFD of the fundamental mode HE_{11} as a function of the fiber diameter d is shown in Figure 1. Since the refractive index of silica is nearly independent of the wavelength and therefore the refractive index difference between the silica core and the air cladding does not

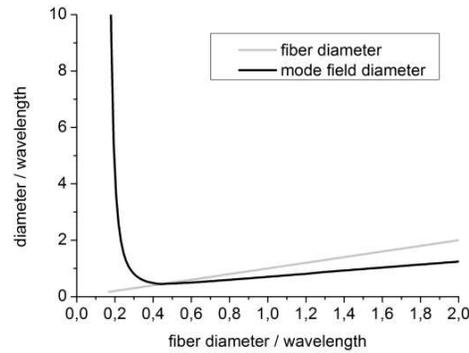


Figure 1: Evolution of the mode field diameter of the fundamental mode HE_{11} of a pure silica fiber in air.

change with wavelength, the calculated MFD values are valid not only for one specific wavelength λ but for specific ratios d/λ . The MFD can be reduced down to a value of $d_{MF} = 0.45\lambda$ at a fiber diameter of $d = 0.44\lambda$. Down to this value the MFD scales quite well with the fiber diameter. For smaller fiber diameters the fiber is now longer able to confine the electromagnetic field due to diffraction effects; it extends rapidly into the cladding and the MFD is enlarged over several orders of magnitude. Within this diameter range increasing losses can be observed accompanying the mode enlargement.

Theory predicts that the fundamental mode of a dielectric optical waveguide is always bound to it and can be guided without additional losses. Particularly with regard to the core size of an optical fiber, it can be arbitrarily small without affecting the guiding abilities concerning the fundamental mode. In experiments, optical waveguides cannot keep this promise. Figure 2 shows the transmission of an optical fiber taper at a wavelength of 1550 nm while its diameter is reduced to sub wavelength scale. It can be seen that transmission stays nearly at maximum down to a diameter of 400 nm. Below this value a sharp drop within a small diameter range can be observed. As threshold diameter d_{th} between lossless and lossy transmission behavior the diameter at the 50% transmission level is observed; in the case investigated, d_{th} amounted to 390 nm/1550 nm = 0.25λ .

This behavior contradicting theory can be explained when realistic assumptions concerning the homogeneity of the waveguide are taken into account since theory accounts for perfect cylindrical fibers only. Small distortions lead to coupling effects between the bound fundamental mode of the fiber with propagation constant β_1 and the radiation modes of the surrounding cladding with propagation constant β_2 . The energy that is transferred to the radiation modes propagates away from the fiber and leads to a decrease in transmission intensity. Based on the length scale criterion for adiabatic fiber tapers [10], a criterion concerning the local taper length scale can be derived [11] that marks the boundary between adiabatic and non-adiabatic values. This boundary can also be regarded in terms of the local taper angle $\Omega = d(\beta_1 - \beta_2)/4\pi$ as shown in Figure 3 which illustrates the dramatic threshold behavior. The fiber taper of Figure 2 was drawn at constant local taper angle of $5 \cdot 10^{-5}$, which leads to a transition from adiabatic to non-adiabatic behavior around a fiber diameter of $d = 0.25\lambda$ as observed in the experiment. This taper angle roughly equals a variation by one atomic layer per micrometer fiber length. But even in the case of a very smooth fiber with a local taper angle matching a radius variation of one atomic layer per meter pushes the threshold diameter down to only $d_{th} = 0.17\lambda$.

3. GENERATION OF PLASMONIC WAVES

When optical waves are to be guided along smaller structures and with smaller MFDs, surface modes at dielectric-metallic interfaces have to be used instead of the volume modes discussed before.

Two conditions have to be fulfilled in the same wavelength region to achieve the excitation of SPPs. First, the propagation constants of the surface and volume modes have to be equal or at least similar to each other, what is referred to as phase-matching condition. Furthermore, the field distributions of the two modes coupling with each other have to be similar, which is the case when the surface mode loses its bound nature and the effective index falls below the refractive index of the fiber cladding. The pure surface mode becomes a hybrid mode. This is called the mode-matching condition.

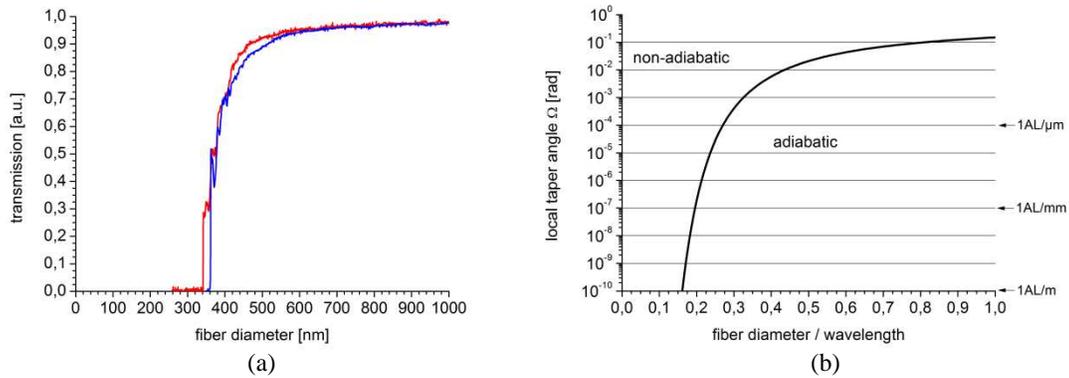


Figure 2: (a) Transmission of two independent optical fiber tapers as a function of the fiber diameter at 1550 nm wavelength. (b) Boundary between adiabatic and non-adiabatic taper angles that indicates the threshold behavior for small fiber diameters, nearly independent from the smoothness of the fiber the experimental transition to non-adiabatic behavior is around $d = 0.2\lambda$.

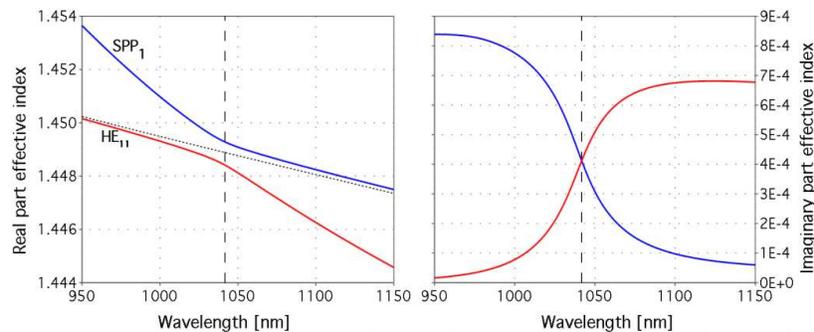


Figure 3: Real and imaginary part of the mode indices near the resonance wavelength for HE_{11} and SPP_1 .

The following example in Figure 3 shows the real and the imaginary part of the first radial order surface mode SPP_1 and the fundamental mode HE_{11} of the fiber. The fiber has a radius of $7.5\ \mu\text{m}$ and is coated with a gold layer having a thickness of 40 nm. As the surrounding dielectric medium, an index matching oil with a nominal refractive index of 1.428 was used.

The dotted line shown in Figure 3 is the fundamental mode of the fiber taper in the absence of the metal. This source mode couples into one or both of the two other modes. In the short wavelength region the source mode initially couples almost completely into the volume mode. Coming closer to the resonance wavelength, this portion decreases, while the amount of light coupled into the surface mode increases. At the resonance wavelength, finally both amounts are close to equal.

On the long-wavelength side of the resonance, the behavior is different. The original surface mode develops to a hybrid mode and finally to a pure volume mode (HE_{11}), into which the source mode couples completely if the wavelength is sufficiently larger than the resonance wavelength. The original volume mode HE_{11} develops into a different mode of the same azimuthal and radial order — the EH_{11} . The resonance wavelength as well as the shape (depth and width) of the resonance dips depends on both the geometrical parameters like the fiber radius or thickness of the gold layer and the refractive index of the surrounding medium. Due to the thickness variation of real gold layers, the experimentally observed resonance curves which appear as dips in the transmission spectra are always broader than the calculated perfect values. Figure 4 shows experimentally obtained transmission spectra for immersion in different index oils. The fiber was tapered to a radius of $5.2\ \mu\text{m}$ and afterwards coated with a gold layer 35.5 nm thick. The measured resonance wavelengths were well in agreement with the calculated ones.

The transversal magnetic (TM) mode with the electric field component perpendicular to the metal-dielectric-interfaces exhibits an increasing effective mode index with decreasing fiber diameter as shown in Figure 4(b). Since the field in the cladding drops as $\exp(-|k_r||r|)$ with $k_r = \sqrt{\beta^2 - k^2}$, mode confinement increases with the reduction of the fiber diameter. The increase in propagation constant is accompanied by an increase in propagation losses. The well-known trade-off in plasmonic waveguides between mode confinement and losses also applies to plasmonic fibers and limits their

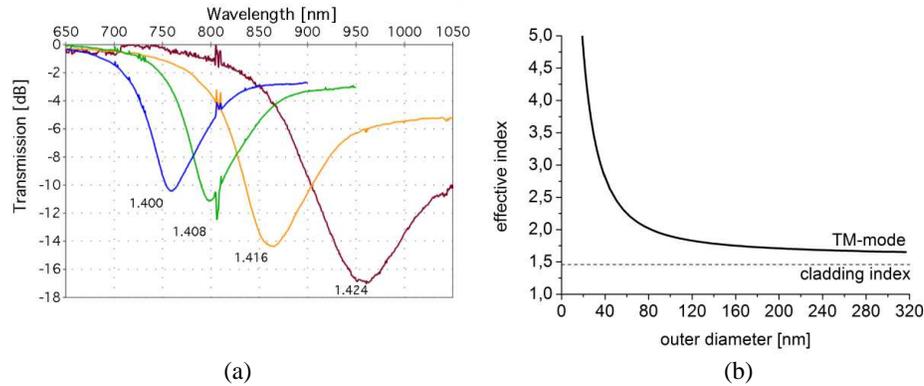


Figure 4: (a) Transmission spectra of a metal-coated fiber taper for several surrounding dielectrics. (b) Calculated evolution of the mode index of the TM mode of a silver coated fiber in index matching oil at 633nm wavelength.

application length to a few tens of micrometers.

4. CONCLUSIONS

Dielectric optical fibers are suited for mode confinement down to a diameter of 0.44λ . Further mode confinement is possible in plasmonic fibers where the propagation constant of the fundamental TM mode increases with the reduction of the fiber diameter.

ACKNOWLEDGMENT

Funding by the Thuringian Ministry of Education, Science and Culture is gratefully acknowledged.

REFERENCES

1. Polynkin, P., A. Polynkin, N. Peyghambarian, and M. Mansuripur, "Evanescent field-based optical fiber sensing device for measuring the refractive index of liquids in microfluidic channels," *Opt. Lett.*, Vol. 30, No. 11, 1273–1275, 2005.
2. Leon-Saval, S. G., T. A. Birks, W. J. Wadsworth, P. S. J. Russell, and M. W. Mason, "Supercontinuum generation in submicron fibre waveguides," *Opt. Express*, Vol. 12, No. 13, 2864–2869, 2004.
3. Balykin, V. I., K. Hakuta, and F. Le Kien, J. Liang, and M. Morinaga, "Atom trapping and guiding with a subwavelength-diameter optical fiber," *Phys. Rev. A*, Vol. 70, No. 1, 011401, 2004.
4. Tong, L., R. R. Gattass, J. B. Ashcom, S. He, J. Lou, M. Shen, I. Maxwell, and E. Mazur, "Subwavelength diameter silica wires for low-loss optical wave guiding," *Nature*, Vol. 426, 6968, 816–819, 2003.
5. Sumetsky, M., Y. Dulashko, P. Domachuk, and B. J. Eggleton, "Thinnest optical waveguide: experimental test," *Opt. Lett.*, Vol. 32, No. 7, 754–756, 2007.
6. Burke, J. J. and G. I. Stegmann, "Surface-polariton-like waves guided by thin, lossy metal films," *Phys. Rev. B*, Vol. 33, No. 8, 5186–5201, 1986.
7. Maier, S. A., "Plasmonics: The promise of highly integrated optical devices," *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 12, No. 6, 1671–1677, 2006.
8. Díez, A., M. V. Andrés, and J. L. Cruz, "Hybrid surface plasma modes in circular metal-coated tapered fibers," *J. Opt. Soc. Am. A*, Vol. 16, No. 12, 2978–2982, December 1999.
9. Themistos, C., B. M. Rahman, M. Rajarajan, K. Kalli, and K. T. V. Grattan, "Characterization of surface-plasmon modes in metal-clad optical waveguides," *Appl. Opt.*, Vol. 45, No. 33, 8523–8530, 2006.
10. Love, J. D., W. M. Henry, W. J. Stewart, R. J. Black, S. Lacroix, and F. Gonthier, "Tapered single-mode fibres and devices. Part 1: Adiabaticity criteria," *IEE Proc.*, Vol. 138, 343–354, 1991.
11. Hartung, A., S. Brueckner, and H. Bartelt, "Limits of light guidance in optical nanofibers," *Opt. Ex.*, Vol. 18, No. 4, 3754–3761, 2010.

First-order Optical Systems: Radon-Wigner Transform Approach

Genaro Saavedra and Walter D. Furlan

Optics Department, Universitat de València, c/ Dr. Moliner 50, E46100 Burjassot, Spain

Abstract— The Radon-Wigner transform is presented as a tool for the description of 1st-order optical systems. The input/output relationships for this phase-space representation are obtained and their application in analysis and design tasks is pointed out.

1. INTRODUCTION

Phase-space representation of signals is a well-established tool for describing physical systems. Although the use of phase-space joint distributions is widely spread, when they are interpreted as joint probability densities some difficulties arise, namely: their conjugate coordinates are non commutative and cannot be simultaneously specified with absolute accuracy, they present complex or negative values, and they may spread in regions of the phase space where either the signal or its Fourier transform vanish. As an alternative, the projections (*marginals*) of the phase-space distributions are strictly positive and they give information about the signal on both phase-space variables. These projections can be formally associated with probability functions, avoiding all interpretation ambiguities associated with the original phase-space distributions. This is the case of the Radon-Wigner transform (RWT), defined by the complete set of projections of the Wigner distribution function in phase-space.

The RWT was first introduced for the analysis and synthesis of frequency modulated time signals and it is a relatively new formalism in Optics [1, 2]. However, it has found several applications in this field during the last years. Among them, there are the analysis of diffraction patterns, the computation of merit functions of optical systems or the tomographic reconstruction of optical fields.

In this contribution, the RWT is presented as a tool for the description of 1st-order optical systems. The input/output relationships for this phase-space representation are obtained and their application in analysis and design tasks is pointed out. In particular, the use of this representation in the efficient computation of diffraction Fresnel patterns is presented in this manuscript.

2. THE RADON-WIGNER TRANSFORM

The Radon transform of a function of (at least) two coordinates is defined as the set of its generalized marginals for different projection angles. When we consider the Wigner distribution function $W_f(x, \xi)$ associated with a given signal $f(x)$, the generalized marginals define the Radon-Wigner transform (RWT) of this signal. Mathematically,

$$RW_f(x_\theta, \theta) = \begin{cases} \int_{-\infty}^{+\infty} W_f\left(x, \frac{x_\theta}{\sin\theta} - \frac{x}{\tan\theta}\right) dx & \text{for } \theta \neq 0, \frac{\pi}{2} \\ \int_{-\infty}^{+\infty} W_f(x_\theta, \xi) d\xi & \text{for } \theta = 0 \\ \int_{-\infty}^{+\infty} W_f(x, x_\theta) dx & \text{for } \theta = \frac{\pi}{2} \end{cases} \quad (1)$$

where x_θ and y_θ are the Cartesian coordinates rotated by the projection angle θ . An alternative expression can be obtained, in terms of the original signal, as follows

$$RW_f(x_\theta, \theta) = \begin{cases} \left| \frac{1}{\sin\theta} \int_{-\infty}^{+\infty} f(x) \exp\left(i\pi \frac{x^2}{\tan\theta}\right) \exp\left(-i2\pi \frac{x_\theta x}{\sin\theta}\right) dx \right|^2 & \text{for } \theta \neq 0, \frac{\pi}{2} \\ |f(x = x_\theta)|^2 & \text{for } \theta = 0 \\ \left| F\left(\xi = x \frac{\pi}{2}\right) \right|^2 & \text{for } \theta = \frac{\pi}{2} \end{cases}, \quad (2)$$

where $F(\xi)$ stands for the Fourier transform of $f(x)$. From Eqs. (1) and (2), the relationships of RWT with other phase-space representations can be established, as summarized in Fig. 1.

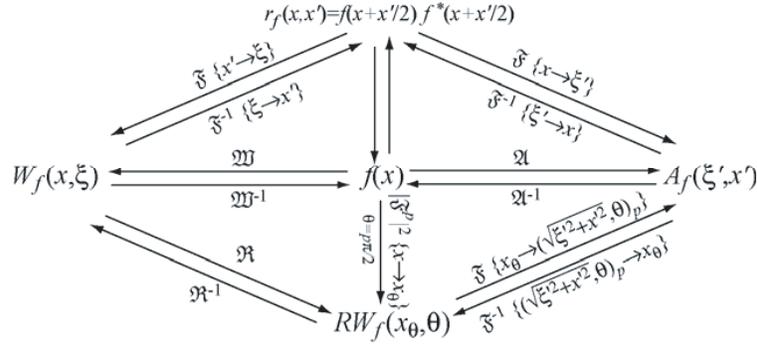


Figure 1: Relationship diagram between the original signal $f(x)$ and different phase-space representations. \mathcal{F} , \mathcal{F}^p , \mathcal{W} , \mathcal{A} and \mathcal{R} stand for Fourier transform, fractional Fourier transform, Wigner distribution function, Ambiguity function and Radon transformation, respectively, while \bullet^{-1} represents the corresponding inverse operator.

3. TRANSFORMATION THROUGH CANONICAL FIRST-ORDER SYSTEMS

The transformation of the RWT of a signal $f(x)$ after passing through a first-order optical system is of particular interest. In this case, the input signal undergoes a canonical transformation defined through four real parameters (a, b, c, d) in such a way that the transformed signal $g(x)$ is given by

$$g(x) = \begin{cases} \frac{1}{\sqrt{ib}} \exp\left(\frac{-i\pi dx'^2}{b}\right) \int_{-\infty}^{+\infty} f(x') \exp\left(\frac{-i\pi ax'^2}{b}\right) \exp\left(\frac{i2\pi}{b} x x'\right) dx' & b \neq 0 \\ \exp\left(\frac{-i\pi cx'^2}{a}\right) \frac{1}{\sqrt{a}} f\left(\frac{x}{a}\right) & b = 0 \end{cases}. \quad (3)$$

Nonabsorbing systems are implicitly considered here, which leads to the constraint $ad - bc = 1$. By applying the definition in Eq. (2), it is straightforward to obtain

$$RW_g(x_\theta, \theta) = \begin{cases} \int_{-\infty}^{+\infty} W_f\left(ax + b\left(\frac{x_\theta}{\sin\theta} - \frac{x}{\tan\theta}\right), cx + d\left(\frac{x_\theta}{\sin\theta} - \frac{x}{\tan\theta}\right)\right) dx & \text{for } \theta \neq 0, \frac{\pi}{2} \\ \int_{-\infty}^{+\infty} W_f(ax_\theta + b\xi, cx_\theta + d\xi) d\xi & \text{for } \theta = 0 \\ \int_{-\infty}^{+\infty} W_f(ax + bx_\theta, cx + dx_\theta) dx & \text{for } \theta = \frac{\pi}{2} \end{cases} \propto RW_f(x_{\theta'}, \theta'), \quad (4)$$

where the mapped coordinates are given by

$$\tan\theta' = -\frac{a \tan\theta - b}{c \tan\theta - d}, \quad x_{\theta'} = \frac{x_\theta}{(a \sin\theta - b \cos\theta)} \sin\theta'. \quad (5)$$

This is a very interesting result, since the values of the new RWT can be obtained from the original RWT at mapped coordinates given by Eq. (5). Special attention is usually paid to the cases $\theta = 0, \pi/2$ since, according to Eq. (2), the modulus squared of the canonical transformation in Eq. (3) and its Fourier transform are then obtained, respectively.

Many useful applications of this result can be addressed. Let us consider, as an example, free space (Fresnel) propagation of a 1D spatial coherent-light distribution $f(x)$, with wavelength λ , along a distance z . Application of the mapping law in Eq. (5) leads to

$$RW_g(x_\theta, \theta) \propto RW_f(x_{\theta'}, \theta'), \quad \tan\theta' = \tan\theta - \lambda z, \quad x_{\theta'} = \frac{x_\theta}{(\sin\theta + \lambda z \cos\theta)} \sin\theta'. \quad (6)$$

Finally, by taking $\theta = 0$ the irradiance diffraction pattern $|g(x)|^2$ is obtained. In terms of the RWT of the diffracting amplitude, we find

$$RW_g(x_0, 0) = |g(x_0)|^2 \propto RW_f(x_{\theta'}, \theta'), \quad \tan\theta' = -\lambda z, \quad x_{\theta'} = \frac{x_0}{\lambda z} \sin\theta', \quad (7)$$

namely, that the irradiance magnitude at any position (x_0, z) in the diffraction volume of the structure $f(x)$ is given by a particular value of its RWT. Thus, this phase-space representation provides all the information, properly rearranged, about any diffraction pattern of the original object. Moreover, since λ appears as a parameter in the mapping law in Eq. (7), this result is also valid for any value of the wavelength of the field. In this way, analysis and design of diffractive systems, even under polychromatic illumination, can be efficiently performed by means of this representation. Many other applications, as spatial-frequency analysis of imaging systems or tomographic reconstruction of complex optical fields, can be successfully attained by use of RWT representation. Furthermore, the optical implementation of the RWT for complex signals opens new possibilities in all-optical efficient optical information processing systems [3].

ACKNOWLEDGMENT

Authors acknowledge the financial support of the Spanish *Ministerio de Ciencia e Innovación* (grant FIS2009-9135) and *Generalitat Valenciana* (grant PROMETEO/2009/077).

REFERENCES

1. Wood, J. C. and D. T. Barry, "Radon transformation of time-frequency distributions for analysis of multicomponent signals," *Proc. Int. Conf. Acoust. Speech Signal Process.*, Vol. 4, 257–261, 1992.
2. Wood, J. C. and D. T. Barry, "Linear signal synthesis using the Radon-Wigner transform," *IEEE Trans. Signal Process.*, Vol. 42, 2105–2111, 1994.
3. Furlan, W. D. and G. Saavedra, "The Radon-Wigner transform in analysis, design and processing of optical signals," *Phase Space Optics: Fundamentals and Applications*, M. Testorf, B. Hennelly and J. Ojeda-Castañeda, eds., McGraw-Hill, New York, 2010.

Wigner Based Analysis of Geometric Related Resolution Degradation and Geometric Super Resolution Configurations

Zeev Zalevsky

School of Engineering, Bar-Ilan University, Ramat-Gan 52900, Israel

Abstract— Geometric related resolution limitations of an imaging system can be divided into two bounds. The first is related to the density of the spatial sampling and the second to the sampling itself which is not ideal since the pixels perform spatial averaging of the light impinging on their area. Geometric super resolving approaches aim to overcome both limitations by proper encoding and decoding of the spatial information or its spectrum prior to being sampled by the detector and by performing over sampling using, e.g., micro scanning.

In this paper we use the Wigner transform to represent the two above mentioned geometric resolution limits and to mathematically analyze and to understand them in the Wigner domain before and after applying geometrical super resolution approaches.

1. INTRODUCTION

Geometric related resolution limitations of an imaging system can be divided into two bounds. The first is related to the density of the spatial sampling, i.e., how densely the analog spatial distribution of the intensity is being sampled by the detector. This is directly affected by the pitch of the pixels in the detection array. According to Nyquist criteria the maximal spectral bandwidth that may be reconstructed is inversely proportional to this pitch. Higher spectral bandwidths are being spectrally overlapped due to under sampling effect.

The second bound is related to the fact that the sampling is not ideal, i.e., it is not done with a two dimensional train of delta functions but rather with pixels performing spatial averaging of the light impinging on their area. The larger the fill factor of each pixel is the less ideal is the sampling. The non ideal sampling is resulted in low pass filtering in the spectral domain.

Geometric super resolving approaches aim to overcome both limitations by proper encoding and decoding of the spatial information or its spectrum prior to being sampled by the detector. The under sampling effect by itself may also be reduced by performing a spatial micro scanning procedure [1] in which a set of images with relative sub pixel shifts is captured. Improved geometrical resolution can be obtained by interlacing and applying digital interpolation.

The second bound which is related to the spatial shape of each sampling pixel is more problematic than the pitch issue. There are several approaches to resolves this difficulty. The most common one is to apply inverse filtering by various means [2–7]. Other approaches involve either spatial or spectral encoding of information while the spatial encoding involves positioning of random binary mask in the intermediate image plane [8, 9] and the spectral encoding uses mask that is positioned in the aperture plane [10, 11].

The idea in using the space encoding random mask is not to modify or reshape the point spread function of the pixel's responsivity but rather to represent the problem of the spatial-blurring function as a global operation over the entire image. This is achieved with the aid of linear algebra, by describing the operation as a vector multiplied by a matrix. The problem is that there are more variables than equations in such a representation. The purpose of the encoding random mask is to add equations (since at the spatial positions where the binary mask blocks the light, we have a priori knowledge of the anticipated readout) and to have an equal number of equations and variables so that the global-matrix-inversion operation can be realized with high mathematical stability and with low sensitivity to the noise in the image. A high-resolution image is obtained after inverting the blurring matrix.

In this paper we use the Wigner transform to represent the two above mentioned geometric resolution limits. Then, we use Wigner to represent the geometric super resolving approaches that involve applying random masks for encoding as well as micro scanning for increasing the spatial sampling density.

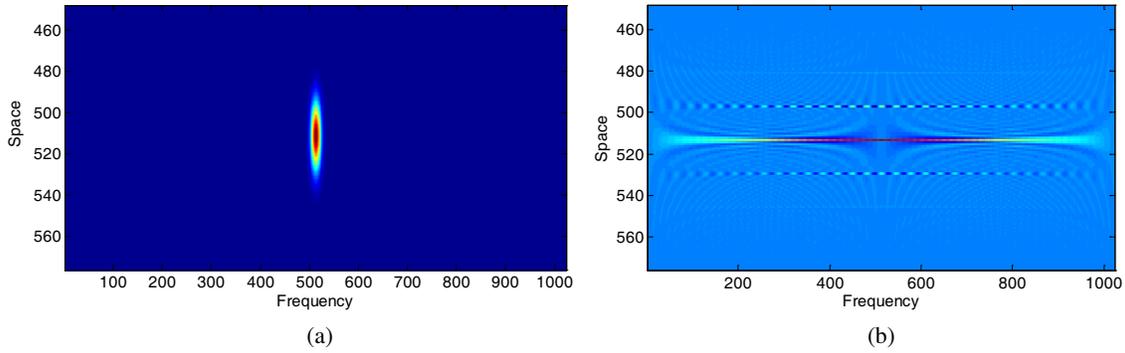


Figure 1: (a) Wigner transform of a Gaussian signal. (b) Wigner transform of under sampled signal.

2. MATHEMATICAL DERIVATION

2.1. Wigner of Sampled Signals

The Wigner distribution of a sampled signal can be computed as follows. We will denote the sampled version of $f(x)$ by $f_s(x)$:

$$f_s(x) = f(x) \cdot \sum_{n=-\infty}^{\infty} \delta(x - n\delta x) = \sum_{n=-\infty}^{\infty} f(n\delta x) \delta(x - n\delta x) \quad (1)$$

where δx is the spatial sampling interval and $\delta(x)$ is the delta function of Dirac. The Wigner distribution of the sampled signal of Eq. (1) is:

$$\begin{aligned} W_{f_s}(x, \nu) &= \int f_s\left(x + \frac{x'}{2}\right) f_s^*\left(x - \frac{x'}{2}\right) \exp(-2\pi i \nu x') dx' \\ &= \sum_n \sum_m f(n\delta x) f^*(m\delta x) \delta\left(x - \left(\frac{n+m}{2}\right)\delta x\right) \exp(-2\pi i \nu (n-m)\delta x) \end{aligned} \quad (2)$$

Let us assume for our numerical simulation a Gaussian signal of:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \Delta x)^2}{2\sigma^2}\right) \quad (3)$$

where σ is the standard deviation of the Gaussian and Δx is its average. We will assume in our Matlab simulation a signal with 1024 pixels while Δx equals to 512 pixels. The standard deviation we chose was 32 pixels. The result of the Wigner simulation is seen in Fig. 1(a). When performing under sampling the resulted Wigner distribution can be seen in Fig. 1(b). There the sampling period δx was of 32 pixels. One may see that the spectral distribution of Fig. 1(b) is indeed extended in comparison to the one of Fig. 1(a) due to the (overlapping) spectral replications. Although over the spectral axis the distribution is very wide, the high resolution information is lost since the different spectral replication overlap and thus superimposed/mixed.

2.2. Wigner of Micro-scanned Signals

Micro scanning by factor of M can resolve the under sampling problem since it increases the density of the spatial sampling. Denoting the micro scanned signal as $g_M(x)$ yields:

$$\begin{aligned} g_M(x) &= \sum_{m=0}^{M-1} \left[\sum_{n=-\infty}^{\infty} f\left(n\delta x - m\frac{\delta x}{M}\right) \delta\left(x - n\delta x - m\frac{\delta x}{M}\right) \right] \\ &= \sum_{n=-\infty}^{\infty} f\left(n\frac{\delta x}{M}\right) \delta\left(x - n\frac{\delta x}{M}\right) \end{aligned} \quad (4)$$

The Wigner distribution of the micro scanned signal equals to:

$$W_{g_M}(x, \nu) = \sum_n \sum_m f\left(\frac{n}{M}\delta x\right) f^*\left(\frac{m}{M}\delta x\right) \delta\left(x - \left(\frac{n+m}{2M}\right)\delta x\right) \exp\left(-2\pi i \nu (n-m)\frac{\delta x}{M}\right) \quad (5)$$

In the simulation of Fig. 2 we used $M = 16$ such that the sampling period was effectively reduced to $\delta x/M = 2$ pixels. In the Wigner distribution of Fig. 2 one may see how the under sampling is being replaced with over sampling since now the replications along the frequency axis are no longer overlapped as before (in the simulated over sampling the distance between spectral replication is larger than the bandwidth of the signal).

2.3. Non Ideal Sampling

The non ideal sampling is basically equivalent to blurring the analog function with the spatial responsivity of the pixels of the detector prior to being ideally sampled. Denoting by $r(x)$ the spatial distribution of the responsivity of the pixels yields the following relation for the not ideally sampled signal $f^{(b)}(x)$:

$$f^{(b)}(x) = \int f(x')r(x-x')dx' \quad (6)$$

After the blurring the signal is ideally sampled which yields:

$$f^{(b)}(x) = \sum_{n=-\infty}^{\infty} f^{(b)}(n\delta x)\delta(x-n\delta x) = \sum_{n=-\infty}^{\infty} \int f(x')r(n\delta x-x')\delta(x-n\delta x)dx' \quad (7)$$

The Fourier transform of the not ideally sampled signal equals:

$$\begin{aligned} F_{f_s}^{(b)}(\nu) &= (F(\nu)R(\nu)) \otimes \frac{1}{\delta x} \sum_n \delta\left(\nu - \frac{n}{\delta x}\right) \\ &= \frac{1}{\delta x} \sum_n \int F(\nu')R(\nu')\delta\left(\nu - \nu' + \frac{n}{\delta x}\right) d\nu' = \frac{1}{\delta x} \sum_n F\left(\nu + \frac{n}{\delta x}\right)R\left(\nu + \frac{n}{\delta x}\right) \end{aligned} \quad (8)$$

where \otimes denotes convolution operation and the capital letters $F(\nu)$ and $R(\nu)$ stand for the Fourier transform of the spatial signals $f(x)$ and $r(x)$:

$$F(\nu) = \int f(x)\exp(-2\pi i x \nu) dx \quad R(\nu) = \int r(x)\exp(-2\pi i x \nu) dx \quad (9)$$

Thus, the non ideal sampling yields low pass filtering in the Fourier domain, being imposed by $R(\nu)$. Therefore, the improvement of anti aliasing being obtained before due to the micro scanning is now being lost due to this low pass filter. The Wigner transform of the not ideally sampled signal equals to:

$$\begin{aligned} W_{f_s^{(b)}}(x, \nu) &= \int F_{f_s}^{(b)}\left(\nu + \frac{\nu'}{2}\right)F_{f_s}^{(b)*}\left(\nu - \frac{\nu'}{2}\right)\exp(2\pi i \nu' x) d\nu' \\ &= \frac{1}{\delta x^2} \int \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} F\left(\nu + \frac{\nu'}{2} + \frac{n}{\delta x}\right)R\left(\nu + \frac{\nu'}{2} + \frac{n}{\delta x}\right) \\ &\quad \cdot F^*\left(\nu - \frac{\nu'}{2} + \frac{m}{\delta x}\right)R^*\left(\nu - \frac{\nu'}{2} + \frac{m}{\delta x}\right)\exp(2\pi i \nu' x) d\nu' \end{aligned} \quad (10)$$

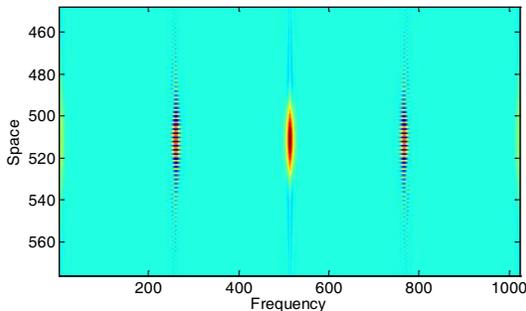


Figure 2: Wigner transform of a sampled signal after performing micro scanning.

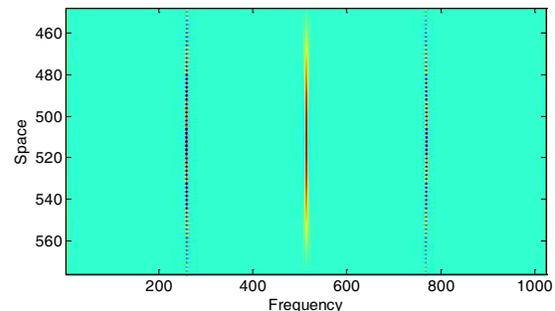


Figure 3: Wigner transform of not ideally sampled signal after performing micro scanning.

Note that in this case we used the definition of the Wigner based upon the Fourier transform of the signals (unlike the definition of Eq. (2) where the spatial distribution of the signal was used to compute the Wigner transformation).

In Fig. 3, we present the numerical simulation of the not ideally sampled signal $f(x)$. One may see how the low pass effect of the non ideal sampling reduces the high frequency content (the replicated spectral information is much narrower along the frequency axis). In the simulation of Fig. 3 we blurred the spatial Gaussian $f(x)$ with spatial *rect* function $r(x)$ having width of 128 pixels and which represents detection pixels having dimension of 128 pixels each. The performed micro scanning was such that $\delta x/M = 2$ pixels. One may see in Fig. 3 that indeed due to the micro scanning the spectral replications are sufficiently separated but due to the non ideal sampling each replication is low passed and thus it contains less spectral frequencies than the original signal.

3. GEOMETRIC SUPER RESOLUTION

The geometric super resolving configuration which overcomes the non ideal sampling of the pixels of the detection array involves addition of a spatially random encoding mask positioned in the intermediate image plane and which encodes the analog light distribution prior to its being sampled and micro scanned (to increase the obtainable spatial sampling density). Usage of spatial random encoding prior to non ideal sampling can recover the loss we obtained in Fig. 3 for the high frequencies. We assume that our high resolution spatial encoding mask is denoted by $p(x)$ yielding the following not ideally sampled signal (after being encoded):

$$f_s^{(b,e)}(x) = \left[\int (f(x') p(x')) r(x - x') dx' \right] \sum_{n=-\infty}^{\infty} \delta(x - n\delta x) \quad (11)$$

and its Fourier transform equals to:

$$\begin{aligned} F_{f_s}^{(b,e)}(\nu) &= \left[\int (F(\nu') P(\nu - \nu')) R(\nu) \right] \otimes \frac{1}{\delta x} \sum_{n=-\infty}^{\infty} \delta\left(\nu - \frac{n}{\delta x}\right) \\ &= [(F(\nu') \otimes P(\nu - \nu')) R(\nu)] \otimes \frac{1}{\delta x} \sum_{n=-\infty}^{\infty} \delta\left(\nu - \frac{n}{\delta x}\right) \end{aligned} \quad (12)$$

where \otimes denotes convolution operation. This results with:

$$\begin{aligned} F_{f_s}^{(b,e)}(\nu) &= \frac{1}{\delta x} \sum_{n=-\infty}^{\infty} \int \int F(\nu'') P(\nu' - \nu'') R(\nu') \delta\left(\nu - \nu' + \frac{n}{\delta x}\right) d\nu' d\nu'' \\ &= \frac{1}{\delta x} \sum_n R\left(\nu + \frac{n}{\delta x}\right) \int F(\nu'') P\left(\nu + \frac{n}{\delta x} - \nu''\right) d\nu'' \end{aligned} \quad (13)$$

By applying the Wigner definition based upon the Fourier transform of signals, we may compute the resulting Wigner distribution of $f_s^{(b,e)}(x)$:

$$\begin{aligned} W_{f_s^{(b,e)}}(x, \nu) &= \frac{1}{\delta x^2} \sum_n \sum_m R\left(\nu + \frac{\nu'}{2} + \frac{n}{\delta x}\right) R^*\left(\nu - \frac{\nu'}{2} + \frac{m}{\delta x}\right) \iiint F(\nu_1) P\left(\nu + \frac{\nu'}{2} + \frac{n}{\delta x} - \nu_1\right) \\ &\quad \cdot F^*(\nu_2) P^*\left(\nu - \frac{\nu'}{2} + \frac{m}{\delta x} - \nu_2\right) \exp(2\pi i \nu' x) d\nu_1 d\nu_2 d\nu' \end{aligned} \quad (14)$$

Since now the spatial signal is encoded before being low passed, the information is not lost but rather spread over the entire spectral domain. Right after, each replication (which contains the full spectral information due to the encoding) after solving the aliasing problem via the micro scanning may be used to decode the resolution information of the original high resolution signal $f(x)$. For the simulation of Fig. 4(a) we used binary random high resolution encoding mask having the Wigner transform seen in Fig. 4(b). The micro scanning over sampling generated sampling period of $\delta x/M = 2$ pixels. One may see how the information is spread over the entire Wigner domain due to the high resolution binary encoding mask.

Since the Wigner definition may either be computed using spatial signals (Eq. (2)) or the Fourier transform of signals (see Eq. (10)), performing the encoding over the Fourier (spectral) domain rather than over the space may result in similar outcome of spreading rather than losing the high resolution information. In the spectral encoding one needs to perform decoding as well in order to recover the mixed information. From Fig. 4 one may see that unlike in Fig. 3, the information is spread over the entire phase space however its spreading is not a simple replication but rather expresses the encoding that was applied. This encoding (and spreading) will also use us in order to perform the decoding and to obtain the information reconstruction.

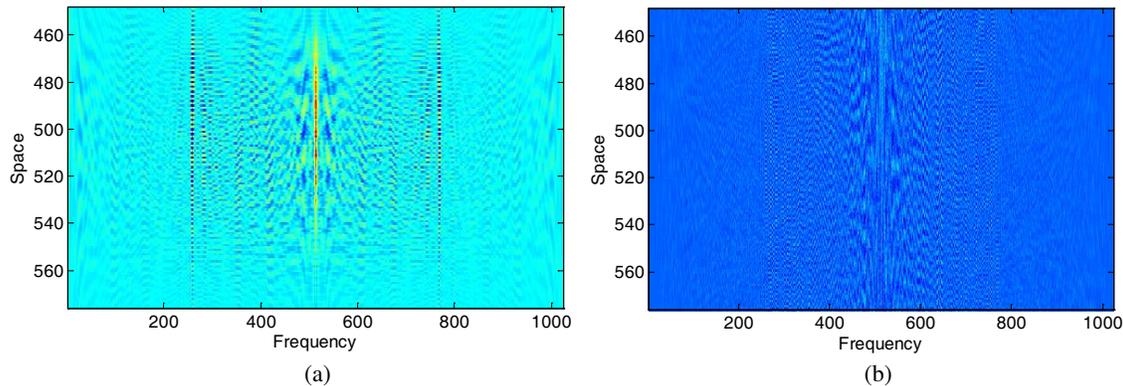


Figure 4: (a) Wigner transform of encoded signal after performing non ideal sampling. (b) Wigner transform of the spatial binary random encoding mask.

4. CONCLUSIONS

In this paper we used the Wigner transform to represent two major geometry related image resolution limiting factors: the density of the spatial sampling grid and the spatial responsivity of each sampling pixel. We analyzed both limitations using the Wigner domain and simulated their outcome. Then we used the Wigner distribution to simulate geometric super resolving techniques.

REFERENCES

1. Fortin, J., P. Chevrette, and R. Plante, "Evaluation of the microscanning process," *Proc. SPIE*, Vol. 2269, 271–279, 1994.
2. Tekalp, A. M., M. K. Ozkan, and M. I. Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space varying image restoration," *Proceedings IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Vol. 3, 169–172, 1992.
3. Kim, S. P., N. K. Bose, and H. M. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 38, 1013–1027, 1990.
4. Bose, N. K., H. C. Kim, and H. M. Valenzuela, "Recursive total least squares algorithm for image reconstruction from noisy, undersampled frames," *Multidimensional Systems and Signal Processing*, Vol. 4, 253–268, 1993.
5. Kim, S. P. and W. Y. Su, "Recursive high-resolution reconstruction of blurred multiframe images," *IEEE Transactions on Image Processing*, Vol. 2, 534–539, 1993.
6. Hardie, R. C., K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Transactions on Image Processing*, Vol. 6, 1621–1633, 1997.
7. Stark, H. and P. Oskoui, "High-resolution image recovery from image-plane arrays using convex projections," *J. Opt. Soc. Am.*, Vol. A6, 1715–1726, 1989.
8. Borkowski, A., Z. Zalevsky, and B. Javidi, "Geometrical super-resolved imaging using non-periodic spatial masking," *J. Opt. Soc. Am.*, Vol. A36, 589–601, 2009.
9. Gehm, M. E. and M. D. Stenner, "Compressive measurement for target tracking in persistent, pervasive surveillance applications," *Proc. SPIE*, Vol. 7468, 2009.
10. Zalevsky, Z., J. Solomon, and D. Mendlovic, "Geometrical super resolution using code-division multiplexing," *Appl. Opt.*, Vol. 42, 32–40, 2005.
11. Ashok, A. and M. A. Neifeld, "Pseudorandom phase masks for superresolution imaging from subpixel shifting," *Appl. Opt.*, Vol. 46, 2256–2268, 2007.

Partially Coherent Ambiguity Functions for Depth-variant Point Spread Function Design

Roarke Horstmeyer¹, Se Baek Oh², Otkrist Gupta¹, and Ramesh Raskar¹

¹Media Lab, MIT, 75 Amherst St., Cambridge, MA 02139, USA

²Department of Mechanical Engineering, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Abstract— The ambiguity function (AF) provides a convenient way to model how a camera with a modified aperture responds to defocus. We use the AF to design an optimal aperture distribution, which creates a depth-variant point spread function (PSF) from a sparse set of desired intensity patterns at different focal depths. Prior knowledge of the coherence state of the light is used to constrain the optimization in the mutual intensity domain. We use an assumption of spatially coherent light to design a fixed-pattern aperture mask. The concept of a dynamic aperture mask that displays several aperture patterns during one image exposure is also suggested, which is modeled under an assumption of partially coherent light. Parallels are drawn between the optimal aperture functions for this dynamic mask and the eigenmodes of a coherent mode decomposition. We demonstrate how the space of design for a 3D intensity distribution of light using partially coherent assumptions is less constrained than under coherent light assumptions.

1. INTRODUCTION

Modifying a conventional camera with an aperture mask and a matched post-processing step can extend imaging capabilities into a variety of domains, potentially facilitating depth estimation, object tracking, extended depth-of-field or even super-resolution. Recently, an iterative “mode-selective” method of designing aperture masks using a sparse set of desired PSF intensity distributions at different planes of focus was presented [1]. The set of PSFs is used to generate a desired ambiguity function (AF), which can completely represent a camera’s response to defocus. The mode-selection algorithm converges to a valid AF solution that minimizes the mean-squared error between its associated PSFs and the input set of desired PSFs. This AF solution directly yields the optimal aperture mask amplitude and phase distribution to place at the camera pupil plane. If the set of desired PSFs are not physically realizable from a thin mask (i.e., do not obey the constraints of propagation), the algorithm converges to a nearby solution that is realizable.

The mode-selection algorithm is among a class of techniques, including phase retrieval [2], phase space tomography [3] and transport-of-intensity [4], which determine the amplitude and phase of a wavefront from multiple intensity distributions at planes along the direction of propagation. Unlike other methods, mode-selection applies a global constraint to the solution set in the mutual intensity domain, based on a singular value decomposition (SVD). In [1], a spatially coherent mutual intensity function J_c is used as a constraint to determine a fixed aperture mask pattern. Following, we present a method to extend the mode-selection algorithm to operate with a dynamic aperture mask, which can display a set of M amplitude and phase distributions over the course of one image exposure. Each mask in the set is determined from a unique orthogonal mode of the coherent mode decomposition of J . These modes can be found from the M largest singular values of the SVD of J . Since the desired AF is comprised of more than one orthogonal mutual intensity mode, it is equivalent to the optimizing the AF and J of a partially coherent system.

2. MODE-SELECTION FOR APERTURE MASK DESIGN

It is well known that the AF can model the response of an imaging system as a polar display of its optical transfer functions (OTFs) at different planes of defocus [5]. Specifically, if we consider the simplified 1D imaging setup in Fig. 1 with a 1D plane wave $U(x)$ at aperture coordinate x , then the 2D AF of the setup is given by,

$$AF(x', u) = \int U\left(x + \frac{x'}{2}\right) U^*\left(x - \frac{x'}{2}\right) e^{-2\pi i x u} dx \quad (1)$$

where x and x' are the space and spatial frequency coordinates at the pupil plane, respectively, u is a second parameter proportional to defocus, and the asterisk represents complex conjugation. The 1D PSF of the camera at different planes of defocus along z can be transformed into a 1D OTF

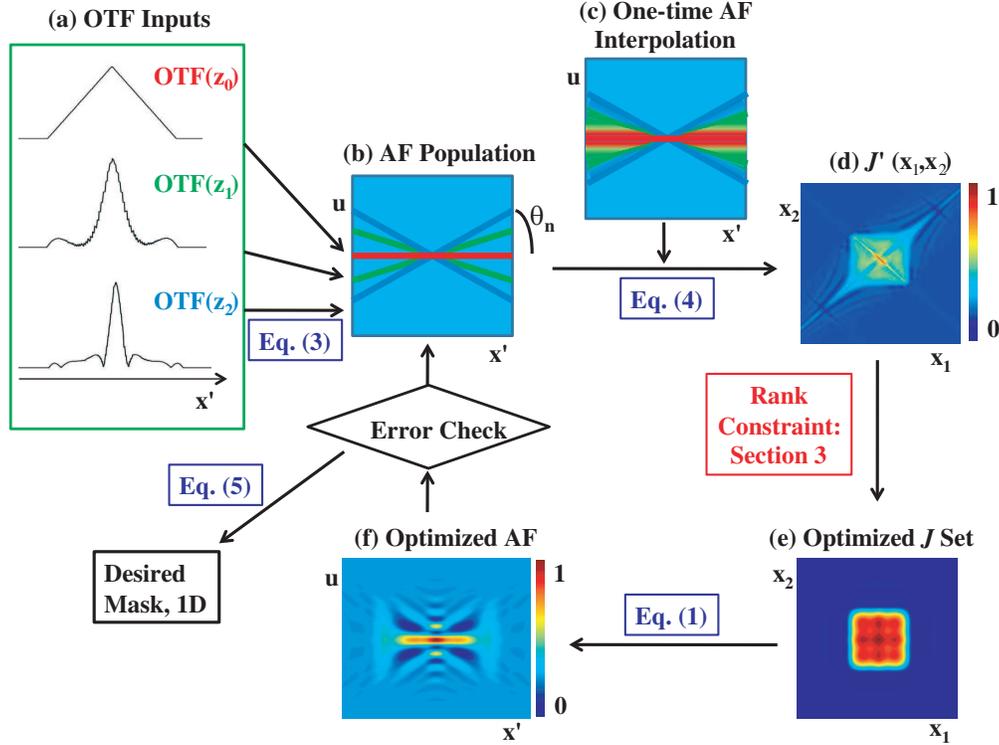


Figure 2: A schematic diagram of the mode-selection algorithm operating in 1D. (a) A set of n desired OTFs (here $n = 3$ for an open aperture), which are determined from desired PSF responses, are used as input. (b) Each OTF populates a slice of the AF, then filled in with a one-time linear interpolation (c). (d) The mutual intensity (J) can be constrained by taking its singular value decomposition shown in (e). (f) An optimized AF is now obtained, which is re-populated with the desired OTF values in (a) along the specific slices in (b). Iteration is stopped at a specified error value, and Eq. (3) is then used to invert the AF into the optimal 1D aperture mask.

Furthermore, from [7], we know that an optimized discrete *coherent* mutual intensity J_c must fulfill a rank-1 condition. A good rank-1 approximation is given by the first singular value of the SVD in Eq. (4),

$$J_c(x_1, x_2) = \sum_{i=1}^1 s_i \lambda_i v_i = \lambda_1 |s_1\rangle \langle v_1|. \quad (5)$$

In other words, to fulfill the coherence constraint implicit in a PSF measurement, we can represent J_c as the outer-product between the first column of $S(s_1)$, and the first row of $V(v_1)$. Since a spatially coherent wave is composed of a single mutual intensity mode, all singular values besides λ_1 are 0. This constraint reduces our redundant 2D phase space representation to the 2 1D vectors s_1 and v_1 , which are equal if J is positive semi-definite.

3.2. Applying a Partially Coherent Constraint

The choice to use the SVD of $J(x_1, x_2)$ to obtain a coherent mutual intensity is based upon the well-known coherent mode decomposition [8], which states that the mutual intensity of a source of any coherence state can be represented as a summation of N mutually orthogonal coherent modes. The SVD provides just such a decomposition into mutually orthogonal components, each with a specific weight λ_i , for any complex 2D matrix. Thus, we can extend the coherence decomposition in Eq. (4) to a given degree of partial coherence by adding up the first M singular values of the SVD of $J(x_1, x_2)$, where $M < N$:

$$J_{pc}(x_1, x_2) = \sum_{i=1}^M s_i \lambda_i v_i = \sum_{i=1}^M U_i(x_1) U_i^*(x_2) \quad (6)$$

From the Eckart-Young Theorem, it is clear $J_{pc}(x_1, x_2)$ is an optimal approximation of $J(x_1, x_2)$, since J_{pc} is the rank- M approximation of J with minimized Euclidean error [9]. Furthermore, since

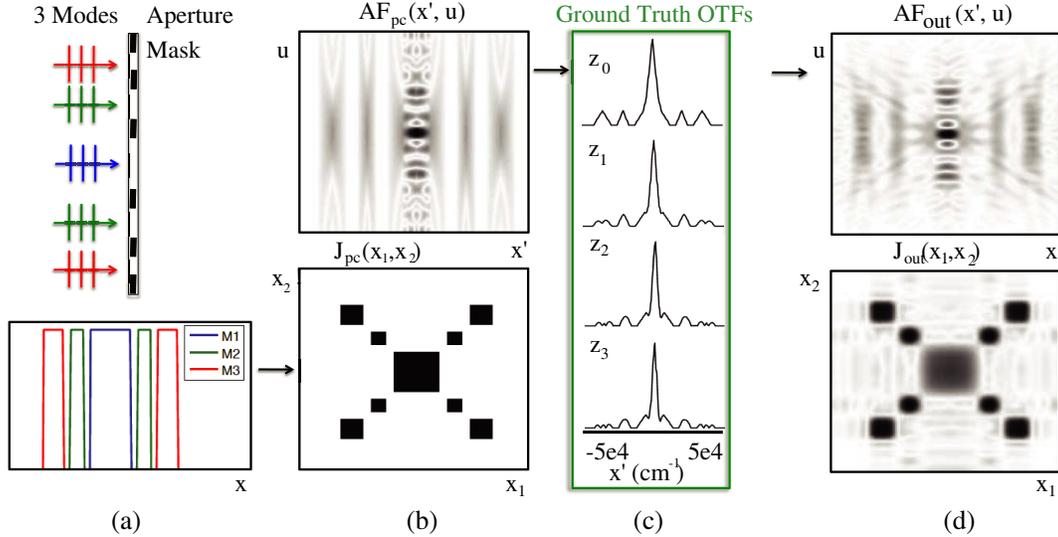


Figure 3: Demonstration of the mode-selection algorithm applied to a known partially coherent input. (a) 3 mutually incoherent plane waves strike an aperture at different locations, creating 3 orthogonal coherent modes. (b) The partially coherent AF_{pc} and mutual intensity J_{pc} (rank 3) for this scenario (maximum is black). (c) 4 OTF slices from AF_{pc} , corresponding to 4 different defocus planes along z , are used as algorithm input. (d) The optimized AF_{out} and J_{out} .

each mode is orthogonal, $J_{pc}(x_1, x_2)$ consists of a sum of M unique, coherent mutual intensities. Implementing this new constraint in the mode-selection process, $J_{pc}(x_1, x_2)$ will be transformed each iteration into an AF of a partially coherent source,

$$AF_{pc}(x', u) = \int J_{pc} \left(x + \frac{x'}{2}, x - \frac{x'}{2} \right) e^{-2\pi i x u} dx = \sum_{i=1}^M \lambda_i AF_i(x', u) \quad (7)$$

Here we also express AF_{pc} as a summation over coherent, orthogonal AF_i 's, which follows from a similar property of the AFs Fourier dual, the Wigner distribution [10]. Each AF_i obeys Eq. (1) for an orthogonal mode $U_i(x)$. Eq. (7) demonstrates that optimizing for a partially coherent AF is equivalent to simultaneously optimizing M coherent, mutually orthogonal AFs that must be added to create a desired set of input PSF intensities. These M coherent AFs will provide a set of M aperture masks at the algorithm's output, each weighted by its associated singular value λ_i . This summation of AFs to achieve a desired response is directly connected to the longstanding problem of OTF synthesis, studied earlier by Marechal [6]. As with synthesizing OTFs, one way to implement the summation in Eq. (7) is to multiplex each coherent mode over time, which has previously been proposed to simulate partially coherent illumination [11]. Specifically, the fixed aperture mask in Fig. 1(a) can be replaced with a dynamic screen, like a spatial light modulator (SLM), which can display the aperture mask associated with each coherent mode for a finite amount of time over the duration of one image exposure. The length of time each mode is displayed will be proportional to its singular value λ_i . We will now demonstrate that multiple aperture masks can offer additional flexibility in creating a desired depth-dependent PSF as compared to a single aperture mask.

4. PARTIALLY COHERENT AMBIGUITY FUNCTIONS FOR DESIGN

To incorporate partial coherence effects into the algorithm described in Section 2, only the rank constraint connecting Fig. 2(d) and Fig. 2(e) must be modified with Eq. (6). For a given input set of PSFs and a desired number of coherent modes M , this simple change will allow mode-selection to find M optimal weighted aperture mask functions. In most cases, these M aperture functions will lead to a depth-dependent PSF that is a closer match to the desired input as compared with the PSF created by a single aperture function.

First, as a demonstration of the mode selection algorithm's ability to accurately converge, performance is tested for a set of ground-truth OTFs that are known to obey the constraints of partially coherent propagation. This is equivalent to testing mode-selection's ability to recreate an entire

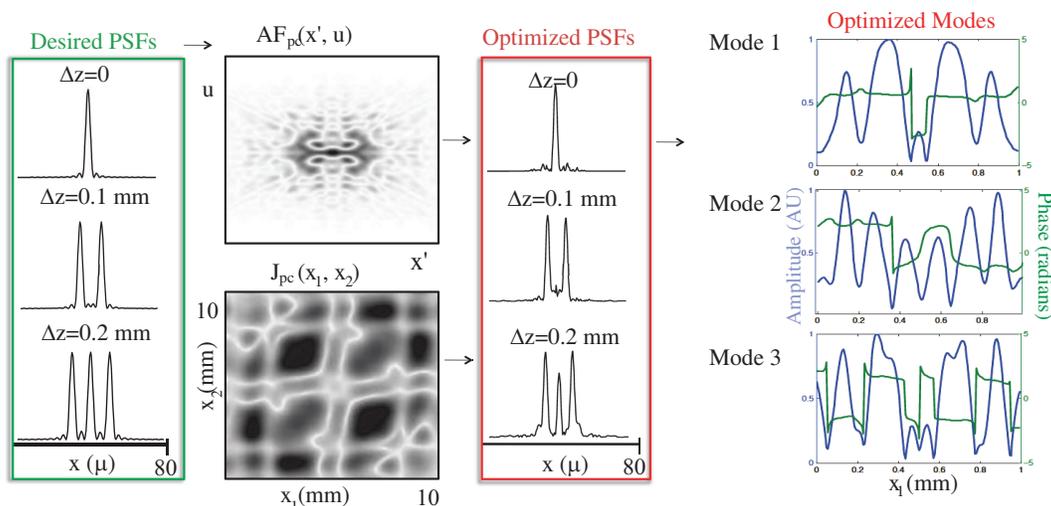


Figure 4: The mode-selection algorithm applied to determining an optimal aperture mask set from a desired depth dependent PSF (left), with $M = 3$. The algorithm yields 3 1D mask functions (right), each for 1 mode of J .

partially coherent AF from a few OTF inputs. Here, each OTF will be a sum of the OTFs produced by a set of known aperture masks at a given defocus plane. Fig. 3(b) displays an example partially coherent J_{pc} and AF_{pc} used to generate 4 input OTFs: in-focus, at $W_{20} = 0.5\lambda$, at $W_{20} = \lambda$ and at $W_{20} = 1.5\lambda$. For a 10 mm mask and a lens with 50 mm focal length, this corresponds roughly to 0.2 mm, 0.4 mm and 0.6 mm of sensor defocus, respectively. A faithful reproduction is achieved after 50 algorithm iterations. The performance metric of mean-squared error (MSE) from desired OTFs is 8×10^{-4} , which is on the same order as coherent-only operation. MSE is defined as the normalized squared difference between model input (i.e., ground truth OTFs) and output. Errors can be attributed to the limited maximum angle of the OTF slices (similar to the limited-angle problem in tomographic reconstruction), as well as a large amount of rapid changes from 3 overlapping central AF cross-terms in this particular example.

Second, the partially coherent rank-selection algorithm is applied to the problem of designing a desired PSF response (Fig. 4). Specifically, a set of desired input PSF intensities that may not obey the constraints of propagation are used as input (i.e., one point turning into two and then three points at three defocus planes). The algorithm iterates to find an optimal rank- M mutual intensity function, which corresponds to a set of M aperture masks that can be shown over time to recreate this desired PSF set. While 3 modes worked well in the example ($MSE = 5.4 \times 10^{-3}$), M can be varied depending upon system specifics to typically (but not always) provide a better estimate for more modes. In this example, 3 modes performs much better than 1 mode, which is used as an example in [1] and offers an $MSE = 0.032$.

5. CONCLUSION

Implementing the mode-selection algorithm under a partially coherent framework allows for additional flexibility in the design of depth-dependent PSFs. Future work will consider the tradeoffs involved with optimizing over a different number of modes, an experimental implementation of a dynamic aperture mask, and alternative ways to constrain the decomposition of J to enhance performance.

ACKNOWLEDGMENT

We would like to thank Zhengyun Zhang for helpful discussions and suggestions. This work has been supported in part by a National Defense Science and Engineering Graduate (NDSEG) fellowship awarded through the Air Force Office of Scientific Research. Ramesh Raskar is supported by an Alfred P. Sloan Research Fellowship and a DARPA Young Faculty Award.

REFERENCES

1. Horstmeyer, R., S. B. Oh, and R. Raskar, "Iterative aperture mask design in phase space using a rank constraint," *Opt. Express*, Vol. 18, 22545–22555, 2010.

2. Fienup, J. R., “Iterative method applied to image reconstruction and to computer generated holograms,” *Opt. Eng.*, Vol. 19, 297–305, 1980.
3. Raymer, M. G., M. Beck, and D. F. McAlister, “Complex wave-field reconstruction using phase-space tomography,” *Phys. Rev. Lett.*, Vol. 72, No. 8, 1137–1140, 1994.
4. Waller, L., L. Tian, and G. Barbastathis, “Transport of intensity phase-amplitude imaging with higher order intensity derivatives,” *Opt. Express*, Vol. 18, 12552–12561, 2010.
5. Brenner, K. H., A. W. Lohmann, and J. Ojeda-Castaneda, “The ambiguity function as a polar display of the OTF,” *Opt. Commun.*, Vol. 44, 323–326, 1983.
6. Goodman, J. W., *Introduction to Fourier Optics*, Chap. 6, McGraw-Hill, 1982.
7. Ozaktas, H. M., S. Yuksel, and M. A. Kutay, “Linear algebraic theory of partial coherence: Discrete fields and measures of partial coherence,” *JOSA A*, Vol. 19, No. 8, 1563–1571, 2002.
8. Wolf, E., M., “New theory of partial coherence in the space-frequency domain. Part I: Spectra and cross spectra of steady-state sources,” *JOSA*, Vol. 72, No. 3, 1982.
9. Hogben, L. M., *Handbook of Linear Algebra*, Chap. 5, Chapman and Hall, New York, 2007.
10. Bastiaans, M. J., “Application of the Wigner distribution function to partially coherent light,” *JOSA A*, Vol. 3, No. 8, 1227–1238, 1986.
11. De Santis, P., F. Gori, G. Guattari, and C. Palma, “Synthesis of partially coherent fields,” *JOSA A*, Vol. 3, No. 8, 1258–1262, 1986.

Complex Amplitude Filters for Extended Depth of Field

J. Ojeda-Castañeda, E. Yezpez-Vidal, and E. Ayala
University of Guanajuato, Salamanca, Guanajuato 36885, México

Abstract— We present a complex amplitude transmittance filter that has a modulation transfer function with low sensitivity to focus errors, and with reduced oscillations. We use linear regression for identifying a tendency line, and in this way we define the inverse filter. Departures of the tendency line are treated as Euclidian distances in hyperspace.

1. INTRODUCTION

Certain nonconventional masks are able to reduce the influence of focus error on the Modulation Transfer Function (MTF); and in this manner, one can extend the depth of field of an optical system [1–12]. When extending the depth of field, there are two main challenges. First, if one wishes to preserve lateral resolution (associated to the full pupil aperture), then one avoids the trivial solution of closing down the pupil. Second, if one wishes to preserve light gathering power, then one must use phase-only masks.

Here, we propose the use of an optical element that uses both amplitude variations and phase variations, for generating a MTF with low sensitivity to focus errors, and with reduced oscillations. The proposed filter preserves lateral resolution, and exhibits a stable MTF as compared with previously proposed filters for extending the depth of field.

In Section 2, we show that by combining the use of a mask with Gaussian amplitude transmittance, it is possible to preserve resolution, while reducing the influence of focus error on the MTF. Furthermore, the Gaussian apodizer reduces the oscillations of the MTF. In Section 3, we discuss the slight reduction in light throughput. In Section 4, we propose the use of tendency curve, of several MTFs. Any departures of the tendency MTF are treated as Euclidian distances in hyperspace. For the sake of clarity our discussion is 1-D; and we assume that the imaging process has unit magnification.

2. THE INFLUENCE OF FOCUS ERRORS ON THE MTF

We denote the generalized pupil function as

$$P(\mu; W) = Q(\mu) \exp \left(i2\pi W \left(\frac{\mu}{\Omega} \right)^2 \right). \quad (1)$$

In Eq. (1) the Greek letter μ represents the spatial frequency; and Ω denotes the cut-off spatial frequency of the pupil aperture. The function $Q(\mu)$ describes the complex amplitude transmittance of the pupil aperture. We use W as a shorthand notation for representing the 1-D version of Hopkins focus error coefficient, $W_{2,0}$, in units of wavelength λ . That is, $W = (W_{2,0}/\lambda)$. The optical transfer function, $H(\mu; W)$, is the normalized version, $(1/N)$, of the auto-correlation of the generalized pupil function $P(\mu; W)$

$$H(\mu; W) = \left(\frac{1}{N} \right) \int_{-\infty}^{\infty} Q \left(\nu + \frac{\mu}{2} \right) Q^* \left(\nu - \frac{\mu}{2} \right) \exp \left(i2\pi \left(\frac{2W\mu}{\Omega^2} \right) \nu \right) d\nu. \quad (2)$$

The MTF is the modulus of Eq. (2). For the cubic phase mask, in Reference [2], the complex amplitude transmittance of the pupil aperture is

$$Q(\mu) = \exp \left(i2\pi a \left(\frac{\mu}{\Omega} \right)^3 \right) \text{rect} \left(\frac{\mu}{2\Omega} \right). \quad (3)$$

In Eq. (3) the letter “ a ” represents the maximum value of the optical path difference, in units of λ , which is introduced by the phase mask. By substituting Eq. (3) in Eq. (2), we obtain that the MTF is

$$|H(\mu; W)| = \left| \left(\frac{1}{2\Omega} \right) \int_{-\infty}^{\infty} \exp \left(i6\pi a \left(\frac{\mu}{\Omega} \right) \left(\frac{\nu}{\Omega} \right)^2 \right) \exp \left(i2\pi \left(\frac{2W\mu}{\Omega^2} \right) \nu \right) \text{rect} \left[\frac{\nu}{2\Omega - |\mu|} \right] d\nu \right|. \quad (4)$$

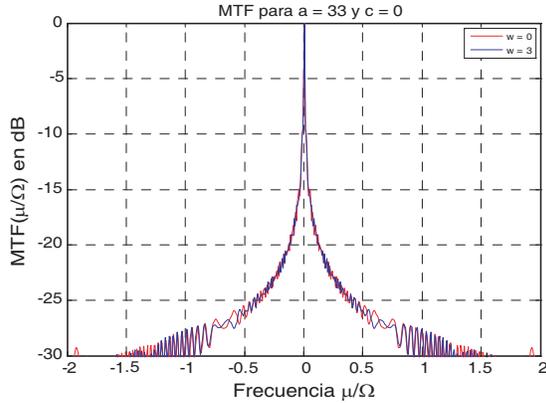


Figure 1: MTF for the cubic phase mask without Gaussian apodizer.

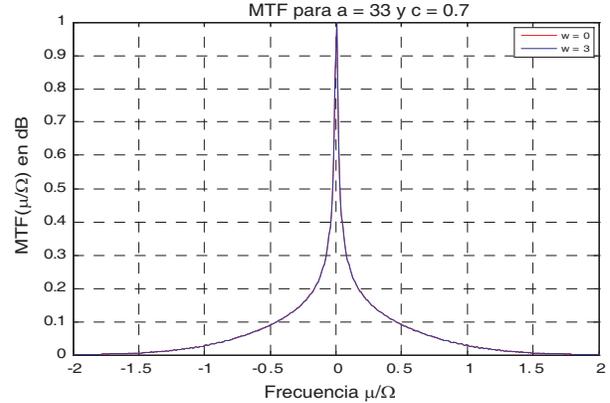


Figure 2: MTF for the cubic phase mask with Gaussian apodizer.

We note that at each value of μ , the integral in Eq. (4) can be expressed in terms of Fresnel integrals. In Figure 1, we show a MTF for the cubic phase mask, if the maximum optical path difference is $a = 33$. For the curve in red color, the coefficient of focus error is $W = 0$, and for the curve in blue color. $W = 3$. It is apparent from Figure 1 that the MTF does not exhibit zero values in its passband. However, the MTF curve shows oscillations around its tendency line. Here, for reducing the oscillations around the tendency line, we propose the use a weak Gaussian filter. Taking into account the presence of the Gaussian apodizer, the new complex amplitude transmittance of the pupil aperture is

$$G(\mu; c) = \exp\left(-2\pi c \left(\frac{\mu}{\Omega}\right)^2\right) Q(\mu). \quad (5)$$

The influence of the Gaussian filter is displayed in Figure 2, where we show a typical MTF if $a = 33$, $c = 0.7$, $W = 0$ for the red color curve, and almost completely on the top of the first curve we have $W = 3$ for the blue color curve.

It is apparent from Figure 2 that the new MTF preserves the low sensitivity to focus errors. But in addition, the new MTF has reduced oscillations around its tendency line. Admittedly, the above results are obtained by sacrificing light throughput, as we discuss next.

3. LIGHT THROUGHPUT

The use the Gaussian mask, in Eq. (5), generates a reduction in light throughput that is represented by the inte

$$T = \left(\frac{1}{2\Omega}\right) \int_{-\Omega}^{\Omega} |G(\mu; W)|^2 d\mu = \int_0^1 \exp[-4\pi c(\mu/\Omega)^2] d\left(\frac{\mu}{\Omega}\right). \quad (6)$$

From the numerical simulations in Figure 2, we suggest the use of an attenuation factor $c \approx 0.25$. Then, the light throughput decreases from unity to 0.49. In other words, by using the Gaussian filter the light throughput reduces by the same amount as when one increases f -number by one step. However, it is worth noting that here resolution is preserved.

4. TENDENCY CURVE AND EUCLIDIAN DISTANCE

In Figure 2, we note that the in focus MTF and the out-of-focus MTF are practically identical. Next, we state in a quantitative manner, the similarity if the two MTFs. To that end, we consider a family set of MTFs. Each member of the family is a MTF with variable focus error. For instance we can have a family, of say, 7 MTFs with $W = 0, 0.5, 1, 1.5, 2, 2.5, 3$.

Now, we recognize that we have a large collection of points that are specified by two coordinates: spatial frequency, and a MTF value. The collection of points can be thought of as the experimental data of a tendency curve. Then, by employing a linear regression technique, one can find a tendency MTF.

The tendency MTF has two useful properties. First, we can use the tendency MTF for evaluating the inverse filter of the family set of MTFs. Second, any departure of a real MTF from the tendency MTF can be associated an error. The sum of the squared errors is a Euclidian distances

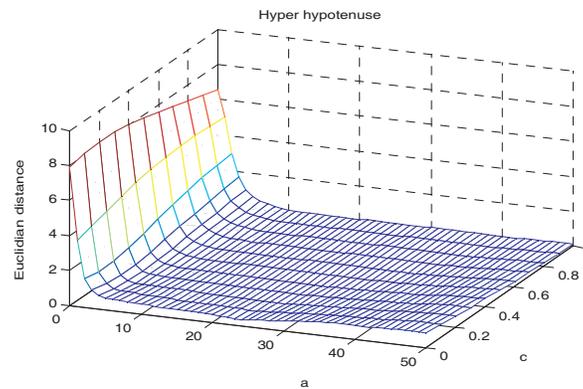


Figure 3: Hyper hypotenuse of the proposed mask.

in hyperspace. This Euclidian distance is hyper hypotenuse, which evaluates the departures of a family set of MTFs from the tendency MTF.

In other words, the hyper hypotenuse evaluates in a quantitative fashion the statement feature of low sensitivity to focus error. And consequently, the hyper hypotenuse is useful for assessing the quality of a mask that generates MTFs with low sensitivity to focus error. In Figure 3, we display the hyper hypotenuse (vertical axis) of the proposed complex amplitude mask, as a function the optical path difference “ a ” in Eq. (3), and the attenuation factor “ c ” in Eq. (5). It is apparent from Figure 3 that small values of “ a ” and “ c ” are needed to reach a steady state in the MTF curves.

5. FINAL REMARKS

We have presented a complex amplitude transmittance filter that combines the use of a cubic phase mask and a Gaussian apodizer. This complex amplitude mask reduces the impact of focus error, as well as the oscillations of the MTF curve.

We have proposed to use several MTFs, with variable focus error, for generating a family set. By using linear regression techniques, this family set can be used to identify a tendency MTF.

The tendency MTF is useful for identifying an inverse filter for the whole family set of MTFs. Furthermore, any departure of a real MTF from the tendency MTF can be related to a Euclidian hyper hypotenuse, for assessing low sensitivity to focus error.

From our numerical simulations, we conclude that the hyper hypotenuse reaches a minimum steady state value, if the parameters of the complex filter are set to the following values $a \geq 10$ and $c \approx 0.25$.

ACKNOWLEDGMENT

We are indebted to E. García-Almanza for useful numerical verifications. We gratefully acknowledge the financial support of ProMeP, SNI and the University of Guanajuato.

REFERENCES

1. Ojeda-Castañeda, J., R. Ramos, and A. Noyola-Isgleas, *Appl. Opt.*, Vol. 27, 2583, 1988.
2. Dowski, Jr., E. R. and W. T. Cathey, *Appl. Opt.*, Vol. 34, 1859, 1995.
3. Chi, W. and N. George, *Opt. Lett.*, Vol. 26, 875, 2001.
4. Saucedo, A. and J. Ojeda-Castañeda, *Opt. Lett.*, Vol. 29, 560, 2004.
5. Mezouari, S., G. Muyo, and A. R. Harvey, *Proc. SPIE*, Vol. 5249, 238, 2004.
6. Castro, A. and J. Ojeda-Castañeda, *Appl. Opt.*, Vol. 43, 3474, 2004.
7. Ares García, J., S. Bará, M. Gomez García, Z. Jaroszewicz, A. Kolodziejczyk, and K. Petelcyc, *Opt. Exp.*, Vol. 16, 18371, 2008.
8. Takahashi, Y. and S. Komatsu, *Opt. Lett.*, Vol. 33, 1515, 2008.
9. Ojeda-Castañeda, J., J. E. A. Landgrave, and C. M. Gómez-Sarabia, *Appl. Opt.*, Vol. 47, E99, 2008.
10. Mouroulis, P., *Opt. Exp.*, Vol. 16, 12995, 2008.
11. Muyo, G., A. Singh, M. Anderson, D. Huckridge, A. Wood, and A. R. Harvey, *Opt. Express*, Vol. 17, 21118, 2009.
12. Ojeda-Castañeda, J., E. Yépez-Vidal, and E. García-Almanza, *Photonics Letters of Poland*, Vol. 2, 2010, to be publish.

Temporal Similarity for Short Pulses

Jorge Ojeda-Castañeda, Cristina M. Gómez-Sarabia, and Helena E. López-Aviléz

University of Guanajuato, Salamanca, Guanajuato 36885, México

Abstract— We use a temporal cross-correlation for analyzing the similarity between the complex amplitude envelopes of two short pulses. We unveil a relationship between temporal similarity and the expectation value, as used in quantum physics. We apply this relationship for analyzing the temporal version of the fractional Talbot effect.

1. INTRODUCTION

For analyzing the propagation of modes in an optical fiber, often it is convenient to compare the near field diffraction patterns against the far field diffraction pattern. One can do a quantitative comparison of similarity, by performing a spatial cross-correlation between two diffraction patterns [1]. Lohmann and Ojeda-Castaneda have developed a definition for Fresnel similarity, which is defined as a cross-correlation between an initial complex amplitude distribution and its Fresnel diffraction pattern [2]. This latter cross-correlation finds practical applications for setting array illuminators [3, 5].

Our aim here is to analyze the temporal similarity between the complex amplitude envelopes of two short pulses. The first short pulse is at the input of a dispersive medium, and the second short pulse is at the output of a dispersive medium. In Section 2, we define a temporal similarity function as a cross-correlation between the complex amplitude envelopes of the two short pulses. And then, in Section 3, we apply this relationship for analyzing the temporal Talbot effect, at fractions of the Talbot length.

2. TEMPORAL SIMILARITY

If one neglects absorption and nonlinear effects, it is common to describe the evolution of the slowly varying complex envelope employing linear system theory tools. Let us consider the schematic diagram in Fig. 1.

At the input of a dispersive medium, the slowly varying complex envelope is represented by the Fourier series

$$u_0(t) = \sum_{m=-\infty}^{\infty} a_m \exp(itm \Omega). \quad (1)$$

Then, at a distance z along the dispersive medium, the slowly varying complex envelope is expressed as

$$u(\tau; z) = \sum_{m=-\infty}^{\infty} a_m H(m \Omega; z; \beta) \exp(i\tau m \Omega). \quad (2)$$

In Eq. (2), τ is the time measured in the proper reference frame of the pulse; we denote as $H(\omega; z; \beta)$ the transfer function at a distance z , inside the dispersive medium. In symbolic fashion, the letter β indicates a dispersion coefficient, in the Taylor series expansion of the dispersion relation around the angular frequency ω_0 .

Next, we use the concept of Fresnel similarity in reference [2] for defining temporal similarity as the cross-correlation function

$$S(\tau; z) = (1/T) \int_{-T/2}^{T/2} u(\tau + t; z) u_0^*(t) dt. \quad (3)$$

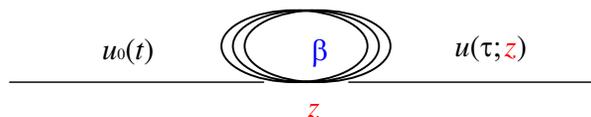


Figure 1: Schematic diagram of the optical setup.

In Eq. (3), $T = (2\pi/\omega)$. By substituting Eqs. (1) and (2) in Eq. (3), we obtain that

$$S(\tau; z) = \sum_{m=-\infty}^{\infty} |a_m|^2 H(m\Omega; z; \beta) \exp(i\tau m \Omega). \quad (4)$$

It is apparent from Eq. (4) that the temporal similarity function is periodic in time. Now, we recognize that at discrete values of time, $\tau = (n/N)T$ with N any natural number, $n = 0, 1, 2, \dots, (N-1)$, the temporal similarity function is

$$S((n/N)T; z) = \sum_{m=-\infty}^{\infty} |a_m|^2 H(m\Omega; z; \beta) \exp(2\pi mn/N). \quad (5)$$

For a first-order dispersion medium, at discrete values of time, the temporal similarity function becomes

$$S((n/N)T; z) = \sum_{m=-\infty}^{\infty} |a_m|^2 \exp[-i2\pi z (\Omega^2 \beta_2 / 4\pi) m^2] \exp(i2\pi mn/N). \quad (6)$$

For a second-order dispersion medium, Eq. (5) takes the form

$$S((n/N)T; z) = \sum_{m=-\infty}^{\infty} |a_m|^2 \exp[-i2\pi z (\Omega^3 \beta_3 / 12\pi) m^2] \exp(2\pi mn/N). \quad (7)$$

Now, it is apparent from Eqs. (6) and (7) that the temporal similarity function is also a periodic function along the z -axis. Of course, the period along the z -axis, say Z_T , is the Talbot length in the respective dispersive medium. That is, for a firstorder dispersion medium, with $\beta_2 > 0$,

$$Z_T = 4\pi / (\Omega^2 \beta_2). \quad (8a)$$

And for a secondorder dispersion medium, with $\beta_3 > 0$,

$$Z_T = 12\pi / (\Omega^3 \beta_3). \quad (8b)$$

In what follows, we discuss some illustrative applications of our previous formulation.

3. FRACTIONAL TALBOT EFFECT

For gaining physical insight, on the above definition and its mathematical expressions, we assume that at the input we have a Dirac comb. For this extreme case, the properties of the temporal similarity function are closely related to the complex amplitude envelope itself. If the input is a Dirac comb, then equation Eq. (5) becomes

$$S((n/N)T; z) = \sum_{m=-\infty}^{\infty} H(m\Omega; z; \beta) \exp(i2\pi mn/N). \quad (9)$$

Hence, the discrete version of the temporal similarity function can be thought of as a discrete impulse response [6]. Next, we note that at fractions of the Talbot length $z = Z_T/N$, for a firstorder dispersion medium, Eq. (9) becomes

$$S((n/N)T; Z_T/N) = \sum_{m=-\infty}^{\infty} \exp(-i2\pi m^2/N) \exp(i2\pi mn/N). \quad (10)$$

And for the a second-order dispersion medium, the Eq. (9) becomes

$$S(nT/N; Z_T/N) = \sum_{m=-\infty}^{\infty} \exp(-i2\pi m^3/N) \exp(-i2\pi mn/N). \quad (11)$$

We note that for describing Fresnel diffraction, Guigay [7] reported a closed formula for evaluating Eq. (10):

$$S((n/N)T; Z_T/N) = \left[\exp(-i\pi/4) / \sqrt{2N} \right] [1 + (-1)^n (i)^N] \exp(-i\pi n^2 / 2N). \quad (12)$$

Table 1: Square modulus of the temporal similarity function.

| N | If $n = 2q$ | If $n = 2q + 1$ |
|---------------|-------------|-----------------|
| 2, 6, 10, ... | 0 | $(2/N)$ |
| 4, 8, 12, ... | $(2/N)$ | 0 |
| 3, 5, 7, ... | $(1/N)$ | $(1/N)$ |

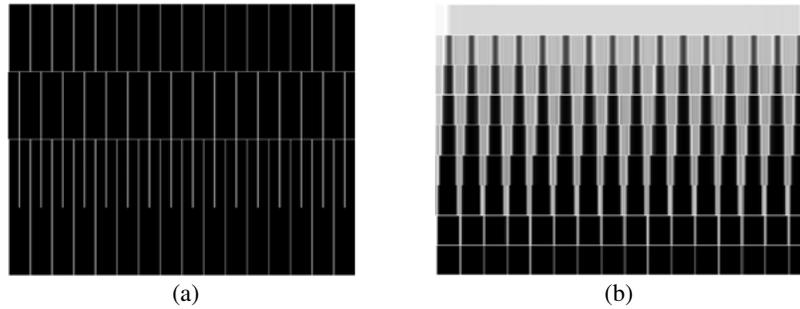


Figure 2: Evolution of the intensity distribution.

As far as we know, there is not an analytical formula for evaluating Eq. (11). Next, in Table 1 we display the values of the square modulus of Eq. (12); namely

$$|S((n/N)T; Z_T/N)^2| = (1/N) [1 + (-1)^n \cos(\pi N/2)]. \quad (13)$$

Along column one, in Table 1, we list three family sets of values of N . A Along column two, we specify the values of if the index n is an integer, even number; while along column three, the index n is an integer, odd number. It is apparent from Table 1 that if N is a an odd integer number, then has a fixed value for any $n = 0 \dots (N - 1)$. This result is in agreement to the fact that the output complex amplitude envelope exhibits $(N - 1)$ interlaced spikes, bet tween two consecutive deltas of the initial Dirac comb.

Hence, at $z = Z_T/N$, the repetition rate increases by a factor N . However, the e increment in repetition rate is associated to a power spectrum reduction by a factor $(1/N)$. In Fig. 2, we illustrate the intensity distributions associated to the evolution of a temporal (horizontal axis) Dirac's comb, as the short pulse propagates along the dispersion medium (z -axis is vertical the direction).

In Fig. 2, along the horizontal (time) axis, the window has 512 pixels, and the period is equal to 32 pixels. In Fig. 2(a), the four pictures depict the evolution from $z = 0, Z_T/4, Z_T/2,$ and Z_T . In Fig. 2(b) the eight pictures depict the evolution from $z = 0$ till $Z_T/64$ in steps of $Z_T/512$. Hence, Fig. 2 illustrates the use of phase-only modulations for incrementing the repetition rate and for setting temporal Talbot array illuminators [6, 8].

4. FINAL REMARKS

We have defined a temporal similarity function, as a temporal cross-correlation between the input complex amplitude envelope, and the output complex amplitude envelope, of the same dispersive medium. For periodic short pulses the temporal similarity function has simple expressions, which show that the temporal similarity function is periodic both in time, and along the z -axis. The period along the z -axis, Z_T , is the Talbot length in the respective dispersive medium.

For a first-order dispersion medium, the discrete version of the temporal similarity function follows the close expression reported by Guigay. We have shown that by using the sampled values of the temporal similarity function, as coefficients, one can express the output complex amplitude envelope as a finite linear superposition of the input complex amplitude envelope. The coefficients measure the similarity between the input complex amplitude envelope and the output complex amplitude envelope.

We are indebted to Adolf W. Lohmann, Pedro Andres, Jesus Lancis and Omel Mendoza-Yero, Emmanuel Yopez and Eloy García for useful discussions on this topic. The support of ProMeP, SNI and the University of Guanajuato is gratefully acknowledged.

REFERENCES

1. Feit, M. D. and J. A. Fleck, "Calculation of dispersion in graded-index multimode fibers by a propagating-beam method," *Appl. Opt.*, Vol. 18, 2843–2851, 1979.
2. Lohmann, A. W. and J. Ojeda-Castañeda, "Fresnel similarity," *Opt. Comm.*, Vol. 249, 397, 2005.
3. Lohmann, A. W., *Optik*, Vol. 79, 41, 1988.
4. Leger, J. R. and G. J. Swanson, *Opt. Lett.*, Vol. 15, 288, 1990.
5. Arrizón, V. and J. Ojeda-Castañeda, *Appl. Opt.*, Vol. 33, 5925, 1994.
6. Ojeda-Castañeda, J., P. Andrés, and O. Mendoza-Yero, *Optical Memory and Neural Networks*, Vol. 18, 260, 2009.
7. Guigay, J. P., *Optica Acta*, Vol. 18, 677, 1971.
8. Ojeda-Castaneda, J. and C. M. Gomez-Sarabia, *Photonics Letters of Poland*, Vol. 2, 2010, to be published.

Conditions for Photon-particle Interactions

Tibor Berceci

Budapest University of Technology and Economics, Hungary

Abstract— Some aspects of photon-particle interactions are investigated. Although that subject has already been treated in many publications there are still some unsolved problems. To generate a specific process by absorbing photons two conditions have to be met: beside the well known criterion that the energy of the photon has to be equal to the energy needed to the specific process another new criterion is that the length of the photon in time has to be equal to the processing time of the specific process. That means: the processing time and the processing energy are strictly connected to each other.

1. INTRODUCTION

The photon-electron interaction has already been treated in many theoretical and experimental papers [1–4]. That phenomenon was always in the center of interest in the area of optics. So far the best explanations come from quantum physics. Spontaneous emission can be calculated very accurately by quantum electrodynamics [5–7]. This way, statistical results are obtained because the process is considered to be stochastic. By other words, it is possible to calculate the probabilities of large number of interactions.

However, the photon emission can also be considered as a singular process. That is utilized in this paper because currently a single photon emission or the emission of a series of individual photons is already feasible [8–10]. Recently a single atomic layer of graphene has been created [11] making possible the investigation of a single photon-particle interaction [12].

The main goal of the present investigation is to establish some basic relationships for the process of photon-particle interactions. Their properties can be described more easily if we consider these processes in a direct band gap semiconductor material. In this paper that method is used [13, 14]. The relationships derived are in good conformity with the well-known properties of photon-particle interactions, however, they provide some new results as well.

2. PHOTON-ELECTRON INTERACTIONS

During the interaction between photons and electrons some energy transfer is carried out. The energy transfer can occur in two directions. When an electron gets from a higher energy level to a lower one, a photon can be radiated. In this case the energy difference of the electron is transferred to the generated photon. On the other hand when a photon is absorbed an electron can get to a higher energy level, the photon energy is then transferred to the electron. The first process is called photon generation, while the second process is called photon absorption.

However, these processes do not occur in every case. When an electron transits from the conductance band to the valence band its energy is reduced either by radiating a photon, or by increasing the temperature of the atom (by other words generating a “phonon”). Similarly, when a photon is absorbed an electron can take the photon energy over and as its consequence either it transits from the valence band to the conductance band; or the absorbed photon energy increases the temperature of the atom (e.g., by generating a “phonon”). Both types of photon-electron interactions have a probability which is less than unity. Presently, these probabilities are determined by statistical evaluation of experimental observations [15] and theoretical studies assuming stochastic processes. Accordingly, the probability of interaction is a characteristic of the specific material.

In order to determine some relationships for the conditions of interactions first we investigate the photon generation in a direct band gap semiconductor based on its well known process [3, 4]. Now, we consider that the generated photon has a specific frequency. The photon frequency, f_{ph} is given by the well-known equation [1, 4]:

$$f_{ph} = E_b/h \tag{1}$$

where E_b is the band gap energy, and h is the Planck constant.

3. WAVE REPRESENTATION OF A PHOTON

The photon is an electromagnetic energy burst which has a frequency given by Eq. (1). The electromagnetic burst has a very short length in time called photon length. The burst can be represented by an electric field having a short sinusoidal time function. A half period of the photon frequency seems to be the best wave representation [16]. This way the electric field component is given by the following time function:

$$A(t) = 0 \quad \text{if} \quad t < 0 \quad (2)$$

$$A(t) = A_0 \sin^2(\omega_{ph}t) \quad \text{if} \quad 0 \leq t \leq T_{ph}/2 \quad (3)$$

$$A(t) = 0 \quad \text{if} \quad t > T_{ph}/2 \quad (4)$$

Here $A(t)$ is the time varying electric field component, A_0 is its amplitude, ω_{ph} is the angular photon frequency: $\omega_{ph} = 2\pi f_{ph}$, t is the time, and T_{ph} is the period time of the photon frequency: $T_{ph} = 1/f_{ph}$.

The time function of the electric field component based on Eqs. (2), (3), (4) is shown in Fig. 1. The time function has a smooth transition from zero to maximum and then back to zero.

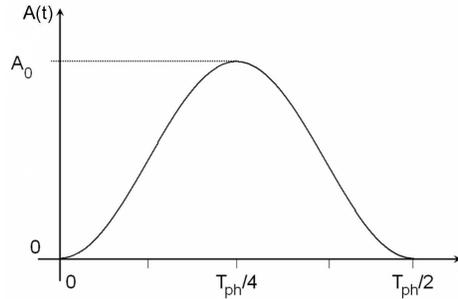


Figure 1: The time function of the electric field component representing a photon.

4. PROCESSING TIME

The time interval during which a process is carried out is called the processing time. The photon generation also has a processing time. That is equal to the transition time of the radiative electron from the conductance band to the valence band.

Based on Eqs. (2), (3), (4) the processing time τ_{ph} of photon generation is equal to the half period time of the photon frequency:

$$\tau_{ph} = \frac{T_{ph}}{2} = \frac{1}{2f_{ph}} \quad (5)$$

The photon generation time is also dependent on the band gap energy. From Eqs. (1) and (5):

$$\tau_{ph} = \frac{h}{2E_b} \quad (6)$$

When the band gap energy is higher the photon generation time is shorter, or by other words the radiated energy burst is shorter. Eqs. (5) and (6) for the photon generation time or processing time make possible to compare the photon properties at different frequencies.

The question can be arisen: is the processing time a characteristic of a specific process? The answer is: the processing time is a well determined fundamental characteristic of a process. That means the processing time is an inherent property of a specific process.

5. MANDELSTAM-TAMM UNCERTAINTY RELATION

When the processing time is determined, the Mandelstam-Tamm time-energy uncertainty relation [17–19] has to be taken into account as follows:

$$\Delta t \Delta E \geq \frac{h}{4\pi} \quad (7)$$

Here Δt is the change in time due to the ΔE change in the energy and h is the Planck constant.

In the present case E_b is the energy change and τ_{ph} is the time change. Applying Eq. (7) for that case we get:

$$\tau_{ph} = \frac{h}{2E_b} \geq \frac{h}{4\pi E_b} \quad (8)$$

Consequently, the approach for the duration of photon generation is in conformity with the Mandelstam-Tamm time-energy uncertainty relation [17].

We presented an approach for the duration of photon generation. Although we consider the photon having a specific generation time and energy it cannot be separated into smaller parts because it has been generated by a single energy transition of a radiative electron in the atom [1, 2].

6. CONDITIONS FOR INTERACTIONS

The photons can interact with the material or more precisely with the particles of the material. Due to this interaction some energy transfer can be carried out. As the photon energy is the smallest energy quantum at the photon frequency, interaction can only occur if the photon energy is equal to the energy necessary for a specific process.

However, every process has a well defined time duration which is called processing time. Therefore, the processing time is a characteristic of a specific process. The energy utilized (or delivered) in a process is called the processing energy. The processing time and the processing energy are strictly connected to each other. That means a specific atomic process has (or needs) a well defined processing energy and processing time. A specific process can be carried out only if the processing energy can be utilized (or delivered) during the processing time. We can say the processing energy and the processing time have to be matched. That relationship has many consequences.

In case of photon generation the transition of an electron from a higher energy level to a lower energy level has to occur in a time period equal to the processing time of photon generation. If the transition time of the electron is not equal to the processing time of photon generation then there is no radiation, the energy increases the temperature of the atom. This way only those excited electrons can radiate whose transition time is equal to the processing time of photon generation which is determined by the photon frequency. Some excited electrons cannot transit from the conduction band to the valence band during the processing time of photon generation due to their thermal vibrations. These excited electrons can generate only phonons increasing that way the temperature of the atom.

In the case of photon-electron interaction we have a similar situation. The photon length in time has to be equal to the processing time (i.e., the transition time) of the electron which is to be taken from the valence band to the conduction band by the interacting photon. However, not all of the absorbed photons can generate charge carriers, or by other words excited electrons. If the electron can not get into the conduction band during the processing time due to its thermal vibrations then the energy of the absorbed photon will increase the temperature of the semiconductor material.

Now the question arises: what is the reason that interaction does not occur in every case. That is due to the atomic vibrations or by other words the temperature of the particle. As a consequence of this relationship, it can be stated: in the case of photon generation the percentage of radiative electron transitions is higher when the temperature is lower. Similarly, when photons are dissipated in a direct band gap semiconductor material the percentage of generating charge carriers is higher when the temperature is lower — assuming the same conditions (same number of photons, etc.) [20].

The conditions for interactions can be explained in a different way as well. An atomic process has an “eigen” resonance frequency [12]. The interaction between the wave and the particle is only possible when the frequency of the wave is equal to the resonance frequency of the specific process. Naturally, the interaction is possible not only at the resonance frequency but also in its close vicinity which is dependent on the bandwidth of the resonance. That behavior is similar to the case when an oscillator is injection locked by an input signal [21, 22]. The resonance type description of interaction has the same meaning as the previous description because the photon frequency determines the photon energy and the processing time of photon generation as well.

As it was mentioned a process can only be carried out if the photon energy is equal to the energy needed for the specific process and the photon length in time is equal to the processing time of the specific process. That means for a specific process a well defined photon frequency is needed. As a good example we can take the photosynthesis in the leaves of the trees or vegetation. If they are in a dark room the photosynthesis is not possible although the room can be warm enough which means heat energy is available. However, the infrared photons have neither sufficient energy nor proper duration for the photosynthesis.

7. PHOTON ENERGY DENSITY

In general the photon energy is used for characterizing the effect of an electromagnetic radiation. However, it is also important to know what the concentration of the photon energy is. This aspect is not considered well enough in the literature yet.

The energy density of a process can be defined as the ratio of the energy involved in the process and the processing time. Therefore the photon energy density is obtained as the ratio of the photon energy and the photon duration or by other words the photon length in time:

$$\rho_{ph} = \frac{E_b}{\tau_{ph}} = 2hf_{ph}^2 \quad (9)$$

Here Eqs. (1) and (5) were taken into account.

The energy density of a photon is proportional to the square of the photon frequency. The energy density is significantly higher at higher frequencies due to the higher energy of the photon and also due to its shorter duration or length in time.

This quadratic equation has been derived for the case when photons are generated in a direct band gap semiconductor material. That statement does not depend on the exact value of the generation time because the same approach is used at each photon frequency.

The photon energy density is an important property of the interaction processes, especially when a photon interacts with a biological cell or tissue. The energy is linearly increasing with the frequency; however, the density of photon energy is increasing with the square of the frequency. This relationship has many consequences, first of all in biology. The danger of electromagnetic radiation is in a quadratic relation with the frequency.

8. LENGTH OF OPTICAL PULSES

The generation of a short optical pulse is a crucial task for many applications. There is a trend to generate shorter and shorter pulses. The question arises: is there a lower limitation on the pulse length? The answer is: yes. Determining that limitation we can assume that all photons of the optical pulse are generated exactly at the same time which would give the possible shortest optical pulse. The photon length in time is given by Eq. (5) and its width between the -6 dB power points is equal to the half processing time. Therefore, we can take that value as the lower limit (τ_{op}) for the optical pulse generation at a specific frequency:

$$\tau_{op} \geq \frac{\tau_{ph}}{2} = \frac{1}{4f_{ph}} \quad (10)$$

That idealistic situation cannot be realized therefore it can be stated: it is not feasible to generate optical pulses shorter than the half of the photon generation time. Consequently, if we want to generate shorter optical pulses then it can only be done at a higher frequency or by other words at a shorter wavelength.

That is in good agreement with the best published results. The shortest optical pulse which was experimentally achieved at the wavelength of 750 nm was 5.5 fs (femtosecond) [23]. At this wavelength the limitation based on Eq. (10) is 0.625 fs. Another extremely wonderful experimental result provided an 80 as (attosecond) optical pulse in the x-ray wavelength band [24]. The limitation given by Eq. (10) is about 8.33 as (attosecond) if the longest wavelength (10 nm) of the x-ray band is used for determining the limitation.

9. SECONDARY RADIATION

In the case of secondary radiation the photon participating in the interaction has a higher energy than the band gap energy. That means when a photon transfers its energy to an electron, the electron will get to a higher energy level than the band gap energy. Then the excited electron gets back to an energy level equal to the band gap energy simultaneously giving its additional energy to the atom increasing this way its temperature. Therefore the frequency of the secondary radiation will be determined by the band gap energy.

Consequently, the frequency of the secondary radiation is smaller than the frequency of the absorbed photon. Due to the above described process, there is some delay between the absorption of a photon and the emission of a secondary photon. In some cases the excited electron is trapped at a higher energy level than the band gap energy for a longer time. Accordingly, the secondary radiation will be delayed significantly.

A similar process occurs when higher frequency (higher energy) photons are used to pump the gain section of a semiconductor laser diode. The pump frequency is always higher than the operation frequency.

Naturally, there are more complex interaction processes, e.g., when two photons interact with a particle or in another case a photon, a phonon and a particle is involved into the interaction. In any case beside the processing energy the processing time has to be considered as well.

10. PASSING TIME

Concerning the nature of time several questions can arise. What is the meaning of the phrase “passing time” or “time is passing”? Is time always passing with the same speed? We can state: time is passing as changes occur in the state of the material, like changes in its energy level, location, motion, composition, etc. To perform such a change, some energy is needed. If the inherent energy of a specific physical process which is utilized to carry out the change in the state of material is higher, the change is carried out in a shorter time. Naturally, the utilization (or delivery) of energy is dependent on external conditions as well, therefore, the same conditions are assumed when the relationship between time and energy is considered.

The relationship between time and energy at atomic level can be applied to many other physical processes. Based on the afore mentioned relationship between time and energy it can be concluded: getting energy from fossil fuels by burning them the duration of the process is not short because the delivered energy from a unit amount of the material is relatively not high. However, getting energy in a nuclear plant from a radioactive material is a much faster process because the delivered energy from a unit amount of the material is rather high. Furthermore, getting energy in a fusion reactor is an extremely fast process because the delivered energy from a unit amount of the material is enormously high. Due to the very fast delivery of a high amount of energy, the control of the fusion process is very cumbersome.

Now the question arises: how can we measure the duration of a process or by other words how can we determine the time as it is passing? We use again the definition: time is passing as changes in the state of material occur (including its energy level). Therefore, we utilize some kind of energy oscillation for that purpose. Earlier a pendulum was applied to measure how time was passing. In that case periodic static and dynamic mechanical energy changes occurred. Presently, a quartz crystal oscillator is used which means periodic mechanical and electrical energy changes. The atomic oscillations provide the most stable periodicity of energy changes. Then in any of the previous cases a counter is applied and as counting proceeds time is passing. According to that method time is passing always with the same speed because the energy fluctuation in the oscillation process is kept constant.

Consequently, it can be stated the processing time which is called “internal time” in the literature [25] is an inherent property of nature. However, the universal time which is called “external time” in the literature [25] has been introduced by us; it is really a time reference.

11. CONCLUSIONS

When light is used to illuminate an object the photons interact with the material of the object. Due to that interaction the atoms or molecules of the object can undergo some changes. However, as it has been pointed out: to achieve these changes we have to consider that each process has a specific processing time and needs a specific energy. Therefore, to generate a specific process by absorbing photons two conditions have to be met: beside the well known criterion that the energy of the photon has to be equal to the energy needed for the specific process another new criterion is that the length of the photon in time has to be equal to the processing time of the specific process. That means: the processing time and the processing energy are strictly connected to each other. If these two conditions are not satisfied the absorbed photons increase the temperature of the material.

Further, it can be stated: time is passing as changes occur in the state of the material, like changes in its energy level, location, motion, composition, etc. To perform such a change, some energy is needed. If the inherent energy of a specific physical process which is utilized to carry out the change in the state of material is higher, the change is carried out in a shorter time.

ACKNOWLEDGMENT

The author acknowledges the European Union Network of Excellence projects called ISIS and NEFERTITI as well as the Hungarian National Scientific Research Foundation (OTKA projects:

T030148, T026557, T034520, CNK 77564) for the continuous support obtained to his research work.

REFERENCES

1. Einstein, A., “Zur quantentheorie der strahlung (on the quantum theory of radiation),” *Physikalische Zeitschrift*, Vol. 18, 121–128, 1917.
2. Bohr, N., “The quantum postulate and the recent development of atomic theory,” *Nature*, Vol. 121, 580–590, 1928.
3. Saleh, B. E. A. and M. C. Teich, *Fundamentals of Photonics*, John Wiley, New York, 1991.
4. Stachel, J. (Editor), *The Collected Papers of Albert Einstein*, Princeton Univ. Press, Princeton, 1989.
5. Wheeler, J. A. and W. H. Zurek, *Quantum Theory and Measurement*, 9–49, Princeton Univ. Press, Princeton, 1984.
6. Yariv, Y., *Quantum Electronics*, 3rd Edition, John Wiley, New York, 1989.
7. Dirac, P. A. M., *The Principles of Quantum Mechanics*, Oxford Univ. Press, London, 1930.
8. Jacques, V., E. Wu, F. Grosshans, F. Treussart, Ph. Grangier, A. Aspect, and J.-F. Roch, “Experimental realization of wheeler’s delayed-choice gedanken experiment,” *Science*, Vol. 315, No. 5814, 966–968, 2007.
9. Darquié, B., M. P. A. Jones, J. Dingjan, J. Beugnon, S. Bergamini, Y. Sortais, G. Messin, A. Browaeys, and P. Grangier, “Controlled single-photon emission from a single trapped two-level atom,” *Science*, Vol. 309, 454–456, 2005.
10. Wilk, T., S. C. Webster, A. Kuhn, and G. Rempe, “Single-atom single-photon quantum interface,” *Science*, Vol. 317, 488–490, 2007.
11. Novoselov, K. S., A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov, “Electric field effect in atomically thin carbon films,” *Science*, Vol. 306, No. 5696, 666–669, 2004.
12. Deacon, R. S., K. C. Chuang, R. J. Nicholas, K. S. Novoselov, and A. K. Geim, “Cyclotron resonance study of the electron and hole velocity in graphene monolayers,” *Phys. Rev. B.*, Vol. 76, No. 8, 2007.
13. Sze, S. M., *Physics of Semiconductor Devices*, 2nd Edition, John Wiley, New York, 1981.
14. Burkhardt, Ch. E. and J. J. Leventhal, *Topics in Atomic Physics*, Springer, New York, 2006.
15. Olshansky, R., C. Su, J. Manning, and W. Powazinik, “Measurement of radiative and nonradiative recombination rates in InGaAsP and AlGaAs light sources,” *IEEE J. Quantum Electron.*, Vol. 20, No. 8, 838–854, 1984.
16. Berceci, T., “Processing time of photon generation,” *PIERS ONLINE*, Vol. 5, No. 6, 541–545, 2009.
17. Mandelstam, L. and I. Tamm, “The uncertainty relation between energy and time in nonrelativistic quantum mechanics,” *Izv. Akad. Nauk. USSR, Fiz.*, Vol. 9, No. 1–2, 122, 1945; *J. Phys. USSR*, Vol. 9, 249–254, 1945.
18. Hilgevoord, J., “The uncertainty principle for energy and time,” *Am. Jour. of Physics*, Vol. 64, 1451–1456, 1996.
19. Busch, B., “The time-energy uncertainty relation,” *Time in Quantum Mechanics*, Ch. 3, Vol. 734, Springer, Berlin/Heidelberg, 2007.
20. Hafenbrak, R., S. M. Ulrich, P. Michler, L. Wang, A. Rastelli, and O. G. Schmidt, “Triggered polarization-entangled photon pairs from a single quantum dot up to 30 K,” *New Journal of Physics*, Vol. 9, No. 315, 1–10, 2007.
21. Adler, R., “A study of locking phenomena in oscillators,” *Proc. IRE*, Vol. 34, 351–357, 1946.
22. Berceci, T., *Nonlinear Active Microwave Circuits*, Elsevier Sc. Publ., Amsterdam, The Netherlands, 1987.
23. Uiberacker, M., Th. Uphues, M. Schultze, A. J. Verhoef, V. Yakovlev, M. F. Kling, J. Rauschenberger, N. M. Kabachnik, H. Schröder, M. Lezius, K. L. Kompa, H.-G. Müller, M. J. J. Vrakking, S. Hendel, U. Kleineberg, U. Heinzmann, M. Drescher, and F. Krausz, “Attosecond real-time observation of electron tunnelling in atoms,” *Nature*, Vol. 446, 627–632, 2007.
24. Krausz, F. and M. Ivanov, “Attosecond physics,” *Review of Modern Physics*, Vol. 81, 163–234, 2009.
25. Uffink, J., “The rate of evolution of a quantum state,” *Note*, 1–4, Department of History and Foundations of Mathematics and Science, University of Utrecht, Utrecht, The Netherlands; *American Journal of Physics*, Vol. 61, 935–936, 1993.

Investigation of Ionospheric Slab Thickness Behaviour over Cyprus during Minimum Solar Activity

Haris Haralambous

Frederick University, 7 Y. Frederickou St., Palouriotisa, Nicosia 1036, Cyprus

Abstract— The ionospheric slab thickness τ is defined as the ratio of the Total Electron Content (TEC) to the F-region peak electron density (NmF2). Hourly values of TEC and NmF2 collected at Cyprus (coordinates: 35°N, 33°E geographic), are used in the study of slab thickness characteristics during a low solar activity period. Climatology of the slab thickness is described by the diurnal and seasonal variations. Examining the variations of τ , during the different seasons of solar minimum, they are generally characterised by higher night-time values compared to the daytime values. Another important feature of the slab thickness is its large day-to-day variability. As τ represents the ratio of TEC to NmF2, the day-to-day variability of τ reflects the combined effect of the variability of the individual parameters.

1. INTRODUCTION

The ionospheric slab thickness (τ) is a parameter which provides information about the nature of the distribution of ionization at a specific location and is defined as the ratio of the vertical Total Electron Content (TEC) measured in TEC units (1TECu = 10^{16} electrons m^{-3}) to the maximum ionospheric electron density in the F-region (NmF2) [1].

$$\tau = \frac{\text{TEC}}{\text{NmF2}} \quad (1)$$

Since the critical frequency of the F2 layer (foF2) is related to the maximum ionospheric F2-layer electron density (NmF2) by

$$\text{NmF2} = 1.24(\text{foF2})^2 \times 10^{10} \text{ electrons} \times m^{-3} \quad (2)$$

slab thickness can be expressed in terms of foF2 as

$$\tau = \frac{\text{TEC}}{1.24(\text{foF2})^2} \times 10^3 \quad (3)$$

where TEC is measured in TECu, foF2 in MHz, and τ in km.

In other words, τ represents the equivalent thickness of an idealised ionosphere which has the same electron content as the actual ionosphere but uniform electron density equal to the maximum electron density (Figure 1). Slab thickness measurements offer substantial information on the shape of the electron density profile, the neutral and ionospheric temperatures/gradients and on the ionospheric composition and dynamics [2, 3] Previously a number of studies have outlined the relevance of ionospheric slab thickness to the vertical scale height [4].

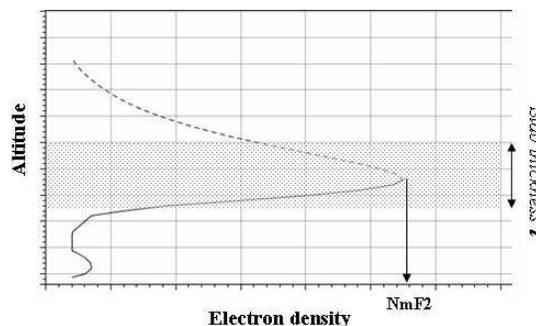


Figure 1: Schematic view of the vertical electron density profile defining slab thickness (τ) and peak density (NmF2).

The electron density profile shown in Figure 1 is comprised by the bottomside part which is measured by the digisonde (continuous line) and a topside part (dashed line) which is modeled [5]. The Cyprus ionospheric station carries out regular vertical incidence soundings with a Digisonde DPS-4D developed by the University of Massachusetts-Lowell [6]. Although the ionograms are automatically scaled for operational use, for the purpose of this study manually corrected hourly values of foF2 were used for the sake of more accurate results. TEC observations have been processed using observables from a GNSS receiver, collocated with the ionospheric sounder, applying Ciralo's method [7]. In fact when GNSS measurements (Global Navigation Satellite System) are used to calculate TEC a percentage has to be attributed to a plasmaspheric contribution in addition to the ionospheric component [8]. The slab thickness database used in this study consists of measurements from January 2009 to August 2010, i.e., covering more than one year of low solar cycle activity period. For the analysis, the data has been sorted according to month and the median value of slab thickness was calculated for each hour.

2. DIURNAL AND SEASONAL BEHAVIOUR

The majority of ionosphere slab thickness studies reported to date are confined to single station observations under limited solar activity variations. From these studies, it has been found that τ shows appreciable diurnal, day-to-day, seasonal, solar and magnetic activity variations with considerable dependence on the location of the observing station. According to Figure 2, where the diurnal profile of monthly median values of slab thickness for each month is shown, the seasonal variation of τ during solar minimum over Cyprus is characterised by a pronounced pre-sunrise peak during winter, autumn and spring months with night-time values of τ higher than the daytime values except during the summer season where the reverse seems to be true. The pre-sunrise peak is a phenomenon closely related to the maintenance of the night-time F layer and can be sufficiently attributed by the lowering of the ionospheric F layer immediately before sunrise to regions of greater

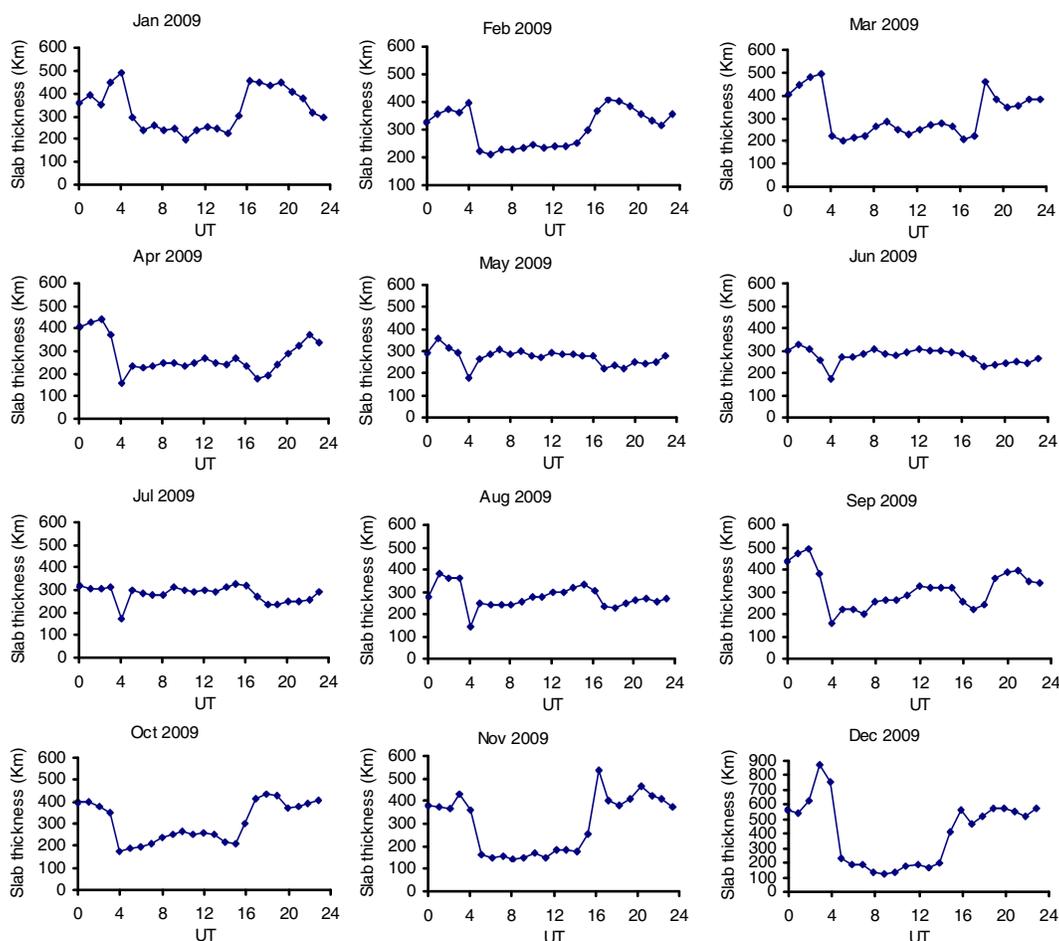


Figure 2: Diurnal profile of monthly median values of slab thickness for each month in 2009.

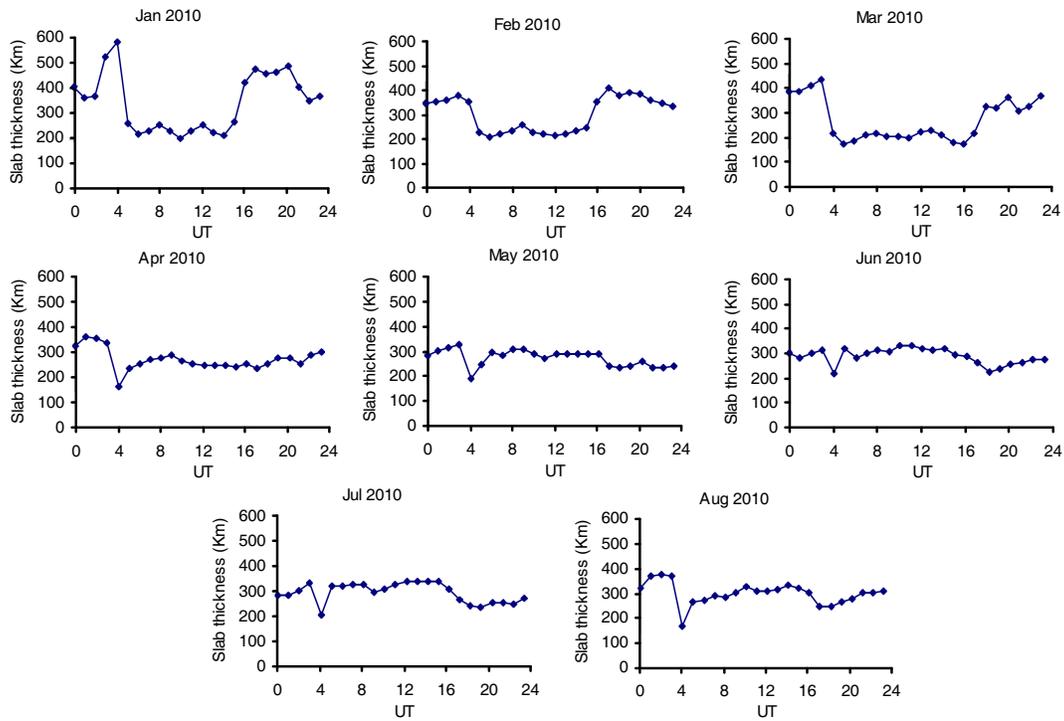


Figure 3: Diurnal profile of monthly median values of slab thickness for each month in 2010.

Table 1: Average ionospheric slab thickness for daytime (08:00–16:00 LT) and night-time (20:00–04:00 LT).

| Month | Average Slab Thickness (Km) | | Average Slab Thickness (Km) | |
|-----------|-----------------------------|------------|-----------------------------|------------|
| | Daytime | Night-time | Daytime | Night-time |
| January | 313 | 354 | 309 | 385 |
| February | 301 | 325 | 279 | 321 |
| March | 276 | 361 | 218 | 320 |
| April | 238 | 334 | 249 | 290 |
| May | 264 | 273 | 271 | 267 |
| June | 280 | 266 | 304 | 292 |
| July | 290 | 278 | 309 | 284 |
| August | 284 | 282 | 301 | 306 |
| September | 282 | 339 | | |
| October | 280 | 312 | | |
| November | 267 | 330 | | |
| December | 316 | 540 | | |

neutral density, leading to increased ion loss due to recombination [3]. The effect is considered to be particularly evident in the bottomside ionosphere that encompasses the density peak. As a result, the decrease in NmF2 and the bottomside density is much faster than the topside ionosphere where the loss rate is lower and thus, an enhancement occurs. In addition, plasmaspheric fluxes can also play a role in the increase. The magnitude of this increase is reportedly larger in low latitudes than in middle latitudes. The pre-sunrise peak is a regular feature reported to appear during solar minimum in all seasons and latitudes, while during solar maximum it is still preserved for high and low latitudes but not evident at middle latitudes [9]. Indeed, from our observations (Figure 2) the pre-sunrise peak seems to gradually diminish from winter to summer. The post-sunset increase in the τ values observed during the different seasons could be explained as being due to the secondary fountain effect caused by the post-sunset occurrence of a strong eastward electric field existing over the equatorial latitudes. The post-sunset enhancement in the τ values observed for the mid-latitude may also be associated with the night-time enhancements in TEC at mid-latitudes which is primarily due to the field aligned plasma flow from the protonosphere to the ionosphere [10].

Several studies in the northern hemisphere indicate that, generally, the slab thickness is greater in summer than in winter [11, 12]. Observations from the southern hemisphere confirm that the daytime slab thickness is greater in summer than in winter [2, 8]. Table 1 gives the mean daytime (08:00–

16:00 LT) and night-time (20:00–04:00 LT) values of ionospheric slab thickness for magnetically quiet days during the period under investigation. It is found that, for magnetically quiet days at solar minimum, the mean night-time values are higher than the mean daytime values for winter, autumn and spring months except during the summer where the reverse is the case.

3. CONCLUSIONS

The diurnal and seasonal behaviour of slab thickness τ for a European low latitude station at Cyprus (coordinates: 35°N, 33°E geographic) during an extremely low solar activity period has been discussed. Examining the variations of τ , during the different seasons of solar minimum, are generally characterised by higher night-time values compared to the daytime values. The 2009 mean values of τ obtained in this study are 282 km during daytime and 334 km during night-time.

REFERENCES

1. Davies, K., *Ionospheric Radio*, Peter Peregrinus Ltd., London, 1990.
2. Titheridge, J. E., “The slab thickness of the mid-latitude ionosphere,” *Planet. Space Sci.*, Vol. 21, 1973.
3. Davies, K. and X. M. Liu, “Ionospheric slab thickness in middle and low latitudes,” *Radio Science*, Vol. 26, 1991.
4. Stankov, S. M. and N. Jakowski, “Topside ionospheric scale height analysis and modelling based on radio occultation measurements,” *J. Atmos. Solar-Terrestrial Phys.*, Vol. 68, No. 2, 134–162, 2006.
5. Reinisch, B. W. and X. Huang, “Deducing topside profiles and total electron content from bottomside ionograms,” *Adv. Space Res.*, Vol. 27, 23–30, 2001.
6. Reinisch, B., et al., “New Digisonde for research and monitoring applications,” *Radio Science*, 2009.
7. Ciraolo, L., “Evaluation of GPS L2-L1 biases and related daily TEC profiles,” *Proc. GPS/Ionosphere Workshop*, 90–97, Neustrelitz, 1993.
8. Breed, A. M., G. L. Goodwin, A. M. Vandenberg, E. A. Essex, K. J. W. Lynn, and J. H. Silby, “Ionospheric total electron content and slab thickness determined in Australia,” *Radio Science*, Vol. 32, No. 4, 1635–1643, 1997.
9. Jayachandran, B., T. N. Krishnankutty, and T. L. Gulyaeva, “Climatology of ionospheric slab thickness,” *Ann. Geophys.*, Vol. 22, 25–33, 2004.
10. Minakoshi, H. and I. Nishimuta, “Ionospheric electron content and equivalent slab thickness at lower mid-latitudes in the Japanese zone,” *Proc. Beacon Satellite Symposium (IBSS)*, 144, Wales, UK, 1994.
11. Spalla, P. and L. Ciraolo, “TEC and foF2 comparison,” *Ann. Geofis.*, Vol. 37, No. 5, 929–938, 1994.
12. Leitingner, L., L. Ciraolo, L. Kersley, S. S. Kouris, and P. Spalla, “Relations between electron content and peak density — regular and extreme behaviour,” *Ann. Geophys.*, Vol. 47, No. 2–3, 1093–1107, 2004.

A Study of Es Layer Characteristics over Cyprus

Photos Vryonides¹, Lefteris Economou², and Haris Haralambous¹

¹Frederick University, 7 Y. Frederickou St., Nicosia 1036, Cyprus

²Intercollege, Ayias Phylaxeos St., Limassol 3507, Cyprus

Abstract— In this study, manually scaled digital ionosonde ionograms obtained in Cyprus (coordinates: 35°N, 33°E geographic) have been analysed to study the diurnal and seasonal occurrence of Es layer during a low solar activity period (January 2009 to August 2010). Processing of ionograms and scaling of Es (Es) parameters was carried out in accordance with the URSI Handbook of Ionogram Interpretation and Reduction [10]. Statistics of Es layer parameters, which include the critical ($f_0\text{Es}$), the layer height ($h_0\text{Es}$) as well as Es layer classification types are also analysed in an effort to identify systematic seasonal and diurnal patterns. The Es diurnal occurrence pattern is generally high during morning hours (starting around 0700 LT) with a strong blanketing character, around noon, diminishes in the afternoon and then reappears again during evening hours, sometimes persisting through the night. Regarding Es classification the most typical types observed over Cyprus are the f , c , and l types.

1. INTRODUCTION

Es is an ionospheric phenomenon classified as an E-region (90–150 km) irregularity characterised by a thin layer of extremely dense ionization patches which can significantly affect radio wave propagation. Es can occur during daytime or nighttime, and its characteristics vary with latitude. These regions can apparently be caused by several mechanisms, and have a wide variety of characteristics. There have been correlated in literature with solar activity, geomagnetic activity and associated to thunderstorms and meteor showers.

There are internationally-agreed standards for the analysis of vertical incidence ionosonde data. These are codified in Piggot and Rawer [10, 11]. Here the parameters determined at each hour are $f_0\text{Es}$, $h'\text{Es}$ and Es type. The term $f_0\text{Es}$ refers to the maximum frequency of an ordinary wave that can be reflected from a Es layer. The maximum frequency is related to the maximum ion concentration. The term $h'\text{Es}$ is the virtual height of the Es layer. Es layer height determination is taken from the flat part of the trace. Because there is little underlying ionization the real height and the virtual height are normally very similar. For Es layers that show retardation effects, there will be greater uncertainty that is indicated by the qualifying letters. Es type is codified into four main classes at mid latitudes: f (flat) is the name given to all Es layers where there is no E layer present (i.e., at night), l (low) occurs below the height of the E layer, h (high) occurs above the height of the E layer, and c (cusp) lies between high and low types, i.e., within the E layer.

Some authors studied the formation mechanisms of the different types of Es using different regional observation data from the ionosphere, but got inconsistent results Es [6] since the ionosphere is complex and has not only global but also regional properties. Although some authors have studied the ionospheric processes and the possible mechanisms of Es [7–9], further studies are needed to arrive to tangible conclusions.

In the present study Es layer behaviour over Cyprus is studied using a modern digital ionosonde (DPS-4D) which has been operating in Cyprus since August 2008. Furthermore a statistical analysis for the occurrence of all types of observed Es including Es- f , Es- l and Es- c over Nicosia region has been undertaken.

2. DATA AND METHOD OF ANALYSIS

In the analysis, we differentiate Es in 3 different types, named flat, low, and cusp and use the subscript letter f , l , c to denote them respectively. Type c is a daytime phenomenon where the radio echo usually is continuous with normal E-layer but later forms a clear cusp near $f_0\text{E}$. Type l is also a daytime type Es which forms below and is disconnected from the normal E-layer. Type f is a flat Es layer where there is no increase of height with increasing frequency and is a nighttime phenomenon. Diurnal and seasonal variation of occurrence of Es and the types observed are given in this section.

The main Es parameter used in the present study is $f_0\text{Es}$ (critical frequency of Es layer). The values collected are averaged to obtain monthly median diurnal variation for the periods representative of Summer (May, June, July and August), Winter (November, December, January and

February) and Equinox (March, April, September and October). The occurrence of Es, expressed as percentage out of the total number of ionogram observations was calculated for different months.

Figure 1 shows the diurnal variation of median frequency foEs (MHz) for the three seasons. It forms about two hours after sunrise, and its critical frequency which is low in the morning increases rapidly as the zenith angle of the sun decreases. The maximum foEs is reached at around noon (local time-UT+2), after which it decreases until evening, when this type of Es disappears. It is noted that the maximum frequency foEs is observed at around local noon for all seasons.

Figure 2 shows the average percentage of Es layer (foEs) during Summer, Winter and Equinox months. It is characterised by two maxima, around 0800–0900 UT in the morning and 15000 UT in the evening for all the three seasons. More pronounced occurrence of foEs is noted during summer months compared to winter and equinox months. The occurrence of Es is high (more than 80%) during 06000 UT to 20000 UT. The double maximum feature is similar with results of Abdu and Batista [3], where they have discussed Es layer phenomena over Cachoeira Paulista (22.6°N, 45°W) which is located in the vicinity of the Brazilian geomagnetic anomaly center. They have pointed out the double maximum feature in the occurrence of Es for summer season being less pronounced during equinox and absent in winter months.

The variation of Es virtual height ($h'Es$) for all seasons is presented in Figure 3. The layer forms in the morning at the almost invariable height of 120 km as this height remains unchanged during the day. The rise in layer height is noted during morning hours around 0700–08000 UT. The Es layer height decreases during local noon but again increases during evening hours. It is interesting to note that the Es layer height decreases dramatically during evening hours from winter and equinox seasons where it remains constant for the summer season. The Es layer height is showing double peak which is similar in the case of occurrence of Es layer (Figure 2). It can also be observed that the effective height of Es varies within the 80–120 km region during this period. The diurnal patterns of f , l , and c types of Es hourly rate occurrence rate for the three seasons are plotted in Figure 4 as a function of local time from January 2009 to August 2010.

The variation of Es layer virtual height ($h'Es$) for all seasons is presented in Figure 3. The layer forms in the morning at the almost invariable height of 120 km as this height remains unchanged during the day. The rise in layer height is noted during morning hours around 0700–08000 UT.

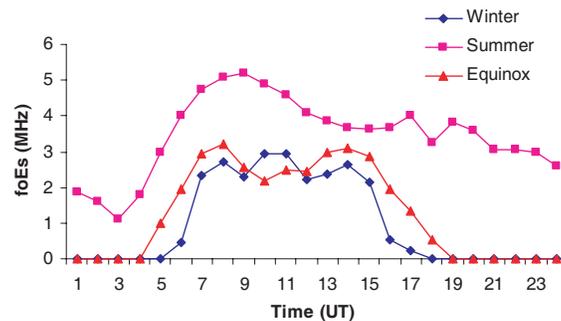


Figure 1: The diurnal variation of average foEs.

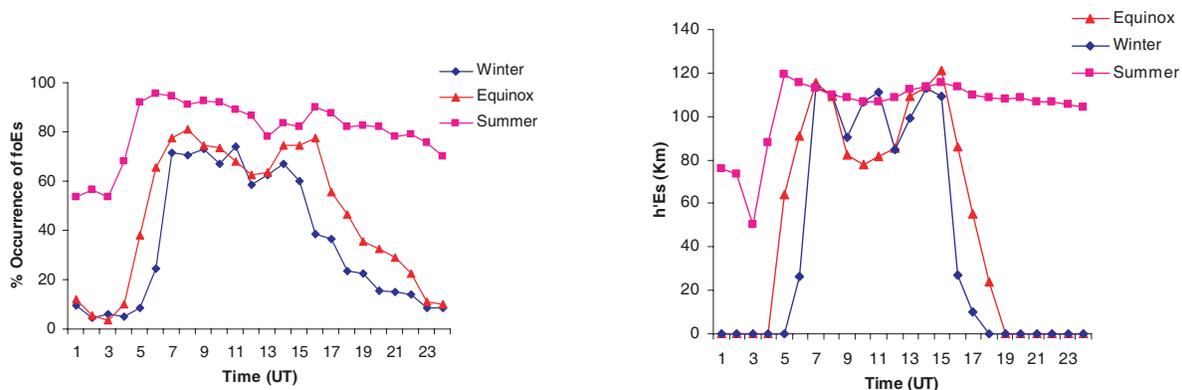


Figure 2: The percentage occurrence of Es.

Figure 3: The diurnal variation of average height.

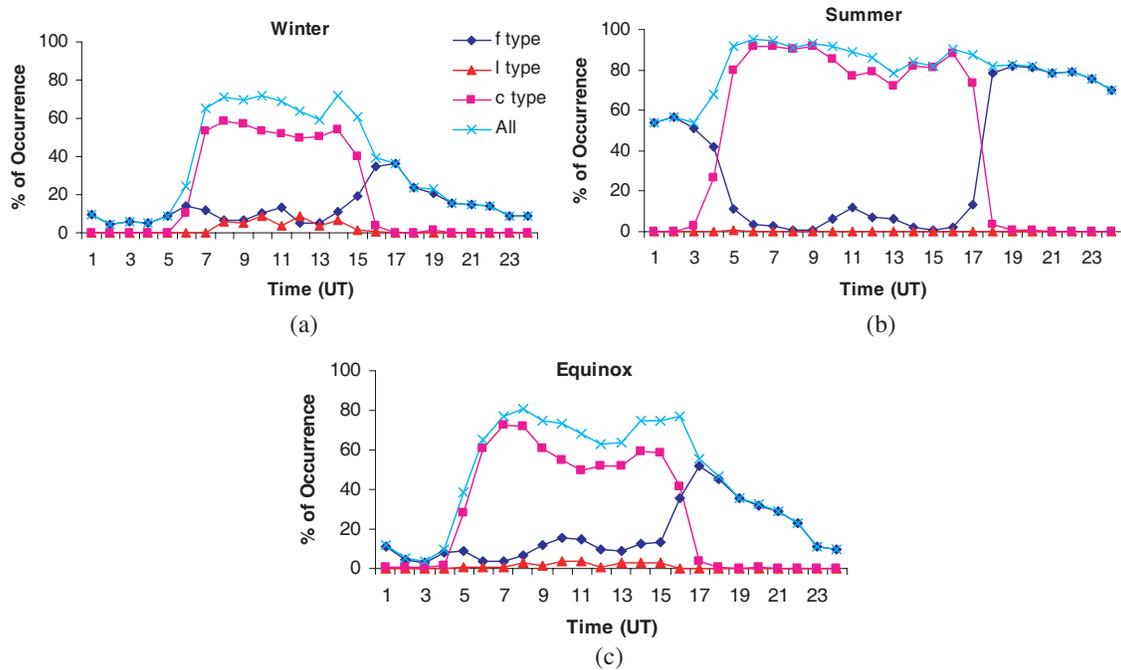


Figure 4: Hourly occurrence rate of different types of Es for (a) Winter, (b) Summer and (c) Equinox.

The Es layer height decreases during local noon but again increases during afternoon hours. It is interesting to note that the Es layer height decreases dramatically during evening hours from winter and equinox seasons where it remains constant for the summer season. The Es height is showing double peak which is similar in the case of occurrence of Es layer (Figure 2). The diurnal patterns of f , l , and c types of Es hourly occurrence rate for the three seasons are plotted in Figure 4 as a function of time. Results for only the main three Es types, l , f , c are determined since the percentage of occurrence for other types is very small. Figure 4 shows the diurnal variations of l , c and f type percentage of occurrence for winter, summer and equinox at low solar activity level. It should be noted that the scaling conventions determine the diurnal variation of some Es types, e.g., l and c have a mainly daytime occurrence while f has a night time occurrence. For the three seasons the c type Es had higher occurrence rate with a double peak at about 10000 UT and 15000 UT.

3. CONCLUSIONS

Using observation data obtained with a digital ionosonde (DPS-4D) at the mid latitude station in Nicosia from January 2009 to August 2010, a statistical analysis for properties of Es has been performed. The analysis undertaken demonstrated clearly that Es is a very frequent phenomenon above Cyprus which primarily occurs during Summer months less in Equinox and less even during Winter months. Especially during the Summer, typically strong ionization around noon completely blankets the F-region. According to the present investigation which coincided with an extended low solar cycle activity period, the most dominant type of Es is the c type occurring throughout the whole year during daytime. The f type is also a frequent type but mostly occurring during night-time.

REFERENCES

1. Sridharan, R., et al., "Winds, wind shears and plasma densities during the initial phase of a magnetic storm from equatorial latitudes," *Journal of Atmospheric and Terrestrial Physics*, Vol. 51, 169–177, 1989.
2. Gupta, S. P., "Features of lower ionosphere during day and night over magnetic equator," *Adv. Space Res.*, Vol. 25, 53–63, 2000.
3. Abdu, M. A. and I. S. Batista, "Es layer phenomena in the Brazilian geomagnetic anomaly: Evidence for a regular particle ionization source," *Journal of Atmospheric and Terrestrial Physics*, Vol. 39, 723–731, 1977.

4. Saksena, R. C. and R. Marwah, “Es occurrence statistics over Indian subcontinent,” *Adv. Space Res.*, Vol. 18, 103–106, 1996.
5. Rishbeth, H., “Long-term changes in ionosphere,” *Adv. Space Res.*, Vol. 20, 2149–2155, 1997.
6. Abdu, M. A., I. S. Batista, P. Muralikrishma, and J. H. A. Sobral, “Long-term trends in Es layers and electric fields over Fontaleza, Brazil,” *Geophys. Res. Lett.*, Vol. 23, 757–760, 1996.
7. Farley, D. T., et al., “The prereversal enhancement of the zonal electric field in the equatorial ionosphere,” *J. Geophys. Res.*, Vol. 91, 13723–13728, 1986.
8. Heelis, R. A., et al., “Electrical coupling of the E and F regions and its effects on F region drifts and winds, planet,” *Planetary and Space Science*, Vol. 22, 743–756, 1974.
9. Woodman, R. F., “Vertical drift velocities and east-west electric fields at the magnetic equator,” *J. Geophys. Res.*, Vol. 75, 6249–6259, 1970.
10. Piggot, W. R. and K. Rawer, *U.R.S.I Handbook of Ionogram Interpretation and Reduction*, 2nd Edition, Report UAG 23, World Data Center A for Solar-Terrestrial Physics, NOAA, Boulder, Colorado, 1972.
11. Piggot, W. R. and K. Rawer, *Revision of Chapters 1–4 of the U.R.S.I Handbook of Ionogram Interpretation and Reduction*, Report UAG-23 A, World Data Center A for Solar-Terrestrial Physics, NOAA, Boulder, Colorado, 1978.

Integral Localized Approximation Description of v -th Order Bessel Beams in the Generalized Lorenz-Mie Theory and Applications to Optical Trapping

L. A. Ambrosio and H. E. Hernández-Figueroa
 School of Electrical and Computer Engineering (FEEC)
 Department of Microwaves and Optics (DMO)
 University of Campinas (Unicamp), Campinas 13083-970, SP, Brazil

Abstract— Theoretical derivations and numerical calculations of the beam-shape coefficients (BSCs) of the generalized Lorenz-Mie theory (GLMT) are presented by adopting the integral localized approximation and assuming, for the first time in the literature, an arbitrary v -th order Bessel beam. Numerical comparisons between our new approach and other time-consuming methods, such as quadratures, are performed, and it is revealed that the integral localized approximation provides a fast and efficient code for numerically evaluating the BSCs of Bessel beams and their associated electromagnetic field components. This new fast and robust approach can be advantageously used in the analysis of scattering problems with Bessel beams, such as in optical trapping systems.

1. INTRODUCTION

The GLMT [1, 2] is an extension of the Lorenz-Mie theory [3] to arbitrary laser beams, being applied in the determination, among others, of the optical forces and torques in optical trapping systems using Gaussian and laser sheet beams. Initially, it presented to be a really time consuming theory because of the use of quadratures [1] or finite series methods [4] and, together with some intrinsic limitations related to the theoretical representation of Gaussian beams, this certainly justified, at least to some extent, the preference for simplified approaches such as ray optics or closed-form solutions in the Rayleigh regime. Even after the introduction of the integral localized approximation (ILA) [5] for speeding up the numerical evaluation of the BSCs, setting up the GLMT as a robust, efficient and fast method in scattering problems involving arbitrary particles and arbitrary incident laser beams, the numerical results for the scattering problem using a Gaussian beam differed significantly from some practical results in optical trapping experiments due, among other factor, to the fact that this type of beam is not an exact solution of Maxwell's equations.

However, during the last years, a lot of interest has been given to laser beams capable of providing multiple trapping of particles. These multi-ringed structures, also known as localized or Bessel beams (BBs), can trap several particles along their high or low intensity rings, depending upon the relative refractive index between the particle and the surrounding medium [6, 7]. Because of their increasing practical interest, a complete theory, which could predict all physical properties of interest (*viz.*, optical force and torque, scattering and internal electromagnetic field components, absorption and extinction cross sections) with great accuracy, is desirable. Although we can again adopt a ray optics approach for particles much larger than the wavelength of the impinging beam, or use single formula for force or torque calculations in the Rayleigh regime, we resent from such a complete formulation.

In this way, this work is devoted to describe, for the first time in the literature, an arbitrary linearly polarized v -th order BB in the framework of the GLMT by adopting the ILA for evaluating their BSCs. We start with a zero-order BB, which presents the same symmetry relations observed for Gaussian beams and whose BSCs can be analytically predicted. The time required for computing these coefficients are compared with other quadrature methods available in the GLMT. Then, we extend our analysis to higher-order BBs, where new symmetry relations are necessary and no closed formula can be found for the BSCs without resorting to integration. Finally, we show some examples of application of our new approach in the calculation of optical forces in optical trapping systems with BBs and spherical particles.

2. BESSEL BEAM ANALYSIS IN GLMT USING THE ILA

Consider an ideal monochromatic zero-order BB propagating with speed c and angular frequency ω_0 parallel to $+z$, with its optical axis displaced $\rho_0 = (x_0, y_0)$ from the z -axis and making an angle

ϕ_w relative to the x -axis. The longitudinal and transverse wave numbers for this BB are given by $k_z = k \cos \theta_a$ and $k_\rho = k \sin \theta_a$, respectively, θ_a being the associated axicon angle and $k = \omega_0/c$ the wave number. Assuming an x -polarized Bessel beam, the radial components E_r of the electric field and H_r of the magnetic field (in spherical coordinates (r, θ, ϕ)) necessary for completely describe the BSCs in the framework of the GLMT are written as

$$E_r = E_0 J_0 \left[\sin \theta_a \sqrt{(kr)^2 \sin^2 \theta + \rho_0^2 k^2 - 2(kr) \rho_0 \sin \theta \cos(\phi - \phi_0)} \right] e^{-j(kr) \cos \theta_a \cos \theta} \sin \theta \cos \phi \quad (1)$$

$$H_r = H_0 \cos \theta_a J_0 \left[\sin \theta_a \sqrt{(kr)^2 \sin^2 \theta + \rho_0^2 k^2 - 2(kr) \rho_0 \sin \theta \cos(\phi - \phi_0)} \right] e^{-j(kr) \cos \theta_a \cos \theta} \sin \theta \sin \phi \quad (2)$$

where $J_0(\cdot)$ is the ordinary Bessel function, E_0 is the electric field strength, $H_0 = E_0/\eta$, η being the intrinsic impedance of the propagating medium. A factor $\exp(j\omega_0 t)$ has been omitted for convenience.

From (1) and (2), the BSCs $g_{n,TE}^m$ and $g_{n,TM}^m$ in the ILA are analytically found [8]. For the $g_{n,TM}^m$'s:

$$g_{n,TM}^m = \begin{cases} j \frac{2n(n+1)}{(2n+1)} J_1(\sin \theta_a (n+1/2)) J_1(\rho_0 k \sin \theta_a), & m = 0 \\ \frac{1}{2} \left(\frac{-2i}{2n+1} \right)^{|m|-1} \left[\begin{array}{l} J_{|m|-1}(\sin \theta_a (n+1/2)) J_{|m|-1}(\rho_0 k \sin \theta_a) + \\ J_{|m|+1}(\sin \theta_a (n+1/2)) J_{|m|+1}(\rho_0 k \sin \theta_a) \end{array} \right], & |m| \neq 0 \end{cases} \quad (3)$$

with similar relations for the $g_{n,TE}^m$'s [8]. Here, $1 \leq n < n_{\max}$ and $-n \leq m \leq n$. Fig. 1 shows that an adequate description of BBs for optical trapping could be achieved by using just a few n and m . For example, for on-axis beams, only $m = 1$ is necessary in (3) [8]. Table 1 compares the numerical values of the BSCs and the time elapsed in their evaluation first using (3) and then using the faster quadrature scheme in the GLMT (called F1), with double-integration over the spherical coordinates θ and ϕ [9], for a particular on-axis BB. One can see that the ILA and the F1 methods agree very well for low n , where the most significant BSCs are actually found. However, the ILA is approximately 1000 times faster, in average. Notice that, as n increases, so does the average time for F1, because our subroutine uses recursive relations for the computation of the spherical Bessel functions entering the GLMT, so that, for example, for $n = 200$, 200 iterations are performed to compute $j_{200}(R)$.

Let us now assume, as before, a $+z$ -propagating x -polarized BB, but now with arbitrary order $\nu > 0$. In this situation, the azimuth symmetry in ϕ is broken due to the factor $\exp(-j\nu\phi)$ presented

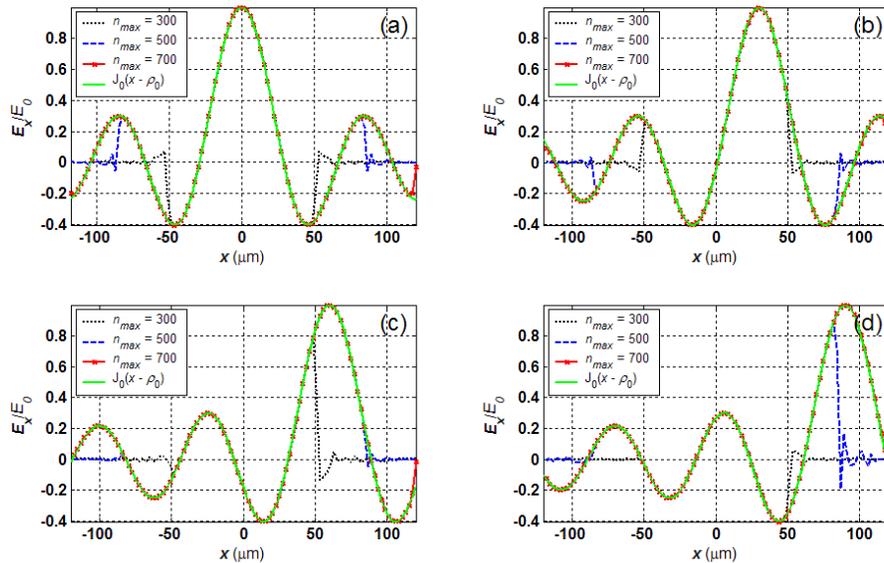


Figure 1: Field intensity profile for an x -polarized Bessel beam with $\lambda = 1064$ nm, $\theta_a = 0.0141$ rad and displaced $\rho_0 =$ (a) 0, (b) 30 μm , (c) 60 μm and (d) 90 μm along x . The accuracy of the ILA for an ordinary Bessel beam depends upon increasing the number of BSCs (n_{\max}). In all cases, $m_{\max} = 15$.

Table 1: Beam-shape coefficients $g_{n,TM}^1$ evaluated by quadrature with double integration (F1) and using the ILA for an on-axis ($\rho_0 = \phi_0 = 0$) zero-order Bessel beam with $\lambda = 1064$ nm and $\theta_a = 0.0141$ rad^a.

| n | ILA (time elapsed, in s) | F1 (time elapsed, in s) |
|-----|--|--|
| 1 | 0.499975 (0.549968×10^{-3}) | $0.499976 + i0.962191 \times 10^{-15}$ (0.512970) |
| 2 | 0.499900 (0.449984×10^{-3}) | $0.499901 + i0.719626 \times 10^{-15}$ (0.531625) |
| 3 | 0.499776 (0.675186×10^{-3}) | $0.499773 - i0.360131 \times 10^{-15}$ (0.546375) |
| 4 | 0.499602 (0.574975×10^{-3}) | $0.499593 - i0.307063 \times 10^{-15}$ (0.593000) |
| 5 | 0.499379 (0.599977×10^{-3}) | $0.499362 + i0.160106 \times 10^{-15}$ (0.593521) |
| 10 | 0.497518 (0.349988×10^{-3}) | $0.497454 + i0.453892 \times 10^{-15}$ (0.624250) |
| 15 | 0.494424 (0.599975×10^{-3}) | $0.494301 - i0.576601 \times 10^{-15}$ (0.702250) |
| 20 | 0.490109 (0.524983×10^{-3}) | $0.489921 - i0.253856 \times 10^{-15}$ (0.734125) |
| 50 | 0.439779 (0.274985×10^{-3}) | $0.439164 - i0.162183 \times 10^{-15}$ (0.968750) |
| 100 | 0.280726 (0.599973×10^{-3}) | $0.279540 - i0.327642 \times 10^{-15}$ (0.139013×10^1) |
| 200 | -0.965629×10^{-1} (0.349975×10^{-3}) | $0.111491 - i0.995058 \times 10^{-13}$ (0.223350×10^1) |
| 400 | 0.201049×10^{-1} (0.268730×10^{-3}) | $0.928932 \times 10^{-3} + i0.336547 \times 10^{-13}$ (0.398363×10^1) |

^a We have chosen to present numbers with six digits, according to the convention adopted by previous authors [5].

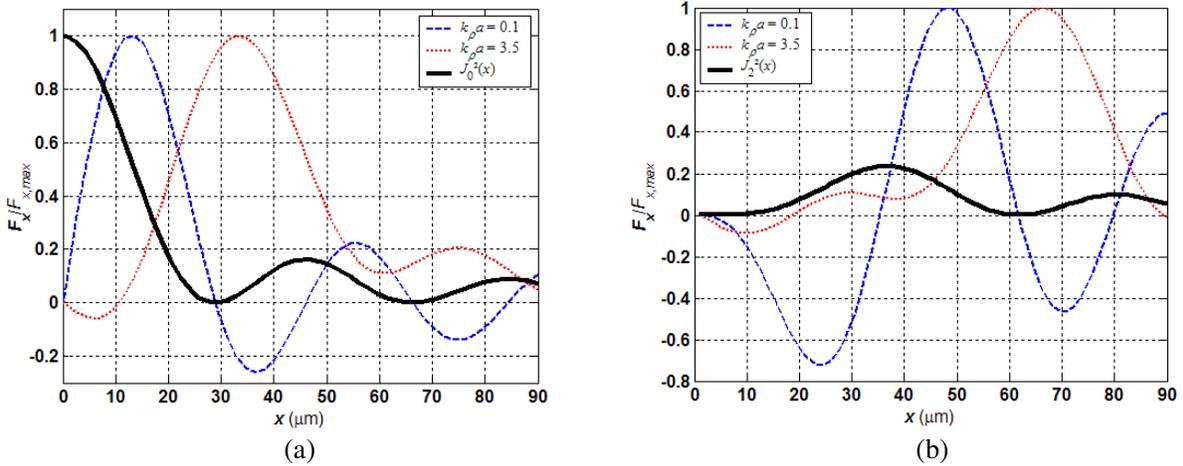


Figure 2: (a) Normalized force profile exerted on a simple dielectric particle of relative refractive index $n_{rel} = 1.1$ and radius a . The field intensity of the zero-order Bessel beam ($\lambda = 1064$ nm and $\theta_a = 0.0141$ rad) is also shown as a solid line. Results coincide almost exactly with previous works based on Rayleigh-Gans approximations [13]. Positive F_x/F_{\max} means an attractive force towards the optical axis of the beam. (b) Same force profile for a Bessel beam with $v = 2$.

in all electromagnetic field components and no analytical solution have been found so far. Thus, the ILA description now leads to a single integral in ϕ that must be numerically solved. Furthermore, due to this asymmetry, we cannot immediately use previous symmetry relations for optimizing the numerical calculation of the BSCs and to increase the time efficiency of the algorithm [10]. Thus, in comparison with zero-order BBs, the total time needed for computing any physical property of the system is increased.

It is possible, however, to develop an alternative scheme in order to have adequate symmetry relations for the BSCs, at least for the particular situation of a linearly polarized higher-order BB. In this case, we can break the complex exponential $\exp(-jv\phi)$ into its real and imaginary parts, thus leading to the following E_x :

$$E_x = E_0 \left\{ \begin{array}{l} \cos v\phi J_v(f(r, \theta, \phi)) e^{-j(kr) \cos \theta_a \cos \theta} \\ j \sin v\phi J_v(f(r, \theta, \phi)) e^{-j(kr) \cos \theta_a \cos \theta} \end{array} \right\} = E_{x,S} + E_{x,A}, \quad (4)$$

where $f(r, \theta, \phi) = \sin \theta \sqrt{(kr)^2 \sin^2 \theta + \rho_0^2 k^2 - 2(kr) \rho_0 \sin \theta \cos(\phi - \phi_0)}$. In (4), we can associate the particular fields $E_{x,S}$ and $E_{x,A}$ with BSCs $g_{n,\{TE, TM\},S}^m$ and $g_{n,\{TE, TM\},A}^m$, each of them having their own symmetry relations. With this, only one of the TM or TE coefficients, together with the condition $m > 0$, is necessary in order to completely determine the whole set of BSCs for each particular field. Of course, the original BSCs are now superpositions of the new S and A BSCs, so that they will have symmetry relations other than those expected for $v = 0$ [11]. In this way, although we now have four set of BSCs, we do not actually need to numerically evaluate all of them. Even for a particular set, the condition $m > 0$ cuts by half the number of coefficients that need to be computed by resolving the integral in ϕ . Finally, we emphasize that, even though this paper deals with linear polarization, all results can be immediately extended to circularly polarized BBs [12].

As an example of application, Fig. 2 shows the optical forces exerted on a simple spherical dielectric particle by a BB with $v = 0$ and $v = 2$. The parameters were chosen so as to coincide with those of [13], for comparison. Because this type of beam exactly satisfies Maxwell's equations, we expect the GLMT to be valid at any optical regime, its only limitation being associated with low-error approximations imposed by the ILA.

3. CONCLUSIONS

For the first time in the literature, the BSCs for a v -th order BB were numerically evaluated using the ILA in the context of the GLMT, valid for all optical regimes. Closed-form solutions were presented for $v = 0$, extending the range of applicability of the ILA to beams other than Gaussian and laser sheets.

Compared to quadrature methods, the fundamental advantage of our approach lies on the incredibly reduced computational time of each BSC. New symmetry relations arise for BBs with $v > 0$ and, although the evaluation of their BSCs demands more computational time, fast computers can perform this task with relative ease.

We have also applied our formalism to compute optical forces in optical trapping systems, comparing their profiles with those already obtained by previous authors. The similitude of results, together with the lower computational time, reinforce the reliability of our approach and strongly suggest the extension of the integral localized approximation to a wider variety of laser beams.

ACKNOWLEDGMENT

The authors wish to thank FAPESP — *Fundação de Amparo à Pesquisa do Estado de São Paulo* — under contracts 2009/54494-9 (L. A. Ambrosio's post doctorate grant) and 2005/51689-2 (CePOF, Optics and Photonics Research Center), for supporting this work.

REFERENCES

1. Gouesbet, G. and G. Gréhan, "Sur la généralisation de la théorie de Lorenz-Mie," *J. Opt. (Paris)*, Vol. 13, 97–103, 1982.
2. Maheu, B., G. Gouesbet, and G. Gréhan, "A concise presentation of the generalized Lorenz-Mie theory," *J. Opt. (Paris)*, Vol. 19, 59–67, 1988.
3. Mie, G., "Beiträge zur optik Trüber medien, speziell kolloidaler Metallösungen," *Ann. Phys.*, Vol. 330, No. 3, 377–445, 1908.
4. Gouesbet, G., G. Gréhan, and B. Maheu, "Expressions to compute the coefficients g_n^m in the generalized Lorenz-Mie theory using finite series," *J. Opt. (Paris)*, Vol. 19, 35–48, 1988.
5. Ren, K. F., G. Gouesbet, and G. Gréhan, "Integral localized approximation in generalized Lorenz-Mie theory," *Appl. Opt.*, Vol. 37, No. 19, 4218–4225, 1998.
6. Garces-Chavez, V., D. McGloin, H. Melville, W. Sibbett, and K. Dholakia, "Simultaneous micromanipulation in multiple planes using a self-reconstructing light beam," *Nature*, Vol. 419, 145–147, 2002.
7. Garces-Chavez, V., D. Roskey, M. D. Summers, H. Melville, D. McGloin, E. M. Wright, and K. Dholakia, "Optical levitation in a Bessel light beam," *Appl. Phys. Lett.*, Vol. 85, No. 18, 4001–4003, 2004.
8. Ambrosio, L. A. and H. E. Hernández-Figueroa, "Integral localized approximation description of ordinary Bessel beams in the generalized Lorenz-Mie theory and applications to optical forces," to be published.

9. Gouesbet, G. and J. A. Lock, “Rigorous justification of the localized approximation to the beam-shape coefficients in generalized Lorenz-Mie theory. I. On-axis beams,” *J. Opt. Soc. Am. A*, Vol. 11, No. 9, 2503–2515, 1994.
10. Ren, K. F., G. Gréhan, and G. Gouesbet, “Symmetry relations in generalized Lorenz-Mie theory,” *J. Opt. Soc. Am. A*, Vol. 11, No. 6, 1812–1817, 1994.
11. Ambrosio, L. A. and H. E. Hernández-Figueroa, “Integral localized approximation description of arbitrary order Bessel beams, symmetry relations and application to optical trapping,” to be published.
12. Polaert, H., G. Gréhan, and G. Gouesbet, “Forces and torques exerted on a multilayered spherical particle by a focused Gaussian beam,” *Opt. Commun.*, Vol. 155, No. 1–3, 169–179, 1998.
13. Rubinov, A. N., A. A. Afanas’ev, I. E. Ermolaev, Y. A. Kurochkin, and S. Y. Mikhnevich, “Localization of spherical particles under the action of gradient forces in the field of a zero-order Bessel beam. Rayleigh-Gans approximation,” *J. Appl. Spectr.*, Vol. 70, No. 4, 565–572, 2003.

Genetic Algorithms Application for the Optimal Design of Magnetic Vagus Nerve Stimulator

M. Chojnowski and J. Starzyński
Warsaw University of Technology, Poland

Abstract— Stimulation of human vagus nerve with use of electromagnetic field is intensively developing since late 1980's. The method pioneered by J. K. Penry and J. C. Dean is based on surgical implantation of electric device into human body and owes its prosperity to very promising results in various therapies. Its contemporary field of clinical applications stretches from epilepsy through depression and anxiety to migraines and still widens. Vagus nerve stimulation (VNS) was the initial trigger for authors to undertake the challenging research towards design of magnetic stimulator. Unlike the already well known electric VNS stimulator, the magnetic one (MVNS) is supposed to operate externally instead of being implanted into patient's body. The realistic, three-dimensional numerical model was prepared to begin with, in order to estimate the feasibility of the concept. One of the problems that emerged soon after the initial calculations were made, was to improve the basic design of the stimulator model. Having a complex, 3D model at hand to deal with a number of calculations required for an inverse problem meant a considerably long time of analysis and high demands for computational capacity. At that point, the need of model simplification became apparent and the most evident measure of bringing it into effect was to reduce the space of analysis into two dimensions. That assumption has conceived and brought to life a 2D model that proved to be less time and capacity consuming, compared to the three-dimensional one. This allowed use of the genetic algorithms for optimization of VNS applicator.

1. INTRODUCTION

Nerve Stimulation (VNS) is employed as approved treatment in drug-resistant epilepsy since late 1990's. It has more recently showed a promising results in treatment-resistant depression, cognitive disorders, anxiety, migraine and obesity [1–3]. The implementation of such treatment is based on electric devices being implanted into patient's body in both: chest and neck area [4]. The fundamental task of that device is to introduce electric impulse causing the action potential of stimulated nerve cells. Knowing that nerve tissues respond to external electric excitation and that they have conductive properties, the idea emerged to employ the external magnetic field for the task [5]. Such an approach would allow to replace an invasive device implantation with external field source as depicted in Figure 1(a), thanks to the current induced in nerve tissue according to the Faraday's law.

2. THE THREE-DIMENSIONAL MODEL

The preliminary calculations of the MVNS concept were run with use of the three-dimensional model developed by Sawicki et al. [6]. The incentive bonus to use that model was that it was designed and aimed towards realistic analysis of electromagnetic field within human body. Authors very soon became aware that while the model allows to estimate the phenomenon thoroughly, it results in a large scale computational problem. The average time needed for solving a particular MVNS problem with use of Intel® Xeon® 2 GHz processor was 13 minutes.

Moreover, the model analyses the electromagnetic field only inside the human body model, omitting the externally set stimulation source. That approach is justified though as the induced currents are greatly smaller than exciting ones, thus they have no influence on the source. However, the eventual conclusion was that using such model for optimization purposes is time consuming and enlarges the problem if more than one stimulating coil is used. In that case the model should be reworked to allow analysis of the coils interaction. That again enforces the need of modelling the body surroundings i.e., the coils and airspace what would result in far more complex computational problem.

Being aware of the constraints that draws from the 3D model, the authors decided to undertake the research towards finding the most relevant 2D model for VNS analysis. In general such model allows considering more aspects with less computation burden comparing to a three dimensional analysis.

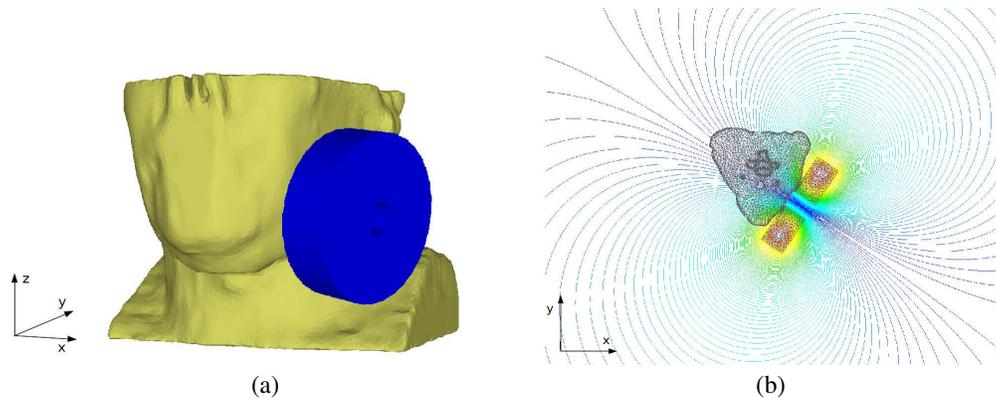


Figure 1: (a) A 3D visualization of the concept of vagus nerve stimulation with an external magnetic field. (b) The simplified 2D model resulting from the cross-section in xOy plane of the 3D model, along with example magnetic potential $|\underline{\vec{A}}|$ distribution.

3. THE TWO-DIMENSIONAL MODEL

The studies with the 3D model had brought an idea of elaborating the simplified model for optimization analyses. The simplification has been concerned because of a numerous calculations that have to be made in order to find the optimal solution.

The most evident measure of bringing the simplification into effect was to reduce the geometric space of analysis into two dimensions. The idea has proven to limit the calculation burden needed for the similar 3D problem and hence reduce the time needed for a particular analysis. Additionally, the 2D model allowed to extend the scope of the analysis to the excitation source as well as to consider the frequency dispersion of both human tissues and excitation source elements (e.g., ferromagnetic core of the stimulating coil).

In order to make alike comparison between above described 3D and 2D models, the assumption was made to use the current density \vec{J}_s as the source value. That is mainly caused by the 3D model approach, where the surroundings of the human body (i.e., neck), in particular the stimulating coil is not being analyzed. It is justified to neglect the eddy currents within the stimulator, knowing that the particular coil windings are isolated. Moreover the eddy currents inducted in the low conducting tissues would not influence the current density distribution within the stimulating coil.

Two dimensional electromagnetic field distribution was described with use of the complex vector magnetic potential $\underline{\vec{A}}$ according to the relation allowing to consider the two fundamental cases:

$$\nabla \cdot \frac{1}{\mu} \nabla \underline{A}_z = \begin{cases} J_{sz} \\ 0 \end{cases} \quad (1)$$

where the right side equals J_{sz} within the coil area and 0 beside it. Such description formulation allows to introduce the electric source by defining the current density value in particular subdomain (area representing the coil windings).

Dirichlet boundary conditions were introduced at the external boundaries of the modelled space. That caused the need to analyze the surroundings to such extent that the magnetic field lines could close their loops without distortion (see Figure 1(b)).

The numerical model created on the above-presented mathematical basis featured four types of materials with different conductivities: soft tissue, bones, vagus nerve and copper. As for the biological tissues, the frequency dispersion was taken into account in accordance with the Cole model [9]. The complexity of the 2D finite elements model was reduced to around 10 000 nodes and 20 000 triangular, three-node elements in comparison to 110 000 nodes and 620 000 tetrahedral, four-node elements used in the 3D model. While using the same computational environment as for the 3D model, the time needed for a single analysis dropped from the average 13 minutes to 20 seconds. Therefore the simplified model wins the time tournament hands down and thus the goal of reducing the time of analysis for optimization purposes was achieved.

4. GENETIC ALGORITHM DESIGN

The general design of the Genetic Algorithm is following:

1. The given number of individuals are being generated for the initial population.
2. The given number of generations are being produced, while for each generation:
 - (a) The given number of the best individuals are saved and transferred to the new generation intact. The selection of the preserved individuals (elite) is being made on the basis of their adaptation (fitness function value).
 - (b) Parents crossover and mutation is made within the selected individuals of the current generation in order to create a new generation.
3. The best adapted individual of all generations is pointed on the basis of its fitness function value.

The fitness function is evaluated on the basis of stimulation source design parameters which are represented by genes. The particular design parameters are described further in Results section. Each individual is represented by the FEM model built on the basis of particular chromosome, i.e., design parameters values. The induced current density value $|\vec{J}|$, probed within the nerve in result of FEM model analysis stands for the fitness function value.

5. GENETIC ALGORITHM IMPLEMENTATION

For the Genetic Algorithm implementation, authors used the ECJ (*Evolutionary Computation in Java*) environment [10], which eliminated the need of implementing the evolutionary basics from scratch. The final GA implementation required developing only a few Java classes to implement the problem-related data processing. With the above-presented genetic algorithm, the creation process of individual in particular generation is independent. The reproduction is made on the basis of the same population, however the crossover, mutation and fitness function evaluation are separate processes for each individual. Such phenomenon gives a natural ability to parallel the calculations [11]. That again gave the authors another chance to speed up the analyses. The ECJ environment was extended by calculations management server, problem repository server and clients implementation to allow the distributed, parallel processing within the computational resources of the Local Area Network (LAN).

A series of simulations were made in order to determine the relationship between the number of calculation nodes (number of processors) and analysis time. The results are presented in Figure 2.

The results shown in the table included within Figure 2 were obtained for optimization task described further under codename *synthesis*. The numerical FEM model had around 120 000 nodes and 240 000 finite elements. The number of nodes and elements are approximate and given for reference. The exact number varied within $\pm 0.7\%$ because the FEM mesh is generated for each genome individually.

The computational resources used in variant 1. was a server equipped with two double-core Intel[®] Xeon[®] 2.0 GHz processors. In the second variant, another server with four AMD Opteron[™] 880 Dual Core 2.4 GHz processors and two workstations with a single Intel[®] Core[™] 2 1.8 GHz processor were added. For the further variants, i.e., 3. and 4., the number of other, hardware-wise identical workstations were complemented to make up for the given number of processors.

| Variant No. | No. of processors | Analysis duration | |
|-------------|-------------------|-------------------|-------------|
| | | [h : m] | [d : h : m] |
| 1. | 4 | 139 : 10 | 5 : 19 : 10 |
| 2. | 16 | 67 : 53 | 2 : 19 : 53 |
| 3. | 32 | 30 : 06 | 1 : 06 : 06 |
| 4. | 64 | 16 : 31 | 0 : 16 : 31 |

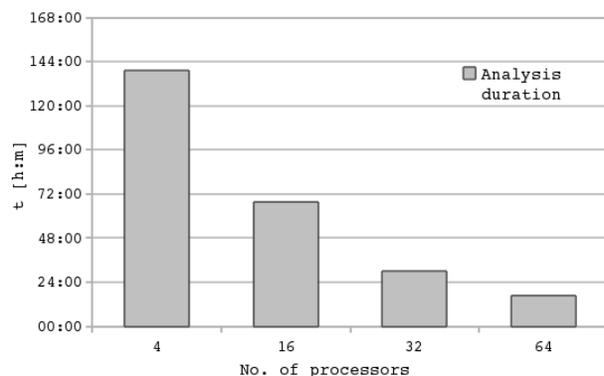


Figure 2: Analysis time related to the number of processors used in the distributed, parallel computation environment.

6. RESULTS

A series of variants with different number of genes (fitness function variables) were examined. The first assumed two genes representing the following parameters:

1. Coil shift in relation to the nerve.
2. The length of the outer side of the coil cross-section, assuming a constant cross-sectional area and that the windings are symmetrical.

The initial model, presented in Figure 3(a), was obtained from the 3D model depicted in Figure 1. The fitness function value for that setup was $|\vec{J}_0| = 0.61997 \text{ A/m}^2$. The optimization process resulted in an improved stimulation system configuration as shown in Figure 3(b) with the fitness function value of $|\vec{J}| = 3.07555 \text{ A/m}^2$. Therefore the current density value induced within the nerve was nearly five times improved: $\frac{|\vec{J}_{opt}|}{|\vec{J}_0|} = 4.96$.

With the second of hereby presented optimization variant, authors assumed that the coil cross-section will be divided into six segments. Each of the segments had a fixed width, resulting from division into equal parts, while their length was variable. As in the first variant, the invariability of a cross-sectional area as well as the windings symmetry were assured. Such assumptions led to the following four design parameters (genes):

1. Coil shift in relation to the nerve.
2. The length of the two outer segments of the coil cross-section.
3. The length of the two middle segments of the coil cross-section.
4. The length of the two inner segments of the coil cross-section.

In that example optimization problem, the initial configuration of stimulating system as shown in Figure 4(a) has given the fitness function value of $|\vec{J}_0| = 0.384875 \text{ A/m}^2$. However, for the setup proposed by GA which is presented in Figure 4(b), the result was $|\vec{J}_{opt}| = 3.03364 \text{ A/m}^2$. That yields the improvement ratio of nearly eight times: $\frac{|\vec{J}_{opt}|}{|\vec{J}_0|} = 7.88$.

The initial coil configuration for the third optimization variant — the *synthesis* — was the same as for the first one. This time though, the winding has been divided into 15 turns. That created total of 30 subsections for the whole coil cross-section. The assumption was made that the source current \vec{J}_s can only occur within no more than 8 turns (16 subsections total for the whole cross-section). Again, the condition was that the windings are symmetrical. The idea of active ($\vec{J}_s \neq 0$) and inactive ($\vec{J}_s = 0$) turns, as well as the windings symmetry can be observed in Figure 5(a). Similarly as in the first and second optimization examples, the coil could change its position.

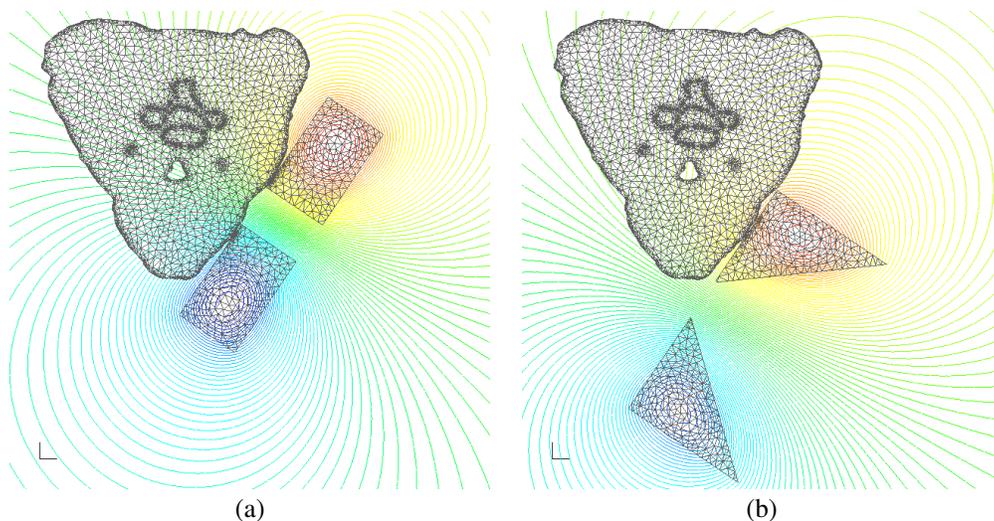


Figure 3: Optimization of stimulating coil with two design parameters (genes). (a) The initial model. (b) The improved resultant model.

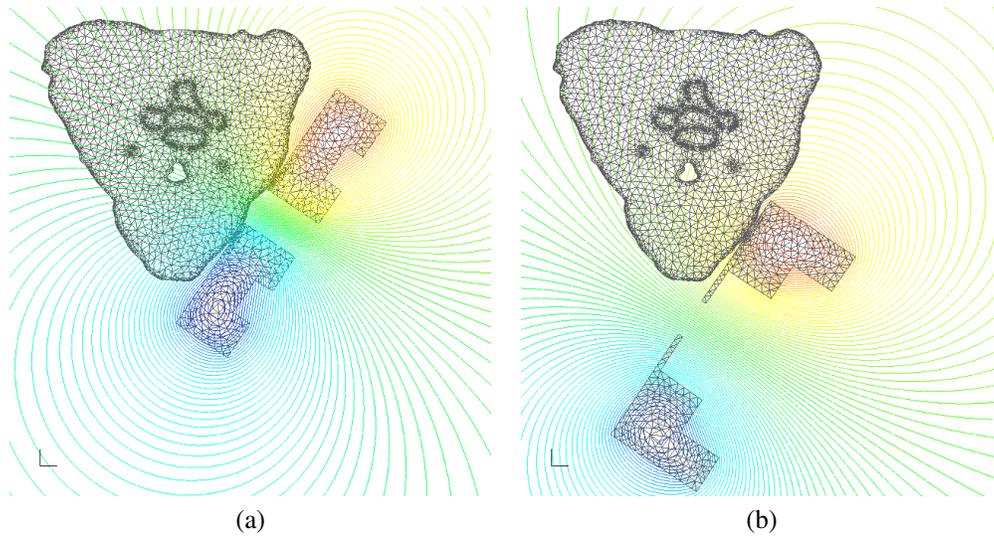


Figure 4: Optimization of stimulating coil with four design parameters (genes). (a) The initial model. (b) The improved resultant model.

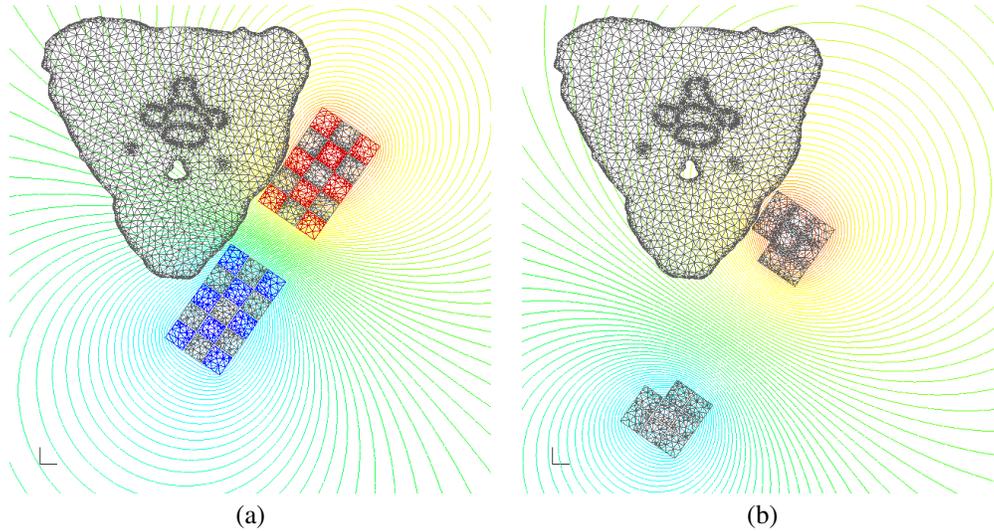


Figure 5: Optimization of stimulating coil with sixteen design parameters (genes) — *synthesis*. (a) The initial model with the maximum allowed symmetrical active turns. (b) The improved resultant model.

All the above-presented assumptions lead to sixteen design parameters (genes):

1. Coil shift in relation to the nerve.
2. Fifteen consecutive variables numbered from 2 to 16 represent the particular turn activity, where:
 - The value of 0 means that the turn is inactive, i.e., $\vec{J}_s = 0$.
 - The value of 1 means that the turn is active, i.e., $\vec{J}_s \neq 0$.

The initial coil configuration as presented in Figure 5(a), yield the fitness function value of $|\vec{J}_0| = 0.225678 \text{ A/m}^2$, while the improved model gives $|\vec{J}_{opt}| = 0.58378 \text{ A/m}^2$. The improvement ratio reaches $\frac{|\vec{J}_{opt}|}{|\vec{J}_0|} = 2.59$.

7. CONCLUSION

The simplified model significantly reduces the time of analysis. It can thus be used for genetic algorithm optimization of VNS stimulator.

ACKNOWLEDGMENT

This work was partially supported by Polish Ministry of Science and Higher Education (research grant No. N N510 148838).

REFERENCES

1. Groves, D. A. and V. J. Brown, “Vagal nerve stimulation: A review of its applications and potential mechanisms that mediate its clinical effects,” *Neuroscience and Biobehavioral Reviews*, Vol. 29, No. 3, 493–500, 2005.
2. Gross, M. and M. Goyal, “Central therapeutic effects of peripheral vagus nerve stimulation,” *American Journal of Electroneurodiagnostic Technology*, Vol. 47, No. 1, 47–52, 2007.
3. Barlas, S., “FDA approves pioneering treatment for obsessive-compulsive disorder,” *Psychiatric Times*, Vol. 26, No. 4, April 8, 2009.
4. Pratap, R., A. Farboud, H. Patel, and P. Montgomery, “Vagal nerve stimulator implantation: The otolaryngologist’s perspective,” *Eur. Arch. Otorhinolaryngol*, Vol. 266, No. 9, 1455–1459, September 2009.
5. Sawicki, B., P. Ponecki, J. Starzyński, and S. Wincenciak, “The quest for the best potential for biomedical eddy currents problems,” *Proceedings of Compumag*, Aachen, Germany, June 2007.
6. Sawicki, B., P. Ponecki, J. Starzyński, and S. Wincenciak, “Scalar potential applied to magnetic stimulation modelling,” *Przegld Elektrotechniczny*, R. 83, Nr. 7–8, 2007.
7. Maccabee, P. J., V. E. Amassian, L. P. Eberle, and R. Q. Cracco, “Magnetic coil stimulation of straight and bent amphibian and mammalian peripheral nerve in vitro: Locus of excitation,” *Journal of Physiology*, Vol. 460, 201–219, 1993.
8. Nagarajan, S. S. and D. M. Durand, “A generalized cable equation for magnetic stimulation of axons,” *IEEE Transactions on Biomedical Engineering*, Vol. 43, No. 3, March 1996.
9. 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, 2271–2293, 1996.
10. Luke, S., L. Panait, G. Balan, S. Paus, Z. Skolicki, E. Popovici, K. Sullivan, J. Harrison, J. Bassett, R. Hubley, A. Chircop, J. Compton, W. Haddon, S. Donnelly, B. Jamil, and J. O’Beirne, “A Java-based evolutionary computation research system,” 2009, <http://cs.gmu.edu/~eclab/projects/ecj/>.
11. Goldberg, D. E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.

Asymmetrical Stripline Based Method for the Electromagnetic Characterization of Metamaterials

S. Gómez, A. Chevalier, and P. Quéffélec

Lab-STICC, UMR-CNRS 3192, UEB, Université de Bretagne Occidentale, France

Abstract— An experimental method for obtaining the effective electromagnetic parameters of metamaterials is presented. The measurement cell consists in an asymmetric stripline which satisfies certain conditions required for the characterization of this type of materials. The advantages of this cell, its electromagnetic analysis and preliminary experimental and simulated results are shown.

1. INTRODUCTION

The metamaterials are artificial composite structures that have electromagnetic properties not observed in their constituent materials or in nature. The most important characteristic is that it is possible to control the response of their permittivity and permeability and adapting their behavior to specific applications. These structures are used in applications involving antennas, filters and couplers due to the possibility of miniaturization which is made by adjusting the value of the refractive index.

The most common way to fabricate this type of materials is by the periodic inclusion of metallic wires inside a dielectric. The shape of these inclusions usually takes the form of Split Ring Resonators (SRR) [1]. The interaction between the electromagnetic fields and the metallic strips produces a resonant behavior that combines two effects: a capacitive response due to the presence of gaps between wires and an inductive effect due to the interaction of the magnetic field with the ring shaped strips. The combination of these two effects produces the enhanced permittivity and permeability and could even result in negative values of these quantities.

Other characteristics of metamaterials are the dispersion, the anisotropy and the heterogeneity. The dispersion: frequency dependence of the permittivity and permeability, implies that in order to describe its behavior a wide band analysis is required. The anisotropy means that the electromagnetic properties of the material are directionally dependent. Finally, due to the heterogeneous character of metamaterials it is necessary to determine if it is possible to utilize homogenization methods for extracting the effective parameters that represent all the structure.

This article presents a new broadband method applied in the characterization of inclusion-based metamaterials SRR type, metasolenoid or similar. It uses an asymmetrical stripline that has been developed in the laboratory for the characterization of magnetic materials [2, 3]. The configuration of the electromagnetic fields in the region where the sample is placed provides an appropriated excitation to the material. Two types of electromagnetic analysis of the stripline are presented. Measured and simulated results for a metasolenoid-type metamaterial are shown.

2. CHARACTERIZATION METHODS

The characterization of the electromagnetic properties of metamaterials is a significant stage in the process of developing new applications and new type of structures. Many factors must be considered for the characterization of metamaterials, as the frequency band, the geometry and the orientation and distribution of the metallic inclusions. This last point is fundamental because the response of the material is determined by the way the electromagnetic field interacts with the inclusions. The enhanced permittivity and permeability result only when the electric and magnetic fields are oriented in one specific direction. In the particular case of a metamaterial type SRR or similar, the enhanced response occurs when the direction of the magnetic field is normal to the surface of the structure, and the electric field is tangent to the spirals. The magnitude of both fields must be constant over all the volume of the metamaterial. The characterization procedure has to provide the correct polarization and distribution of the EM field and take into account the geometry and the orientation of the sample inclusions.

There are some experimental procedures already used for extracting the constitutive parameters of metamaterials. The main objective of these methods is to obtain the S -parameters of a structure charged of material and then determining the permittivity and permeability using different analytical inverse procedures. Free space methods allow us to perform measurements in high frequencies.

They are not recommended for frequencies below 10 GHz, because big samples and different setups are required depending on the frequency band of interest [4]. Other methods based on resonant cavities provide accurate results, but only at a specific frequency, which is not convenient because of the dispersive nature of the metamaterials [5, 6]. The transmission line based methods, such as waveguide [7] or microstrip line [8] could be used in a wide frequency band and do not require big samples for frequencies above 3 GHz, but their main disadvantage is that they do not present the field distribution that produces the expected behavior of the metamaterial. A coaxial line is not practical for these cases because it is necessary to adapt the geometry of the metamaterial to a cylindrical shape.

3. STRIPLINE METHOD

The method proposed in this work uses an asymmetrical stripline that has been developed in our laboratory for the broadband characterization of magnetic materials [2, 3]. The configuration of electromagnetic fields is well adapted for metamaterials characterization. In-situ characterization is allowed because it reproduces an electromagnetic environment close to the one met in practice (use of metamaterial in planar technologies). The measurement cell is composed by a central conductor and two ground planes (Figure 1). The conductor strip is closer to the inferior ground plane to concentrate the most part of the energy in this region where the sample is set. In this area the electric field is parallel to the inclusions and the magnetic field is perpendicular to them. The amplitude of both fields is constant for the entire sample and it is possible to consider the propagation of a quasi-TEM mode. To ensure homogeneous field distribution to avoid the edge effects in the particular case of metamaterials, the width of the sample should not exceed the width of the central conductor. This is why the cell could not be compared to a real microstrip line. The Figure 1 present the electromagnetic field distribution in the cell containing a sample of material.

Using a Vectorial Network Analyzer the S -parameters of the propagation structure are measured. The effective parameters of the material are obtained from this measurements after an electromagnetic analysis of the cell. We use two different theoretical approaches for calculating ϵ_{eff} and μ_{eff} based on the quasi-static approximation.

The first one is a variational procedure in which the cross section of the cell is divided in horizontal layers (air, material and air gaps) representing the different media of the structure (Figure 2(c)). The variational theory consist in homogenizing the transversal section of the line and representing it with an effective permeability and permittivity (Figure 2(b)). Using Green's functions it is possible to extract the capacitance and inductance per unit length of the line and the theoretical values of the effective parameters could be represented in terms of the loaded and unloaded C and L. This first stage corresponds to the direct problem. Using Nicolson/Ross technique it is possible to extract the ϵ_{eff} and μ_{eff} from the reflection and transmission coefficient of the cell (S -parameters). Then, for obtaining the intrinsic parameters of the material from the measurements, an inverse problem is applied. The ϵ and μ from de sample are obtained by an optimization procedure that matches the theoretical and measured values.

This model considers that the length of the sample is infinite in the transversal direction. This assumption could not be applied in the case of metamaterials since the configuration of the electromagnetic fields in the region which is not covered by the central strip is not appropriated for metamaterials's measurement (Figure 2(a)). A correction that allows to obtain the constitutive parameters of the metamaterial should be applied as an additional step for this method.

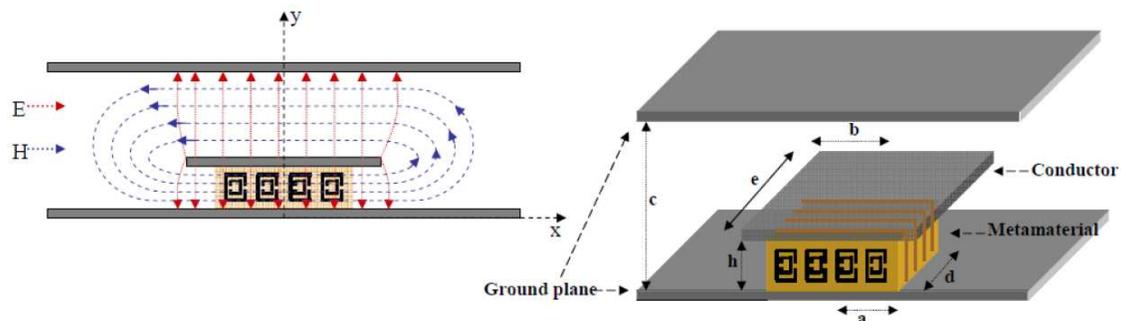


Figure 1: Asymmetrical measurement cell.

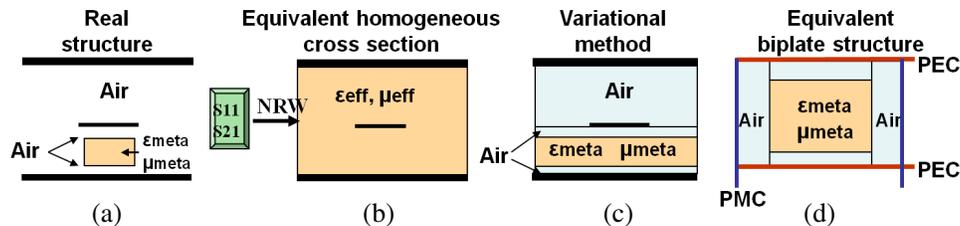


Figure 2: Cross section of the cell and different electromagnetic representations.

The second approach is a quasi-static method based in transmission line theory. Due to the geometry of the cell (concentration of energy below the central conductor), the cross section could be represented as a two conductor structure with boundary conditions as shown in Figure 2(d). We can assume then that only the quasi-TEM mode is propagated. The transmission line analysis allows to represent the S -parameters in terms of the capacitance and inductance per unit length of the line. The inverse problem gives an expression that directly relates the values of the effective permeability and permittivity to the reflection and transmission parameters and no optimization procedure is needed [2].

There is still a correction that must be done, due to the presence of air gaps appearing between the sample and the ground plane and between the metamaterial and the central conductor. For this correction an Effective Medium Approximation (EMA — Wiener's law) is applied. The final expressions are given in Equations (1) and (2).

$$\mu_r = \frac{\omega\mu_0 h_a(1 - R) - Z_0 W/(\gamma)(1 + R)}{\omega h_s \mu_0 (R - 1)} \quad (1)$$

$$\epsilon_r = \frac{h_s(\gamma^2)}{\omega^2 \mu_0 \epsilon_0 (h_s \mu_r + h_a) - h_a \gamma^2}, \quad (2)$$

where ω is the angular frequency, h_a is the height of the air gap, h_s is the height of the sample, Z_0 is the characteristic impedance of the cell, R is the reflection coefficient, γ is the propagation constant and W is the width of the strip conductor.

The last step to retrieve the intrinsic values, taking into account the air over the central conductor is then achieved using the microstrip line model proposed by Hammerstad.

4. ELECTROMAGNETIC SIMULATIONS

Electromagnetic simulations of the measured metasolenoid were made using the commercial software Ansoft HFSS FEM-based simulator (Figure 3(b)). The conditions that must be considered in the simulation of metamaterials are the same mentioned before for measurements: A correct polarization of electromagnetic fields and constant amplitude over all the volume of the metamaterial. Also in this case it is important to define the appropriated boundary conditions and exciting ports such as an infinite periodic structure is represented.

One slab of material was considered. The boundary conditions applied were periodic electric and magnetic conductors (PEC and PMC), and wave ports were used as excitations (Incident TEM mode). This first simulation gave us a general idea of how is the frequency response of the metamaterial, and applying Nicolson Ross expressions obtained from [9] it was possible to obtain the permittivity and permeability from the S -parameters.

5. RESULTS

Figure 3(a) presents the measurement cell and the sample of metamaterial. Figure 4 compares the permittivity and permeability values obtained for the sample using the stripline measurement cell from the two retrieval procedures (variational and transmission line analysis). The theoretical study and the measurements were performed between 10 MHz and 6 GHz. The frequency response of the metamaterial was completely characterized and it was possible to observe the resonant behavior of the electric and magnetic parameters. Also the two retrieval procedures show good agreement and this can be confirmed comparing with the results obtained from the FEM-based electromagnetic simulation. The two different values of permittivity obtained from the TL analysis show the utility of applying mixing laws and effective models due to the heterogeneous nature of the cross section

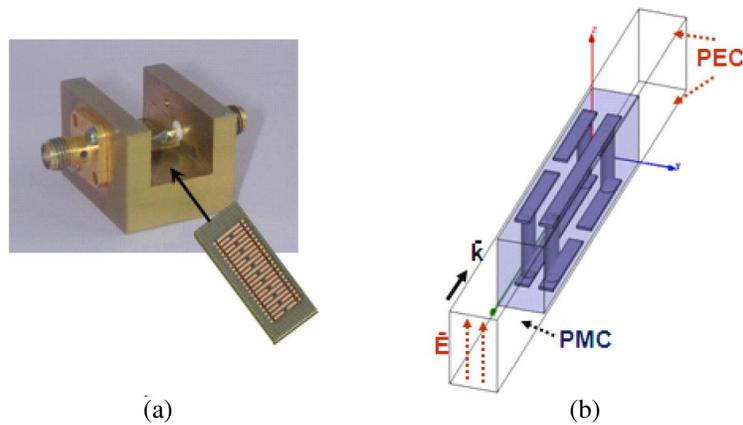


Figure 3: Stripline and unit cell of metamaterial.

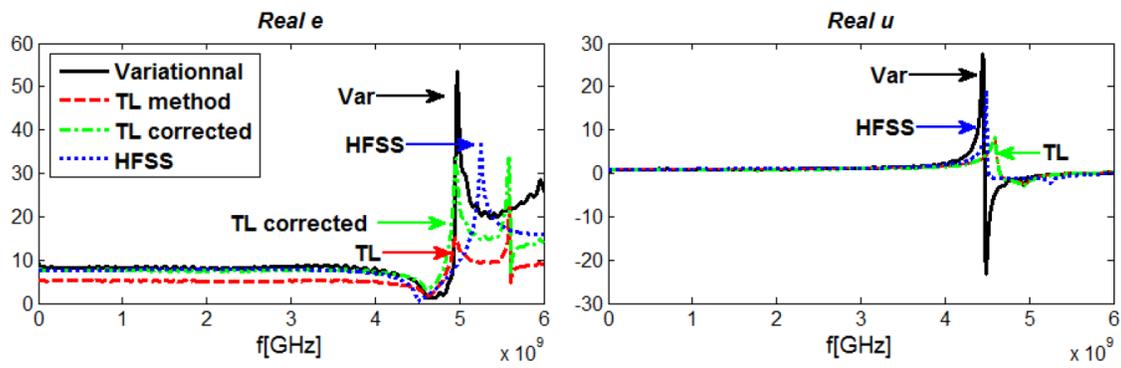


Figure 4: Real part of the permittivity and permeability of metasolenoid-type sample. Comparison between the data given by the two electromagnetic analysis of the cell and HFSS software respectively.

of the cell. Taking into account the presence of air gaps allow us to extract the intrinsic values and to approach the results obtained with the two analysis methods and the FEM-based simulations.

6. CONCLUSION

This work presents the extension of a characterization method applied to the representative measurement of the permittivity and permeability of metamaterials. The complex electromagnetic parameters ϵ and μ of a sample were obtained measuring the S -parameters of an asymmetrical stripline measurement cell and applying analytical procedures. The results were validated applying two different analysis of the cell and comparing the measured data with accurate electromagnetic simulations.

REFERENCES

1. Smith, D., D. C. Vier, T. Koschny, and C. Soukoulis, "Electromagnetic parameter retrieval from inhomogeneous metamaterials," *Phys. Rev.*, Vol. 71, 036617(1–11), 2005.
2. Quéffélec, P., S. Mallgol, and M. Le Floch, "Automatic measurement of complex tensorial permeability of magnetized materials in a wide microwave frequency range," *IEEE Trans. Microwave Theory Tech.*, Vol. 50, 2128–2134, 2002.
3. Salahun, E., P. Quéffélec, M. Le Floch, and P. Gelin, "A broadband permeameter for in situ measurements of rectangular samples," *IEEE Trans. Magn.*, Vol. 37, 2743–2745, 2001.
4. Gregor, R., C. G. Parazzoli, K. Li, B. Koltenbah, and M. Tanielian, "Experimental determination and numerical simulation of the properties of negative index of refraction materials," *Opt. Express*, Vol. 11, 688–695, 2003.
5. Buell, K. and K. Sarabandi, "A method for characterizing complex permittivity and permeability of meta-materials," *Proceeding of IEEE Antennas and Propagation Society International Symposium*, Vol. 2, 408–411, 2002.

6. Chen, L., C. K. Ong, and B. Tan, “Cavity perturbation technique for the measurement of permittivity tensor of uniaxially anisotropic dielectrics,” *IEEE Trans. Instrum. Meas.*, Vol. 48, 1023–1030, 1999.
7. Chen, H., J. Zhang, Y. Bai, Y. Luo, L. Ran, Q. Jiang, and J. A. Kong, “Experimental retrieval of the effective parameters of metamaterials based on a waveguide method,” *Opt. Express*, Vol. 14, 12944–12949, 2006.
8. Yousefi, L., H. Attia, and O. M. Ramahi, “Broadband experimental characterization of artificial magnetic materials based on a microstrip line method,” *Progress In Electromagnetics Research*, Vol. 90, 1–13, 2009.
9. Chen, X., T. Grzegorzcyk, B. Wu, J. Pacheco, and J. Kong, “Robust method to retrieve the constitutive effective parameters of metamaterials,” *Phys. Rev.*, Vol. 70, 016608(1–7), 2004.

Design and Fabrication of Random Optical Surfaces by a Modified Speckle-based Method

V. Brissonneau¹, L. Escoubas², G. Soriano³, F. Flory², G. Maire³, and G. Berginc¹

¹Thalès Optronique S.A., 2 Avenue Gay-Lussac, 78995 Elancourt Cedex, France

²Institut Matériaux Microélectronique Nanosciences de Provence — IM2NP CNRS UMR 6242
Aix-Marseille University, Campus de Saint-Jérôme, Avenue Escadrille Normandie Niemen — Service 231
F-13397 Marseille Cedex 20, France

³Institut Fresnel, Aix-Marseille University, Campus de Saint-Jérôme
Avenue Escadrille Normandie Niemen, Marseille, France

Abstract— Random optical surfaces are of primary interest for several applications such as anti-reflective surfaces, scattering surfaces allowing enhanced performances of photonic sensors, exalted absorption of solar cells, stealth, etc. . . . In this paper, we describe how random optical surfaces with tuned statistical properties can be modeled and fabricated using a speckle pattern recorded in a photoresist. The speckle pattern is obtained by the scattering of a laser beam from a diffractive element such as a piece of tracing paper. This study is mainly focused on the correlation function of the surface which is governed by the correlation function of the scattered intensity recorded in the photoresist. By adjusting the field amplitude distribution of the laser beam, one is able to tune the correlation function of the surface. The original method described by Gray [1], allows the generation of samples with Gaussian correlation. A modified experimental Gray method, comprising a spatial light modulator (SLM) to control the laser beam shape is proposed to fabricate the random optical surfaces having modified statistics: an argon laser beam shaped by the SLM is scattered by a diffusing element on the photoresist deposited at the surface of an optical substrate. Scattering diagrams of the first realized surfaces are compared.

1. INTRODUCTION

Rough surfaces show interesting properties over wide fields. Random optical rough surfaces scattering can be used to achieve enhancement of an imagery system by increasing their depth of focus [2] using an integrated diffuser. Random rough surfaces can also modify the spatial light repartition, leading to light uniformisation [3] or defined spatial shape [4]. Periodic structured surfaces are known to allow a decreasing reflectivity at an interface [5, 6] with a variable efficiency depending on their size and shape [7]. Random rough surfaces show a similar effect but with a wider spectral and angular range [8]. Using such surfaces on a solar cell [9] can reduce reflectivity and so enhance the absorption. We present here a way to generate random optical rough surfaces using a photofabrication method based on the speckle of a diffused laser beam. The method first used by Gray [1] allows to tune the statistical properties of the generated surfaces by modifying parameters of the exposure bench, such as the position of optical components or the number of exposures.

2. EXTENDED GRAY METHOD

The photofabrication of random rough surfaces [10, 11] is mostly based on the method by P. F. Gray [1] published in 1978. The method consists of insulating a photosensitive layer with a speckle pattern as shown in the Figure 1(a). The varying intensity creates a varying height profile in the photosensitive material after exposure development of the photoresist. This speckle pattern is itself obtained by illuminating a diffractive element with an enlarged laser beam. We present here the relationship between the beam shape on the diffractive element and the photofabricated roughness.

We consider a rough diffractive element of height $z = h(\mathbf{r})$ with h a stochastic process. We assume that h is stationary and centered: $\langle h \rangle = 0$. Now this surface is illuminated at pulsation ω by a beam under an incidence angle θ_0 . Defining the wavenumber, $K = \frac{\omega}{c} = \frac{2\pi}{\lambda}$ the incident wavevector is written $\mathbf{K}_0 = \mathbf{k}_0 - q_0 \hat{\mathbf{z}}$ with $\mathbf{K}_0^2 = K^2$. The scattered far-field in the direction defined by the scattering wavevector $\mathbf{K} = \mathbf{k} + q \hat{\mathbf{z}}$ has the shape of a spherical wave of complex amplitude proportionnal to the scattering amplitude $s(\mathbf{k}, \mathbf{k}_0)$.

Following the Kirchhoff theory, this scattering amplitude is written as the surface integral

$$s(\mathbf{k}, \mathbf{k}_0) = \frac{K(\mathbf{k}, \mathbf{k}_0)}{q + q_0} \frac{1}{4\pi^2} \int_{\mathbb{R}^2} g(\mathbf{r}) e^{-i[(\mathbf{k}-\mathbf{k}_0)\cdot\mathbf{r} + (q+q_0)h(\mathbf{r})]} d\mathbf{r} \quad (1)$$

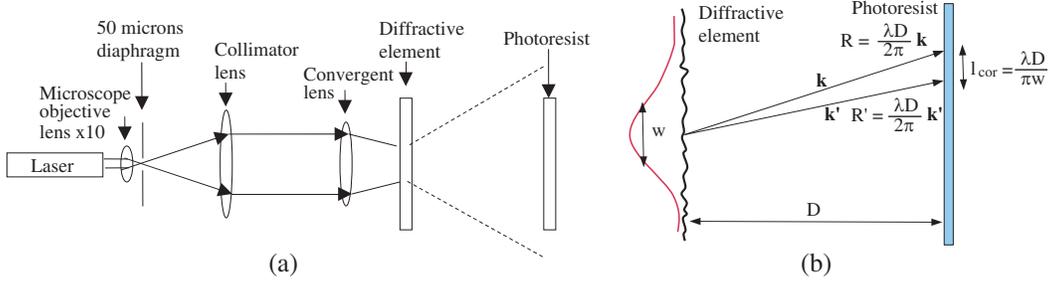


Figure 1: The Gray method (a) principle, (b) beam size and correlation radius on photoresist.

with g modeling the footprint of the illuminating beam on the rough surface and K a factor that depends on the nature of the scattering medium ($z < h(\mathbf{r})$) and on the incident and scattering polarizations. Being independent of the roughness, the term $\frac{K(\mathbf{k}, \mathbf{k}_0)}{q+q_0}$ is first discarded.

If the roughness $\eta(\mathbf{R})$ ¹ obtained on the photosensitive layer after insolation and development of the photoresist is proportional to the optical intensity, defined as $i(\mathbf{k}) = |s(\mathbf{k}, \mathbf{k}_0)|^2$ then the photofabricated roughness correlation is deduced from the intensity autocorrelation function $C(\mathbf{k} - \mathbf{k}') = \langle i(\mathbf{k})i(\mathbf{k}') \rangle$ which compares the intensities in the scattered directions \mathbf{k} and \mathbf{k}' .

If the diffractive element is very rough at the scale of the optical wavelength, then its speckle is fully developed and s is a circular complex Gaussian variable, so that the intensity correlation

$$C(\mathbf{k} - \mathbf{k}') = \langle i(\mathbf{k}) \rangle \langle i(\mathbf{k}') \rangle + |\langle s(\mathbf{k}, \mathbf{k}_0) s^*(\mathbf{k}', \mathbf{k}_0) \rangle|^2 \quad (2)$$

roughly varies as the square of the modulus of the scattering amplitude correlation.

Now, as the beam illuminating the diffractive element is much larger than the characteristic horizontal scale of that element, the scattering amplitude correlation is written

$$\langle s(\mathbf{k}, \mathbf{k}_0) s^*(\mathbf{k}', \mathbf{k}_0) \rangle = \tilde{L} \left(\frac{\mathbf{k} + \mathbf{k}'}{2} - \mathbf{k}_0, q + q_0, q' + q_0 \right) \mathcal{F}[g^2](\mathbf{k} - \mathbf{k}') \quad (3)$$

with Fourier transforms

$$\mathcal{F}[g^2](\mathbf{k}) = \frac{1}{4\pi^2} \int_{\mathbb{R}^2} g^2(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r} \quad (4)$$

$$L(\mathbf{r} - \mathbf{r}', q, q') = \langle e^{-i(qh(\mathbf{r}) - q'h(\mathbf{r}'))} \rangle = \int_{\mathbb{R}^2} \tilde{L}(\mathbf{k}, q, q') e^{+i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} d\mathbf{k} \quad (5)$$

At the end, the correlation of the intensity scattered by the diffractive element varies as the square of the modulus of the Fourier transform of the footprint intensity function $g^2(\mathbf{r})$ of the beam on the diffractive element (Figure 1(b)).

For a Gaussian beam of waist w at the level of the diffractive element,

$$g^2(\mathbf{r}) = I_0 e^{-2(|\mathbf{r}|/w)^2} \Rightarrow C(\mathbf{k}) \propto e^{-\frac{(|\mathbf{k}|w)^2}{4}} \quad (6)$$

Finally, at point \mathbf{R} of the photosensitive layer placed at a distance D of the diffractive element so that the paraxial relation $\mathbf{k} = K \frac{\mathbf{R}}{D}$ holds, the photofabricated roughness correlation appears to be Gaussian

$$C(\mathbf{R} - \mathbf{R}') = \langle \eta(\mathbf{R}) \eta(\mathbf{R}') \rangle C(\mathbf{0}) e^{-|\mathbf{R} - \mathbf{R}'|^2 / \ell^2} \quad \ell = \frac{\lambda D}{\pi w} \quad (7)$$

with a very simple expression for the correlation radius ℓ .

3. PHOTOFABRICATION

The optical bench (Figure 2(a)), consists of an Argon Ion Laser working at 363.8 nm with a maximum output power of 200 mW. The beam is enlarged by using an afocal system. The expanded beam is then shaped by a spatial light modulator (SLM). This SLM, known as DMD, Digital Micromirror Device (Vialux GmbH/Texas Instruments), is an array of 1920 by 1080 micromirrors.

¹ \mathbf{r} indicates the point of height $h(\mathbf{r})$ on the diffractive element while $(\mathbf{R}, \eta(\mathbf{R}))$ is on the photoresist.

Modulation is done by switching each micromirror direction. Two active directions can be defined, an “ON” position, where the beam is reflected in the axis of the optical bench, and an “OFF” position where the beam is sent out of the optical bench. Thus, a binary modulation is allowed. Assuming an image projection during a defined time, one is able to realize a gray-scale modulation by adjusting the “ON”/“OFF” ratio. Eight binary images are successively projected. Thus, during the projection time, a 256 gray-scale (8-bits) image is reproduced. The modulated beam passes through a convergent lens and a diffusing element placed just behind the lens.

Firstly, an exposure was made without using the modulation system. The modulation system was replaced by a plane mirror. The sample was exposed during 46 s using a 100 mW laser beam emitting at 363.8 nm. The theoretical and experimental correlations are considered in the Figure 4 in the paragraph 4.

The experimental correlation function will depend on the parameters of the exposure bench such as the correlation function of the diffusing element and the distance between the sample and the diffusing element. The first experiments were realized with the same diffusing element and the same distances between optical components. The exposure time is adjusted so that the photoresist sizing energy threshold, minimal energy where the photoresist is shaped, is obtained.

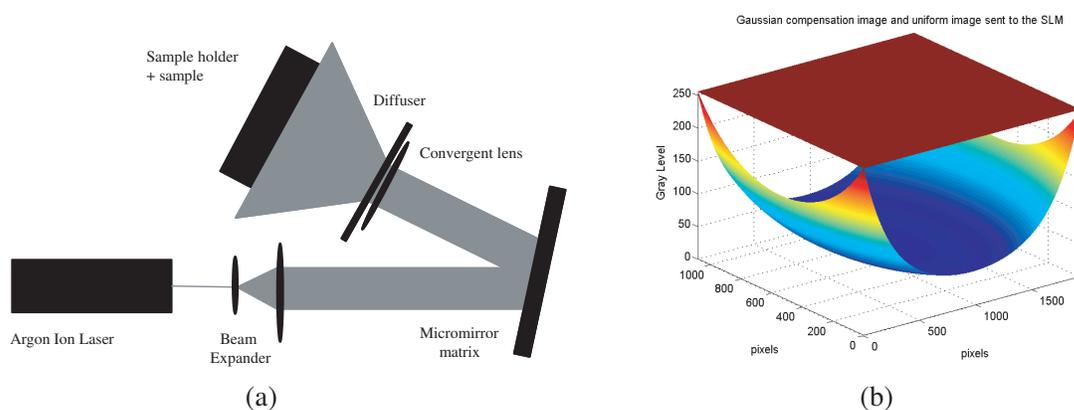


Figure 2: Photofabrication (a) experimental bench, (b) “Gaussian compensation” and “mirror” images displayed with DMD.

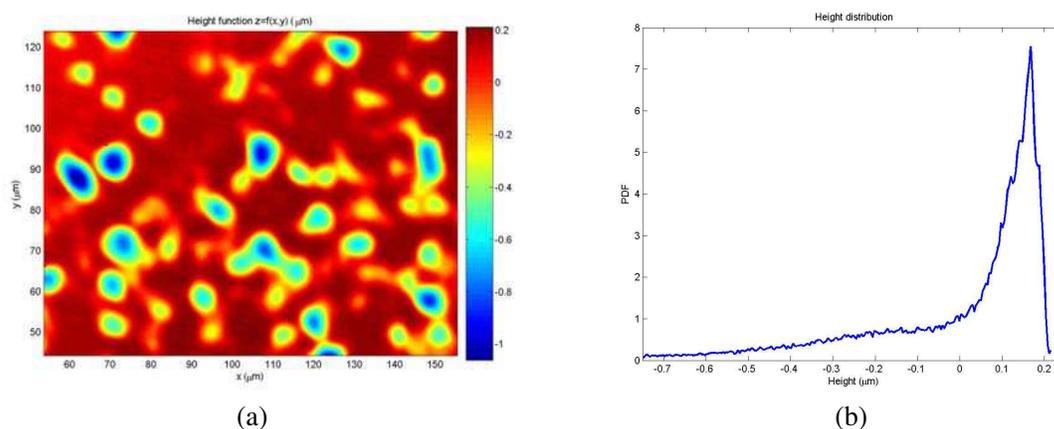


Figure 3: Interferometric microscope measurement, sample realized with micromirror and “mirror” image (a) map of height, (b) the height distribution.

For each photofabricated sample that are here shown, we realized a single exposure, so that the fabricated surfaces height distribution are exponential. The height distribution is shown in the paragraph 4.

The first exposure using micromirrors was realized using the DMD working as a mirror, each micromirror being on the “ON” position. Considering the diffraction of the DMD, only the first order beam is used, the cumulated exposure time was increased to 128 s.

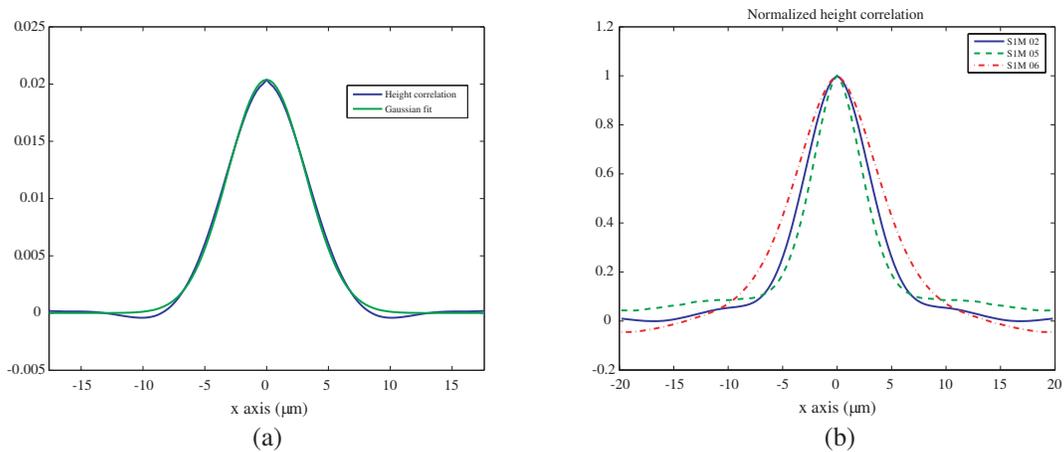


Figure 4: Measured correlation of photofabricated samples (a) without micromirror matrix (comparison of theory and experiment), (b) correlation with samples realized using micromirror matrix and a Gaussian beam (S1M-02) and a Gaussian compensation as explained in the Figure 2(b) (S1M-05 and S1M-06).

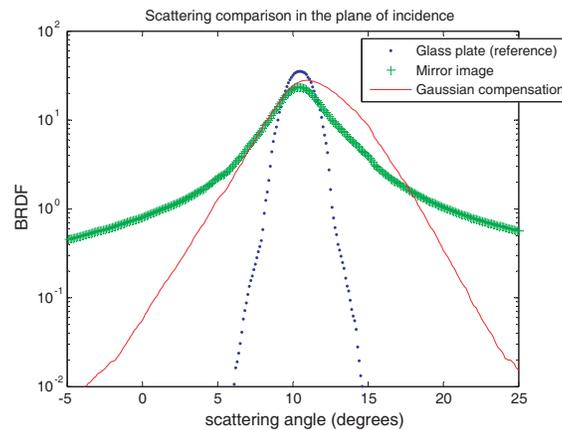


Figure 5: Scattering from realized samples in the plane of incidence.

We then modify the beam's spatial repartition so that we approach an uniform intensity. For this we display an inverted image of the Gaussian beam (Figure 2(b)).

As explained, we choose to use this image in order to realize a compensation of the intensity, so that we have a uniform intensity distribution on the diffuser. In further studies, it will allow us to apply more complex patterns freeing us of the Gaussian laser beam distribution.

4. OPTICAL CHARACTERIZATION

The fabricated surfaces are characterized by an interferometric microscope. This microscope can take a complex image of the roughness, with height information in the measured phase, allowing us to retrieve both height correlation and height distribution. An image of the height and a plot of the height distribution appear on Figure 3 for the sample fabricated with a "mirror" image displayed on the matrix. The intensity of a fully developed speckle is exponentially distributed. The height distribution shown on the Figure 3(b) is clearly exponential. Considering that the sample was realized with a simple exposure, it agrees with Gray [1]. The height root mean square of this sample is around 250 nm, while it is about 50 nm for the sample realized with the "Gaussian compensation" image.

The measured correlation of samples realized with and without the micromirror matrix are shown in the Figure 4. As planned, we can see that the correlation function is modified.

Difference between samples S1M-05 and S1M-06 can be explained by the roughness of these samples. The height root mean square of the S1M-05 sample is about 35 nm while S1M-06 sample one is about 50 nm.

We then consider the light scattered by the realized surfaces. Samples are metalized with

a 60 nm deposited gold layer. The scattering measurement is realized with a goniophotometer REFLET (STIL SA). Figure 5 shows the scattering in the plane of incidence for three samples. The blue curve is a reference consisting of a glass substrate covered with photoresist and the gold cover. The green one is realized with a “mirror” image (every mirror is “ON”). The red curve corresponds to a sample fabricated using the Gaussian compensation image (Figure 2(b)). These measures are realized at the wavelength 600 nm with a ten degrees incidence angle, so that were able to measure the specular part of the scattered beam.

As shown on the scattering diagram represented in Figure 5, the angular scattering of the “compensated” sample is much lower than the one of the “mirror” image.

5. CONCLUSION

In this paper, we present random rough optical surfaces photofabricated using the speckle of a modified laser beam. We show that with a modification of the beam pattern on the scattering element we are able to modify the speckle and so the photofabricated roughness. This paves a way to tuning the statistics of this kind of surface, with the possibility to shape the scattering pattern.

ACKNOWLEDGMENT

These studies were funded by a Délégation Générale à l’Armement (DGA) grant No. 2008 34 0033.

REFERENCES

1. Gray, P. F., “A method of forming optical diffusers of simple known statistical properties,” *Optical Acta*, Vol. 25, No. 8, 765–775, 1978.
2. Garcia-Guerrero, E. E., E. R. Mendez, and H. M. Escamilla, “Design and fabrication of random phase diffusers for extending the depth of focus,” *Optics Express*, Vol. 15, 910–923, 2007.
3. Mendez, E. R., E. E. Garcia-Guerrero, H. M. Escamilla, A. A. Maradudin, T. A. Leskova, and A. V. Shchegrov, “Photofabrication of random achromatic optical diffusers for uniform illumination,” *Applied Optics*, Vol. 40, 1098–1108, 2001.
4. Mendez, E. R., T. A. Leskova, A. A. Maradudin, M. Leyva-Lucero, and J. Munoz-Lopez, “The design of two-dimensional random surfaces with specified scattering properties,” *Journal of Optics A — Pure and Applied Optics*, Vol. 7, 141–151, 2005.
5. Bouffaron, R., L. Escoubas, V. Brissonneau, J. J. Simon, G. Berginc, P. Torchio, F. Flory, and P. Masquet, “Spherically shaped micro-structured antireflective surfaces,” *Optics Express*, Vol. 17, 21590–21597, 2009.
6. Bouffaron, R., L. Escoubas, J. Simon, P. Torchio, F. Flory, G. Berginc, and P. Masquet, “Enhanced antireflecting properties of micro-structured top-flat pyramids,” *Optics Express*, Vol. 16, 19304–19309, 2009.
7. Escoubas, L., R. Bouffaron, V. Brissonneau, J. J. Simon, G. Berginc, F. Flory, and P. Torchio, “Sand-castle biperiodic pattern for spectral and angular broadening of antireflective properties,” *Optics Letters*, Vol. 35, No. 9, 1455–1457, 2010.
8. Chyan, J. Y., W. C. Hsu, and J. A. Yeh, “Broadband antireflective poly-Si nanosponge for thin film solar cells,” *Optics Express*, Vol. 17, 4646–4651, 2009.
9. Rockstuhl, C., S. Fahr, K. Bittkau, T. Beckers, R. Carius, C. Ballif, and F. Lederer, “Comparison and optimization of randomly textured surfaces in thin-film solar cells,” *Optics Express*, Vol. 18, No. 103, 335–342, 2010.
10. Michel, T., A. A. Maradudin, and E. Mendez, “Enhanced backscattering of light from a non-gaussian random metal surface,” *J. Opt. Soc. Am. B*, Vol. 6, 2438–2446, 1989.
11. Maradudin, A. A., *Light Scattering and Nanoscale Surface Roughness (Nanostructure Science and Technology)*, Springer, 2007.

A New Multi-ring SRR Type Metamaterial Design with Multiple Magnetic Resonances

O. Turkmen^{1,2}, E. Ekmekci^{1,3}, and G. Turhan-Sayan¹

¹Department of Electrical and Electronics Engineering
Middle East Technical University, Ankara, Turkey

²Department of Electronics and Telecommunications Engineering
Kocaeli University, Kocaeli, Turkey

³Department of Electronics and Communication Engineering
Suleyman Demirel University, Isparta, Turkey

Abstract— In this study, we propose a new SRR unit cell design for multi-band metamaterial applications. The suggested unit cell structure consists of N number of concentric split rings to obtain magnetic resonances at N distinct frequencies. The value of each distinct resonance frequency can be adjusted by changing design parameters such as metal widths and gap distances for each ring as well as ring-to-ring separations. The suggested multi-band structure is simulated using CST Microwave Studio. Effective medium parameters of the resulting multi-ring SRR type metamaterials are estimated by a retrieval algorithm.

1. INTRODUCTION

Metamaterials are specially designed periodic structures which can show unique properties such as having negative values of permeability and/or negative values of permittivity over finite frequency bands. Theoretical aspects and many important applications of metamaterials in microwave, terahertz and optic regions have been investigated in detail in a vast amount of publications for the last decade [1–6]. Split ring resonator (SRR) type magnetic resonators are among the most popular metamaterial structures having negative permeability over narrow frequency bands. Although various forms of SRR structures have been found useful in narrowband applications, research on metamaterials has also been focused recently on the design of multiband and/or frequency tunable metamaterials.

In this study, a new N -ring SRR unit cell design is introduced for multi-band metamaterial applications. For a given substrate material, design parameters are the side lengths and widths of metal strips, gap distances for each ring, and the separation distances between the rings. As a proof of concept, several multi-band SRR arrays are designed and simulated in this paper for three different cases (for $N = 1, 2$ and 3) by using CST Microwave Studio. Complex transmission and reflection characteristics (i.e., the complex S-parameters S_{21} and S_{11}) of the proposed SRR arrays are obtained by CST, and then they are used to extract the effective medium parameters μ_{eff} and ε_{eff} of the designed metamaterials to verify the nature of resulting resonances. The basic retrieval procedure given in [7] is used for parameters estimation.

2. DESIGN AND SIMULATIONS

The schematic view of the proposed multi-ring unit cell for the case of $N = 3$ is shown in Figure 1 together with the excitation details where propagation direction is in the direction of x -axis, incident

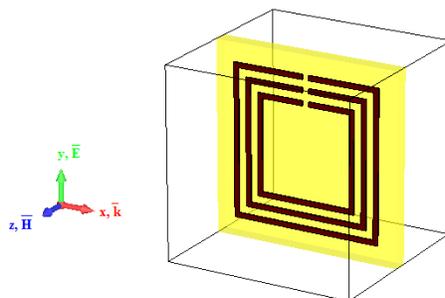


Figure 1: Simulation setup and the excitation.

H field is perpendicular to the SRR plane (i.e., in the direction of z -axis) and the incident E field is perpendicular to the gap containing edges of the SRR rings (i.e., in the direction of y -axis). A cubic computational region with a side length of 7.5 mm is used during the simulation procedure as shown in Figure 1. The PEC type boundary conditions are applied at the boundary surfaces perpendicular to the E field while the PMC type boundary conditions are applied at the boundary surfaces perpendicular to the H field. Remaining boundaries are defined as the input and output ports. Using this setup, single-ring, two-ring and three-ring SRR arrays are designed and simulated on a planar substrate with the relative permittivity of 4.4 and the loss tangent of 0.001. Metallic inclusions are made of copper with the thickness of 0.035 mm and the conductivity of 5.8×10^7 S/m. Dimensions of the substrate in the x , y and z directions are 7.5 mm, 7.5 mm and 0.6 mm respectively.

Each one of the SRR arrays simulated in the first step of this work are composed of only one type of single-ring square-shaped unit cells shown in Figure 2 where the side lengths (L) of the SRR metal rings are chosen to be 6 mm, 5 mm and 4 mm for the SRR-A, SRR-B and SRR-C type unit cells, respectively. The same gap distance $g = 0.25$ mm and the same metal strip width $w = 0.2$ mm are used in these three different unit cells.

In the next step, a two-ring SRR unit cell is designed by combining the SRR-A and SRR-B unit cell topologies which are aligned in the gap-to-gap configuration as shown in Figure 3(a). Similarly,

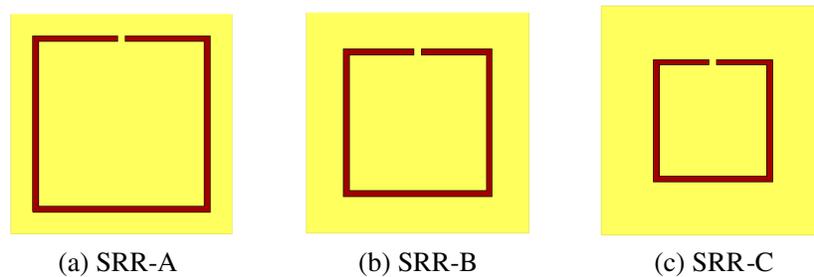


Figure 2: Front view of the single-ring square-shaped SRR unit cells with different parameters.

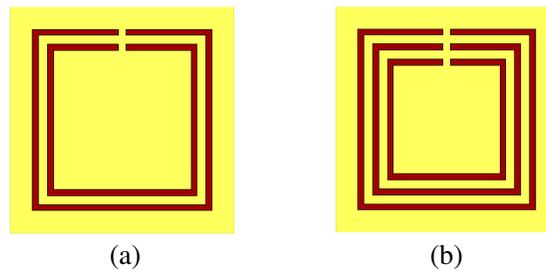
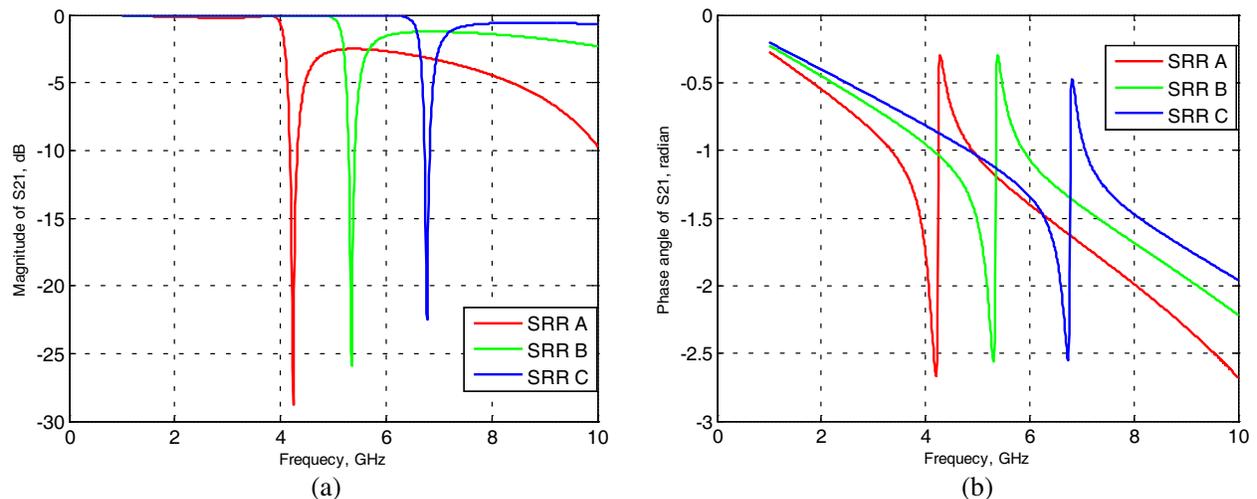


Figure 3: Multi-ring SRR unit cell structures with two and three-rings.



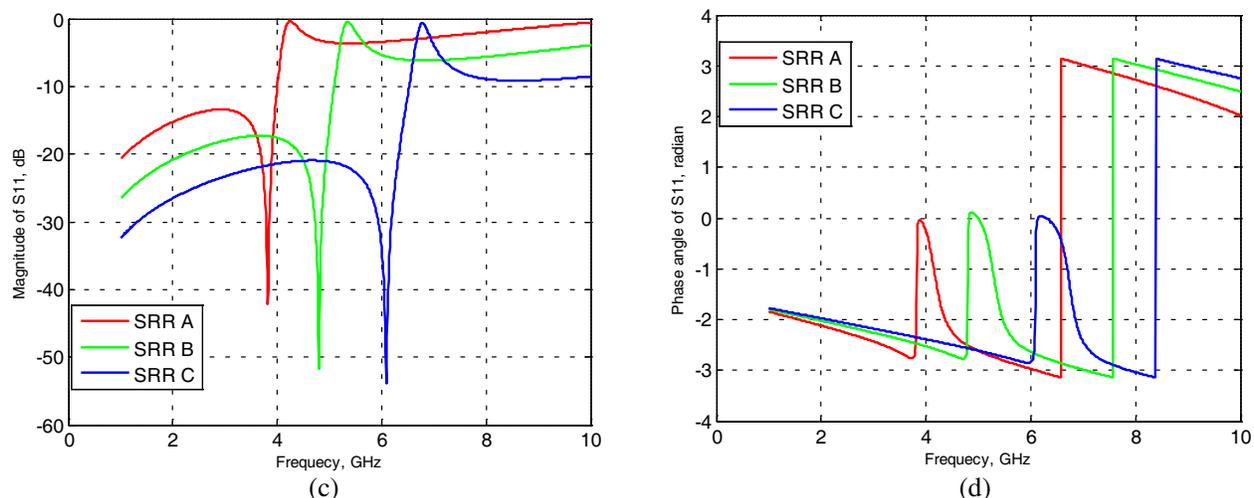


Figure 4: Transmission and reflection spectra of the single-ring SRR arrays shown in Figure 2. (a) Magnitude of S_{21} . (b) Phase angle of S_{21} . (c) Magnitude of S_{11} . (d) Phase angle of S_{11} .

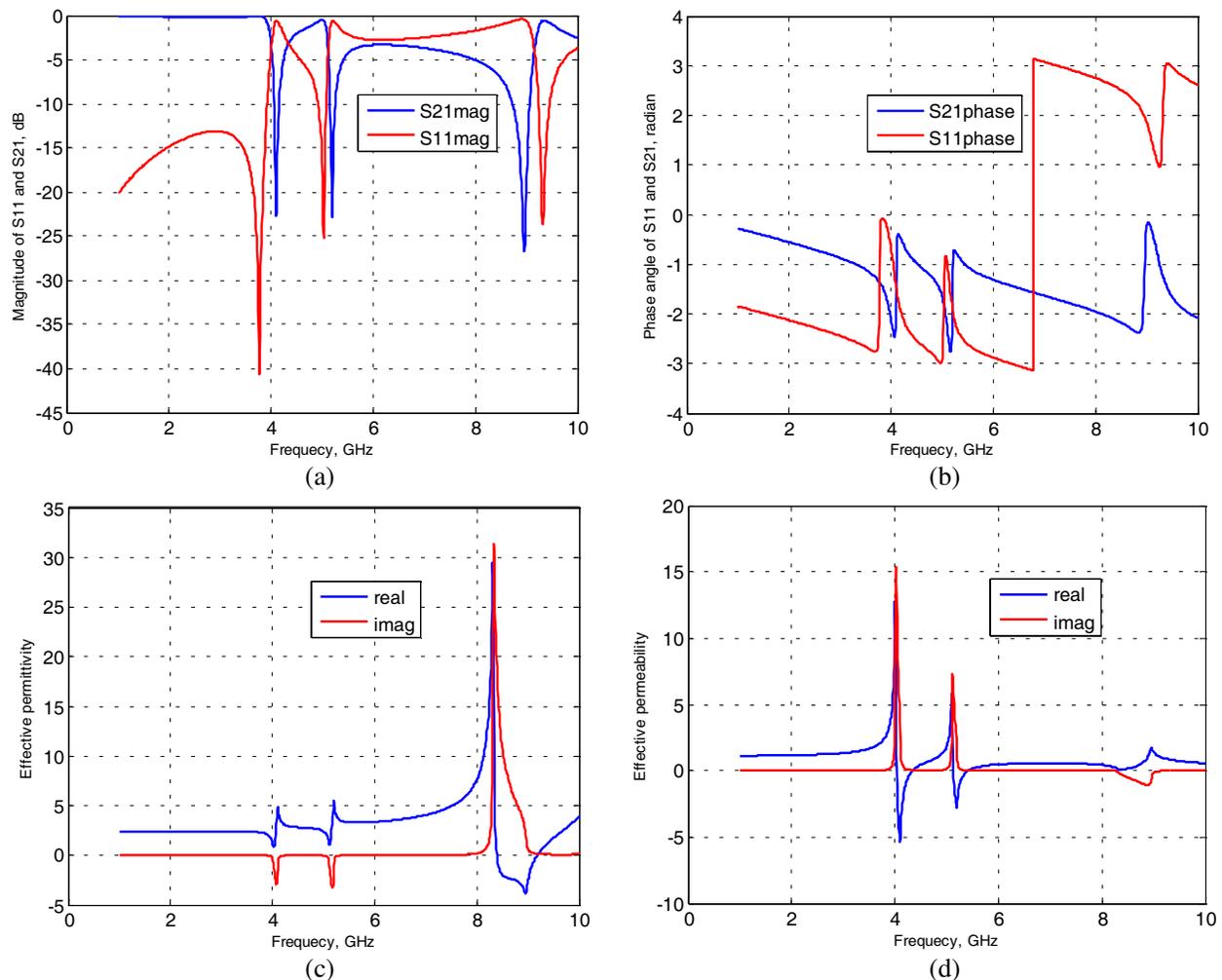


Figure 5: Transmission and reflection spectra, and the effective medium parameters for the two-ring SRR array. (a) Magnitudes of S_{21} and S_{11} . (b) Phase angles of S_{21} and S_{11} . (c) Real and imaginary parts of ϵ_{eff} . (d) Real and imaginary parts of μ_{eff} .

a three-ring SRR unit cell is also designed by combining all three types of unit cells SRR-A, SRR-B and SRR-C as given in Figure (3b). In both of these configurations, the distance between the rings

is chosen to be $s = 0.3$ mm.

3. RESULTS

Complex S -parameters S_{21} and S_{11} computed for the single-ring array topologies SRR-A, SRR-B and SRR-C revealed magnetic resonance frequencies at 4.24 GHz, 5.34 GHz, and 6.77 GHz, respectively. The magnitude and phase plots for these transmission and reflection spectra are shown in Figure 4(a) through 4(d). Increase in the side length (and hence in the overall length) of the metal ring results in an increase of the self inductance [8] leading to a decrease in the LC resonance frequency of the resonator, as expected.

Next, the magnitude and phase spectra of the S_{21} and S_{11} parameters are computed for the two-ring SRR topology of Figure 3(a). Resulting plots are given in Figure 5 together with the plots for the real and imaginary parts of the retrieved parameters, effective permittivity and effective permeability. As seen in Figure 5, the two-ring SRR array structure has three distinct resonances over the range from 1 GHz to 10 GHz. Two of those frequencies at 4.01 GHz and 5.19 GHz are magnetic resonances and the last one at 8.96 GHz is an electric resonance.

As shown in Figure 6, on the other hand, the three-ring SRR array has four distinct resonances. Three of them (at 4.1 GHz, 5.05 GHz and 6.53 GHz) are magnetic resonances and the one at 8.95 GHz is an electric resonance. These results demonstrate that a desired number of magnetic resonances can be realized by selecting the number of SRR rings within the limits of geometrical constraints. It is also worth mentioning that resonance frequencies can also be adjusted by changing the design parameters g , w and s .

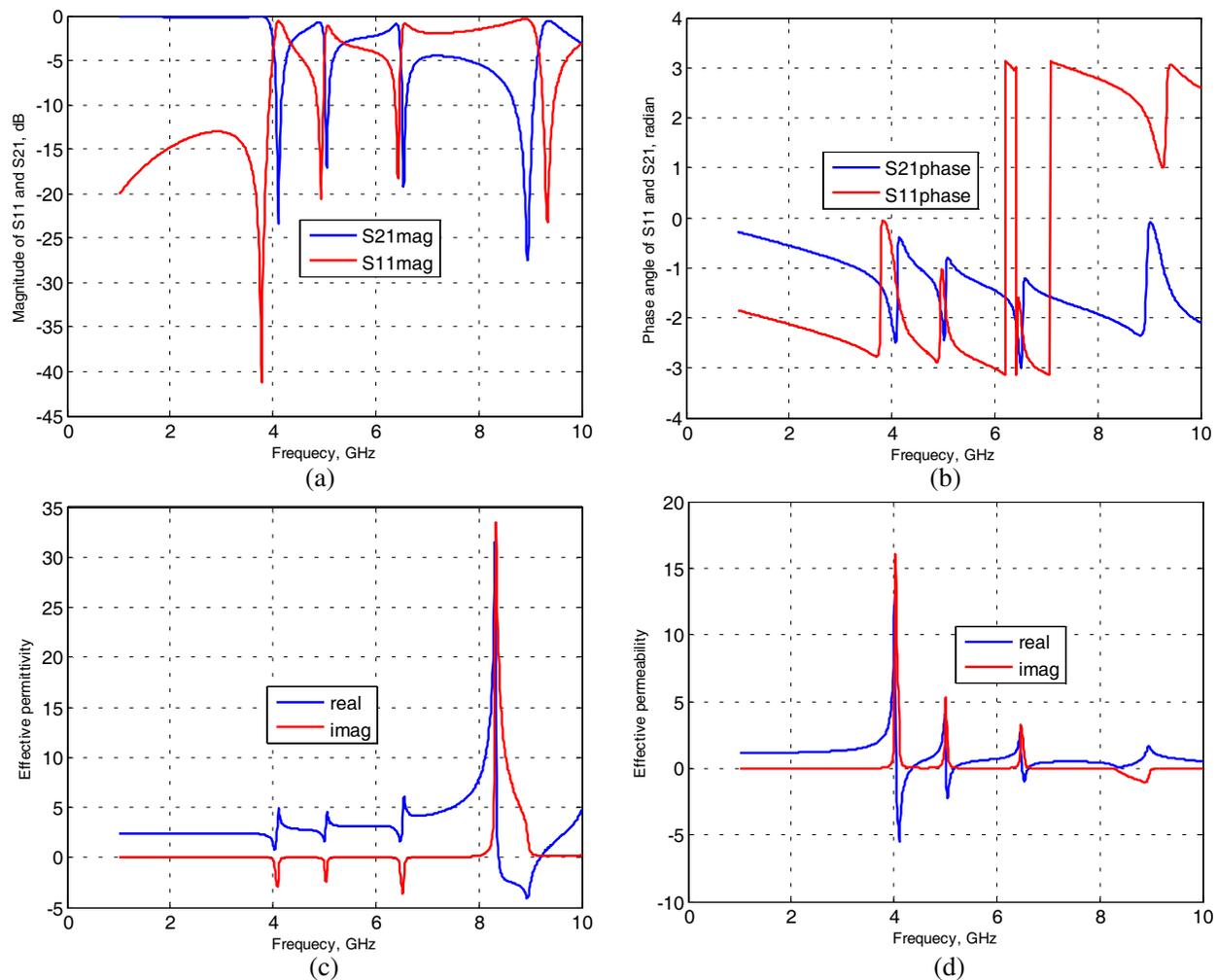


Figure 6: Transmission and reflection spectra, and the effective medium parameters for the three-ring SRR array. (a) Magnitudes of S_{21} and S_{11} . (b) Phase angles of S_{21} and S_{11} . (c) Real and imaginary parts of ϵ_{eff} . (d) Real and imaginary parts of μ_{eff} .

4. CONCLUSIONS

The possibility of multiple magnetic resonances with negative permeability bands are demonstrated for the suggested multi-ring SRR topologies, where the number of resonances is determined by the number of concentric rings. Also, this topology has the flexibility of adjusting the resonance frequencies by changing the design parameters such as the gap width, metal width and inter-ring distances. It should be noted that this multi-ring SRR structure is also capable of displaying multiple electrical resonances under proper excitation conditions. It is believed that the proposed multi-ring SRR design will provide an electrically small and easy-to-fabricate alternative to the present multi-band metamaterial structures [9, 10]. As a future work, proposed SRR structures will be fabricated and the simulation results will be verified by experimental results. Equivalent circuit model of the multi-ring SRR array structure will also be investigated.

REFERENCES

1. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of ϵ and μ ," *Sov. Phys. — Usp.*, Vol. 10, 509–514, 1968.
2. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, "Low-frequency plasmons in thin wire structures," *J. Phys., Condens. Matter*, Vol. 10, 4785–4809, 1998.
3. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE Trans. Microw. Theory Tech.*, Vol. 47, No. 11, 2075–2084, 1999.
4. Smith, D. R., W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, "Composite medium with simultaneously negative permeability and permittivity," *Phys. Rev. Lett.*, Vol. 84, No. 18, 4184–4187, 2000.
5. Ekmekci, E. and G. Turhan-Sayan, "Comparative investigation of resonance characteristics and electrical size of the double-sided SRR, BC-SRR and conventional SRR type metamaterials for varying substrate parameters," *Progress In Electromagnetics Research B*, Vol. 12, 35–62, 2009.
6. Soukoulis, C. M., T. Koschny, J. Zhou, M. Kafesaki, and E. N. Economou, "Magnetic response of split ring resonators at terahertz frequencies," *Phys. Stat. Sol. B*, Vol. 244, 1181–1187, 2007.
7. Ghodgaonkar, D. K., V. V. Varadan, and V. K. Varadan, "Free-space measurement of complex permittivity and complex permeability of magnetic materials at microwave frequencies," *IEEE Trans. Instrum. Meas.*, Vol. 39, No. 2, 387–394, 1990.
8. Kafesaki, M., T. Koschny, R. S. Penciu, T. F. Gundogdu, E. N. Economou, and C. M. Soukoulis, "Left-handed metamaterials: Detailed numerical studies of the transmission properties," *J. Opt. A: Pure Appl. Opt.*, Vol. 7, 12–22, 2005.
9. Ekmekci, E., K. Topalli, T. Akin, and G. Turhan-Sayan, "A tunable multi-band metamaterial design using micro-split SRR structures," *Opt. Express*, Vol. 17, No. 18, 16046–16058, 2009.
10. Ekmekci, E. and G. Turhan-Sayan, "Single loop resonator: Dual-band magnetic metamaterial structure," *Electron. Lett.*, Vol. 46, No. 5, 324–325, 2010.

Active Earth Observation from Unmanned Aerial System

C. E. Lin¹, Y.-C. Huang¹, Y.-H. Lai¹, Y.-L. Yeh¹, C.-C. Cheng², and C.-C. Nien³

¹Department of Aeronautics and Astronautics, National Cheng Kung University, Tainan 701, Taiwan

²EOL, Industrial Technology Research Institute, Chu-Tong, Hsin-Chu 310, Taiwan

³ISTC, Industrial Technology Research Institute, Chu-Tong, Hsin-Chu 310, Taiwan

Abstract— Earth observation may also be carried by low altitude photo acquisition using flying vehicles. The unmanned aerial system (UAS) includes unmanned aerial vehicle (UAV), flight operation and control, payloads for image acquisition, and communication for data downlink. This paper presents a UAS implementation project using high resolution camera set to capture earth images in remote operation. The high resolution camera set includes a wide angle camera and a twin camera. The fast pan-tilt mechanics turns the mirror by switching twin cameras to perform a 25 times zoom-in. The adopted UAV is remotely operated or autopilot to the targets. The earth observation is implemented from low altitude UAV flights by setting target waypoints. The proposed UAS is applied in disaster surveillance before rescue force is evolved.

1. INTRODUCTION

Earth observation from satellite is quite efficient to search for wide area scenario [1, 2]. However, for local area surface instantaneous changes, it might be too long to wait until next available satellite orbit [1]. Satellite images may also be influenced by cloud. Infrared or radar image acquisition technologies are developed to improve satellite earth observation [2]. In disaster surveillance, the ground situation observation might always conflict with poor weather that satellite may not be useful at all.

Civilian unmanned aerial vehicle (UAV) systems can provide useful means for ground surveillance data acquisition from low altitude [3–5]. UAV application for earth observation takes most considerations in its function capability and performance. Conceptually, UAV does not necessary to consider large or small. Its capability is dependent on payloads and acquisition [5]. Mission planning and autopilot flight control are two key functions to support the important payloads for image acquisition and data transmission [6, 7]. Most technologies are quite mature at present, except the communication media is limited from available frequency spectrum.

In disaster surveillance and rescue search applications, UAV requires high performance to overcome poor weather conditions of strong head winds and rain with heavier payloads. Under such consideration, a UAV carries delicate cameras for territorial search being a practical solution [4]. Different types of payloads are designed into UAVs to search under adverse conditions from low altitudes. Two contrasting applications in poor weather earth monitoring require enough thrust to resist high headwinds for cruising and homing, and enough drag and lift for image acquisition payloads at lower speed. UAV design needs to include more parameters in implementation.

For earth observation purposes, digital cameras for image acquisition from low altitude become the most important payload on UAVs. To guide flight manipulation, a wide angle camera is used to view the heading scene of environment awareness. A precision camera set is located in the main cabin to view straight downward to the target area. Zooming control mechanics is necessary to change observe effect on different spots.

This paper presents a recent development in unmanned aerial system for disaster surveillance applications by integrating delicate camera sets for ground image acquisition. Some preliminary test results on the high payload mission UAV are presented and demonstrated in this paper. The proposed Cardinal UAV system is working on continuous modification and verification for better performance to suit for surveillance mission operations.

2. SYSTEM DESIGN

2.1. Cardinal UAV Ce-71

The Cardinal UAV (Ce-71) system has surveillance mission demand for large cabin space to accept different payload and high thrust power for higher payload, as shown in Figure 1.



Figure 1: Cardinal UAV Ce-71.



Figure 2: Ground computer interface.

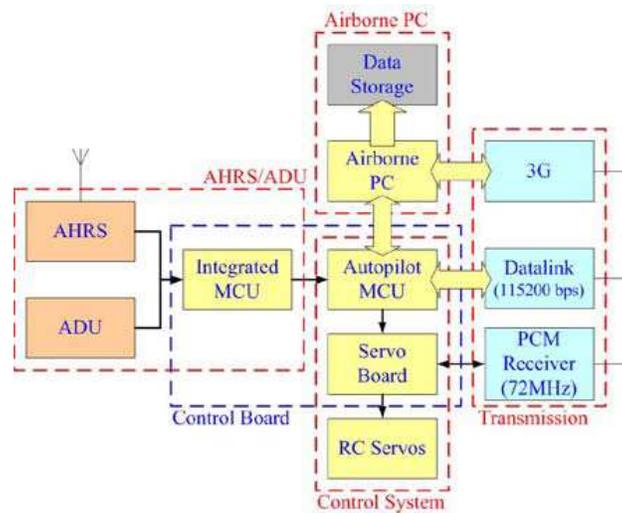


Figure 3: Flight control system.

2.2. UAV Flight Operation

As an academic research, the flight control system is an important task of practice. The flight control system uses multiple microprocessors integrated with MEMS sensors as the data acquisition subsystem and flight control subsystem in coordination with the airborne computer via powerful data link communication from 900 MHz, 2.4 GHz and mobile 3G.

To determine the UAV dynamic motion in missions, the ground control center on a laptop computer is set up to receive downlink messages and uplink flight control commands [7]. The navigation waypoints for each mission are uplinked onto the airborne computer through 900 MHz and 3G systems, and displayed on the ground computer as shown in Figure 2.

2.3. Flight Control Avionics

The flight control avionics is composed of an airborne PC, microprocessor-based flight control and data acquisition with three communication sub-systems. The airborne PC carries a flight monitor and operates via 3G communication to receive navigation waypoints and further commands. The designed flight control system is shown in Figure 3.

2.4. Communication

The communication system includes 900 MHz for wide range two-way flight control, 2.4 GHz for video downlink, and 3G mobile communication for uplink commands/downlink flight surveillance [8]. Since academic research is not permitted with high power communication to cover wide range operation, all available VHF/UHF resources have coverage limit by environment interference to result in capability deficiency.

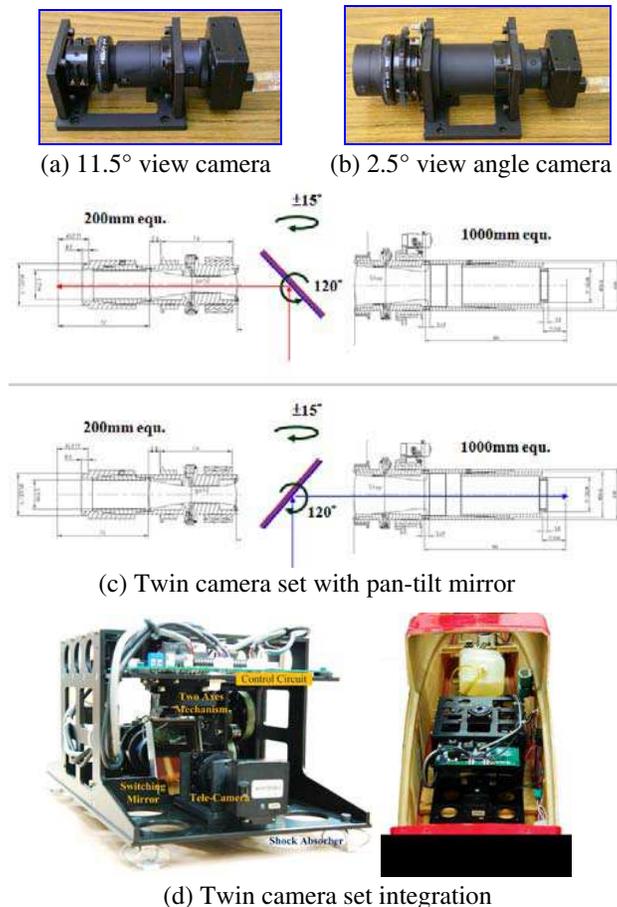
3. CAMERA PAYLOADS FOR EARTH OBSERVATION

Two camera sets are installed on the Cardinal 71 for surveillance image acquisition. The front camera can be a wide angle camera or web-cam heading 5 degrees forward, as shown in Figure 4. The second lens is a high resolution tele-camera with fast zooming mechanics in the main cabin. The pan-tilt zooming mechanics control takes 0.5 second to switch from normal lens into 25 time tele-focus lens. It is used to capture high resolution images whenever the wide angle camera captures scene of need. In the surveillance mission, the tele-focus lens is switched into continuous target image acquisition.

In the flight tests, a 25 time twin camera set is used to search for the ground target and shoot photos. The twin camera set is constructed with two lenses of 200 mm and 1000 mm (135 format equivalent) on each side of pan-tilt reflecting mirror, as shown in Figures 5(a) and (b). The twin cameras have their FOV (field of view) angles as 11.5° and 2.5° , respectively, at each side of the fixture. The pan-tilt switching mirror operation is shown as Figure 5(c). This camera is viewed from the UAV bottom and its pan-tilt mechanism, as shown in Figure 5, reflects input images in either direction of the two lenses.



Figure 4: Front camera arrangement.



(d) Twin camera set integration

Figure 5: Fast zooming twin camera set.



Figure 6: Successive photos at 1 Hz shutter.

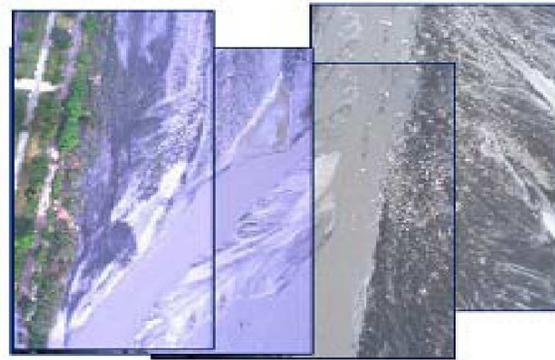


Figure 7: Riverbed watch in successive photos.

4. TEST AND VERIFICATION

Cardinal UAV family has completed test flights for function verification as well as image acquisition from cameras with data downlink. The Cardinal is installed with payloads of IMU, ADU, AHRS, GPS, airborne computer, cameras, communications, antennas, and balance weight to 42 kg for tests.

When the autopilot is activated, the airborne computer will seek for the nearest waypoints that have been uplinked from the ground control computer on the memory. Each waypoint will be “flying-by” following the proposed approaching control algorithm.

In the demonstration, when the target is found by the front camera, the twin camera at normal focus will see the target instantly. The twin camera set can be triggered to turn pan-tilt mirror to reflect the image into 1000 mm lens as a zoom-in mechanism to catch the target more closely. The twin camera requires some operational practice for image acquisition in corresponding to flight control. The onboard installation is shown in Figure 8. The front camera, as Figure 4, plays at the pioneer position to view the spot. The twin camera is triggered in two ways, one by the target GPS position, or the other by manual operation from the front camera.

The camera shooting can be quick enough to capture huge amount of image data, as long as the memory can accept. After the mission, the image can be successively scissored into a spot view figure. Figure 6 and Figure 7 show the successive zoom-in photo at 1 Hz shutter. In image acquisition on the target area, the UAV will fly in a predefined contour such as zig-zag or circle to cover broader area for dense photo shootings.

Figure 6 shows a landscape of a farm and house. The photos can be successively accumulated to show a real situation in very detail.

For the homeland security and disaster surveillance, images over some main bridges and their riverbed can offer important support for disaster forecast and damage assessment. Riverbed watch can be carried out constantly by routine and a double-check watch after storms to get better understanding near main bridges for disaster prevention. Figure 7 shows a zoom-in photo taken for riverbed image acquisition. The Cardinal UAV can take photos by back and forth flights above the target area.

5. CONCLUSION

The Cardinal UAV system has been developed with series of flight tests with different payloads in low altitude. The path planning and the ground computer interface are very flexible design for mission assignment and flight operation in autopilot. The navigation waypoints can be uploaded via available communication, e.g., 900 MHz or 3G promptly. The tests verify that 3G mobile communication is very reliable to communicate between the UAV and the ground computer in the plain area where 3G mobile coverage is good. Otherwise, 900 MHz or 2.4 GHz systems are viable with appropriate ground antenna tracking.

Cameras are the most important payload on the Cardinal system. In the demonstration, the twin camera set has successfully verified the image acquisition capability and effectiveness. The ground images are captured from normal lens and tele-lens by pan-tilt switched zooming within 0.5 seconds.

To extend the UAV into long haul operation, fuel and power are two fatal keys to solve. It

is not operationally reliable or weight effective to use extra large battery packs to supply extra longer power demand. How can a UAV endure long haul operation becomes another challenge to the proposed UAV system. A motorcycle DC generator is selected to mount on engine to supply steady power source, while a structure fuel tank design in the middle wing section for 13 liter extra fuel is modified. Together with high capacity fuel tank and DC power generator, a long haul mission UAV can be viable for longer missions.

The image quality is good to present earth observation from the UAV flights. Ground photos can be controlled to shoot right spots according to the uplink GPS waypoints. Function command can be implemented on the airborne computer to carry on the image acquisition process. The overall tests verify that UAV system for earth observation is quite useful and efficient to local areas surveillance.

ACKNOWLEDGMENT

This work is jointly supported from National Science Council under project NSC99-2221-E006-056 and Industrial Technology Research Institute (ITRI) for the 2009 “E2 Project”.

REFERENCES

1. Geoscience Australia, “Earth observation and satellite imagery,” Australia Government, November 2010, available from web, <http://www.ga.gov.au/remote-sensing/>.
2. Earth Remote Sensing Data Analysis Center, “Earth observation satellite,” November 2010, available from web, http://gds.aster.ersdac.org.jp/gds_www2002/seminar_e/e.o.s.e/set.
3. Eisenbeiss, H., “A mini unmanned aerial vehicle (UAV): System overview and image acquisition,” *International Workshop on Processing and Visualization using High-Resolution Imagery*, Thailand, November 2004.
4. Herwitz, S. R., et al., “Imaging from an unmanned aerial vehicle: Agricultural surveillance and decision support,” *Computers and Electronics in Agriculture*, Vol. 44, 49–61, 2004.
5. Coifman, B., M. McCord, M. Mishalani, and K. Redmill, “Surface transportation surveillance from unmanned aerial vehicles,” *Proc. of the 83rd Annual Meeting of the Transportation Research Board*, 2004.
6. Bortof, S. A., “Path planning for UAVs,” *Proceedings of the American Control Conference*, 364–368, Chicago, June 2000.
7. Beard, R. W., T. W. McLain, M. Goodrich, and E. P. Anderson, “Coordinated target assignment and intercept for unmanned air vehicles,” *IEEE Trans. on Robotics and Automation*, Vol. 18, 911–922, 2002.
8. Lin, C. E., C. W. Hsu, Y. S. Lee, C. C. Li, S. F. Tai, and W. J. Kang, “Verification of UAV flight control and surveillance using mobile communication,” *AIAA Journal of Aerospace Computing, Information and Communication*, Vol. 1, No. 4, 189–197, April 2004.

Undersampled Digitally Heterodyned SFGPR with Variable Sampling Frequency

D. Adirosi¹, G. Alberti², and G. Galiero²

¹Thales Alenia Space Italia, Italy

²Consortium for Research on Advanced Remote Sensing Systems — CO.R.I.S.T.A., Italy

Abstract— In this paper, an innovative architecture of an undersampled digitally heterodyned stepped frequency GPR is presented. Typically SFGPR perform a quadrature downconversion in the analog domain to detect targets; the main disadvantages of this approach can be eliminated by performing it in the digital domain. In order to perform such digitally downconversion, the echoes of the step frequencies transmitted are subsampled with a sampling frequency variable on a step by step basis according to rules described hereafter. This approach allows to greatly simplify both the transmitting and receiving chains of the SFGPR by eliminating a frequency conversion stage.

The key aspects of architecture presented and its advantages are discussed.

1. INTRODUCTION

It is well known that the capability of electromagnetic waves to propagate beyond the physical discontinuities of propagation media makes it possible to exploit them to investigate internal features of dielectric bodies. From this property, an endless number of practical applications have been arisen, ranging from medical prospecting to detection of mines, nondestructive testing of industrial items and GPR applications.

In this paper, a new architecture of an Undersampled Digitally Heterodyned SFGPR with variable sampling frequency is presented. This design starts by the technical knowledge and on field experience achieved during the design of a SFGPR developed in the framework of the ARCHEO project; it was funded by the Italian Ministry of the scientific research and industry whose main aim was the development of tools to aid the archeologists in their on field researches [1].

The key aspects and the advantages of the architecture herein presented are discussed.

2. PROPOSED ARCHITECTURE

The architectures of SFGPRs presented in literature can be grouped into two main categories: homodyne and heterodyne [2].

An homodyne SFGPR detects the targets by mixing the received echoes with a copy of the transmitted signal. In order to extract both the amplitude and phase informations from the received signal, the mixing is performed by using a quadrature mixer and both the in- (I) and quadrature (Q) phase outputs are low pass filtered to get the final I and Q values. The well known disadvantages of an homodyne approach are the following: quadrature demodulator errors due to phase and amplitude gain imbalances, flicker noise and temperature drift of the DC values at In-phase and Quadrature outputs.

An heterodyne SFGPR detects the targets by performing a quadrature downconversion of a constant intermediate frequency (IF) obtained from the received signal by means of a downconversion stage. Because this IF is greater than 0 Hz, this architecture alleviates the problems related to flicker noise and drift of DC values with the temperature. Furthermore the problems related to the quadrature demodulation can be solved if it is performed in the digital domain instead of the analog one by direct acquisition of the IF signal.

The SFGPR architecture presented in this paper (shown in Figure 1) has not only the mentioned advantages of the heterodyne GPRs but also peculiar ones related to the choices made and described hereafter. A brief description of the main blocks is given below:

- **Transmitting Chain** (Tx Chain): the signal to be transmitted is generated by means of a DAC, its output is low pass filtered and amplified prior to be sent to the transmitting antenna. The clock of the DAC is provided by a frequency generation unit;
- **Receiving Chain** (Rx Chain): the signal received by the antenna is amplified by an LNA and provided to a large band ADC; the signal is undersampled with sampling frequencies provided

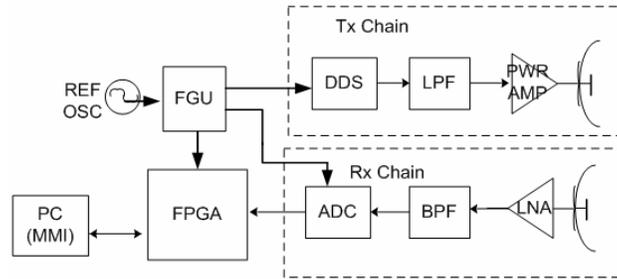


Figure 1: SFGPR architecture presented in this paper.

by the FGU. The samples acquired are transferred to an FPGA whose aim is to perform the quadrature downconversion in the digital domain; the local oscillator of this downconverter is chosen according to the frequency of the transmitted signal and to the planned sampling frequency; its phase takes into account the possibility of spectral inversion of undersampled signals also;

- **Frequency Generation Unit (FGU):** aim of this unit is to generate all the frequencies employed in the SFGPR from a very low phase noise master oscillator: DAC and ADC clocks, FPGA reference clock.

The main advantages of the presented architecture are:

- Absence of a synchronism chain, generally used in SFGPR to get a phase reference of the transmitted signal [1];
- Simplified RF front end: both Tx and Rx chains are substantially constituted by an amplifier and a filter; this simplification is allowed by the undersampling of the received echoes;
- Simplified Frequency Generation Unit ;
- Substantial reduction of the power consumption and weight due to the great simplification brought in the RF front end;
- Flexible architecture with respect of the possible frequency bands to use. In fact the constraints on which bands to use are mainly due to DAC and ADC 3-dB bandwidths; nowadays are present on the market DACs capable to generate frequencies even in the second and third Nyquist zone as well as ADC with input bandwidth as large as 3 GHz. This allows the employment of the same SFGPR in large band of investigation frequencies and with very large synthesized bandwidth. This peculiarity of the architecture presented will be even more enhanced in the near future with the advent of integrated circuits with increased input/output bandwidths as the trend of the last years has indicated.

3. DIGITAL RECEIVER

The achievement of the mentioned goals has been possible by moving SFGPR complexities from the analog domain into the digital one.

The digital section implemented in the FPGA includes a synchronism detector circuit also. In fact because the sampling frequencies are different between them and they are not integer multiple of the FPGA reference clock frequency, it is necessary a circuit that determine when both are phase aligned. Because the characteristic of the FGU is to generate the clocks of the ADC, DAC and FPGA reference clock in a phase locked way and because they are integer multiple of the same reference clock (a very low phase noise master oscillator), it is possible to recover the phase reference in each frequency step by means of the calibration phase usually performed in GPR.

The mentioned synchronism detector circuit has been designed by starting from phase detector circuits of digital PLLs [3] and adapted to the needs of the presented architecture. The design has been made in VHDL and simulations have been performed in order to test its behaviour by means of standard VHDL simulation tools. Many detector circuits, other the one referenced, were considered. The results of the simulations performed and the simplicity of its design drove the choice to the circuit shown in Figure 2. It is based on a circuit presented in [3] but with a D-type register instead of a JK one. The simulations performed on this circuit showed that it is better to use a reference clock (clk_ref) whose frequency is lower than the ADC clocks to be used; the clk_ref used in these simulations were 10 MHz and 100 MHz. These values were chosen in that

they represent typical values for instrumentation reference clocks and they are easily available from different manufacturers with very low phase noise jitter values. In Figure 2, the result of one of the mentioned simulations is shown.

4. SIMULATIONS AND RESULTS

As explained before, the knowledge of the sampling frequencies (f_s) to use as well as the frequencies to be generated in each step must be known in advance to correctly program both the FGU and the digital local oscillator of the quadrature downconverter. Software tools have been developed in order to perform automatically such choice once the input parameters (step frequency, first frequency to generate, number of steps, ...) are provided. In Figure 3, it is shown a snapshot of the man machine interface of the frequency planner with the input parameters used for the simulation results shown hereafter. From the figure are evident not only the values of the two sampling frequencies used but also the intermediate frequencies (f_{IF}) resulting from the undersampling process for each f_s to be used (white and red lines); this sw tool choose the suitable sampling value for each frequency step (green line in Figure 3) that comply with the following rules: try to keep the same f_s of the previous step (typically is the lower one); the resulting f_{IF} must be at least $B_{IF}/2$ away both from DC and $f_s/2$. If any of the foreseen step frequencies to be generated does not comply with the mentioned rules it is underlined by writing the value in the F_Trouble shown in Figure 3. By using a sort of brute force approach (running several times the frequency planner for different combinations of step and sampling frequencies couples) the following values were chosen for the two sampling frequencies: 40 MHz and 58 MHz.

The whole post-processing chain, from the undersampling of the step frequencies to their quadrature demodulation and processing to synthesize simulated targets has been performed by developing an ad hoc simulator in a high level language (LabView™). This simulator use as input the data generated by the frequency planner (Figure 3). The inversion algorithm used to synthesize targets is the Walton modified described in detail in [2].

In order to proof the validity of the post processing chain adopted it is sufficient to provide the delay and the relative amplitude of each target to be simulated. Figure 4 reports the results of two simulations obtained by assuming the presence of two targets at a delay of 7 ns and 12 ns respectively and having the same relative amplitude or in a relationship of 1/2. The power shown in Figure 4 are normalised with respect to the highest value.

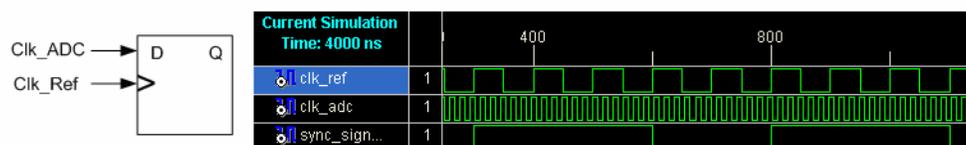


Figure 2: Synchronism Circuit adopted with one of the simulation results obtained.

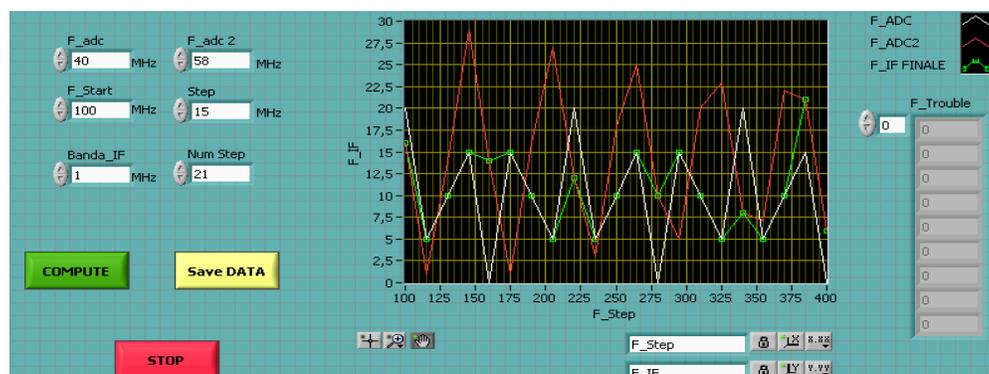


Figure 3: Snapshot of the man-machine interface of the frequency planner.

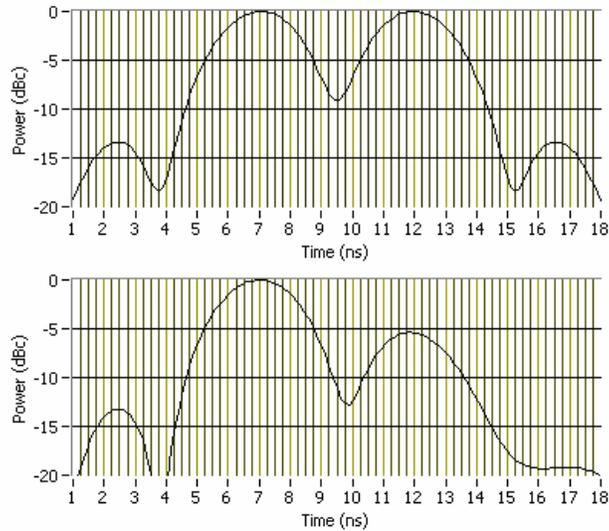


Figure 4: Simulation results for two targets: up — equal amplitude; down — second target 1/2 of the first one.

5. CONCLUSION

This paper presented a new SFGPR architecture based on the experience achieved in previous design. The design presented has the following advantages:

- Absence of a synchronism chain;
- Simplified RF front end;
- Simplified Frequency generation unit;
- Substantial reduction of the power consumption and weight.

The achievement of the mentioned goals has been possible by moving GPR complexities from the analog domain into the digital one. Future developments are: identification of an algorithm to choose optimum parameters set with a new frequency planner tool; implementation of the post-processing chain in VHDL to compare simulation results; implementation of the digital receiver within an FPGA to perform real time test.

REFERENCES

1. Alberti, G., L. Ciofaniello, G. Galiero, R. Persico, and M. Sacchetti, "An Italian experience on stepped frequency GPR," *PIERS Proceedings*, Pisa, Italy, March 2004.
2. Noon, D. A., "Stepped-frequency radar design and signal processing enhances ground penetrating radar performance," Ph.D. Thesis, Department of Electrical & Computer Engineering, University of Queensland, Australia, 1996.
3. Best, R. E., *Phase-locked Loops: Design, Simulation and Applications*, McGraw-Hill Professional, 2007.

Using of Multi-angular Radiometric Measurements for Short Wind Wave Parameters Estimate

M. N. Pospelov, A. V. Kuzmin, and I. N. Sadovsky
Space Research Institute, Russia

Abstract— The paper presents some results of the experiment CAPMOS'05 performed on an oceanographic research platform in the Black Sea. The platform located 600 m off shore was equipped with a set of contact and remote sensors. Conventional contact sensors were used for direct measurements of atmosphere and sea boundary layer parameters (wind speed and direction, air temperature, water temperature and salinity profiles, etc.) whereas microwave and IR radiometers were used for remote measurements of surface temperature and wave parameters. In particular, microwave Ka-band radiometer measurements were applied for gravity-capillary wave spectrum retrieval using the original techniques based on angular measurements. The spectrum component evolution under unstable wind conditions was investigated. It was demonstrated that the spectral components in the vicinity of the curvature maximum are the most sensitive to the wind velocity variations.

1. INTRODUCTION

Short gravity and gravity-capillary waves play a very important role in ocean-atmosphere interaction, affecting the momentum exchange through wind waves generation and dissipation. At the same time, short waves affect the electromagnetic waves emission and scattering from the sea surface, and this effect is used in satellite radiometers and scatterometers for remote measurements of winds over ocean. The relations between wind, waves and emitted/scattered signal are extremely complicated and can hardly be described unambiguously by any theoretical model. Therefore experimental measurements of the ocean-atmosphere interaction parameters under various meteorological conditions are of high importance.

In an ideal case, an experiment should combine synchronous and collocated remote sensing (radiometer, radar, etc.) and in situ measurements. The oceanographic platforms are of particular importance because they provide a unique opportunity of long-term measurements of sea and atmosphere parameters in a fixed point using various kinds of sensors, both remote sensing and contact instruments.

The paper presents the results of the experiment CAPMOS'05 performed on an offshore oceanographic platform in the Black Sea in June 2005. The experiment aimed at air-sea coupling investigations by means of direct and remote measurements was carried out in frames of the project CAPMOS (“Combined Active/Passive Microwave Measurements of Wind Waves for Global Ocean Salinity Monitoring”) sponsored by INTAS (International Association for the promotion of cooperation with scientists from the New Independent States of the former Soviet Union). The project joined several research teams from Russia, Ukraine, Denmark and Italy experienced in experimental study of ocean and atmosphere. The major goal of the experiments was to compare the results of synchronous active and passive microwave measurements of waved sea surface, focusing on the ocean wave spectrum retrieval.

The task to measure the wave parameters in the open sea is a considerable challenge, especially for short gravity-capillary waves. Traditional wave gauges are not usable for very short wave measurements because of disturbances introduced by a gauge itself. The problem may be solved by applying remote techniques, either optical or microwave radar. A number of various models of wind wave spectrum have been proposed to date by various authors; [1–3] are only some of them. It is worthy of note that most of these models agree, more or less, at the range of long waves, whereas at the range of short gravity-capillary waves the disagreement can be of an order of magnitude. This problem calls for further experimental efforts, which would allow gathering more data on the short wave parameters under a variety of environmental conditions.

The experiment CAPMOS'05 was carried out on a research oceanographic platform located about 600 m to the south of Crimea coast near Katsiveli, Ukraine. The sea depth at the site is 28 to 32 m, so the deep water and long fetch conditions were ensured for prevailing winds from the south, south-east and south-west. The platform has several working levels. The lowest deck at 4 meters above the surface is used for various instruments installation. At the main deck at 14 meters

above the sea level the living rooms and laboratories are situated. The mast on the roof is used for meteorological equipment installation. More details about the experiment can be found in [4].

The platform was equipped with a set of contact and remote sensing instruments. Conventional contact sensors were used for direct measurements of atmosphere and sea boundary layer parameters (wind speed and direction, air temperature, water temperature and salinity profiles, etc.) whereas microwave and IR radiometers were used for remote sensing measurements of surface temperature and wave parameters.

The list of radiometers used during the experiment included S-band radiometer (*V*-pol.), K- and Ka-band polarimeters (3 Stokes parameters), W-band radiometer (*V*- and *H*-pol.), and thermal infrared band (IR) radiometer. All the radiometers were mounted on an automatic rotator, which permitted to change the angle of observation in both azimuth and elevation planes. The rotator with radiometers was mounted on a boom 4 meters long, to reduce the reflected radiation from the platform itself as well as a wave pattern distortion. On the same boom, a mast with meteorological instruments and a wavegauge were installed.

The measurements were carried out continuously 24 h/day. The automatic rotator was controlled by a computer, so any desirable algorithm of rotation could be determined. Typical algorithm of rotation consisted of vertical scanning from 20 to 153 degrees from nadir and backward at 0.2 rpm at six fixed azimuths from 60 to 240 degrees stepped by 36 degrees. The back swing in azimuth was performed at a fixed incidence angle about 70 degrees. Duration of the full scanning cycle was about 25 minutes. In such a way, sea-surface brightness from nearly nadir to horizon and atmosphere brightness from horizon to nearly zenith was registered at various frequency bands and polarizations. In addition, an azimuth dependence of the sea-surface brightness arising due to surface wave anisotropy can be estimated.

In the following sections the results of applying a novel approach for the short gravity-capillary wave spectrum retrieval from radiometric measurements are described. The dynamics of retrieved spectral components under unstable wind is traced.

2. ALGORITHM OF WAVE SPECTRUM RETRIEVAL

An idea of radiometric measurements usage for surface wave spectrum retrieval was first proposed in [5], where it was suggested to apply multifrequency radiometry. Another approach was proposed in [6], where angular radiometric measurements were realized. Both these approaches are based on the so-called “critical phenomena”, or Etkin-Kravtsov effect, which is, in fact, a phenomenon of resonant thermal microwave radiation from a rough surface. The resonance conditions look as follows:

$$\rho^2 + 2\rho \sin \theta \cos \varphi - \cos^2 \theta = 0 \quad \rho = n \frac{\lambda}{\Lambda} \quad (1)$$

where $\lambda = 2\pi/k$ is electromagnetic wavelength, k is electromagnetic wave number, $\Lambda = 2\pi/K$ is surface wavelength, K is surface wave number, θ is incidence angle, φ is azimuth angle between surface K and electromagnetic k wave vectors, $n = \pm 1, \pm 2, \dots$ is an order of the resonance. It is evident that for different relations between surface and electromagnetic wave length the resonance occurs at different angles. Fig. 1 illustrates the resonant increasing of periodically uneven water surface emissivity (in arbitrary units at *Y*-axis). Left panel shows the emissivity dependence on the ρ parameter at various incidence angles (that corresponds to the algorithm of the spectrum retrieval proposed in [5]), whereas right panel shows the similar dependence on the incidence angle for various ρ parameters (corresponding to the algorithm proposed in [6]). The emissivity change is proportional to the surface wave amplitude squared, and respective increasing of the brightness temperature may exceed tens of Kelvins for the wave amplitude about 1 mm.

The real ocean surface may be considered as a mixture of harmonic waves with various lengths and amplitudes. The value of the microwave brightness resonant change depends on the intensity of the respective harmonic in the continuous spectrum of surface waves. This is a physical reasoning for the radiometric measurements applying for the surface wave spectrum parameters retrieval. The principal consideration should be kept in mind that we are talking about surface waves with the length of the same order as an electromagnetic wave length, i.e., we consider gravity-capillary waves.

In our study we follow the approach proposed by Trokhimovski [6]. An algorithm of the spectrum parameters retrieval from angular radiometric measurements reduces an inverse problem to the direct one (for more detailed algorithm description see [7]). First, the brightness contrast produced

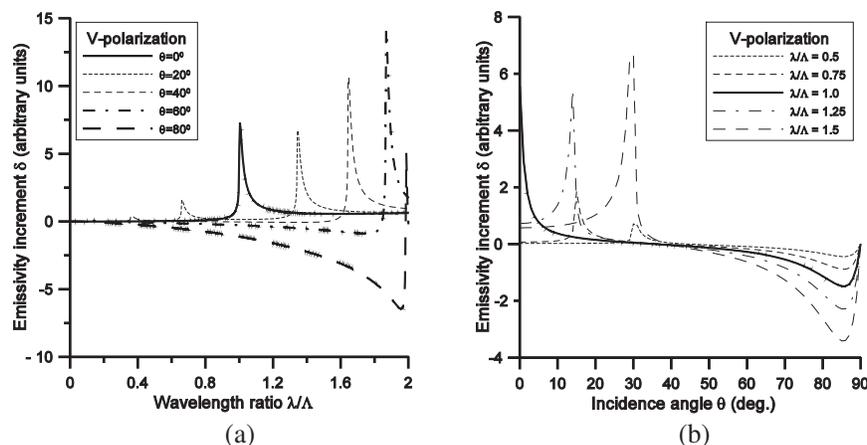


Figure 1: Rough surface emissivity increment (in arbitrary units) as a function of: (a) electromagnetic to surface wavelength ratio; (b) incidence angle.

by the waves with any particular (randomly defined) parameters of spectrum is computed, and then the combination of the spectrum parameters is chosen that ensures the best fit of the computed and the measured experimental contrasts.

To solve the direct problem, a two-scale model of the sea surface is used. The brightness contrast (i.e., the difference of the brightness temperatures of the waved and the smooth surface) produced by the long (relative to electromagnetic wavelength) waves is computed using Kirchhoff approximation, whereas short wave effect is computed using the small perturbation method. Using this approach, the following parameters of the surface wave spectrum may be retrieved: the form of the curvature spectrum $B(K)$ in the range of wave numbers $K_L < K < 2k$; the long wave slope variance σ , which is a kind of an integral parameter characterizing the waves with numbers $K < K_L$. Here K_L is a boundary wave number dividing long and short waves (in our calculations it was chosen $K_L = 0.05k$).

3. EXPERIMENTAL RESULTS

The consecution of the experimental data processing and spectrum parameters retrieval during the CAPMOS'05 was as follows. First, the experimental brightness temperatures were averaged over several successive scans, to reduce the fluctuations. Then the brightness temperature T_{Sm} of the smooth water surface of the same temperature and salinity was computed. And a brightness contrast produced by the waves was computed by subtracting T_{Sm} from the experimental data. Further, the parameters of the wave spectrum (slope variance of long waves σ and a piecewise-linear approximation for gravity-capillary waves curvature B) was randomly defined and a direct problem for these set of parameters was solved until the best fit between computed and measured contrast would be achieved.

Figure 2 shows an example of wave parameters retrieval from Ka-band polarimeter data. In the left panel the spectrum of short wave curvature is plotted. In the same figure three different models of spectrum [1–3] are plotted for a comparison. It is evident that no one of the model spectrum corresponds to the retrieved one. Reasonable agreement is achieved only with the spectrum [1] computed for the wind speed twice as much as it was really observed. The probable reason for such disagreement may be that the model spectrum is a result of various empiric data synthesizing, which generally speaking are related to the ideal conditions of a fully developed sea that was evidently not the case in our experiment. Furthermore, different models differ each from other, and nowadays no one may be given an absolute credence to.

In the right panel of the same figure a retrieved slope variance of long waves is plotted as a function of a boundary wave number K_L , which separates short and long waves in a two-scale model used for calculations. Again, a disagreement with known models is evident. As an indirect evidence of the retrieval quality, the wavegauge data are plotted in the same figure (asterisk symbols). Though these data of the contact measurements are related to longer wavelengths, the results of retrieval fit them rather well after extrapolation.

Figure 3 shows dynamics of gravity-capillary wave spectrum under variable wind conditions. Left panel shows an example of the retrieved curvature spectrum of gravity-capillary waves whereas right panel shows evolution of wind speed and synchronous evolution of several retrieved spectrum

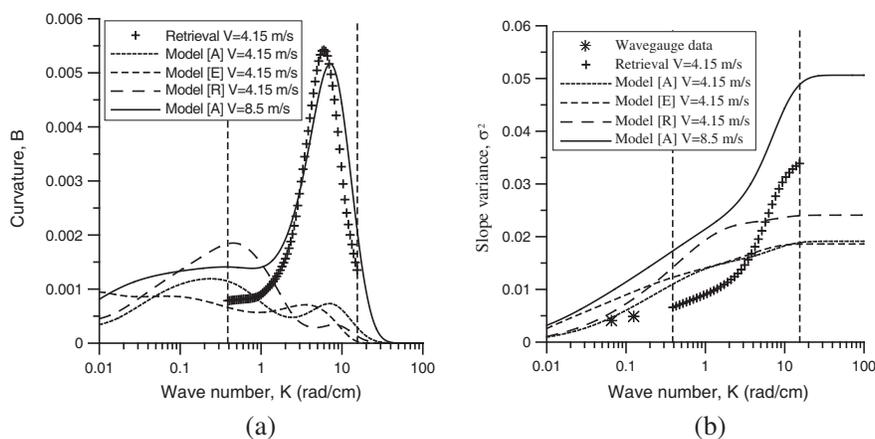


Figure 2: (a) Curvature spectrum of gravity-capillary waves retrieved from Ka-band radiometric measurements (crosses); lines — various models of spectrum $A = [1]$, $E = [2]$, $R = [3]$; the vertical lines indicate the range of wave numbers $0.05k < K < 2k$ where the curvature spectrum was retrieved. (b) Slope variance of long waves as a function of a boundary wave number K_L ; asterisks — wavegauge data.

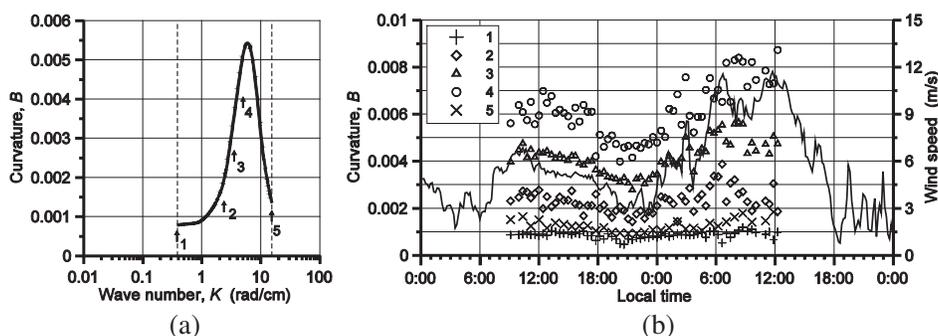


Figure 3: (a) Curvature spectrum at mean wind speed 4 m/s; the numbers 1 to 5 indicate the position of various spectral components. (b) Evolution of the spectral components with the wind speed; solid line — wind speed, symbols 1 to 5 correspond to the spectral components in the left plot.

components. It is evident that the spectral components near the curvature maximum are the most sensitive to the wind speed variations.

4. CONCLUSIONS

A novel approach for the gravity-capillary wave spectrum estimate from the angular radiometric measurements permits to trace the spectrum parameter evolution under unstable winds. The curvature spectrum as well as long wave slope variance followed the variations of surface wind. The spectral components in the vicinity of the wave curvature maximum (about 6 rad/cm) reveal maximal sensitivity to wind velocity variations.

ACKNOWLEDGMENT

The study was supported by INTAS project 03-51-4789 and RFBR project 09-02-00780-a.

REFERENCES

1. Apel, J. R., "An improved model of the ocean surface wave vector spectrum and its effects on radar backscatter," *J. Geophys. Res.*, Vol. 99, 16269–16291, 1994.
2. Elfouhaily, T., B. Chapron, K. Katsaros, and D. Vandemark, "A unified directional spectrum for long and short wind-driven waves," *J. Geophys. Res.*, Vol. 102, 15781–15796, 1997.
3. Romeiser, R., W. Alpers, and V. Wismann, "An improved composite surface model for the radar backscattering cross section of the ocean surface 1. Theory of the model and optimization/validation by scatterometer data," *J. Geophys. Res.*, Vol. 102, 25237–25250, 1997.

4. Komarova, N. Y., F. De Biasio, A. S. Kuznetsov, M. N. Pospelov, and S. Zecchetto, “Combined direct and remote sensing measurements of wave parameters at the off-shore research platform in the Black Sea,” *This Issue*.
5. Irisov, V. G., Y. G. Trokhimovski, and V. S. Etkin, “Radiothermal spectroscopy of the ocean surface,” *Sov. Phys. Docl.*, Vol. 32, No. 11, 914–915, 1987; *Doklady Akad. Nauk SSSR*, Vol. 297, 587–589, 1987.
6. Trokhimovski, Y. G., “Gravity-capillary wave curvature spectrum and mean-square slope retrieved from microwave radiometric measurements (Coastal Ocean Probing Experiment),” *J. Atmosph. and Oceanic Techn.*, Vol. 17, No. 9, 1259–1270, 2000.
7. Kuzmin, A. V. and M. N. Pospelov, “Retrieval of gravity-capillary spectrum parameters by means of microwave radiometric techniques,” *IEEE Trans. Geosci. Remote Sensing*, Vol. 43, No. 5, 983–989, 2005.

Development of Circularly Polarized Synthetic Aperture Radar (CP-SAR) Onboard Small Satellite

Josaphat Tetuko Sri Sumantyo

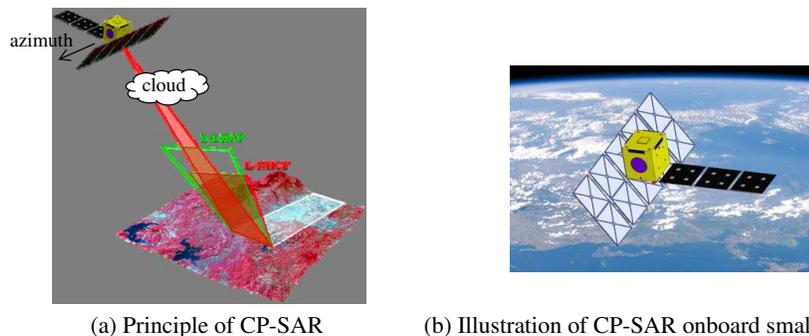
Center for Environmental Remote Sensing, Chiba University
1-33, Yayoi-cho, Inage-ku, Chiba-shi 263-8522, Japan

Abstract— Synthetic Aperture Radar (SAR) is a multipurpose sensor that can be operated in all-weather and day-night time. Conventionally, the SAR sensor has been operated in linear polarization with limited retrieved information. In this research, we are developing Circularly Polarized Synthetic Aperture Radar (CP-SAR) onboard small satellite to retrieve the physical information of Earth surface. Comparing to linear polarized systems, the CP-SAR sensor has the advantage of compactness and low power requirement, since the transmission of CP microwave is not affected by the Faraday rotation effect in the ionosphere. In this paper, the CP-SAR onboard UAV is also introduced.

1. INTRODUCTION

Synthetic Aperture Radar (SAR) is well-known as a multi-purpose sensor that can be operated in all-weather and day-night time. Recently, many missions of SAR sensors are operated in linear polarization (HH , VV and its combination) with limited retrieved information. The characteristic of the conventional SAR sensor, especially for satellite onboard sensor is bulky, high power, sensitive to Faraday rotation effect etc. We proposed the Circularly Polarized Synthetic Aperture Radar (CP-SAR) onboard small satellite that to be launched in 2014 to retrieve the physical information of Earth surface [1]. Fig. 1 shows the illustration of CP-SAR onboard small satellite called μ SAT CP-SAR that is being developing in Josaphat Microwave Remote Sensing laboratory (JMRS), Center for Environmental Remote Sensing, Chiba University, Japan. Table 1 shows the specification of the μ SAT CP-SAR, and Fig. 2 shows each sensor installed on the pre-deploying condition of μ SAT CP-SAR. Fig. 3 shows the μ SAT CP-SAR system that is composed by attitude control system (ACS), CDS (command and data handling system), EPS (electrical Power Subsystem), and CMS (communication subsystem). The detail components as listed in Fig. 3, CDS composed by on-board computer (OBC), telemetry and command unit (TCU) and mission data storage unit (MDU). ACS is composed by electromagnetic torque (EMT), GPS receiver (GPSR), sun sensor (SS) and magnetometer (MAG). EPS is composed by battery charge regulator (BCR), power control unit (PCU) and power distribution unit (PDU). Finally, CMS is composed by S-band transmitter (STX), S-band receiver (SRX) and X-band transmitter (XTX).

In this research, the CP-SAR sensor is employing the elliptical wave propagation and scattering phenomenon by radiating and receiving the elliptically polarized wave, including the special polarization of it as circular and linear polarizations. In this mission we employ three microwave sensors as main sensors, there are CP-SAR, GPS-SAR and GPS-RO, as shown in Fig. 4. CP-SAR is as active sensor that could transmit and receive the L band chirp pulses for land deformation monitoring. GPS-SAR is an experimental passive SAR sensor. This mission plans to investigate the possibility to receive the GPS pulse and process it to retrieve the SAR image. GPS Radio



(a) Principle of CP-SAR

(b) Illustration of CP-SAR onboard small satellite

Figure 1: CP-SAR onboard small satellite.

Table 1: Specification of CP-SAR onboard small satellite.

| Parameter | |
|------------------------|---|
| Launched date | 2014 |
| Altitude | 500 ~ 700 km |
| Inclination angle | 97.6 degrees |
| Frequency / wavelength | 1.27 GHz (L Band) / 24 cm |
| Polarization | TX: RHCP+LHCP RX: RHCP+LHCP |
| Gain / Axial ratio | > 30 dBic / < 3 dB (main beam) |
| Off-nadir angle | 29 degrees (center) |
| Swap width | 50 km |
| Spatial resolution | 30 m |
| Peak power | 90 ~ 300 W (PRF 2000 ~ 2500 Hz, Duty 6%: average 5.6 W) |
| Bandwidth | Chirp pulse: 10 MHz |
| Platform size | 1 m × 1 m × 1 m |
| Weight | < 100 kg |
| Antenna size | Elevation 2.0 m × Azimuth 5.0 m |
| Sensors | Main sensor: CP-SAR Sub sensors: GPS SAR, GPS Radio Occultation (RO), cameras |

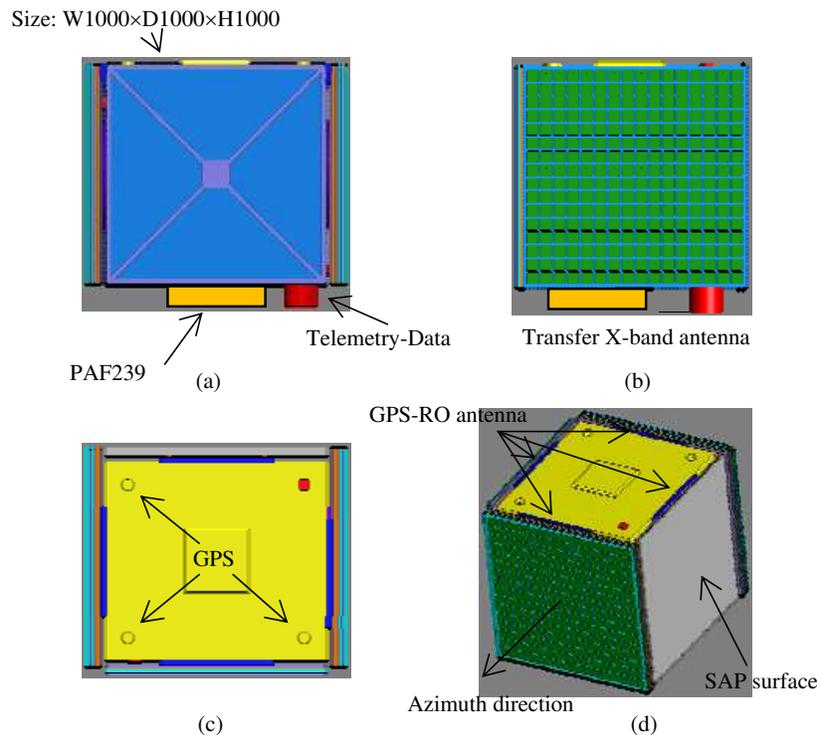
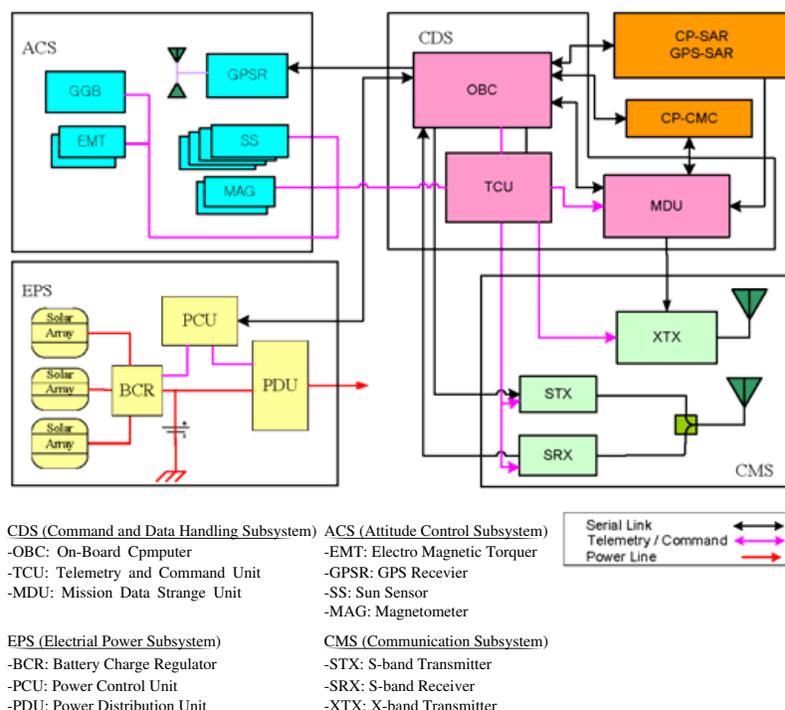
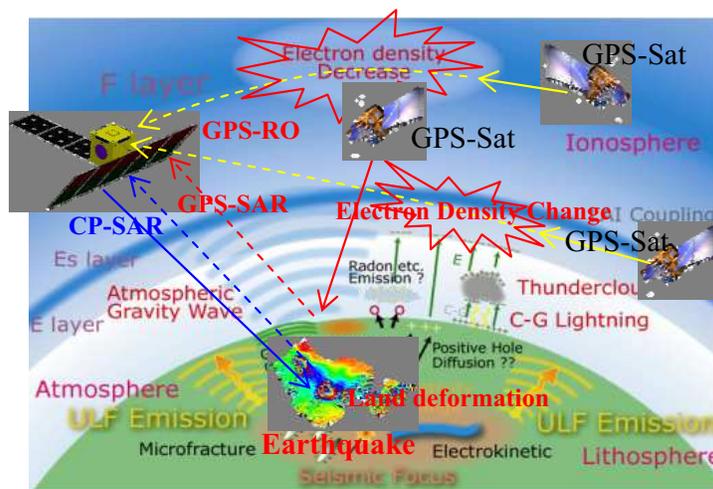


Figure 2: Pre-deploying CP-SAR onboard small satellite: (a) range direction view, (b) azimuth direction view, (c) top view, and (d) bird eye view.

occultation (GPS-RO) is an experimental four unit of patch array antenna sensor to receive the GPS signal and process it to retrieve the conditions of electron in the ionosphere to investigate the coupling of electron density change and land deformation. In the future, this coupling is used to predict the earthquake activity and build the early warning system in Asian countries. The sensor

Figure 3: Block diagram of μ SAT CP-SAR system.Figure 4: μ SAT CP-SAR mission.

is designed as a low cost, light, low power or safe energy, low profile configuration to transmit and receive left-handed circular polarization (LHCP) and right-handed circular polarization (RHCP), where the transmission and reception are both working in RHCP+LHCP as shown in Fig. 1(a). Then these circularly polarized waves are employed to generate the axial ratio image (ARI). This sensor is considered not depending to the platform posture, and it is available to avoid the effect of Faraday rotation during the propagation in Ionosphere. Therefore, the high precision and low noise image is expected to be obtained by the CP-SAR. For this purpose, we are also developing the CP-SAR onboard unmanned aerial vehicle (CP-SAR UAV) for ground testing of this sensor. In this paper, we also introduce the specification and basic system of our CP-SAR UAV development.

2. μ SAT CP-SAR MISSION

The main mission of this μ SAT CP-SAR is to hold (1) the basic research on elliptically or circularly polarized scattering and its imaging technique, and (2) its application development.

In the basic research, we investigate the elliptical (including circular and linear polarizations) scattering wave from the Earth surface, circularly polarized interferometric technique (CP-InSAR),

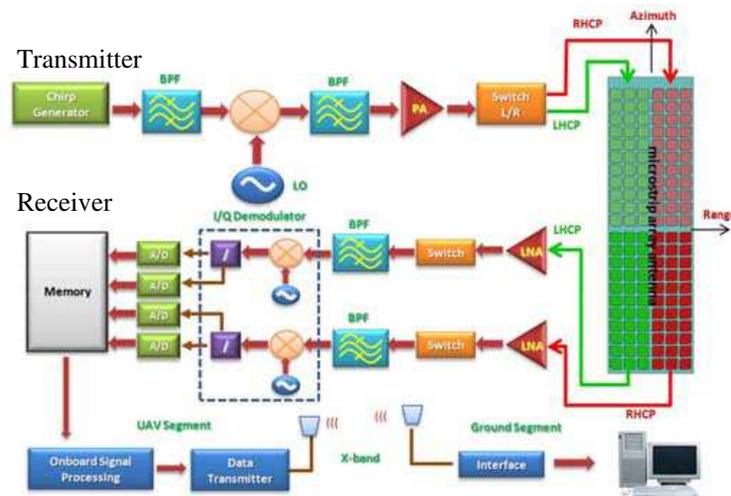


Figure 5: Block diagram of Circularly Polarized Synthetic Aperture Radar (CP-SAR) System.

axial ratio image (ARI) generation etc. We hold the analysis and experiment on circularly polarized wave scattering on vegetation, snow, ice, soil, rock, sand, grass etc to investigate the elliptical scattering wave.

In experiment of CP-InSAR, we will hold some experiments to compare the InSAR technique by using CP and Linear waves. This technique will be implemented to extract the tree trunk height, DEM etc by using the elliptical polarization wave. The axial ratio image (ARI) will be extracted by using the received RHCP and LHCP wave, then this image used to investigate the relationship between the characteristics of ARI and vegetation, soils, snow, ice etc. The image of tilted angle as the response of Earth surface also will be extracted to mapping the physical information of the surface, i.e., geological matters, contour, tree trunk structure and its characteristics, snow-ice classification etc.

In application development, CP-SAR sensor will be implemented for land cover mapping, disaster monitoring, Cryosphere monitoring, oceanographic monitoring etc. Especially, land cover mapping will classify the forest and non-forest area, estimation of tree trunk height, mangrove area monitoring, Arctic and Antarctic environment monitoring etc. In disaster monitoring, CP-SAR sensor will be employed for experiment of CP Differential InSAR in earthquake area, monitoring of volcano activity, forest fire and flood monitoring etc. In snow and ice monitoring, this sensor will be employed for monitoring of ice berg, glacier, snow and ice characteristic investigation. The monitoring of oil spill, inner wave etc will be done for Oceanographic monitoring. Especially in the disaster monitoring, this mission has main target as shown in Fig. 4.

3. CIRCULARLY POLARIZED SYNTHETIC APERTURE RADAR (CP-SAR) SYSTEM

Figure 5 shows the circularly polarized synthetic aperture radar (CP-SAR) system block. Basically, this system is composed by transmitter, receiver, onboard signal processing, data transmitter and ground segment. The transmitter is composed by chirp generator, bandpass filter (BPF), local oscillator (LO), power amplifier (PA), switch for RHCP and LHCP, and two panel of CP microstrip array antenna (LHCP and RHCP panels). Then the generated chirp pulses are transmitted by using these antennas.

The receiver is composed by two panels of CP microstrip array antenna (LHCP and RHCP), low noise amplifier (LNA), switch to reduce the coupling between antenna of transmitter and receiver, bandpass filter (BPF), I/Q demodulator to generate the in-phase and quadrature data (phase), analog/digital (A/D) converter, temporary memory for data processing, onboard signal processing (OSP) unit, data transmitter system. In the OSP unit, the SAR signal processing as range compression, corner turn, azimuth compression etc is processed to produce the SAR image. Then this image will be sent to the ground segment by using X band data transmitter. This data is received by ground segment for further process.

As shown in Fig. 6, the CP-SAR image is generated by using the received LHCP and RHCP pulses. The images of axial ratio (R_w), cylindrical ratio (ε_w), and tilted angle (τ_w) could be derived

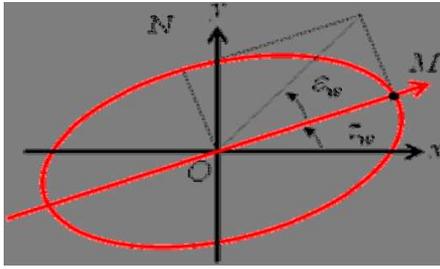


Figure 6: Geometrical illustration of circularly polarized wave.



Figure 7: Josaphat experimental CP-SAR onboard Unmanned Aerial Vehicle (JX-1).

by using Equations (1)–(3).

$$R_w = \frac{E_{RO} + E_{LO}}{E_{RO} - E_{LO}} \quad (1)$$

$$\varepsilon_w = \cot^{-1}(-R_w), \quad -45^\circ \leq \varepsilon \leq 45^\circ \quad (2)$$

$$\tau_w = \frac{\delta_d}{2}, \quad 0 \leq \tau_w \leq 180^\circ \quad (3)$$

where E_{RO} and E_{LO} are intensity of each pixel in right handed and left handed CP-SAR image.

4. CP-SAR ONBOARD UNMANNED AERIAL VEHICLE (CP-SAR UAV)

In this research, the CP-SAR onboard unmanned aerial vehicle (CP-SAR UAV) as shown in Fig. 7 is developed for C-SAR ground testing before it is installed on the small satellite. The platform called Josaphat Laboratory Experimental UAV (JX-1) has 25 kg of payload availability for various microwave sensors (CP-SAR, GPS SAR, and GPS RO) and optic sensors (visible cameras). The operation altitude is 1,000 m to 4,000 m for various CP-SARs ground experiments. The three linear polarization P-, L- and X-bands SAR sensors are also installed on the tail wing of UAV as shown in Fig. 8(a).

The concept of CP-SAR UAV is shown in Fig. 9, then the size and sensor position are shown in Fig. 8. The specification of CP-SAR sensor for UAV (Table 2) is frequency 1.27 GHz, ground resolution 1 m, pulse length 3.9 to 23.87 ms, pulse bandwidth 61.14 to 244.69 MHz, off nadir angle 40° to 60° , swath width 1 km, antenna size 1.5 m \times 0.4 m for LHCP and RHCP, azimuth beamwidth 6.77° , range beamwidth 29.78° , antenna radiation efficiency $>80\%$, PRF 1000 Hz, and peak power 8.65 W (1 km) to 94.38 W (4 km). The CP-SAR has receiver antenna composed by LHCP and RHCP antenna. The data retrieved by LHCP and RHCP antenna is employed to generate the axial ratio image. This image is used to retrieve the physical information of Earth surface, i.e. soil moisture, biomass, Cryosphere, agriculture, ocean dynamics, land deformation, disaster monitoring, digital elevation model etc. In this UAV, we also install the linearly or horizontally polarized SAR (LP-SAR) in frequency P-, L-, and X-Bands. The Linearly polarized SAR data will be compared with CP-SAR data, and employ for some applications.

4.1. Transmitter Subsystem

The CP-SAR installed in UAV system is almost same with the system for small satellite, where a transmitter is a subsystem which sends microwave signal to the target or object. The main functions of the transmitter are to generate the desired waveform, amplify it and radiated into space by the antenna. The transmitter subsystem consists of a chirp generator, filter (BPF), up-converter, a power amplifier (PA), a local oscillator (LO), and a switch.

The chirp generator is purposed to generate chirp signal form 0 until 250 MHz. This chirp signal then will be filtered by using bandpass filter to reduce the noise bandwidth and eliminate harmonic introduced by the mixer. The modulator is the device which modulates the output signal from chirp generator onto the carrier frequency. In this system, the up-converter is used to shift the RF frequency from 1145 MHz until 1395 MHz (1.27 GHz as the center frequency) by using Local Oscillator. The output power of transmitter then amplified at 25 W ($\cong 54.77$ dBm). In the last part of this subsystem are placed two circularly polarized antennas which have right and left handed circular polarization. The antenna which will transmit the radar signal can be selected by using a swich.

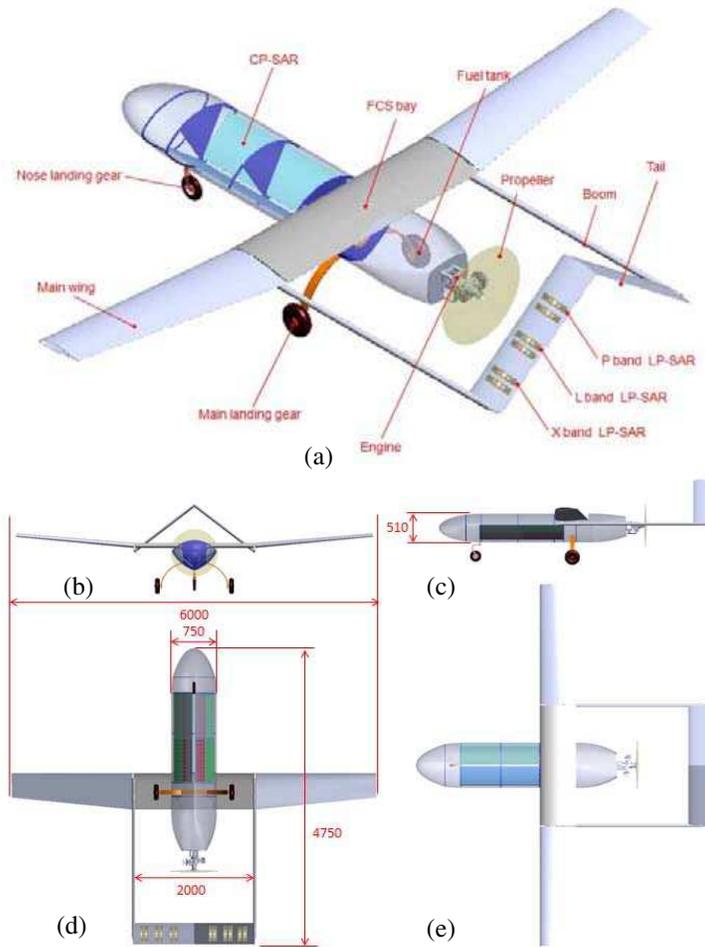


Figure 8: Size and sensor position of CP-SAR UAV: (a) bird eye view and components in CP-SAR UAV system, (b) front view, (c) side view, (d) bottom view, and (e) top view. Size in millimeter.

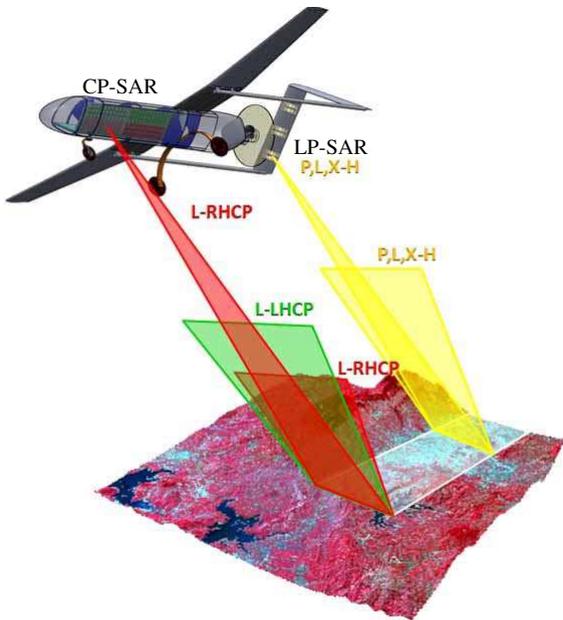
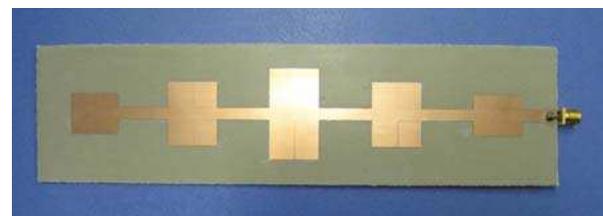


Figure 9: The concept of our CP-SAR onboard UAV.



(a) linearly polarized antenna



(b) circularly polarized antennas

Figure 10: Circularly polarized antennas developed in our laboratory.

4.2. Receiver subsystem

The receiver is a subsystem which receives echoes signals of the radar transmission via an antenna. The function of a radar receiver is to amplify the received signal and filter them in a manner that will provide the maximum difference between desired echoes (ground) and undesired interference.

On the receiver subsystem the hardware are composed by two similar part to process signal form LHCP and RHCP antennas simultaneously. The receiver subsystem consists of a low noise amplifier (LNA), a band pass filter (BPF), two I/Q demodulators, 4 channels analog to digital converter (ADC) and data recorder (memory).

The signals emitted by transmitter will be received by using both LHCP and RHCP antennas in receiver subsystem. The very low power signal captured by antenna is amplified by using low noise amplifier. The bandpass filter is re-implemented to assure only the wanted radar signal is processed. The down-converter mixer (I/Q demodulator) is used to split filtered signal into I and Q component. The out puts of I/Q demodulator are ready to be sampled and digitized by analog to digital converter (ADC) unit.

The clock unit controls and manages the timing for many components in the sensor such as the chirp generator, the signal processor component and also the frequency generator. This timing

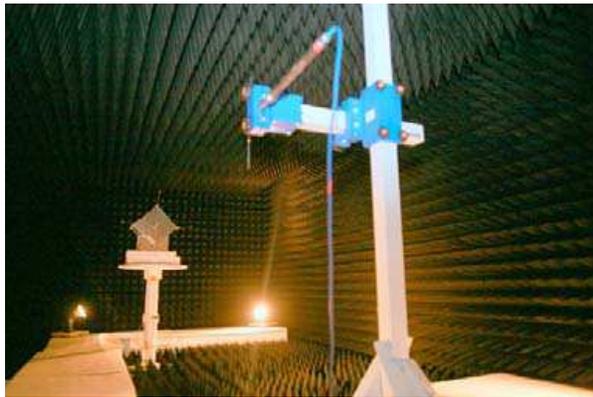


Figure 11: Ground test of small satellite in our laboratory.

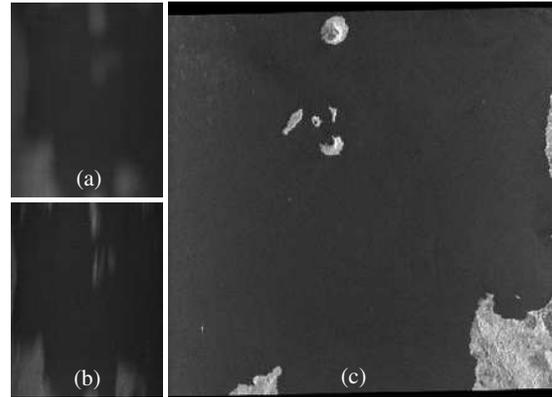


Figure 12: Output image of CP-SAR image signal processing C Simulated using JERS-1 SAR image (a) raw data, (b) range compressed image, and (c) azimuth compressed image: Krakatau volcano island, Indonesia).

Table 2: Parameters of CP-SAR onboard UAV.

| Parameters | Value |
|--------------------|------------------------------------|
| Altitude | 1–4 km |
| Frequency | 1.27 GHz (L-Band) |
| Polarization | Tx: RHCP + LHCP Rx: RHCP + LHCP |
| Image Size | 50 km ² |
| Pulse Length | 3.9 up to 23.87 μ s |
| Pulse Bandwidth | 61.14 up to 244.69 MHz |
| Off Nadir | 40° up to 60° |
| Resolution | \approx 1 m |
| Swadth width | 1 km |
| PRF | 1000 Hz |
| Peak Power | 8.65 W up to 94.38 W |
| SNR | 15 dB |
| Data Take Duration | \approx 31.70 minutes |

function is correlated to transmitting and receiving timing of the chirp pulse.

Figure 10 shows some circularly polarized microstrip or patch antenna that developed in our laboratory. Fig. 11 shows the ground test of small satellite system in our laboratory. Fig. 12 shows the output of CP-SAR image signal processing, even the test image employed the linear polarized image, JERS-1 SAR image.

5. CONCLUSION AND FUTURE RESEARCH

In this paper, we proposed the circularly polarized Synthetic aperture radar (CP-SAR) onboard small satellite. The CP-SAR is designed as the small, light in weight and low power consumption system. The CP-SAR sensor is developed to radiate and receives circularly polarized wave. This CP-SAR sensor will be applicable for land cover, disaster monitoring, snow cover, and oceanography mapping, etc. We also proposed the CP-SAR onboard UAV for ground experiment at Shikabe airport (Japan) and Sulaiman airport (Indonesia).

ACKNOWLEDGMENT

The authors would like to thank the Japan Society for the Promotion of Science (JSPS) for Grant-in-Aid for Scientific Research 2007-Young Scientist (A) (No. 19686025); National Institute of Information and Communication Technology (NICT) for International Research Collaboration Research Grant; Chiba University COE Start-up Programme “Small Satellite Institute for Earth Diagnosis”; and The Japan Society for The Promotion of Science (JSPS) Japan-East Network of Exchange for Students and Youths (JENESYS) Programme.

REFERENCES

1. Sri Sumantyo, J. T., et al., “Development of circularly polarized synthetic aperture radar onboard microsatellite,” *PIERS Proceedings*, 382–385, Beijing, China, March 23–27, 2009.

System Aspects of Mutual Coupling in Reconfigurable Active Phased Array

S. Celentano¹, L. Infante¹, S. Sabatini¹, M. R. Toma¹, and T. Johansson²

¹SELEX Sistemi Integrati, Italy

²SAAB Electronic Defence Systems, Sweden

Abstract— In a MAESA (Multifunction/multirole Active Electronically Steered Array) system performing simultaneously different TX and RX functions with different sub-apertures there are unwanted effects due to the interferences of a transmitting sub-aperture on the receiving one. This type of interference is called “mutual coupling”. The sub-apertures dedicated to different activities (Radar, ECM, ESM, and COM) can interfere among one and another by means of i) direct coupling due to direct free-space radiation and to substrate-conducted waves; ii) reflection from platform structure; iii) clutter reflections due to the close returns from environment. An analysis of the direct coupling effect has been carried out to determine if two given system functions can be executed with an acceptable amount of mutual interference or not. The scope of this paper is to present the mutual interference model and evaluation method for the interference levels between TX and RX functions serving as a knowledge base for future design of MAESA applications.

1. INTRODUCTION

Electronic Warfare (EW) systems, radar systems and microwave communication (COM) systems, typically share the same operating frequency band. In most of the current applications, these functions are implemented by separate systems, each of them using their own hardware and software, each of them managed by their own “Control Manager”. The reciprocal electromagnetic interference among functions presents serious limitations to the integration of such systems on a platform.

A multifunction/multirole system based on an adaptive/reconfigurable antenna array, such as a MAESA system, is conceived to integrate different RF functions of Radar, EW and COM systems, sharing the same physical hardware and software resources as much as possible. In a system like this, it would be an advantage if it is possible to perform simultaneous transmission and reception for different functions. The scope of the analysis is to determine if two given functions can be executed with an acceptable amount of mutual interference or not. The result of the analysis is a crucial result within a MAESA concept, since the implementation of multi-functionality requires a good understanding and description of the interference between functions.

The interference between the simultaneous transmission and reception of different functions which occurs in a MAESA system is carried out in this paper. The analysis of the mutual interference is performed through the following steps: (1) evaluate if saturation occurs in the different stages of the receiver from TRM to A/D converter; (2) evaluate the interfering signal at the output of the receiver when no saturation occurs; (3) analyse the consequences of the interference signal in terms of performances evaluation for the selected system functions. The paper is organised as follows. In Section 2, we describe the mutual coupling model between sub-arrays used by transmitting and receiving functions operating simultaneously on MAESA system. In Section 3, we devise the method to evaluate the interference power to (i) verify if saturation occurs at RX channel compromising the possibility to perform the receiving function employing it; (ii) evaluate the loss of performances on the receiving function due to simultaneous execution of transmitting functions. Section 4 deals with the analyses of the mutual coupling effects on the radar and ESM performances. Finally conclusions are given in Section 5.

2. MUTUAL COUPLING MODEL

A full active phased array aperture with sub-array architecture is assumed for our analysis. The aperture is built with wide-band transmit/receive modules (TRMs) performing transmitting functions in the band 4.5–18 GHz and receiving functions in the band 2–18 GHz. The TRMs are grouped in sub-arrays in order to move the complexity of the TX/RX chain (waveform generator, wide band receiver (WBR), high dynamic range receiver (HDR), digital beam-forming network (DBFN)) from the single TRM to sub-arrays level. The sub-arrays can be allocated in dynamic way to perform

different activities (Radar, EW, and COM), so that each function (Radar, ESM, ECM, COM) can be executed using different parts of the same aperture (sub-apertures). The number of sub-arrays grouped together into a sub-aperture is optimized depending on the activity to be performed and it can change over the scenario evolution time according to the function performance requirements. The minimum sub-aperture is given by the sub-array dimensions. In this way, according to the MAESA concept, the whole aperture is used by the different functions in antenna sharing (as well as in time sharing) and this is possible due to the modular architecture of the MAESA antenna (at sub-arrays level) which makes it reconfigurable in real time adapting to the scenario evolution. An other view of the MAESA concept is that different apertures are used for the simultaneous execution of different system functions.

The electromagnetic mutual coupling (isolation) among sub-arrays of the same aperture has been derived from the following analytical model. There are a number of reports available to provide information on mutual coupling coefficients ([1–3]). Let suppose that we want to compute the interfering power at the input of the TRM tagged by indexes (p, q) in a receiving sub-array (SR) when all elements (m, n) of sub-array (ST) are transmitting as depicted in Fig. 1. To this end, we can use the following expression:

$$C_{ST,(pq)\in SR} = \frac{\sum_{(mn)\in ST} a_{nm} \cdot C_{mn,pq}^{f_1} \cdot e^{-j(\Psi_{mn}^x(f_1, \theta_{ST}, \phi_{ST}) + \Psi_{mn}^y(f_1, \theta_{ST}, \phi_{ST}))}}{\sum_{(mn)\in ST} a_{nm}} \quad (1)$$

Assuming the overall array dimension large when compared to the operative wavelength we can assume that “all” the array elements see an external homogenous environment. This assumption is valid for almost all the array elements with the exception of some edge elements located near the array truncation. Under these hypotheses we can state that all the active elements of the transmitting sub-array perform an inter-element mutual coupling behaviour similar to the center element. To this end by using the invariance of mutual coupling coefficients with respect to mutual translation of the position of the elements [1] it is possible to sum all the inter-element mutual coupling coefficients $C_{mn,pq}^{f_1}$ related to the distance $(p-m, q-n)$ of the particular element (m, n) of the sub-array ST. To complete the model we consider also the eventual amplitude taper a_{mn} and the phase steering $\Psi_{mn}^x(f_1, \theta, \phi)$ and $\Psi_{mn}^y(f_1, \theta, \phi)$ of the transmitting sub-array. In fact, the mutual coupling between two different sub-apertures varies with relative sub-apertures scanning angles. By considering the free-space wave number $k = 2\pi/\lambda$ it is possible to express the phase term as a function of the scanning angles (θ_{ST}, ϕ_{ST}) and element positions (x_{mn}, y_{mn}) as follows:

$$\Psi_{mn}^x(f_1, \theta_{ST}, \phi_{ST}) = kx_{mn} \sin \theta_{ST} \cos \phi_{ST} \quad H\text{-plane} \quad (2a)$$

$$\Psi_{mn}^y(f_1, \theta_{ST}, \phi_{ST}) = ky_{mn} \sin \theta_{ST} \sin \phi_{ST} \quad E\text{-Plane} \quad (2b)$$

The sub-array to sub-array coupling for the receiving sub-array SR is computed by the following expression:

$$C_{STvsSR} = \sum_{(pq)\in SR} C_{ST,(pq)\in SR} b_{pq} e^{-j(\Psi_{pq}^x(f_1, \theta_{SR}, \phi_{SR}) + \Psi_{pq}^y(f_1, \theta_{SR}, \phi_{SR}))} \quad (3)$$

In (3), the coupling coefficients over the sub-array ST given by (1), are scaled by the eventual amplitude taper b_{mn} and the phase steering $\Psi_{pq}^x(f_1, \theta_{SR}, \phi_{SR})$ and $\Psi_{pq}^y(f_1, \theta_{SR}, \phi_{SR})$ of the receiving sub-array SR and then summed over each TRM of indexes (p, q) in the sub-array SR.

3. MODEL FOR INTERFERENCE EVALUATION

The information on the mutual coupling coefficients, derived in an analytical way in Section 2, is used to evaluate the effects of mutual interference inside the MAESA concept. The latter mainly based on the assumption to perform the simultaneous transmission and reception of different functions in the system.

The interference level due to TX sub-arrays that are transmitting simultaneously at some RX activities could cause the saturation of the channels used by the receiving functions. The saturation can occurs at different stages of the receiving chain from TRM to A/D converter. If saturation occurs, the receiver should be regarded as non-functional and the function using that receiver as not working. Otherwise, if the receiving channel is available, i.e., no saturation occurs, it can be

used taking into account possible performance degradation due to interference level. For discussion on saturation it is worth to refer to Fig. 2 where a sub-array front-end of an MAESA system is shown together with the receiver and digital beam forming (DBF) functions. For sake of generality, it is assumed that the sub-array can feed in parallel two separate receivers: High Dynamic Range (HDR) and Wide Band (WBR). All the sub-arrays that contribute to HD receivers are connected in DBFN 1. All the sub-arrays that contribute to WB receivers are connected in DBFN 2.

The scope of the developed model is to evaluate, for each RX aperture/sub-aperture the interference power over each RX sub-array that is part of the RX aperture/sub-aperture itself: (i) at the input of the sub-array front-end; (ii) at the output of the RX channel, due to TX sub-arrays that are transmitting simultaneously at the RX activities. To compute each of them, three contributions are considered: (1) interference power due to TX sub-arrays inside the same aperture; (2) interference power due to TX sub-arrays inside apertures arranged side-by-side (co-planar antennas); (3) interference power due to TX sub-arrays inside antennas arranged around a perfectly conducting vertical corner (at a 90° corner antennas). The interference power is equal to the input power P_{TX} , as seen through the mutual coupling coefficients MC, where the MC coefficients are computed, for each of the above mentioned contributions, according to the equations previously described in Section 2. To compute the interference power \bar{P}_{RF} at the input of the RX sub-array front-end the following equation is used:

$$\bar{P}_{RF} = \int_0^{\infty} C^2(f) \cdot S(f) df \tag{4}$$

where $C(f)$ = mutual coupling (MC) coefficients and $S(f)$ = power spectrum of the interfering signal at the output of its transmitting aperture. The integral for computation of the interfering power \bar{P}_{RF} can be simplified as:

$$\bar{P}_{RF} = P_{Peak} \cdot C^2(f_{TX}) \tag{5}$$

where P_{Peak} is the peak power for each of the transmitting functions and the computation of the MC coefficients is considered at $f = f_{TX}$. Over each RX sub-array, the model computes separately, the interference power due to each TX function, considering the proper MC coefficient, summing the contribution of all the TX sub-arrays assigned to the TX function itself. If the receiver chain is operating in the linear range, the interfering power \bar{P}_{If} at the output of the receiver can be estimated as:

$$\bar{P}_{If} = \int_0^{\infty} C^2(f) \cdot S(f) \cdot G(f) df \tag{6}$$

where: $C(f)$ = MC coefficients, $S(f)$ = power spectrum of the interfering signal at the output of its transmitting aperture and $G(f)$ = power gain of filter in receiver. The Eq. (6) gives an estimate of the interference in terms of an average noise level. The influence on system functions can then

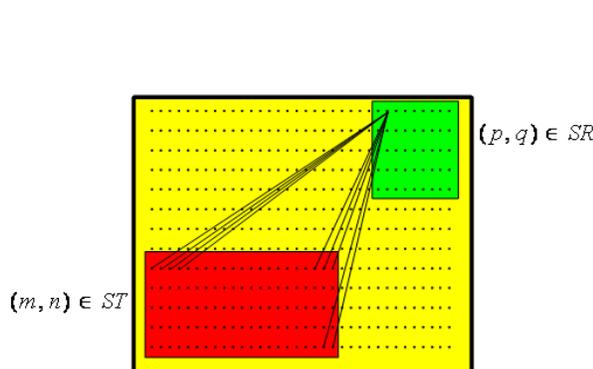


Figure 1: Sub-aperture subdivision for inter-aperture mutual coupling.

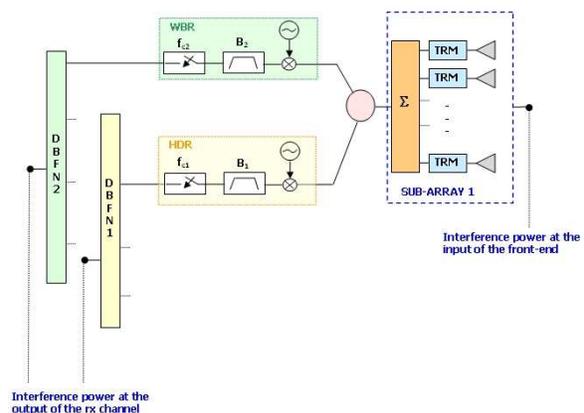


Figure 2: Interference power evaluated at different points of an active phased array with sub-arrays architecture.

be estimated for each function, e.g., in terms of range reduction for radar and ESM or BER (Bit Error Rate) for communication.

We have used a single filter as representative of the selectivity of the whole receiver that is an acceptable assumption. An expression derived from [4] for the power gain $G(f)$ of a filter is:

$$G(f) = \frac{\alpha \cdot G_{\max}}{1 + \left| \frac{1}{W_{RX}} \cdot \left(\frac{f}{f_{RX}} - \frac{f_{RX}}{f} \right) \right|^{2 \cdot N}} \quad (7)$$

where: G_{\max} = in-band gain of the filter, f = frequency, f_{RX} = frequency where the filter is tuned, W_{RX} = relative bandwidth of filter, N = order of the filter, α = normalizing factor to allow the integral of $G(f)$ to be equal to G_{\max} .

The HDR receiver has $N = 6$ pole filter, while the WB receiver filter has a typical order $N = 10$.

The integral for computation of the interfering power \bar{P}_{If} at the output of the receiver can be simplified as:

$$\bar{P}_{If} = C^2(f_{RX}) \cdot \int_0^{\infty} S(f) \cdot G(f) df = C^2(f_{RX}) \cdot P_{TX} \quad (8)$$

$$P_{TX} = \int_0^{\infty} S(f) \cdot G(f) df \quad (9)$$

To compute the interference power at the output of the RX-channel, the mutual coupling coefficients have to be computed considering $f = f_{RX}$. The effective computation of the interference power requires the definition of the transmitted spectrum $S(f)$, for this reason an expression for each function is obtained when the integral (9) is solved:

- Radar function

$$\begin{cases} P_{TX} = P_{peak} \cdot \frac{f_{prf}}{\pi^2 \cdot |f_{TX} - f_{RX}|^2} \cdot B_{RX}, & \text{if } |f_{TX} - f_{RX}| > B_{TX} \\ P_{TX} = P_{peak} & \text{otherwise} \end{cases} \quad (10)$$

where f_{TX} is the transmitter carrier frequency, f_{RX} is the center frequency of the receiver filter, f_{prf} is the pulse repetition frequency and P_{peak} is the transmitter peak power.

- ECM function

$$\begin{cases} P_{TX} = \frac{1}{2\pi^2} \sigma_n^2 B_n \frac{B_{RX}}{|f_{TX} - f_{RX}|^2 - \left(\frac{B_{RX}}{2}\right)^2}, & \text{if } 2|f_{TX} - f_{RX}| > B_{RX} \\ P_{TX} = P_{peak} = \frac{1}{2} \pi^2 \cdot \sigma_n^2 \cdot B_n, & \text{otherwise} \end{cases} \quad (11)$$

where f_{TX} is the transmitter frequency, f_{RX} is the centre frequency of the receiver filter and σ_n^2 is the total noise power.

- COM function

$$\begin{cases} P_{TX} = P_{peak} \cdot \frac{1}{\tau_s \pi^2 \cdot |f_{TX} - f_{RX}|^2} \cdot B_{RX}, & \text{if } |f_{TX} - f_{RX}| > B_{TX} \\ P_{TX} = P_{peak} & \text{otherwise} \end{cases} \quad (12)$$

where, f_{TX} is the transmitter carrier frequency, f_{RX} is the centre frequency of the receiver filter, τ_s is the symbol time and P_{peak} is the transmitter peak power. The interference power values, computed as described before, can be exploited to: (i) have the whole interference power over the RX function at IF, that is then used in the processing of the function to evaluate the loss of performances of the RX function itself; (ii) verify saturation of RX channels.

3.1. Whole Interference Power over the RX Function at IF

The interference power at the output of the RX channel is considered. For each RX function, first of all, each RX sub-array, that is part of the RX aperture/sub-aperture assigned to the RX function itself, is considered. Over each RX sub-array, the model computes separately the interference power due to each TX function considering the proper MC coefficient and summing the contribution of all the TX sub-arrays assigned to the TX function itself. Then, to have the complete interference over the RX function, the contributions of all the RX sub-arrays that are assigned to the RX function itself are summed.

3.2. Saturation of RX Channels

For each RX sub-array, both the interference power at the input of the front-end and the interference power at the output of the RX channel are considered. The first one is considered to verify TRM saturation, the second one to verify receiver saturation. TRM saturation occurs when one or more interference power values at the input of the front-end, due to each TX function, are higher than TRM saturation threshold dependent on the frequency. The receiver saturation occurs when the total interference on the RX function is higher than a proper receiver threshold dependent on the type of receiver used to perform the RX function (e.g., WBR or HDR). In conclusion, the whole RX channel is considered saturated when (i) the TRMs that are part of the RX sub-array are saturated, or (ii) the receiver is saturated.

4. EFFECTS OF INTERFERING SIGNAL ON SYSTEM FUNCTION PERFORMANCE

This section deals with the analyses of the mutual coupling effects on the radar and ESM performances. The considered mutual coupling occurs when during the radar reception or ESM activity other MAESA functions (i.e., TX COMM or ECM) are simultaneously transmitting using a part of the same aperture. The analysis is based on the general assumption that the interfering power can be assumed as noise-like type.

4.1. Sensitivity Degradation for Radar Function in Presence of External Noise-like Disturbance

To evaluate the sensitivity degradation of our system, we consider a system dimensioned to fit the requirements. In nominal conditions, without external disturbance, the performances in terms of the maximum range R_{\max} are given by the well-known equation: $R_{\max} = \sqrt[4]{\frac{P_T B \tau A^2 \sigma}{L 4 \pi \lambda^2 S N R_0 N}}$. When our system interferes with an external disturbance, the performance will be lower than the nominal case and the equation becomes: $R = \sqrt[4]{\frac{P_T B \tau A^2 \sigma}{L 4 \pi \lambda^2 S N R_0 (N+D)}}$ where D is the power due to the disturbance.

Degradation is then normalized as follows $r = \frac{R}{R_{\max}} = \sqrt[4]{\frac{N}{N+D}} = \sqrt[4]{\frac{1}{1+D/N}}$ where $N = K T_0 F B$ is the thermal receiver noise power. The performance evaluation model for radar function is based on the previous formula. The parameter D is the interfering power at the output of the receiver (see Section 3). Inverting the latter equation, we can derive the disturbance-to-noise power ratio (D/N) as a function of the range reduction factor r : $\frac{D}{N} = \frac{(1-r^4)}{r^4}$. This formula is drawn in Fig.3. It shows how the maximum radar range decreases as the disturbance-to-noise power ratio increases.

4.2. Sensitivity Degradation for ESM Function in Presence of External Noise-like Disturbance

This section deals with the analysis of the mutual coupling effect on the ESM performances. The same analytic approach carried out above for the radar function is applied for the ESM performances assessment in presence of an external disturbance. Change only slightly the analytical expressions because the propagation occurs at a single way. ESM range performance reduced in presence of an external noise-like disturbance (D) is $R = \sqrt{\frac{ERP_{ref} A}{4 \pi S N R_0 (N+D)}}$. Let be considered the following

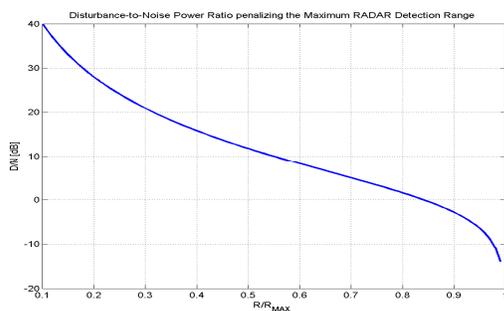


Figure 3: Relationship between the disturbance-to-noise power ratio (D/N) and the radar performance degradation.

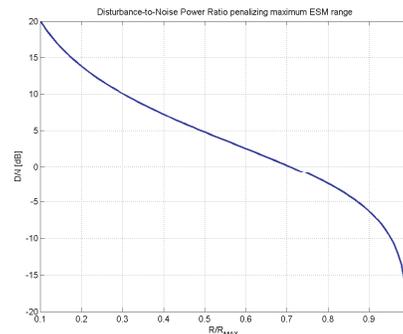


Figure 4: Relationship between the disturbance-to-noise power ratio (D/N) and the ESM performance degradation.

degradation factor (r): $r = \frac{R}{R_{\max}} = \sqrt{\frac{N}{N+D}} = \sqrt{\frac{1}{1+D/N}}$. Inverting the above formula we can derive the disturbance-to-noise power ratio (D/N) as a function of the range reduction factor r : $\frac{D}{N} = \frac{(1-r^2)}{r^2}$. This formula is drawn in Fig. 4 showing how the ESM range decreases as the disturbance-to-noise power ratio (D/N) increases.

5. CONCLUSION

An analysis of the system impact from the mutual interference with respect to antenna coupling effects has been performed and a functional description of model for calculation of the interference levels between TX and RX system functions operating simultaneously has been given.

ACKNOWLEDGMENT

The research activity presented in this paper has been carried out within the M-AESA program, funded by the Italian and Swedish MoDs, which involves the companies Elettronica S.p.A., SELEX Sistemi Integrati and SAAB Electronic Defence Systems.

REFERENCES

1. Lehtreck, L. W., "Effects of coupling accumulation in antenna arrays," *IEEE Trans. Antennas and Propagation*, No. 1, 31–36, 1968.
2. Pozar, D. M., "The active element pattern," *IEEE Trans. Antennas and Propagation*, No. 42, 1176–1178, 1994.
3. Galindo, V. and C. P. Wu, "On the asymptotic decay of coupling for infinite phased arrays," *Proceedings of the IEEE*, 31–36, 1968.
4. Matthaei, G., L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks, and Coupling Structures*, Artech Microwave Library, ISBN-10: 0890060991.

An Interpretation of Maxwell Equation by Using the Formalism of Gravitational Waves

Patrick Vaudon

Xlim, Mixt Research Unit CNRS-University No. 6172, 123 rue Albert Thomas, Limoges 87000, France

Abstract— Maxwell equations are known to be fully compatible with the special theory of relativity, but have no representation in terms of space-time perturbations, in opposition with gravitationnal waves which are described as the propagation of small perturbations in the metric of space-time.

We propose in this paper to interpret Maxwell equations and electromagnetic waves by using the formalism now well established to describe gravitationnal waves. After deriving a momentum-energy tensor T^{ij} for electromagnetic source, we can apply Einstein's equations in their linearised form. We obtain a wave propagation equation on the metric coefficients which is usually used to explain the propagation of gravitationnal waves. Since this propagation equation is identical as the electromagnetic one, we examine how electromagnetic waves can be interpreted as the propagation of small perturbations in the space-time metric. We show that in such an interpretation, the electromagnetic potential plays an equivalent rôle as the gravitationnal potential, and then, can be represented by the metric coefficients. We give the explicite relation between electromagnetic potentials and the metric coefficients which are obtained in this approach and the whole electromagnetic energy associated to a point charge.

1. INTRODUCTION

These last years, very important research efforts have been made in order to try to detect gravitationnal waves. [1, 2] This has lead to develop many explanations on this difficult subject, and today, this kind of waves is well understood. A lot of high level courses give many details to describe how they are derived from gravitation equations [3, 4]. Gravitationnal waves are solutions of these equations in a linearised frame and appear as small perturbations of the space-time metric which propagate in a flat space time described by a Minkowsky metric.

On the other side, electromagnetic waves travel trough space-time like gravitationnal waves, but have no interpretation in terms of space-time curvature. Some works devopped in this way are around Maxwell-Einstein equations which are built by introducing the momentum-energy tensor of the electromagnetic field in the second member of gravitation equations [5–8]. However these equations cannot allow to retrieve Maxwell's ones since they don't include the electromagnetic sources which are known to be charges and currents.

We propose in this paper to introduce in Maxwell-Einstein equations a new momentum-energy tensor which contains the electromagnetic sources, instead of the previous one which contains only the energy of the radiated electromagnetic field. Then we examine the main consequences of such a proposition, and particularly how Maxwell equations can be retrieved in Einstein's one, and how the electromagnetic potential can be linked to the curvature of space time.

2. A MOMENTUM-ENERGY TENSOR FOR ELECTROMAGNETIC SOURCES.

When we examine the momentum energy tensor used to derive gravitationnal waves, we see that it is constructed with moving masses. In a similar manner, if we want to give a momentum-energy tensor T^{ij} for electromagnetic waves, we assume that it will be constructed with moving charges, more precisely with the help of the four-vector current density J^i which is a tensor of first rank.

The momentum-energy tensor in Einstein equations is a rank two tensor. With the idea of retrieving Maxwell equations, we shall assume in a first approximation that all terms are nuls, except those which are presents in the four-vector current density J^i , and we put in a concise form:

$$T^{ij} = \psi c J^i \delta_i^j \quad (1)$$

where c is the light speed, δ_i^j the mixte covariant tensor of Kronecker, and ψ a convenient constant which must allow to respect the dimensions equation. We underline that the summation rule doesn't apply to the right term of Eq. (1): it only means that all terms of the T^{ij} tensor are nul, except on the diagonal ($i = j$) where we find the four-vector current density J^i .

The main property of a momentum-energy tensor concerns its divergence which must be equal to zero.

$$\frac{\partial T^{ij}}{\partial x^i} = \psi c \frac{\partial (J^i \delta_i^j)}{\partial x^i} = 0 \quad (2)$$

Since all terms are nul for i different from j , it remains only:

$$\frac{\partial J^i}{\partial x^i} = 0 \quad (3)$$

which is a usual assumption for small current elements, when we derive electromagnetic fields from Maxwell equations.

3. THE GRAVITATIONNAL WAVES EQUATION

Starting with the gravitation equation:

$$R^{ij} - \frac{1}{2} R g^{ij} = \chi T^{ij} \quad \chi = \frac{8\pi G}{c^4} \quad (4)$$

we assume that the metric coefficients g^{ij} may be considered as small perturbations h^{ij} of the Minkowsky metric η^{ij} :

$$g^{ij} = \eta^{ij} + h^{ij} \quad (5)$$

Since the perturbations h^{ij} are infinitesimal, there introduction in the above Einstein equation allows important simplifications which may be found in courses on gravitationnal waves [3]. Omitting further details, it is found:

$$\frac{\partial^2 h_j^k}{\partial x^i \partial x^k} - \frac{\partial^2 h}{\partial x^i \partial x^j} + \frac{\partial^2 h_i^k}{\partial x^j \partial x^k} - \square(h_{ij}) - \eta_{ij} \frac{\partial^2 h^{pq}}{\partial x^p \partial x^q} + \eta_{ij} \square(h) = \frac{16\pi G}{c^4} T_{ij} \quad (6)$$

where the square symbol represents the d'Alembert operator, and h the trace of the h^{ij} tensor.

By choosing a suitable coordinate system, and a suitable change of variable, this last equation may be put in a simple form:

$$\square(\bar{h}_{ij}) = -\frac{16\pi G}{c^4} T^{ij} \quad (7)$$

together with:

$$\frac{\partial(\bar{h}_i^j)}{\partial x^j} = 0 \quad (8)$$

and

$$\bar{h}^{ij} = h^{ij} - \frac{1}{2} \eta^{ij} h \quad (9)$$

The condition (8) is called the harmonic gauge, and the variable change (9) introduces the \bar{h}^{ij} metric coefficients which will be named the modified metric coefficient in the following.

4. THE ELECTROMAGNETIC POTENTIALS EQUATION

Since Eq. (9) represents a wave equation, and since we have defined a momentum-energy tensor for electromagnetic sources, we can try to use this equation to retrieve Maxwell equations. This can be done by recalling that Maxwell equations are equivalent to the electromagnetic potentials equation:

$$\square(A^i) = -\mu_0 J^i \quad (10)$$

together with the Lorenz gauge.

The identification between (7) and (10) gives a relation between the electromagnetic potentials and the modified metric coefficients:

$$A^i \delta_i^j = \frac{\mu_0 c^3}{16\pi G \Psi} \bar{h}^{ij} \quad (11)$$

The electromagnetic potential must verify the Lorenz gauge:

$$\frac{\partial A^i}{\partial x^i} = 0 \quad (12)$$

By using the relations (11) and (8), this condition implies:

$$\frac{\partial(A^i \delta_i^j)}{\partial x^i} = \frac{\mu_0 c^3}{16\pi G \Psi} \frac{\partial \bar{h}^{ij}}{\partial x^i} = 0 \quad (13)$$

which ensures that the harmonic gauge is satisfied for the modified metric coefficients.

We conclude that we can retrieve Maxwell equations through the space-time description of general relativity, in a linearised frame analogous to the frame used to define the gravitationnal waves. As in this last case, the electromagnetic potentials are represented, with the help of a multiplicative constant, by the metric coefficients.

5. THE COMPLETE DETERMINATION OF THE ELECTROMAGNETIC MOMENTUM-ENERGY TENSOR

The previous relations don't allow to found the explicite value of the constant ψ . After some trials, the best way to obtain this constant seems to derive the radiated energy in two manner: first, like gravitationnal waves by using the space-time formalism, and second by using Maxwell equations.

To obtain the radiated energy for gravitationnal waves is a difficult work, which has been proposed the first time by Einstein himself. No more details can be given in a short paper, but will be found in most of courses on gravitationnal waves. Omitting further details, we found for a dipole located at the origin, oriented along the Ox axis and of dipolar momentum p :

$$\Delta P_s(t, r) = \frac{G}{4\pi r^2 c^3} \psi^2 \left(\frac{d^2 p(t - r/c)}{dt^2} \right)^2 \quad (14)$$

ΔP_s is the surfacic density of radiated power at a distance r of the origin on the Oz axis. The same problem, solved with Maxwell equations give:

$$\Delta P_S(t, r) = \frac{\mu_0}{(4\pi r)^2 c} \left(\frac{d^2 p(t - r/c)}{dt^2} \right)^2 \quad (15)$$

and by identification, the constant ψ is found to be:

$$\psi = \frac{1}{\sqrt{4\pi\epsilon_0 G}} \quad (16)$$

The electromagnetic momentum-energy tensor is now completely determined:

$$T^{ij} = \frac{c}{\sqrt{4\pi\epsilon_0 G}} J^i \delta_i^j \quad (17)$$

By analogy with the momentum-energy tensor of moving masses, the component T^{00} is assumed to contain the whole electromagnetic energy density of a point charge q , and we found this energy W_e after integration:

$$W_e = q \frac{c^2}{\sqrt{4\pi\epsilon_0 G}} = qV_p \quad (18)$$

where V_p is known to be the Planck potential in the Planck units. Alternatively, for an electron, this energy may be written by using the Planck energy E_p , and we obtain:

$$W_e = \sqrt{\alpha} \cdot E_p = \sqrt{\alpha} c^2 \sqrt{\frac{c\hbar}{G}} \quad (19)$$

where α represents the fine structure constant from Sommerfeld.

6. CONCLUSION

A momentum-energy tensor for electromagnetic sources has been proposed in this paper. It may play an analog role as the momentum energy tensor for gravitationnal sources. By substitution in Einstein equations, we have obtained the following equation:

$$R^{ij} - \frac{1}{2} R g^{ij} = \frac{\chi c}{\sqrt{4\pi\epsilon_0 G}} J^i \delta_i^j \quad (20)$$

where $J_i \delta_i^j = J^i$ for $i = j$, and $J_i \delta_i^j = 0$ for i different from j .

When it is linearised, we found that this equation agrees exactly with Maxwell equations for the diagonal terms. This allows to link the electromagnetic potentials to the metric coefficients, in the same manner as for gravitational phenomena.

This approach leads to a simple analogy between the expression of gravitational energy and electromagnetic energy in terms of Planck units:

$$W_e = \frac{q}{q_p} E_p \quad W_m = \frac{m}{m_p} E_p \quad (21)$$

where m_p is defined as the Planck masse: $m_p = \sqrt{\frac{c\hbar}{G}}$, and q_p as the Planck charge: $q_p = \sqrt{4\pi\epsilon_0 c\hbar}$.

A simple coherence is verified when we examine the modulus of the force between two identical charges or two identical masses. If this ratio is equal to one, it means that these two forces have the same amplitude:

$$\frac{F_e}{F_m} = \frac{\frac{q^2}{4\pi\epsilon_0 r^2}}{G \frac{m^2}{r^2}} = \frac{q^2}{m^2} \psi^2 = 1 \quad (22)$$

We may then suppose, that they are derived from a same quantity of energy, electromagnetic for charges, and gravitational for masses. From (21), we obtain:

$$\frac{W_e}{W_m} = \frac{qm_p}{mq_p} = \frac{q}{m} \frac{1}{\sqrt{4\pi\epsilon_0 G}} = \frac{q}{m} \psi = 1 \quad (23)$$

Since the whole electromagnetic energy of the charge has been obtained by using the gravitation equation, these elements of coherence seem to show a possible interpretation of Maxwell equation by using gravitational formalism.

REFERENCES

1. <http://hal.archives-ouvertes.fr/docs/00/28/06/10/PDF/in2p3-00280610.pdf>.
2. <http://www.ligo.caltech.edu/>.
3. Hobson, M. P, G. P. Efstathiou, and A. N. Lasenby, *General Relativity: An Introduction for Physicists*, Cambridge University Press, New York, 2006.
4. http://en.wikipedia.org/wiki/Gravitational_wave.
5. Lovelock, D., “A spherically symmetric solution of the Maxwell-Einstein equations,” *Communications in Mathematical Physics*, Vol. 5, No. 4, 257–261, 1967.
6. Dolan, P., “A singularity free solution of the Maxwell-Einstein equations,” *Communications in Mathematical Physics*, Vol. 9, No. 2, 161–168, 1968.
7. Caricato, G., “Sur le problème de CAUCHY intrinsèque pour les équations de Maxwell-Einstein dans le vide,” *Annales de l'I. H. P.*, Section A, Tome 11, No. 4, 373–392, 1969.
8. McGuire, P. and R. Ruffini, “Some magnetic and electric monopole one-body solutions of the Maxwell-Einstein equations,” *Phys. Rev.*, Vol. 12, 3019–3025, 1975.

Velocity Curl and Spin in Electromagnetic Fields

Zi-Hua Weng

School of Physics and Mechanical & Electrical Engineering
Xiamen University, Xiamen 361005, China

Abstract— The paper discusses impact of the velocity curl and the spin angular velocity on the movement in weightless state in the gravitational and electromagnetic fields, by means of the features of the algebra of octonions. In the gravitational field with the electric-like and magnetic-like components, the object occupies the acceleration and angular velocity. In the electromagnetic field with the electric field intensity and magnetic flux density, the charged particle possesses the spin angular velocity and acceleration-like quantity. The research claims the angular velocity can be detected directly, but the spin angular velocity can not be yet in the gravitational field or the electromagnetic field.

1. INTRODUCTION

J. C. Maxwell [1] described first the electromagnetic theory with the algebra of quaternions, which was invented by W. R. Hamilton in 1843 [2]. Nowadays the gravitational field and electromagnetic field [3] both can be described by the algebra of quaternions, although these two fields are quite distinct from each other. Combining the quaternion space of the electromagnetic field with that of the gravitational field, the octonion space [4] can illustrate simultaneously the features of gravitational field and electromagnetic field.

In the octonion space, the radius vector and the integral of field potential can be combined together to become the compounding radius vector, which can be considered as the vector in the function space. And the related space is called as the compounding space [5]. In this compounding space we can define similarly the field potential, field strength, radius vector, velocity, velocity curl, and angular velocity etc. It is found that the velocity curl has an influence on the movement state of object. And the weightless condition is equivalence to that the compounding field strength is equal to zero.

In the compounding space, the gravitational strength possesses two components. One is the electric-like component, which is the expansion of Newtonian gravitational strength [6] and is similar to the electric field intensity, while the other is to the magnetic flux density. In the gravitational field, two component equations of weightlessness can be derived from the zero compounding field strength. In the weightless state of the gravitational field, it is equal to zero not only the vector sum of the acceleration and the electric-like component of gravitational strength, but also the vector sum of the angular velocity and the magnetic-like component.

Similarly two component equations for the weightless-like state of the electromagnetic fields can be derived from the zero compounding field strength. One is the zero vector sum of the electric field intensity and the acceleration-like quantity, the other is the zero vector sum of the magnetic flux density and the velocity curl. This velocity curl can be considered as the spin angular velocity with the non-quantum explanation. The research claims that the spin angular velocity [7] is necessary when the charged particle is on the weightless-like state.

2. VELOCITY CURL IN GRAVITATIONAL FIELD

Making use of the features of the algebra of quaternions, we may obtain some kinds of compounding physical quantities in the gravitational field. And the definition of compounding physical quantity can draw out the weightless state and the velocity curl in the gravitational field.

2.1. Quaternion Radius Vector

In the quaternion space, the basis vector is $\mathbb{E}_g = (1, \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$, and the coordinates are r_0, r_1, r_2 , and r_3 , while the radius vector is $\mathbb{R}_g = \Sigma(r_i \mathbf{i}_i)$. The latter can be combined with the quaternion physical quantity $\mathbb{X}_g = \Sigma(x_i \mathbf{i}_i)$ to become the compounding radius vector $\bar{\mathbb{R}}_g$,

$$\bar{\mathbb{R}}_g = \mathbb{R}_g + k_{rx} \mathbb{X}_g, \quad (1)$$

where r_i and x_i are all real; $r_0 = v_0 t$; $x_0 = a_0 t$; v_0 is the speed of light; a_0 is the scalar potential of gravitational field; t denotes the time; $k_{rx} = 1/v_0$ is the coefficient; and the \mathbb{X}_g is the integral of gravitational potential; $i = 0, 1, 2, 3$, with $\mathbf{i}_0 = 1$.

In other words, the $\bar{\mathbb{R}}_g$ can be considered as the radius vector in the quaternion compounding space, with the quaternion basis vector $(1, \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$. In the quaternion compounding space, the velocity $\mathbb{V}_g(v_0, v_1, v_2, v_3)$ is combined with the gravitational potential $\mathbb{A}_g(a_0, a_1, a_2, a_3)$ to become the compounding velocity $\bar{\mathbb{V}}_g(\bar{v}_0, \bar{v}_1, \bar{v}_2, \bar{v}_3)$,

$$\bar{\mathbb{V}}_g = \mathbb{V}_g + k_{rx}\mathbb{A}_g, \quad (2)$$

where the $\mathbf{a} = \Sigma(a_j \mathbf{i}_j)$; $j = 1, 2, 3$.

The above states the velocity \mathbb{V}_g are connected with the gravitational potential \mathbb{A}_g . Not only the field potential but also the velocity has an influence on the movement state, such as the movement in weightless state in the gravitational field.

Table 1: The octonion multiplication table.

| | 1 | \mathbf{i}_1 | \mathbf{i}_2 | \mathbf{i}_3 | \mathbf{I}_0 | \mathbf{I}_1 | \mathbf{I}_2 | \mathbf{I}_3 |
|----------------|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 1 | 1 | \mathbf{i}_1 | \mathbf{i}_2 | \mathbf{i}_3 | \mathbf{I}_0 | \mathbf{I}_1 | \mathbf{I}_2 | \mathbf{I}_3 |
| \mathbf{i}_1 | \mathbf{i}_1 | -1 | \mathbf{i}_3 | $-\mathbf{i}_2$ | \mathbf{I}_1 | $-\mathbf{I}_0$ | $-\mathbf{I}_3$ | \mathbf{I}_2 |
| \mathbf{i}_2 | \mathbf{i}_2 | $-\mathbf{i}_3$ | -1 | \mathbf{i}_1 | \mathbf{I}_2 | \mathbf{I}_3 | $-\mathbf{I}_0$ | $-\mathbf{I}_1$ |
| \mathbf{i}_3 | \mathbf{i}_3 | \mathbf{i}_2 | $-\mathbf{i}_1$ | -1 | \mathbf{I}_3 | $-\mathbf{I}_2$ | \mathbf{I}_1 | $-\mathbf{I}_0$ |
| \mathbf{I}_0 | \mathbf{I}_0 | $-\mathbf{I}_1$ | $-\mathbf{I}_2$ | $-\mathbf{I}_3$ | -1 | \mathbf{i}_1 | \mathbf{i}_2 | \mathbf{i}_3 |
| \mathbf{I}_1 | \mathbf{I}_1 | \mathbf{I}_0 | $-\mathbf{I}_3$ | \mathbf{I}_2 | $-\mathbf{i}_1$ | -1 | $-\mathbf{i}_3$ | \mathbf{i}_2 |
| \mathbf{I}_2 | \mathbf{I}_2 | \mathbf{I}_3 | \mathbf{I}_0 | $-\mathbf{I}_1$ | $-\mathbf{i}_2$ | \mathbf{i}_3 | -1 | $-\mathbf{i}_1$ |
| \mathbf{I}_3 | \mathbf{I}_3 | $-\mathbf{I}_2$ | \mathbf{I}_1 | \mathbf{I}_0 | $-\mathbf{i}_3$ | $-\mathbf{i}_2$ | \mathbf{i}_1 | -1 |

2.2. Velocity Curl

In the compounding space, the compounding potential $\bar{\mathbb{A}}_g(\bar{a}_0, \bar{a}_1, \bar{a}_2, \bar{a}_3)$ is defined as

$$\bar{\mathbb{A}}_g = \mathbb{A}_g + K_{rx}\mathbb{V}_g, \quad (3)$$

where $K_{rx} = 1/k_{rx}$.

The compounding strength $\bar{\mathbb{B}}_g(\bar{b}_0, \bar{b}_1, \bar{b}_2, \bar{b}_3)$ is defined from the above,

$$\bar{\mathbb{B}}_g = \diamond \circ \bar{\mathbb{A}}_g = \mathbb{B}_g + K_{rx}\mathbb{U}_g, \quad (4)$$

where the \diamond denotes the octonion multiplication in Table 1; the operator $\diamond = \Sigma(\mathbf{i}_i \partial_i)$, $\nabla = \Sigma(\mathbf{i}_j \partial_j)$, $\partial_i = \partial/\partial r_i$; the velocity $\mathbb{V}_g = v_0[\diamond \circ \mathbb{R}_g - \nabla \cdot \{\Sigma(r_j \mathbf{i}_j)\}]$, the velocity curl $\mathbb{U}_g = \diamond \circ \mathbb{V}_g$; the field potential $\mathbb{A}_g = \diamond \circ \mathbb{X}_g$; the field strength $\mathbb{B}_g = \diamond \circ \mathbb{A}_g$; the scalar $\bar{b}_0 = \partial_0 \bar{a}_0 - \Sigma(\partial_j \bar{a}_j)$.

To simplify succeeding calculation, we choose the gauge condition $\bar{b}_0 = 0$ in the paper. And then the gravitational strength is written as $\bar{\mathbb{B}}_g = \bar{\mathbf{g}}/\bar{v}_0 + \bar{\mathbf{b}}$,

$$\bar{\mathbf{g}}/\bar{v}_0 = \partial_0 \bar{\mathbf{a}} + \nabla \bar{a}_0, \quad \bar{\mathbf{b}} = \nabla \times \bar{\mathbf{a}}, \quad (5)$$

where $\bar{\mathbf{a}} = \Sigma(\bar{a}_j \mathbf{i}_j)$; $\bar{\mathbf{b}} = \Sigma(\bar{b}_j \mathbf{i}_j)$.

The above equations state the gravitational field in the compounding space possesses not only the field potential and field strength but also the velocity and the velocity curl. The gravitational strength $\bar{\mathbb{B}}_g$ and velocity curl \mathbb{U}_g both have an influence on the movement in weightless state.

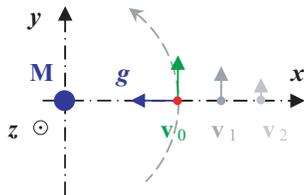


Figure 1: The object is moving in the weightless states around the M with some different velocities \mathbf{v}_0 , \mathbf{v}_1 , and \mathbf{v}_2 , under the gravitational field with the electric-like component \mathbf{g} in different orbits.

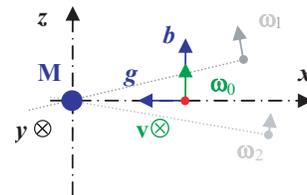


Figure 2: The object is orbiting in the weightless state around the M with the velocity \mathbf{v} and angular velocities $\bar{\omega}_0$, $\bar{\omega}_1$, and $\bar{\omega}_2$, under the gravitational field with the \mathbf{g} and magnetic-like component \mathbf{b} .

In the quaternion compounding space, the weightless state can be derived directly from that the compounding strength is equal to zero, that is, $\bar{\mathbf{g}} = 0$ and $\bar{\mathbf{b}} = 0$,

$$\mathbf{g} + k_{rx}\mathbf{a}_g = 0, \quad \mathbf{b} + k_{rx}\mathbf{u}_g = 0, \quad (6)$$

where $\mathbf{g}/v_0 = \partial_0\mathbf{a} + \nabla a_0$ is the electric-like component of gravitational strength and is the expansion of Newtonian gravitational strength; $\mathbf{a}_g/v_0 = \partial_0\mathbf{v} + \nabla v_0$ is the acceleration; $\mathbf{b} = \nabla \times \mathbf{a}$ is the magnetic-like component of gravitational strength; $\mathbf{u}_g = \nabla \times \mathbf{v}$ is the velocity curl, with $\mathbf{v} = \Sigma(v_j\mathbf{i}_j)$ and $\mathbf{r} = \Sigma(r_j\mathbf{i}_j)$. In most cases there is $\bar{v}_0 \approx v_0$. In the planar polar coordinates, the velocity is $\mathbf{v} = \vec{\omega} \times \mathbf{r}$, with $\vec{\omega}$ being the angular velocity. And then we have the relation $\nabla \times \mathbf{v} = 2\vec{\omega}$.

The above means that the object possesses the acceleration \mathbf{a}_g and the velocity curl \mathbf{u}_g in the gravitational field with the \mathbf{g} and \mathbf{b} . In the gravitational field, the object must be in the movement state that there is either the linear motion or the rotation. And in the weightless state, the object should have the angular velocity, even in the linear movement.

Table 2: The multiplication of the physical quantity in the octonion space.

| definition | meaning |
|-----------------------------|--|
| $\nabla \cdot \mathbf{a}$ | $-(\partial_1 a_1 + \partial_2 a_2 + \partial_3 a_3)$ |
| $\nabla \times \mathbf{a}$ | $\mathbf{i}_1(\partial_2 a_3 - \partial_3 a_2) + \mathbf{i}_2(\partial_3 a_1 - \partial_1 a_3) + \mathbf{i}_3(\partial_1 a_2 - \partial_2 a_1)$ |
| ∇a_0 | $\mathbf{i}_1\partial_1 a_0 + \mathbf{i}_2\partial_2 a_0 + \mathbf{i}_3\partial_3 a_0$ |
| $\partial_0 \mathbf{a}$ | $\mathbf{i}_1\partial_0 a_1 + \mathbf{i}_2\partial_0 a_2 + \mathbf{i}_3\partial_0 a_3$ |
| $\nabla \cdot \mathbf{A}$ | $-(\partial_1 A_1 + \partial_2 A_2 + \partial_3 A_3)\mathbf{I}_0$ |
| $\nabla \times \mathbf{A}$ | $-\mathbf{I}_1(\partial_2 A_3 - \partial_3 A_2) - \mathbf{I}_2(\partial_3 A_1 - \partial_1 A_3) - \mathbf{I}_3(\partial_1 A_2 - \partial_2 A_1)$ |
| $\nabla \circ \mathbf{A}_0$ | $\mathbf{I}_1\partial_1 A_0 + \mathbf{I}_2\partial_2 A_0 + \mathbf{I}_3\partial_3 A_0$ |
| $\partial_0 \mathbf{A}$ | $\mathbf{I}_1\partial_0 A_1 + \mathbf{I}_2\partial_0 A_2 + \mathbf{I}_3\partial_0 A_3$ |

3. SPIN IN ELECTROMAGNETIC FIELD

By means of the property of the algebra of octonions, we may obtain some kinds of compounding physical quantities in the electromagnetic and gravitational fields simultaneously. And the definition of compounding physical quantity can deduce the weightless state, velocity curl and spin angular velocity in the electromagnetic and gravitational fields.

3.1. Octonion Radius Vector

In the octonion compounding space, the basis vector \mathbb{E} consists of the basis vectors \mathbb{E}_g and \mathbb{E}_e . The basis vector $\mathbb{E}_e = (\mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3)$ is the basis vector of the quaternion space for the electromagnetic field, while the \mathbb{E}_g is for the gravitational field. Meanwhile the basis vector \mathbb{E}_e is independent of the \mathbb{E}_g , with $\mathbb{E}_e = \mathbb{E}_g \circ \mathbf{I}_0$. These two basis vectors are combined together to become the basis vector for the octonion space,

$$\mathbb{E} = (1, \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3). \quad (7)$$

In the octonion space for the gravitational field and the electromagnetic field, the radius vector $\mathbb{R} = \Sigma(r_i\mathbf{i}_i) + k_{eg}\Sigma(R_i\mathbf{I}_i)$ can be combined with the octonion quantity $\mathbb{X} = \Sigma(x_i\mathbf{i}_i) + k_{eg}\Sigma(X_i\mathbf{I}_i)$ to become the compounding radius vector $\bar{\mathbb{R}}$,

$$\bar{\mathbb{R}} = \mathbb{R} + k_{rx}\mathbb{X}. \quad (8)$$

where k_{eg} is the coefficient, $\mathbb{R}_e = \Sigma(R_i\mathbf{I}_i)$.

In other words, the $\bar{\mathbb{R}}$ can be considered as the radius vector in the octonion compounding space, with the basis vector \mathbb{E} . And the velocity $\mathbb{V} = \Sigma(v_i\mathbf{i}_i) + k_{eg}\Sigma(V_i\mathbf{I}_i)$ can be combined with the field potential $\mathbb{A} = \Sigma(a_i\mathbf{i}_i) + k_{eg}\Sigma(A_i\mathbf{I}_i)$ to become the velocity $\bar{\mathbb{V}} = \Sigma(\bar{v}_i\mathbf{i}_i) + k_{eg}\Sigma(\bar{V}_i\mathbf{I}_i)$ in the octonion compounding space,

$$\bar{\mathbb{V}} = \mathbb{V} + k_{rx}\mathbb{A}, \quad (9)$$

where $\bar{v}_i = v_i + k_{rx}a_i$; $\bar{V}_i = V_i + k_{rx}A_i$; $\mathbb{V}_e = \Sigma(V_i\mathbf{I}_i)$. And the field potential \mathbb{A} consists of the gravitational potential $\mathbb{A}_g = \Sigma(a_i\mathbf{i}_i)$ and the electromagnetic potential $\mathbb{A}_e = \Sigma(A_i\mathbf{I}_i)$.

3.2. Spin Angular Velocity

In the octonion compounding space, the potential $\bar{\mathbb{A}} = \Sigma(\bar{a}_i \mathbf{i}_i) + k_{eg} \Sigma(\bar{A}_i \mathbf{I}_i)$ is defined as,

$$\bar{\mathbb{A}} = \mathbb{A} + K_{rx} \mathbb{V}, \quad (10)$$

where $\bar{a}_i = a_i + K_{rx} v_i$; $\bar{A}_i = A_i + K_{rx} V_i$. \mathbf{A}_0 is the scalar potential of electromagnetic field.

The compounding strength $\bar{\mathbb{B}} = \Sigma(\bar{b}_i \mathbf{i}_i) + k_{eg} \Sigma(\bar{B}_i \mathbf{I}_i)$ is defined from the field potential $\bar{\mathbb{A}}$,

$$\bar{\mathbb{B}} = \diamond \circ \bar{\mathbb{A}} = \mathbb{B} + K_{rx} \mathbb{U}, \quad (11)$$

where $\bar{b}_i = b_i + K_{rx} u_i$, $\bar{B}_i = B_i + K_{rx} U_i$; the velocity $\mathbb{V} = v_0[\diamond \circ \mathbb{R} - \nabla \cdot \{\Sigma(r_j \mathbf{i}_j)\}]$, the velocity curl $\mathbb{U} = \diamond \circ \mathbb{V}$; the field potential $\mathbb{A} = \diamond \circ \mathbb{X}$; the field strength $\mathbb{B} = \diamond \circ \mathbb{A}$. The compounding strength $\mathbb{B} = \mathbb{B}_g + k_{eg} \mathbb{B}_e$ consists of the gravitational strength $\mathbb{B}_g = \Sigma(b_i \mathbf{i}_i)$ as well as the electromagnetic strength $\mathbb{B}_e = \Sigma(B_i \mathbf{I}_i)$. $\mathbf{A}_0 = A_0 \mathbf{I}_0$ is the scalar potential of electromagnetic field.

To simplify following calculation, we choose two gauge conditions $\bar{b}_0 = 0$ and $\bar{B}_0 = 0$. As a result, the gravitational strength $\bar{\mathbb{B}}_g = \bar{\mathbf{g}}/\bar{v}_0 + \bar{\mathbf{b}}$ is similar to that in Eq.(5), while the electromagnetic strength is written as $\bar{\mathbb{B}}_e = \bar{\mathbf{E}}/\bar{v}_0 + \bar{\mathbf{B}}$, with $\bar{\mathbf{A}} = \Sigma(\bar{A}_j \mathbf{I}_j)$ and $\bar{\mathbf{B}} = \Sigma(\bar{B}_j \mathbf{I}_j)$,

$$\bar{\mathbf{E}}/\bar{v}_0 = \partial_0 \bar{\mathbf{A}} + \nabla \circ \mathbf{A}_0, \quad \bar{\mathbf{B}} = \nabla \times \bar{\mathbf{A}}. \quad (12)$$

The above equations state the electromagnetic field in the compounding space possesses not only the field potential \mathbb{A}_e and field strength \mathbb{B}_e but also the velocity \mathbb{V}_e and the velocity curl \mathbb{U}_e . The electromagnetic strength \mathbb{B}_e and the velocity curl \mathbb{U}_e both have an influence on the weightless-like movements of charged particle in the electromagnetic field.

In the octonion compounding space, the condition of the weightless state is equivalence to that the compounding strength is equal to zero, that is $\bar{\mathbb{B}} = \bar{\mathbb{B}}_g + k_{eg} \bar{\mathbb{B}}_e = 0$, or

$$\bar{\mathbb{B}}_g = 0, \quad \bar{\mathbb{B}}_e = 0. \quad (13)$$

Similarly to Eq. (6), the weightless-like state is equivalence to that the compounding strength is equal to zero in the electromagnetic field, that is, $\bar{\mathbf{E}} = 0$ and $\bar{\mathbf{B}} = 0$,

$$\mathbf{E} + k_{rx} \mathbf{a}_e = 0, \quad \mathbf{B} + k_{rx} \mathbf{u}_e = 0, \quad (14)$$

where $\mathbf{E}/v_0 = \partial_0 \mathbf{A} + \nabla \circ \mathbf{A}_0$ is the electric field intensity; $\mathbf{a}_e/v_0 = \partial_0 \mathbf{V} + \nabla \circ \mathbf{V}_0$ is the acceleration-like quantity; $\mathbf{B} = \nabla \times \mathbf{A}$ is the magnetic flux density; $\mathbf{u}_e = \nabla \times \mathbf{V}$ is the velocity curl, with the $\mathbf{V} = \Sigma(V_j \mathbf{I}_j)$ and $\mathbf{A} = \Sigma(A_j \mathbf{I}_j)$. Similarly in the planar polar coordinates, the velocity is written as the $\mathbf{V} = -\vec{\omega}_e \times \mathbf{r}$, with the $\vec{\omega}_e$ being the physical quantity similar to the angular velocity in the gravitational field. Therefore we have the relation $\nabla \times \mathbf{V} = -2\vec{\omega}_e$.

In the electromagnetic field with the \mathbf{E} and \mathbf{B} , the above means that the charged particle has the acceleration-like quantity \mathbf{a}_e and the velocity curl \mathbf{u}_e . And that the charged particle should be in the movement state that there is either the linear-like movement or the rotation-like locomotion. And in the movement in weightless-like state (or the forceless state), the charged particle must have the 'angular velocity' $\vec{\omega}_e$, even in the linear-like movement. Accordingly the $\vec{\omega}_e$ can be considered as the spin angular velocity of the charged particle.

In the paper, the physical quantity with the basis vector \mathbb{E}_g can be detected directly, but that with the \mathbb{E}_e can not be yet. According to this viewpoint, the velocity \mathbf{v} , the angular velocity $\vec{\omega}$, the electric-like strength \mathbf{g} , and the magnetic-like strength \mathbf{b} etc can be detected for being in the quaternion space \mathbb{E}_g of the gravitational field. But the velocity \mathbf{V} , the spin angular velocity $\vec{\omega}_e$,

Table 3: The movement state in the electromagnetic and gravitational fields.

| field strength | acceleration and velocity curl | forceless state |
|---------------------------------------|---|-----------------------|
| electric-like strength \mathbf{g} | acceleration \mathbf{a}_g | weightless state |
| magnetic-like strength \mathbf{b} | velocity curl \mathbf{u}_g | weightless state |
| electric field intensity \mathbf{E} | acceleration-like quantity \mathbf{a}_e | weightless-like state |
| magnetic flux density \mathbf{B} | velocity curl \mathbf{u}_e | weightless-like state |

the electric field intensity \mathbf{E} , and the magnetic flux density \mathbf{B} etc can not be detected directly for being in the quaternion space \mathbb{E}_e of the electromagnetic field.

However, the multiplication of two physical quantities with \mathbb{E}_e can be detected, because this multiplication possesses the basis vector \mathbb{E}_g in terms of the Table 1. For example, the velocity \mathbf{V} and the magnetic flux density \mathbf{B} both can not be detected, but the Lorentz force $\mathbf{F} = -q\mathbf{V} \times \mathbf{B}$ can be detected directly, because the \mathbf{F} is the physical quantity with the vector basis \mathbb{E}_g , with the q being the electric charge. Further the angular velocity $\vec{\omega}$ can be detected directly, while the spin angular velocity $\vec{\omega}_e$ can not be yet for being the physical quantity with the basis vector \mathbb{E}_e .

4. CONCLUSIONS

Making use of the features of quaternions, we obtain some compounding physical quantities for the gravitational field. The compounding physical quantity can draw out the weightless state and velocity curl in the gravitational field. The object possesses the acceleration \mathbf{a}_g and velocity curl \mathbf{u}_g in the gravitational field with the \mathbf{g} and \mathbf{b} . In the gravitational field, the object is in the movement state of linear movement with the rotation, and has the angular velocity $\vec{\omega}$ in the movement in weightless state.

By means of the property of the algebra of octonions, we acquire some compounding physical quantities for the electromagnetic and gravitational fields simultaneously. And the compounding physical quantity can deduce the weightless state, the velocity curl, and the spin angular velocity for the electromagnetic field. The charged particle possesses the acceleration-like quantity \mathbf{a}_e and the velocity curl \mathbf{u}_e in the electromagnetic field with the \mathbf{E} and \mathbf{B} . In the electromagnetic field, the charged particle is in the movement state of either the linear-like motion or the rotation-like locomotion. In the movement in weightless-like state, the charged particle has the 'angular velocity' $\vec{\omega}_e$, which is considered as the spin angular velocity of the charged particle in the electromagnetic field. The spin angular velocity $\vec{\omega}_e$ can not be detected for being the physical quantity with the basis vector \mathbb{E}_e , although the multiplication of two physical quantities with the \mathbb{E}_e can be detected directly in the octonion compounding space.

It should be noted that the study for velocity curl and spin angular velocity examined only one simple case in the gravitational and electromagnetic fields. Despite its preliminary character, this study can clearly indicate the field strength, velocity, velocity curl, and spin angular velocity have the influence on the movement in weightless state in the gravitational field or electromagnetic field. For the future studies, the related research will concentrate on only the impact of the spin angular velocity on the movement state of charged particle in the electromagnetic field.

ACKNOWLEDGMENT

This project was supported partially by the National Natural Science Foundation of China under grant number 60677039.

REFERENCES

1. Maxwell, J. C., *A Treatise on Electricity and Magnetism*, Dover Publications, New York, 1954.
2. Hamilton, W. R., *Elements of Quaternions*, Longmans, Green & Co., London, 1866.
3. Heaviside, O., "A gravitational and electromagnetic analogy," *The Electrician*, Vol. 31, 281–282, 359, 1893.
4. Cayley, A., *The Collected Mathematical Papers*, Johnson Reprint Co., New York, 1963.
5. Weng, Z.-H., "Wave equations in electromagnetic and gravitational fields," *PIERS Proceedings*, 971–975, Cambridge, USA, July 5–8, 2010.
6. Newton, I., *The Mathematical Principles of Natural Philosophy*, trans. A. Motte, Dawsons of Pall Mall, London, 1968.
7. Uhlenbeck, G. E. and S. Goudsmit, "Spinning electrons and the structure of spectra," *Nature*, Vol. 117, No. 2938, 264–265, 1926.

Theorem for the Identity of the $L(c, n)$ and $\hat{L}(\hat{c}, \hat{n})$ Numbers and Its Application in the Theory of Waveguides

Georgi Nikolov Georgiev¹ and Mariana Nikolova Georgieva-Grosse²

¹Faculty of Mathematics and Informatics
University of Veliko Tirnovo “St. St. Cyril and Methodius”, BG-5000 Veliko Tirnovo, Bulgaria
²Meterstrasse 4, D-70839 Gerlingen, Germany

Abstract— It is proved numerically that if for the arbitrary real numbers c and \hat{c} , and for the natural ones n and \hat{n} it is fulfilled: $c = \hat{c}$ and $n = \hat{n}$, then the positive real numbers $L(c, n)$ and $\hat{L}(\hat{c}, \hat{n})$, connected with the positive purely imaginary, resp. real zeros of the Kummer function $\Phi(a, c; x)$ of specially selected complex, resp. real parameters and variable, coincide. This statement is formulated as a Theorem for the identity of the $L(c, n)$ and $\hat{L}(\hat{c}, \hat{n})$ numbers. A physical interpretation of the result in case $c = \hat{c} = 3$ is made: a curve of equation, written in terms of any of the latter is found which separates the domains of propagation of the normal and slow rotationally symmetric TE modes in the circular waveguide with magnetized azimuthally in negative (clockwise) direction ferrite.

1. INTRODUCTION

The $L(c, n)$ numbers [1, 2] are positive real ones, defined as the common limits of some infinite sequences of numbers whose terms are proportional to the positive purely imaginary zeros of a complex Kummer confluent hypergeometric function $\Phi(a, c; x)$ [3] of chosen in an appropriate way parameters and variable [1, 2]. The notion $\hat{L}(\hat{c}, \hat{n})$ number is determined likewise [4]. However, the generating function and the zeros, used now, are real. To distinguish both cases, it is accepted to introduce hats “^” above all the symbols, relevant to the second class of numbers. The quantities c and \hat{c} (the second parameters of Kummer functions) are arbitrary real numbers, while n and \hat{n} (the numbers of the zeros in question of the same) — natural ones.

In this study a theorem is proved numerically, asserting that if $c = \hat{c}$ and $n = \hat{n}$, it is valid: $L(c, n) \equiv \hat{L}(\hat{c}, \hat{n})$. The two possibilities $c \neq l$, resp. $\hat{c} \neq \hat{l}$ ($l = \hat{l} = 0, -1, -2, \dots$) and $c = l$, resp. $\hat{c} = \hat{l}$ (in which $\Phi(a, c; x)$ is not defined), are considered. It is shown that the theory of the circular waveguide, entirely filled with azimuthally magnetized ferrite, propagating rotationally symmetric TE modes [1, 3–5] may serve as a field of employment of the outcome obtained.

2. $L(c, n)$ AND $\hat{L}(\hat{c}, \hat{n})$ NUMBERS

Definition 1: The $L(c, n)$ numbers in which c is an arbitrary real and n is a natural number, are finite positive real ones, specified in the following way [2]:

i) Provided $c \neq l$, $l = 0, -1, -2, \dots$:

$$L(c, n) = \lim_{k_- \rightarrow -\infty} K_-(c, n, k_-) = \lim_{k_- \rightarrow -\infty} M_-(c, n, k_-) \quad (1)$$

where $K_-(c, n, k_-) = |k_-| \zeta_{k_-, n}^{(c)}$, $M_-(c, n, k_-) = |a_-| \zeta_{k_-, n}^{(c)}$ and $\zeta_{k_-, n}^{(c)}$ is the n th positive purely imaginary zero in x of the Kummer confluent hypergeometric function $\Phi(a_-, c; x)$ [3], ($n = 1, 2, 3, \dots$) with $a_- = c/2 - jk_-$ — complex, $x = jz$, z — real, positive and k_- — real, negative, (c, n — fixed).

ii) On the understanding that $c = l$ and ε is an infinitesimal positive real number:

$$L(c, n) = \lim_{\varepsilon \rightarrow 0} L(l - \varepsilon, n) = \lim_{\varepsilon \rightarrow 0} L(l + \varepsilon, n + 1) = L(2 - l, n). \quad (2)$$

Furthermore, it is valid, too [2]:

$$L(1 + l, n) = L(1 - l, n), \quad (3)$$

$$\lim_{\varepsilon \rightarrow 0} L(l + \varepsilon, 1) = 0, \quad (4)$$

$$\left(\begin{array}{c} L(0.5, n) \\ L(1.5, n) \end{array} \right) = \left(\begin{array}{c} (2n - 1)^2 \\ (2n)^2 \end{array} \right) \left(\frac{\pi}{4} \right)^2. \quad (5)$$

Definition 2: The $\hat{L}(\hat{c}, \hat{n})$ numbers in which \hat{c} is an arbitrary real and \hat{n} is a natural number, are finite positive real ones, determined, as follows [4]:

i) If $\hat{c} \neq \hat{l}$, $\hat{l} = 0, -1, -2, \dots$:

$$\hat{L}(\hat{c}, \hat{n}) = \lim_{\hat{k}_- \rightarrow -\infty} \hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-) = \lim_{\hat{k}_- \rightarrow -\infty} \hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-) \tag{6}$$

where $\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-) = |\hat{k}_-|^{\zeta_{\hat{k}_-, \hat{n}}^{(\hat{c})}}$, $\hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-) = |\hat{a}_-|^{\zeta_{\hat{k}_-, \hat{n}}^{(\hat{c})}}$ and $\zeta_{\hat{k}_-, \hat{n}}^{(\hat{c})}$ is the \hat{n} th positive real zero in \hat{x} of the Kummer confluent hypergeometric function $\Phi(\hat{a}_-, \hat{c}; \hat{x})$ [3] with $\hat{a}_- = \hat{c}/2 + \hat{k}_-$ — real, negative, \hat{x} — real, positive and \hat{k}_- — real, negative, ($\hat{n} = 1, 2, \dots, \hat{p}$, $\hat{p} = \text{abs}[\hat{a}_-]$ in case $\hat{c} > 0$ or $\hat{n} = 1, 2, \dots, \hat{s}$, $\hat{s} = \hat{p} - \hat{q}$, $\hat{p} = \text{abs}[\hat{a}_-]$, $\hat{q} = \text{abs}[\hat{c}]$, $\hat{q} = 1, 2, \dots, \hat{p} - 1$ when $\hat{c} < 0$ and $\hat{a}_- < \hat{c} < 0$), ($[\hat{a}_-]$ denotes the largest integer less or equal to \hat{a}_-), (\hat{c}, \hat{n} — fixed).

ii) Stipulating that $\hat{c} = \hat{l}$ and $\hat{\varepsilon}$ is an infinitesimal positive real number:

$$\hat{L}(\hat{c}, \hat{n}) = \lim_{\hat{\varepsilon} \rightarrow 0} L(\hat{l} - \hat{\varepsilon}, \hat{n}) = \lim_{\hat{\varepsilon} \rightarrow 0} L(\hat{l} + \hat{\varepsilon}, \hat{n} + 1) = \hat{L}(2 - \hat{l}, \hat{n}). \tag{7}$$

Moreover, it is implemented, as well [4]:

$$\hat{L}(1 + \hat{l}, \hat{n}) = \hat{L}(1 - \hat{l}, \hat{n}), \tag{8}$$

$$\lim_{\hat{\varepsilon} \rightarrow 0} L(\hat{l} + \hat{\varepsilon}, 1) = 0, \tag{9}$$

$$\begin{pmatrix} \hat{L}(0.5, \hat{n}) \\ \hat{L}(1.5, \hat{n}) \end{pmatrix} = \begin{pmatrix} (2\hat{n} - 1)^2 \\ (2\hat{n})^2 \end{pmatrix} \left(\frac{\pi}{4}\right)^2. \tag{10}$$

3. BASIC THEOREM

3.1. Formulation

Theorem 1: On condition that $c = \hat{c}$ and $n = \hat{n}$ where c and \hat{c} are arbitrary real, and n and \hat{n} — natural numbers, the $L(c, n)$ numbers in the sense of **Definition 1** and the $\hat{L}(\hat{c}, \hat{n})$ ones in that of **Definition 2**, are identical. Accordingly, it holds:

$$L(c, n) \equiv \hat{L}(\hat{c}, \hat{n}). \tag{11}$$

If beside the above assumptions it is fulfilled also that $c = \hat{c} \neq l$ ($l = \hat{l}$) and $k_- = \hat{k}_-$ is large negative, it is true:

$$\zeta_{k_-, n}^{(c)} \approx \zeta_{\hat{k}_-, \hat{n}}^{(\hat{c})}, \tag{12}$$

$$K_-(c, n, k_-) \approx M_-(c, n, k_-) \approx \hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-) \approx \hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-). \tag{13}$$

(The meaning of all symbols is as in **Definitions 1** and **2**.)

3.2. Corrolaries

i) The $L(c, n)$ ($\hat{L}(\hat{c}, \hat{n})$) numbers ($c = \hat{c} \neq 0, -1, -2, \dots$, $n = \hat{n} = 1, 2, 3, \dots$), could approximately be calculated by the expression: $L(c, n) \approx |k_-|^{\zeta_{k_-, n}^{(c)}} \approx |\hat{k}_-|^{\zeta_{\hat{k}_-, \hat{n}}^{(\hat{c})}} \approx |a_-|^{\zeta_{k_-, n}^{(c)}} \approx |\hat{a}_-|^{\zeta_{\hat{k}_-, \hat{n}}^{(\hat{c})}} \approx K_-(c, n, k_-) \approx M_-(c, n, k_-) \approx \hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-) \approx \hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-)$ taking for granted that the numerical equivalents of the zeros $\zeta_{k_-, n}^{(c)}$ ($\zeta_{\hat{k}_-, \hat{n}}^{(\hat{c})}$) of function $\Phi(a_-, c; x)$ ($\Phi(\hat{a}_-, \hat{c}; \hat{x})$) or of the numbers $K_-(c, n, k_-)$ ($\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-)$) or $M_-(c, n, k_-)$ ($\hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-)$) for large negative k_- (\hat{k}_-), are figured out beforehand.

ii) The values of zeros $\zeta_{k_-, n}^{(c)}$ ($\zeta_{\hat{k}_-, \hat{n}}^{(\hat{c})}$) of the function $\Phi(a_-, c; x)$ ($\Phi(\hat{a}_-, \hat{c}; \hat{x})$) ($c = \hat{c} \neq 0, -1, -2, \dots$, $n = \hat{n} = 1, 2, 3, \dots$) for large negative k_- (\hat{k}_-) could at a rough guess be counted from the formula:

$$\begin{aligned} \zeta_{k_-, n}^{(c)} \approx \zeta_{\hat{k}_-, \hat{n}}^{(\hat{c})} \approx L(c, n) / |k_-| &= \hat{L}(\hat{c}, \hat{n}) / |\hat{k}_-| \approx L(c, n) / |a_-| = \hat{L}(\hat{c}, \hat{n}) / |\hat{a}_-| \approx K_-(c, n, k_-) / |k_-| \\ &\approx \hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-) / |\hat{k}_-| \approx M_-(c, n, k_-) / |a_-| \approx \hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-) / |\hat{a}_-|, \end{aligned} \tag{14}$$

assuming that the $L(c, n) (\hat{L}(\hat{c}, \hat{n}))$ numbers or $K_-(c, n, k_-) (\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-))$, $M_-(c, n, k_-) (\hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-))$ are known.

iii) The zeros $\zeta_{k_{-,1},n}^{(c)}$ and $\zeta_{k_{-,2},n}^{(c)}$ of $\Phi(a_-, c; x)$ for equal c and n , corresponding to any two large negative real numbers $k_{-,1}$ and $k_{-,2}$, satisfy the relations: i) $\zeta_{k_{-,1},n}^{(c)} \approx 10^{-h} \zeta_{k_{-,2},n}^{(c)}$ (if $|k_{-,1}| = 10^h |k_{-,2}|$ and h is a positive or negative integer or zero); ii) $k_{-,1}/k_{-,2} \approx \zeta_{k_{-,2},n}^{(c)}/\zeta_{k_{-,1},n}^{(c)}$; iii) $\zeta_{k_{-,1},n}^{(c)}/k_{-,2} \approx \zeta_{k_{-,2},n}^{(c)}/k_{-,1}$; iv) $|k_{-,1}| \zeta_{k_{-,1},n}^{(c)} \approx |k_{-,2}| \zeta_{k_{-,2},n}^{(c)} \approx L(c, n)$; i) $|a_{-,1}|/|a_{-,2}| \approx \zeta_{k_{-,2},n}^{(c)}/\zeta_{k_{-,1},n}^{(c)}$; ii) $\zeta_{k_{-,1},n}^{(c)}/|a_{-,2}| \approx \zeta_{k_{-,2},n}^{(c)}/|a_{-,1}|$; iii) $|a_{-,1}| \zeta_{k_{-,1},n}^{(c)} \approx |a_{-,2}| \zeta_{k_{-,2},n}^{(c)} \approx L(c, n)$ ($|a_{-,1}|$ and $|a_{-,2}|$ are the moduli of first parameter a of Φ , conforming with $k_{-,1}$ and $k_{-,2}$). (In points iv) and vii) $L(c, n)$ is one and the same.) It is true, as well: $\zeta_{k_-,n}^{(c)}/|k_-| \approx L(c, n)/|k_-|^2$ and $\zeta_{k_-,n}^{(c)}/|a_-| \approx L(c, n)/|a_-|^2$. Putting hats “^” on all quantities or on the ones, relevant to $k_{-,1}(k_{-,2})$ only, the expressions remain true.

3.3. Numerical Proof

The proof of statements (12) and (13) of Theorem 1 for $c = \hat{c}$ — positive integers and arbitrary real numbers (positive and negative, save for $c = l$ ($\hat{c} = \hat{l}$)) and $n = \hat{n} = 1$, is illustrated in Tables 1 and 2. It follows from the coincidence of the underlined digits in the second and fifth columns, and of the ones, marked by bold face type in the third, fourth, sixth and seventh columns, resp. The number of identical digits is a function of $c = \hat{c}$, $n = \hat{n}$ and $k_- = \hat{k}_-$. It increases when $|k_-|(|\hat{k}_-|)$ grows. The truthfulness of relation (11) in this case is a consequence of that of Eq. (13), combined with Definitions (1) and (6). The validity of formula (11) for $c = \hat{c}$ and $n = \hat{n}$ provided $c = l$ ($\hat{c} = \hat{l}$) is substantiated, comparing the results in Table 6 [2] and Table 3 [4], which indicate in the vicinity of points $l = \hat{l}$, assuming that $\varepsilon = \hat{\varepsilon} \rightarrow 0$, it is true: $L(l - \varepsilon, n) = \hat{L}(\hat{l} - \hat{\varepsilon}, \hat{n})$, $L(l + \varepsilon, n + 1) = \hat{L}(\hat{l} + \hat{\varepsilon}, \hat{n} + 1)$, or $L(l, n) = \hat{L}(\hat{l}, \hat{n})$. Graphical and numerical outcomes for the $L(c, n) (\hat{L}(\hat{c}, \hat{n}))$ numbers could be found in Refs. [2] and [4]. In case $k_- = \hat{k}_- = -1.10^0, -1.10^1, -1.10^2, \dots, -1.10^{10}$ and $n = \hat{n} = 1$, $K_-(c, n, k_-) (\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-))$ yields $L(c, n) (\hat{L}(\hat{c}, \hat{n}))$ with an accuracy up to 0, 2, 4, 6, 7, 10, 12, 14, 16, 18, 20 (0, 2, 3, 6, 8, 10, 12, 14, 16, 18, 20) correct decimal places, if $c = \hat{c} = 1$. For $c = \hat{c} = 3$ and $c = \hat{c} = 5$ the number of the latter drops to 0, 1, 3, 4, 6, 9, 10, 13, 15, 16, 18 (0, 0, 2, 5, 7, 8, 11, 13, 14, 17, 19) and 0, 0, 2, 4, 5, 8, 9, 12, 13, 16, 18 (0, 0, 2, 4, 6, 8, 10, 11, 14, 15, 17), resp. The entire parts of $K_-(c, n, k_-) (\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-))$ and $L(c, n) (\hat{L}(\hat{c}, \hat{n}))$ do not differ for $k_- = \hat{k}_- \geq -1.10^0, k_- = \hat{k}_- \geq -1.10^1$ when $c = \hat{c} = 1$ and 3, resp. while if $c = \hat{c} = 5$, then $k_- \geq -1.10^1$ and $\hat{k}_- \geq -1.10^2$.

Table 1: Numbers $\zeta_{k_-,n}^{(c)}$, $K_-(c, n, k_-)$, $M_-(c, n, k_-)$, $\hat{\zeta}_{\hat{k}_-, \hat{n}}^{(\hat{c})}$, $\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-)$ and $\hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-)$ for $c = \hat{c} = 1$, $c = \hat{c} = 3$, $n = \hat{n} = 1$ and large negative $k_- = \hat{k}_-$.

| k_- | $\zeta_{k_-,n}^{(c)}$ | $K_-(c, n, k_-)$ | $M_-(c, n, k_-)$ | $\hat{\zeta}_{\hat{k}_-, \hat{n}}^{(\hat{c})}$ | $\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-)$ | $\hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-)$ |
|-----------|-----------------------------|----------------------|----------------------|--|--|--|
| $c = 1$ | | | | | | |
| -1.10^4 | (<u>−4</u>) 1.44579 64896 | 1.44579 64896 | 1.44579 64914 | (<u>−4</u>) 1.44579 64919 | 1.44579 64919 | 1.44572 42021 |
| -2.10^4 | (<u>−5</u>) 7.22898 24523 | 1.44579 64905 | 1.44579 64909 | (<u>−5</u>) 7.22898 24551 | 1.44579 64910 | 1.44576 03461 |
| -4.10^4 | (<u>−5</u>) 3.61449 12267 | 1.44579 64907 | 1.44579 64908 | (<u>−5</u>) 3.61449 12270 | 1.44579 64908 | 1.44577 84184 |
| -6.10^4 | (<u>−5</u>) 2.40966 08178 | 1.44579 64907 | 1.44579 64908 | (<u>−5</u>) 2.40966 08179 | 1.44579 64908 | 1.44578 44425 |
| -8.10^4 | (<u>−5</u>) 1.80724 56134 | 1.44579 64907 | 1.44579 64907 | (<u>−5</u>) 1.80724 56134 | 1.44579 64908 | 1.44578 74545 |
| -1.10^5 | (<u>−5</u>) 1.44579 64907 | 1.44579 64907 | 1.44579 64907 | (<u>−5</u>) 1.44579 64907 | 1.44579 64907 | 1.44578 92618 |
| $c = 3$ | | | | | | |
| -1.10^4 | (<u>−4</u>) 6.59365 40623 | 6.59365 40623 | 6.59365 41365 | (<u>−4</u>) 6.59365 41513 | 6.59365 41513 | 6.59266 51031 |
| -2.10^4 | (<u>−4</u>) 3.29682 70478 | 6.59365 40957 | 6.59365 41142 | (<u>−4</u>) 3.29682 70590 | 6.59365 41179 | 6.59315 95939 |
| -4.10^4 | (<u>−4</u>) 1.64841 35260 | 6.59365 41040 | 6.59365 41086 | (<u>−4</u>) 1.64841 35274 | 6.59365 41096 | 6.59340 68475 |
| -6.10^4 | (<u>−4</u>) 1.09894 23509 | 6.59365 41056 | 6.59365 41076 | (<u>−4</u>) 1.09894 23513 | 6.59365 41080 | 6.59348 92667 |
| -8.10^4 | (<u>−5</u>) 8.24206 76326 | 6.59365 41061 | 6.59365 41073 | (<u>−5</u>) 8.24206 76344 | 6.59365 41075 | 6.59353 04765 |
| -1.10^5 | (<u>−5</u>) 6.59365 41063 | 6.59365 41063 | 6.59365 41071 | (<u>−5</u>) 6.59365 41072 | 6.59365 41072 | 6.59355 52024 |

Table 2: Numbers $\zeta_{k-,n}^{(c)}$, $K_-(c, n, k_-)$, $M_-(c, n, k_-)$, $\hat{\zeta}_{\hat{k}_-, \hat{n}}^{(\hat{c})}$, $\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-)$ and $\hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-)$ for $c = \hat{c} = -4.5108794236$, $c = \hat{c} = 16.0127935846$, $n = \hat{n} = 1$ and large negative $k_- = \hat{k}_-$.

| k_- | $\zeta_{k-,n}^{(c)}$ | $K_-(c, n, k_-)$ | $M_-(c, n, k_-)$ | $\hat{\zeta}_{\hat{k}_-, \hat{n}}^{(\hat{c})}$ | $\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-)$ | $\hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-)$ |
|---------------------|----------------------|------------------------|------------------------|--|--|--|
| $c = -4.5108794236$ | | | | | | |
| -1.10^4 | (-3) 1.31617 44806 | 13.16174 48063 | 13.16174 51411 | (-3) 1.31617 45417 | 13.16174 54172 | 13.16471 39695 |
| -2.10^4 | (-4) 6.58087 25177 | 13.16174 50354 | 13.16174 51191 | (-4) 6.58087 25941 | 13.16174 51881 | 13.16322 94642 |
| -4.10^4 | (-4) 3.29043 62732 | 13.16174 50927 | 13.16174 51136 | (-4) 3.29043 62827 | 13.16174 51308 | 13.16248 72689 |
| -6.10^4 | (-4) 2.19362 41839 | 13.16174 51033 | 13.16174 51126 | (-4) 2.19362 41867 | 13.16174 51202 | 13.16223 98789 |
| -8.10^4 | (-4) 1.64521 81384 | 13.16174 51070 | 13.16174 51122 | (-4) 1.64521 81396 | 13.16174 51165 | 13.16211 61856 |
| -1.10^5 | (-4) 1.31617 45109 | 13.16174 51087 | 13.16174 51120 | (-4) 1.31617 45115 | 13.16174 51148 | 13.16204 19700 |
| $c = 16.0127935846$ | | | | | | |
| -1.10^4 | (-2) 1.00084 07454 | 100.08407 45422 | 100.08410 66203 | (-2) 1.00084 10995 | 100.08410 99513 | 100.00397 86416 |
| -2.10^4 | (-3) 5.00420 43910 | 100.08408 78206 | 100.08409 58401 | (-3) 5.00420 48336 | 100.08409 66729 | 100.04403 10234 |
| -4.10^4 | (-3) 2.50210 22785 | 100.08409 11402 | 100.08409 31451 | (-3) 2.50210 23338 | 100.08409 33533 | 100.06406 05292 |
| -6.10^4 | (-3) 1.66806 81959 | 100.08409 17550 | 100.08409 26460 | (-3) 1.66806 82123 | 100.08409 27385 | 100.07073 75226 |
| -8.10^4 | (-3) 1.25105 11496 | 100.08409 19701 | 100.08409 24713 | (-3) 1.25105 11565 | 100.08409 25234 | 100.07407 61114 |
| -1.10^5 | (-3) 1.00084 09207 | 100.08409 20697 | 100.08409 23905 | (-3) 1.00084 092 ⁴² | 100.08409 24238 | 100.07607 92942 |

4. APPLICATION IN THE THEORY OF WAVEGUIDES

The propagation of normal TE_{0n} (slow $\hat{T}E_{0\hat{n}}$) modes of phase constant β ($\hat{\beta}$) in the circular waveguide of radius r_0 , uniformly filled with lossless remanent ferrite of azimuthal magnetization, described by a Polder permeability tensor with nonzero diagonal components equal to unity and off-diagonal ones $j\alpha$ and $-j\alpha$, $\alpha = \gamma M_r / \omega$, (γ — gyromagnetic ratio, M_r — remanent magnetization, ω — angular frequency of the wave), and a scalar permittivity $\varepsilon = \varepsilon_0 \varepsilon_r$, is subject to the equations [1, 2, 4, 5]:

$$\Phi(a, c; x_0) = 0, \tag{15}$$

$$\Phi(\hat{a}, \hat{c}; \hat{x}_0) = 0. \tag{16}$$

In Eq. (15), $a = c/2 - jk$, $c = 3$, $x_0 = jz_0$, $k = \alpha\bar{\beta}/(2\bar{\beta}_2)$, $\bar{\beta}_2 = (1 - \alpha^2 - \bar{\beta}^2)^{1/2}$, $z_0 = 2\bar{\beta}_2\bar{r}_0$, (k , z_0 — real, $-\infty < k < +\infty$, $z_0 > 0$, $0 < |\alpha| < 1$). It is satisfied provided $\bar{\beta}_2 = \zeta_{k,n}^{(c)}/(2\bar{r}_0)$ that determines the eigenvalue spectrum of the fields examined. Here $\bar{\beta} = \beta/(\beta_0\sqrt{\varepsilon_r})$, $\bar{\beta}_2 = \beta_2/(\beta_0\sqrt{\varepsilon_r})$, and $\bar{r} = \beta_0 r_0 \sqrt{\varepsilon_r}$, ($\beta_2 = [\omega^2 \varepsilon_0 \mu_0 \varepsilon_r (1 - \alpha^2) - \beta^2]^{1/2}$ — radial wavenumber and $\beta_0 = \omega \sqrt{\varepsilon_0 \mu_0}$, — free space phase constant). Fig. 1 shows the domain in which the normal TE_{01} wave may be supported for positive ($\alpha_+ > 0$) ferrite magnetization. The relevant normalized (barred) $\bar{\beta}(\bar{r}_0)$ — phase characteristic with α as parameter, computed, following the procedure, described earlier [1], are presented in this picture by solid lines. The domain mentioned (portrayed by light green) is restricted by the side of lower frequencies through the $\alpha = 0$ solid curve, corresponding to a dielectric-loading and is unlimited from above. Fig. 2 indicates that the modes for negative ($\alpha_- < 0$) magnetization are transmitted in a bilaterally bounded area. The $\bar{\beta}(\bar{r}_0)$ — dashed phase characteristics end at an En_{1-} — envelope dotted line of equation $\bar{r}_{0en-} = L(c, n)/[|\alpha_{en-}|(1 - \alpha_{en-}^2)^{1/2}]$, $\bar{\beta}_{en-} = (1 - \alpha_{en-}^2)^{1/2}$, (α_{en-} is a parameter).

In Eq. (16), $\hat{a} = \hat{c}/2 + \hat{k}$, $\hat{c} = 3$, $\hat{x}_0 = 2\hat{\beta}_2\hat{r}_0$, $\hat{k} = \hat{\alpha}\hat{\beta}/(2\hat{\beta}_2)$, $\hat{\beta}_2 = [\hat{\beta}^2 - (1 - \hat{\alpha}^2)]^{1/2}$, $\hat{\beta} = \hat{\beta}/(\beta_0\sqrt{\varepsilon_r})$, $\hat{\beta}_2 = \hat{\beta}_2/(\beta_0\sqrt{\varepsilon_r})$ and $\hat{r}_0 = \beta_0\hat{r}_0\sqrt{\varepsilon_r}$. All quantities coherent with the slow waves are real and are distinguished by hats “ ^ ”. They propagate, if $\hat{\beta}_2 = \hat{\zeta}_{\hat{k}_-, \hat{n}}^{(\hat{c})}/(2\hat{r}_0)$, assuming $\hat{k}_- < -1.5$, i.e., for $\hat{\alpha}_- < 0$ only [4]. The $\hat{\beta}^{(1)}(\hat{r}_0^{(1)})$ — dashed phase curves of the $\hat{T}E_{0\hat{n}}^{(1)}$ mode (the one of the two possible waves), sustained for $|\hat{\alpha}_-^{(1)}| < 1$, are depicted in Fig. 2. They possess a lower limit — the $\hat{E}n_{1-}$ — envelope of equation: $\hat{r}_{0en-}^{(1)} = \hat{L}(\hat{c}, \hat{n})/[\{|\hat{\alpha}_{en-}^{(1)}|[1 - (\hat{\alpha}_{en-}^{(1)})^2]^{1/2}\}]$, $\hat{\beta}_{en-}^{(1)} = [1 - (\hat{\alpha}_{en-}^{(1)})^2]^{1/2}$, ($\hat{\alpha}_{en-}^{(1)}$ is a parameter). The region in which the slow wave referred to might get excited is narrowed from below and termless beyond the frequency range above. (Throughout the paper the subscripts

“+” (“−”) and “ $en-$ ” label the quantities, coming up with the positive (negative) magnetization, resp. and to the envelopes and the superscript (1) — corresponding to the first slow $\hat{T}E_{0\hat{n}}^{(1)}$ mode.) Since $c = \hat{c} = 3$, in view of Theorem 1 (Eq. (11)) it follows that if $\alpha_{en-} \equiv \hat{\alpha}_{en-}^{(1)}$ and $n = \hat{n}$, it follows that $\bar{r}_{0en-} \equiv \bar{r}_{0en-}^{(1)}$, i.e., the En_{1-} — and $\hat{E}n_{1-}$ — lines coincide. Thus, in case of negative magnetization the latter separates the area of propagation in two sub-domains: a left one in which propagation of a normal mode is observed and a right-hand one in which a slow wave is transmitted.

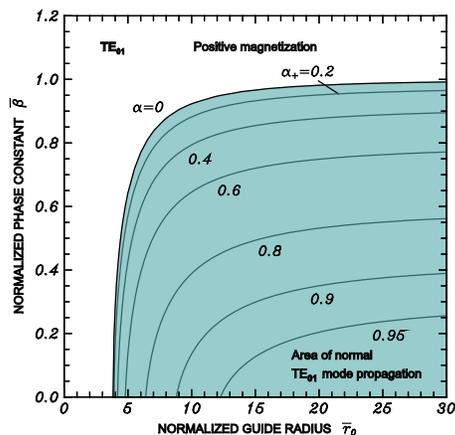


Figure 1: Area of normal TE_{01} mode propagation in case of positive magnetization of the azimuthally magnetized circular ferrite waveguide.

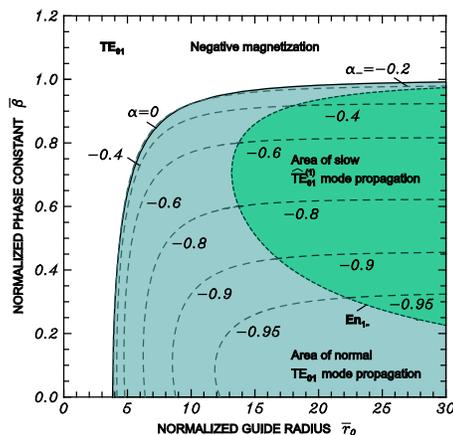


Figure 2: Areas of normal TE_{01} and slow $\hat{T}E_{01}^{(1)}$ mode propagation in case of negative magnetization of the azimuthally magnetized circular ferrite waveguide.

5. CONCLUSION

It is demonstrated numerically that the positive real numbers $L(c, n)$ and $\hat{L}(\hat{c}, \hat{n})$ (the limits of specially constructed sequences of numbers with terms, involving the zeros of certain complex, resp. real Kummer functions) coincide, if their parameters c and \hat{c} (arbitrary real), and n and \hat{n} (restricted natural numbers) are the same. It is shown that the peculiarities of the rotationally symmetric TE modes transmission in the circular waveguide, entirely filled with azimuthally magnetized ferrite are directly connected with this mathematical outcome. From a physical point of view its validity in case $c = \hat{c} = 3$ means existence of a curve in the phase portrait of configuration in case of negative magnetization of the load at which the propagating mode is transformed from normal to slow.

ACKNOWLEDGMENT

We express our gratitude to our mother Trifonka Romanova Popnikolova and to our late father Nikola Georgiev Popnikolov for their self-denial and for their tremendous efforts to support all our undertakings.

REFERENCES

- Georgiev, G. N. and M. N. Georgieva-Grosse, “A new property of the complex Kummer function and its application to waveguide propagation,” *IEEE Antennas Wireless Propagat. Lett.*, Vol. 2, 306–309, December 2003.
- Georgiev, G. N. and M. N. Georgieva-Grosse, “The $L(c, n)$ numbers and their application in the theory of waveguides,” *Proc. Int. Conf. Days Diffr. 2008 DD’08*, 44–57, St. Petersburg, Russia, June 3–6, 2008.
- Tricomi, F. G., *Funzioni Ipergeometriche Confluenti*, Edizioni Cremonese, Rome, Italy, 1954.
- Georgiev, G. N. and M. N. Georgieva-Grosse, “Theory of the $\hat{L}(\hat{c}, \hat{n})$ numbers and its application to the slow wave propagation in the circular ferrite waveguide,” *PIERS Proceedings*, 976–980, Cambridge, USA, July 5–8, 2010.
- Georgieva-Grosse, M. N. and G. N. Georgiev, “Advanced studies of the differential phase shift in the azimuthally magnetized circular ferrite waveguide,” *PIERS Proceedings*, 841–845, Cambridge, USA, July 5–8, 2010.

Electromagnetic Sources and Observers in Motion VI — New Motional Optics

S. E. Wright

Moor Lane Laboratory, ECASS Technologies Ltd., Kirkburton, Huddersfield, HD8 0QS, UK

Abstract— Paper V is the first of two papers, given at this symposium, where the arguments for revising Einstein’s relativity are discussed. In this paper VI, the new motional implications resulting from the theory revision are considered. Einstein’s Special Relativity (SR) has a non causal ether-less aspect, based on no propagation medium (ether), predicting motional properties that have not been measured. It also has a medium based aspect that does predict measured properties. However this medium based aspect appear to be incomplete, it does not distinguish between source and observer motion, whose motional properties are clearly measurable. The revised theory extends the Lorentz transform to include these motions, which now completes the radiation process, removing ambiguities and paradoxes present in Einstein’s SR. Worked examples demonstrate that the motional effects are generated by direct contact of individual source and observer motion with respect to the propagation medium, in direct opposition to Einstein’s concept of relativity. The theory illustrates the use of an instantaneous source and observer event time transform and compares integrated times for arbitrary source and observer flight paths through the universe. The revised theory predicts that material time travel is fallacious (non causal), but space craft can travel at effective speeds greater than that of light, with actual speeds less than the speed of light in the medium.

1. INTRODUCTION

For motion, away from large masses, the observed time (rate of happening) t_o , in terms of the source time t_s , is determined from the EM Motional Analysis (EMMA) in reference [1]. The segment of integrated or accumulative observer event time τ_o , in terms of the integrated source event time τ_s , for arbitrary source and observer flight paths through the universe, is given by:

$$\tau_o = \int [\epsilon_s \alpha_s^{-1} \epsilon_o^{-1} \alpha_o]_{t_p} d\tau_s, \quad t_o = \tau_o + R_o/c, \quad R_o = \alpha_o R_p \tag{1}$$

t_p outside the square bracket indicates that quantities inside the bracket are to be evaluated in the stationary propagation medium time. When the source and observer are not in contact initially, then there will be an initial propagation time delay (retarded time), which occurs before the observations are seen. R_o and R_p are initial propagation distances in the moving and stationary frames, respectively, as illustrated in Figure 1. This figure has been made possible through the common reference of the medium. Clear distinction is made between source and observer motion, relative to this medium.

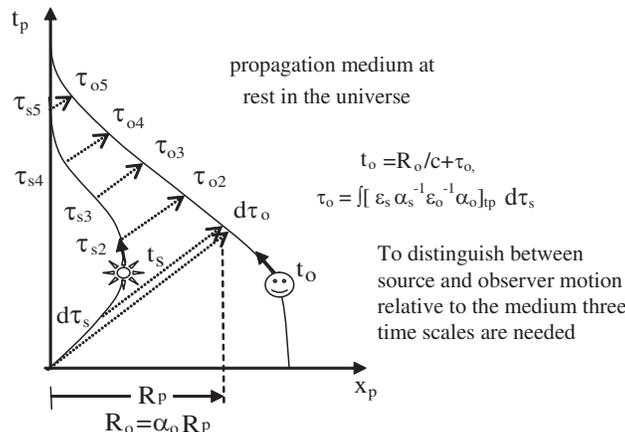


Figure 1: Space-time diagram. Relationship between source time t_s and observer time t_o for motion relative to a stationary propagation medium of time t_p .

The vertical and horizontal axes represent absolute time t_p and space x_p , in the medium respectively. Corresponding times, between the source emission event time τ_s and observer reception

event time τ_o are connected by propagation light paths. These paths represent light moving forward with time, from source to observer, going from left to right in the direction of the arrows, whose gradient is $1/c$ (45° if x_p is in light years). Figure 1 is valid generally for arbitrary (constant and accelerative) motion. Whereas, Einstein's inertial frame cannot distinguish between source and observer motion, or between constantly moving and stationary systems. Therefore, Figure 1 cannot be constructed, according to Einstein's relativity.

For intricate source and observer flight paths, the observed event time is a complex function of the source time, speed and direction at the source emission point, the observer speed and direction at the observer reception position and the propagation distance between them. Interchanging the source and the observer, will give different observations. The light paths then go from right to left and have a negative gradient, giving a totally different observed time history for the same flight paths. The motional history of the source and observer is therefore indelibly imprinted in the observed time history. In these particular flight paths, the τ_s and τ_o time curves start at the same time but at a different place. They end at a different time, but at the same place. The directional Doppler time change, affected by $\varepsilon_s \varepsilon_o^{-1}$, in Equation (1), is caused through the wave dynamics including the generation, transmission and reception of the EM waves. The omni-directional Lorentzian time change, determined by $\alpha_s^{-1} \alpha_o$, is caused through the kinematic contraction (both time and space) in the moving frame, through motion relative to the stationary propagation medium. The figures validity and authority is based on absolute source and observer motion with respect to the propagating medium. From Equation (1), the instantaneous motional event time transform becomes:

$$\mathcal{K}_t = (\tau_o/\tau_s) = (\varepsilon_s \alpha_s^{-1})(\varepsilon_o^{-1} \alpha_o) = \mathcal{K}_{\varepsilon_s, \alpha_s} \mathcal{K}_{\varepsilon_o, \alpha_o} \quad (2)$$

The individual source and observer event time transforms, at their respective emission and reception times and positions in the medium, are labelled $\mathcal{K}_{\varepsilon_s, \alpha_s} = \varepsilon_s \alpha_s^{-1}$, and $\mathcal{K}_{\varepsilon_o, \alpha_o} = \varepsilon_o^{-1} \alpha_o$. The kinematic operators are $\alpha_s = (1 - M_s^2)^{1/2}$, $\alpha_o = (1 - M_o^2)^{1/2}$, and the dynamic operators $\varepsilon_s = 1 - M_s \cos \sigma_s$, $\varepsilon_o = 1 + M_o \cos \sigma_o$, where $M_s = s/c$ and $M_o = o/c$. The directional angles σ_s and σ_o are made with the directions of motion and a straight line joining the source emission and observer reception positions. The total observer event time τ_o , at the end of any given flight paths, in terms of the integrated source event time τ_s , depends on the initial propagation distance R_o and on how \mathcal{K}_t (the individual motional operators ε_o^{-1} , α_o , ε_s , and α_s^{-1}) changes with time during the source and observer journeys through space. In this medium based theory, time changes instantly at the source emission and observer reception times in the medium, according to their speeds at that instant, whether it is constant or arbitrary motion. At low speeds $\alpha_s \approx \alpha_o \approx 1$, $\alpha_s^{-1} \alpha_o \approx 1$, and inline motion $\sigma_o = \sigma_s = 0$, Equation (2) gives: $\tau_o/\tau_s = \varepsilon_s \varepsilon_o^{-1} = (1 - M_s \cos \sigma_s)/(1 + M_o \cos \sigma_o) \approx 1 - (M_s \cos \sigma_s - M_o \cos \sigma_o) \approx 1 - (M_s - M_o)$. The effect reduces to a relative velocity between the source and observer, a result relativists would not argue with.

2. EVENT TIMES

Considerable speculation has arisen through time travel in recent years. One should first ask what is meant by time? Time is considered here to be a means of measuring and comparing the order and speed of events. For time to be causal (time predictable), it must be a solution of the wave equation based on a reference medium. Stationary source and observer clock times are defined to move forward at a constant rate, the same as that in the propagation medium. However, the source event time has to be transmitted from the source to the observer, via the propagation medium, before it can be realised at the observer, in accordance with cause and effect. Thus, the observed event time in terms of the source time, given by Equation (2), depends on Doppler's dynamic and Lorentz's kinematic source and observer motion with respect to the medium. This can slow down or speed up the observed source events. Time is therefore an observed variant effect, dependent on source and observer motion with respect to the medium. Without a medium, there is no physical mechanism, no wave propagation, no observations, no direction of time and no cause and effect, allowing imaginary events to occur before the cause.

The instantaneous EM motional event time transform \mathcal{K}_t , given by Equation (2), defines the direction of time. It has positive values representing events going forward and negative values going backwards in time, as illustrated in Figure 2. As a simple example, consider a moving observer, with the source at rest, $\alpha_s = \varepsilon_s = 1$, M is now positive for a receding observer. The time transform $\tau_o/\tau_s = \mathcal{K}_{\varepsilon_o, \alpha_o} = \alpha_o \varepsilon_o^{-1}$, is given by the solid curve in Figure 2. It shows a range of observer speeds, including greater than the speed of light, $|M| > 1$. It can be seen that the kinematic time

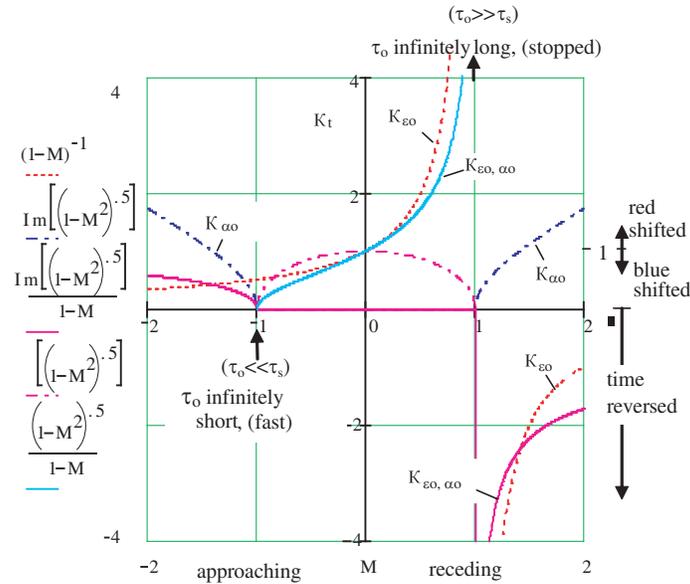


Figure 2: Instantaneous event time transforms for observer speed greater than that of light $|M_o| > 1$ relative to a stationary medium and source $M_s = 0$, for inline motion ($\sigma_o = \sigma_s = 0$).

transform $\mathcal{K}_{\alpha o}$ (chain dotted) is positive and symmetrical about $M = 0$. It has singularities (zeros) at $M = \pm 1$, indicating infinitely short (fast) observed source time. For $|M| > 1$, $\mathcal{K}_{\alpha o}$ is again finite, showing observer time expanding (slowing), but it is now curiously imaginary, which is interpreted as an oblique time running at 90° to our own. The kinematic time transform has no possibility of becoming negative, i.e., enabling time to go backwards.

On the other hand, the dynamic time transform $\mathcal{K}_{\epsilon o}$ is not symmetrical about $M = 0$. It has an infinity for the receding observer at $M = +1$, indicating infinitely long observed source expansion (stopping of time). So for $M < 1$, observed time slows and at $M = 1$ time stops. For $M > 1$ the time transform becomes negative indicating the observer time can now run backwards, leaving the present and going back to the past. In the approaching observer case, M negative, there is no negation of the motional time transform and therefore no going back in time. The observed source time will speed up, become infinitely fast at $M = -1$ and then become finite again for $M < -1$. Thus, only for a receding observer, there is a causal potential to go back in time and observe the past, but not to participate or interfere with it. One cannot go back in time and terminate his mother's life, when she was a child, as this would result in him or her not being born, which is clearly absurd.

Unfortunately, at the speed of light, mass appears to become infinite making this speed questionable. Therefore, any prediction for speeds greater than the speed of light must be speculative. However, if this critical speed could be avoided or bypassed, then the properties again become well behaved (finite and causal). Perhaps a traveller could go back to the past, see a broken cup reassemble itself, or more macabre, see relatives resurrecting from their graves. Perhaps even for $M < 1$, a cusp in a non-linear (curved) flight path could also give glimpses of the past. Whether or not super-'lightic' speeds are possible, one can at least predict their properties, according to this revised theory. Visual time travel-visually going back to the past is certainly causal. Material (substance transmission) time travel into the future or past is not possible, it is not in agreement with causality. To actually participate, change the source events as they happen, the observer must come in physical contact (interact) with the source, at the source event time, it cannot affect events, separated by distance or time. By definition this rules out any kind of active participatory time travel. Thus, causally it is possible to observe past events, but not to participate or interfere with them, because they have already happened. It is not possible (non causal) to observe or participate in source events in the future, because they have not yet occurred.

Through motion or visiting a large mass, a traveller can slow his/her time down compared with a stationary system time and return to its future, actually participating in its present. This is quite feasible, but it is not reversible time travel. The process is one-way, one cannot reverse the flow of time in this situation. All that can be achieved is to alter the rate of ageing (time slowing) of one system compared to another. This also could be true within black holes and connecting worm holes, if they are accessible. Space will shrink to practically nothing and time will almost stop, allowing

great distances in the stationary frame to be achieved in little time in the moving frame. This assumes worm holes can be exited at the other end of the journey to realise any benefit. However, this is still not time travel, it is just getting there quicker and ageing less.

3. ACCUMULATIVE TIMES

Having a propagation medium allows both the source and observer to be plotted on the same reference diagram, distinguishing between their motions and observations, which the ether-less SR cannot do. Figure 1 compares the total accumulative (integrated) source and observer event times, for arbitrary motion, compared with the time in the stationary propagation medium. In the case of accumulative discrete source and observer event time segments $\tau_{s1}, \tau_{s2}, \tau_{s3}$ etc. and $\tau_{o1}, \tau_{o2}, \tau_{o3}$ etc., respectively, Equation (1) becomes:

$$t_o = R_o/c + \tau_o = \alpha_o R_p/c + \left[\frac{\epsilon_s \alpha_o}{\alpha_s \epsilon_o} \Delta\tau_s \right]_{\tau_{s1}, \tau_{o1}}^{\tau_{s2}, \tau_{o2}} + \left[\frac{\epsilon_s \alpha_o}{\alpha_s \epsilon_o} \Delta\tau_s \right]_{\tau_{s2}, \tau_{o2}}^{\tau_{s3}, \tau_{o3}} + \left[\frac{\epsilon_s \alpha_o}{\alpha_s \epsilon_o} \Delta\tau_s \right]_{\tau_{s3}, \tau_{o3}}^{\tau_{s4}, \tau_{o4}} + \left[\frac{\epsilon_s \alpha_o}{\alpha_s \epsilon_o} \Delta\tau_s \right]_{\tau_{s4}, \tau_{o4}}^{\tau_{s5}, \tau_{o5}} \text{ etc.} \quad (3)$$

As a simple example to illustrate the use of the theory, Figure 3 is a space-time diagram showing a moving source of speed ‘s’ moving in a straight line with respect to the stationary medium and observer. The y and x axes are in years and light years (LY), respectively. The figure illustrates a source making an excursion of $Ex = 3$ LY, 6 LY there and back, at a motional Mach number $M = s/c = 0.6$. The motion is direct ($\sigma = 0$) from a stationary observer at the origin in the stationary propagating medium. The observer is moving in time but not in space (vertical line in diagram). The light path communicating the turning point of the source to the observer, indicated by the dotted line, goes from right to left.

The gradient of the source flight path in light years is $ct_p/x_p = c(t_p/x_p) = c/s = 1/M = 1.66$ or 59° and the gradient of the light path $= c/c = 1$ or 45° , also $\alpha_s = [1 - (M)^2]^{1/2} = 0.8$. The source flight path relative to the medium, $x_p = 3$ LY, corresponds to a stationary observer of time t_o in the propagating medium of time t_p where $t_o = t_p = x_p/Mc = 3 \times 1.66 = 5$ years. Now in terms of the source event time τ_s (hypotenuse) in the moving frame, $\tau_s = \alpha_s \tau_p = 0.8 \times 5 = 4$ years. Or more precisely $\tau_s = (-\tau_p^2 + (x_p/c)^2)^{1/2} = (-5^2 + 3^2)^{1/2} = (-25 + 9)^{1/2} = (-16)^{1/2} = j4$ years, the quantity is in fact imaginary. Note the imaginary triangle ratios 3, 4 and 5. In this kind of geometry, Pythagoras’s Theory does not work. The longest side in the right angle triangle is not the hypotenuse, it’s the opposite. From Equation (3), the event time changes τ_p in the propagation medium, for a source time change of $\tau_s = 8$ years (4 years forward, 4 years reverse), becomes for a

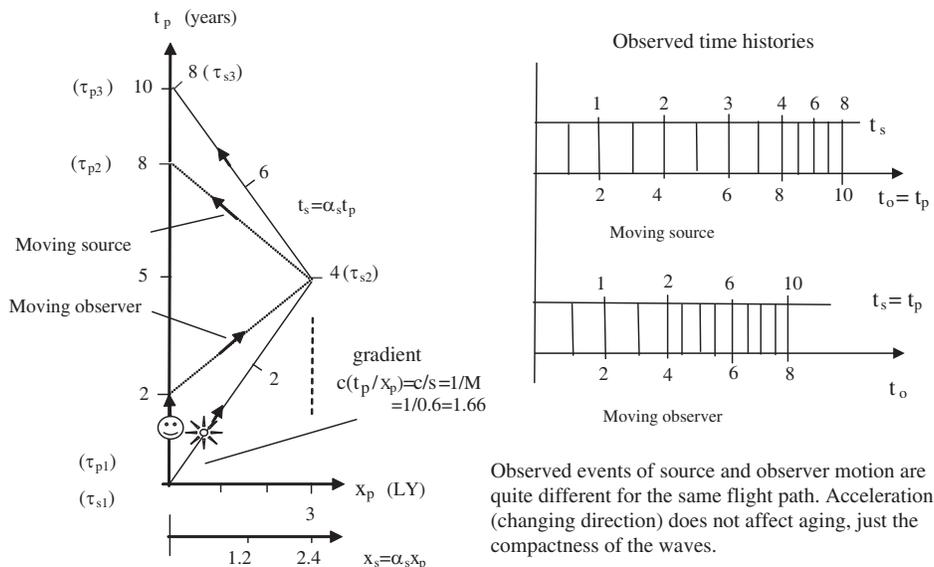


Figure 3: Moving source, stationary observer and moving observer, stationary source. $M = 0.6$, on a 6 LY excursion relative to a stationary propagating medium.

stationary observer $o = 0$, $\varepsilon_o = 1$, $\alpha_o = 1$, $\tau_o = \tau$, and zero separation distance $R_p = 0$.

$$\begin{aligned} \tau_p &= \left[\left(\frac{\varepsilon_s}{\alpha_s} \right)^+ \Delta\tau_s \right]_{\tau_{s1}}^{\tau_{s2}} + \left[\left(\frac{\varepsilon_s}{\alpha_s} \right)^- \Delta\tau_s \right]_{\tau_{s2}}^{\tau_{s3}} = \left[\frac{1+0.6}{0.8} \right] 4 + \left[\frac{1-0.6}{0.8} \right] 4 \\ &= (2) \times (4) + (0.5) \times (4) = 8 + 2 = 10 \end{aligned} \quad (4)$$

Thus $\tau_s = 4$ and 8 years on the right side of Figure 3 become, from Equation (4), $\tau_p = 8, 10$ years on the left-hand side of the figure (4 years expanded into 8 and then 4 years compressed into 2). These times are compared in the adjacent diagram. The corresponding stationary frame ($\tau_p = 10$ years, $x_p = 6$ LY) becomes in the moving frame ($\tau_s = \alpha_s \tau_p = 0.8 \times 10 = 8$ years, $x_s = \alpha_s x_p = 0.8 \times 6 = 4.8$ LY), i.e., the same distance in the real world is covered in less time in the moving frame, or more distance for the same time, i.e., the astronauts get there quicker in the moving frame. Across frames, the hybrid Mach number, from Reference [1], for a Mach number of $M = 0.6$, becomes $M^* = x_p/\tau_s c = x_p/\tau_p \alpha_s c = M/\alpha_s = 0.6/0.8 = 0.75$. The shrinking of distance or space in the moving frame is of little consequence in this case. However, astronauts should rotate regularly to avoid becoming flat in one direction. It could provide a painless slimming cure.

It is therefore legitimate to measure distance x_p in the familiar stationary frame and the slower time τ_s in the moving astronaut's frame. This allows the astronaut to travel vast distances in the familiar world, in his/hers slower time in the moving frame. For example, if $\alpha = 0.01$, $M \approx 0.999$, $M^* = M/\alpha \approx 100M$, i.e., the astronaut travels at effectively 100 times the speed of light, and yet retains his sub-'lightic' speed in the medium.

Inter-changing the source and observer, i.e., the observer moves and the source is stationary, the light paths now run from left to right, giving a different set of observations, as illustrated on the right hand figure. The moving source gets younger relative to the stationary observer. Whereas, the stationary source now gets older, viewed from the moving observer. This is in contrast to Einstein's ether-less SR, which can only deal with relative motion between systems, not individual motions relative to the medium. According to Einstein's concept of relativity, inter-changing the source and observer, for the same flight paths should make no difference to the observations, which is clearly not the case. Einstein claimed that either system can appear younger, observed by the other, which is clearly absurd.

This example resolves the famous travelling twin paradox. Arbitrarily, without a medium to reference motion, the relativists decide the twin who turns round to come home (accelerates) is the one that ages less. However, there is nothing unique regarding acceleration. Aging depends on the instantaneous velocity integrated over time i.e., the velocity profile over the flight path. Turning round does not affect the aging process, it affects the compacting of the wave only. It is now abundantly clear from Figure 3 who goes travelling and who sees what, without paradox. More complex source and observer flight paths are considered in reference [2]. Without a propagation medium, these flight path constructions are not possible.

4. CONCLUSION

All optics in the revised theory are based on sources and observers moving directly with respect to the propagation medium, which Einstein's concept of relativity cannot recognise. According to the revised theory, it is not possible (non causal) to travel materially into the future or past. It is possible (causally) only to travel visually to the past, but not to interfere or change it. It is possible to slow one's time down through material transport, by physically moving at a high speed or visiting a massive body. However, again this is not reversible time travel, it is just relatively controlling the rate of ageing. However, it is possible to travel at effective speeds faster than light whilst maintaining sub-'lightic' speeds in the propagation medium.

REFERENCES

1. 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.
2. Wright, S. E., *Problems with Einstein's Relativity*, Trafford Publishing, 2010.

Electromagnetic Sources and Observers in Motion V — A Revised Theory of Relativity

S. E. Wright¹

Moor Lane Laboratory, ECASS Technologies Ltd., Kirkburton, Huddersfield, HD8 0QS, UK

Abstract— This theory is in accord with those aspects of Einstein’s Special Relativity (SR) that are medium based and have been measured. However, it appears that the ether-less aspects of SR are non causal and cannot be measured. The revised theory re-establishes the propagation medium (ether) and predicts measured differences between source and observer motion, which Einstein’s ether-less SR cannot do. The lack of medium, and its associated ether-less relativity, has not been measured or verified theoretically. Researchers claiming the verification have usually verified aspects of Lorentz’s medium based transform (LT), not the medium’s absence. Although Einstein denied the medium’s presence, his 1905 motional electrodynamics is in fact based on the medium, in agreement with the LT. However, Einstein’s SR appears to be incomplete; it does not identify or predict the complete radiation process of sources and observers in motion. The medium based LT is extended, to include these mechanisms, making classical and modern motional optics compatible again. Maxwell’s equations and Einstein’s inertial frames are, as usual, event invariant, but the light propagation time is variant, which was not considered in Einstein’s ether-less SR. The Michelson and Morley experiment is invariant in both respects, inferring the medium can only move with the system. Reinstating the medium, and distinguishing between source and observer motion, now completes the radiation process absent in Einstein’s SR. This suggests, with the other conclusions in the papers mention below an electrical theory of everything. This paper is the fifth in a series of papers given at PIERS, this last year, two in Xi’an China and two in Cambridge USA. A companion paper VI is also given at this symposium, it considers the application of the new motional optics predicted by the revised theory.

1. THREE ASPECTS OF RELATIVITY

Einstein’s Special Relativity (SR) [1] has puzzled scientists over the last 100 years. The root of the confusion seems to arise through Einstein believing there was no propagation medium (ether), but then used a medium in his analysis. Einstein’s SR predicts properties that have been observed, resulting for example in Einstein’s famous $E = mc^2$. There are properties that have been measured, such as distinction between source and observer motion, which SR cannot predict. There are properties that Einstein’s SR predicts that have not been measured. This dilemma appears to have created three aspects of SR:

(A). Unobserved ether-less aspect. Einstein claimed that his SR is based on no propagation medium leading to his concept of relativity. Einstein appeared not to realize that the wave equation cannot be solved without a reference (medium). Without a solution there can be no predicted time, or observed sequence of events, defining causality. To simulate the medium’s removal, Minkowski’s [2], oblique transform axes were used. This implied simultaneity, no absolute time and space, no causality, no observerbility, no distinction between source and observer motion or between stationary and constantly moving systems, and no direction of time, where the future can be observed before it occurs. None of these ether-less, non causal, oblique axes claims appear to have been measured or shown to have any scientific foundation.

(B). Incomplete medium based aspect. Amazingly, in spite of his rejecting the medium, Einstein’s analysis (motional electrodynamics), actually solves the wave equation based on a reference medium. The use of the propagation medium appears to render Einstein’s ether-less aspect of SR as redundant, and confirms the ether-less claims made above as erroneous. Einstein’s SR therefore appears to be medium based, not ether-less as he claimed, satisfying the rectangular axes, medium based Lorentz Transform (LT) [3], not Minkowski’s ether-less, oblique axes transform [2]. The medium based SR predicts causal motional properties, including Lorentz’s time and space contraction, which Einstein never acknowledged. Even so, Einstein’s medium based analysis is incomplete.

(C). Complete medium based aspect. To describe the complete radiation, propagation and reception processes, to distinguish between stationary and constantly moving systems, and to

¹Professor Wright contact details are: selwyn.wright@nrlworld.com.

distinguish between observed differences between source and observer motion, the LT has to be extended to include source and observer motion with respect to the medium. This complete theory includes the rectangular axes medium based predictions of LT. However, it predicts additional new observed properties discussed in a companion Paper VI at this conference [4], which now appears to account for all known observations.

2. SUPPORT DETAILS

Newton [5], based his mechanics on the assumption that time and space can be judged with respect to an absolute reference (propagation medium or ether) at rest in the universe. Maxwell [6] developed his EM field equations based on a propagation medium. Also, Lorentz [3] based his transform (LT) on the medium's existence. Whereas Einstein [1], in his Special Relativity (SR), decided that there was no need for a medium to transmit electromagnetic (EM) waves (light). The Michelson and Morley experiment (MMX) [7] initially appeared to support Einstein's ether-less claim. However, considerable data, some of which is documented in reference [8], establishes the medium's existence. The medium's rejection is discussed in reference [9]. Here the rejection causes fundamental problems, including the inability to transmit light, no distinction is made between source and observer motion and there is no causal prediction of time. A revised medium based theory, described in reference [10], completes the radiation process. *This new theory is unlikely to be in error because:* (i) It's simply an extension of the well established LT. (ii) It's in accord with those aspects of Einstein's SR that are based on the LT. (iii) It recognises the necessity of the existence of the propagation medium to make causal predictions. (iv) It removes those aspects of SR that are not supported by the LT (ether-less, irrational and non causal aspects). (v) It extends the LT to distinguish between measured differences of sources and observers in motion with respect to the propagation medium.

Galileo [11], was the first to formally suspect, that there is no difference in events (what's seen, observed physics) in a constantly moving fixed source-observer system, compared to a stationary one. *However, this is not true when one takes into account the light propagation time.* Although within the moving system, Galileo was right, the observed events are invariant (source and observer Doppler effects [12] cancel exactly), the propagation speed and time do change. This is the situation for all classical waves, such as sound in air, vibrations in structures and waves on water. They all require a medium and are solutions of the wave equation, they are therefore causal (time predictable). Optics (modern physics) are basically no different, they also obey the same physical rules, except there is an additional Lorentzian effect. Here time and space contract through motion relative to the medium. This is by exactly the same ratio, now making the speed of light invariant in the electromagnetic case. The invariant physics (Galileo) and the speed of light (Lorentz), embodied in Einstein's two postulates, maintain observed events in Einstein's inertial frame and Maxwell's equations, as invariant, the same as classical systems. But the individual source and observer Doppler effects become identical and the propagation time perpendicular to the direction of motion becomes invariant.

However, contrary to Einstein's claim, the propagation time in the direction of motion remains asymmetric and variant, exactly the same as classical wave systems. This asymmetric propagation time variation is measurable. It was first verified by Sagnac [13], is confirmed daily in satellite communication systems, discussed in [8]. These rational EM wave predictions are part of the revised causal theory. In contrast, Einstein believed that systems were fundamentally different from classical ones — they did not require a medium. Without complete information, and in spite of Newton, Lorentz and Maxwell's support for the medium, *Einstein took a leap of faith to assume there was no medium.* It appears that the medium does in fact exist and that the ether-less relativity concept has not been observed. Also, that the measured optics of sources and observers in motion correspond to motion relative to the medium the same as all other wave propagating systems.

The new medium based theory extends the LT to describe (for the first time) both source and observer motion. It distinguishes between moving source (satellite) motion angles, GPS [14], and moving observer (telescope) motion angles, Bradley [15]. It also predicts asymmetric propagation time delays in satellite communications, Saburi et al. [26], which SR cannot predict. The theory removes the ether-less aspect of SR, such as simultaneity (equal propagation times upstream and down), and reciprocity (exchanging sources and observers make no difference optically), which have not been measured. The new theory supports Fresnel's [16] light convection, transverse Doppler, changes in time, space, mass, momentum and Einstein's energy equation. It supports relativistic

quantum mechanics, quantum electrodynamics, chromodynamics and the Standard Model in particle physics. The theory is also in accord with general relativity; here the medium inclusion leads to Schwarzschild [17], metric and event horizon. Assuming the medium to be homogeneous and isotropic, the Robertson-Walker metric in cosmology is obtained. There is nothing in any of these theories to doubt the existence of the medium. In fact none of the motional properties can be accounted for causally, without a medium. *This new theory appears to restore a rational universe, common sense and predicts new (measured) motional optics, in addition to those predicted by the LT.*

An extensive study into EM source and observer motion has been documented in the PIER papers [8–10]. In summary, the main points from this investigation are:

- a) Indisputable data shows that the propagation medium for the transmission of light exists.
- b) Considerable evidence indicates that the medium moves with gravitational bodies (Earth and Sun).
- c) Einstein's ether-less claim is a simulation created through Minkowski's oblique transform axes.
- d) These ether-less, oblique axes aspects of SR have not been measured, they have no physical reality.
- e) A new medium based, rectangular axes, source and observer motion theory has been developed.

It is the ether-less aspect of SR that appears to make relativity incomprehensible. Its properties are non causal (irrational, non predictable) and non observable (cannot be observed or measured). The new medium based theory completes the analysis for systems in motion generally, with respect to the medium, restoring the continuity between classical and modern physics. There appears to be nothing mysterious about light propagation, it has a medium and satisfies the wave equation exactly like all other wave forms. Because of the additional Lorentz's time and space contraction, the EM wave equation includes a few additional features compared with the classical theory, otherwise it is causal and of the same form. Einstein's SR was a good attempt to fit existing data into his theory at that time. *However, Einstein's conclusion that there is no medium, when clearly data shows there is one, and the fact that he used a medium in his own analysis, is a major inconsistency.* The first rational step to make sense and progress in understanding EM sources and observers in motion is to reject the irrational notion that there is no propagation medium.

3. MEDIUM BASED UNIVERSE

The propagation medium, for the transmission of EM waves (light), was passionately supported by Lorentz and more formally by Poincaré [18]. These pioneers predicted that time and space contracted through motion relative to the medium. However, the medium upon which these theories are based, gradually disappeared from EM theory since Einstein declared its obsolescence. *The apparent lack of a medium meant that the medium based classical physics could no longer describe observations of EM systems in motion* (modern physics), providing a mismatch between the two theories. This apparent difference was suggested through the Michelson and Morley experiment (MMX). Here light propagation, inexplicably at first, was found to be symmetrical in and against the direction of the Earth's motion through the universe. The propagation time should have been asymmetrical, for motion relative to a 'stationary' medium, according to classical physics.

In an attempt to explain the propagation symmetry and the invariant speed of light, Einstein introduced his ether-less aspects of SR. Here Einstein made a series of remarkable assumptions, without any definitive proof or observation. He assumed that there was nothing in space to propagate the light; the waves and their energy were somehow propagated metaphysically. Unlike regular wave propagation asymmetry, Einstein talked about inertial frames where he believed motion could not be detected. In an attempt to satisfy the MMX, Einstein's assumed the propagation time upstream and down remained symmetrical (light simultaneity), irrespective of motion. Against all known physics, Einstein concluded that there was an unidentified light transmission mechanism that produced simultaneity without a medium. Although stationary EM systems were accepted as having a medium, being a solution of the wave equation and causal, *Einstein now supported the inconsistency that EM systems lost their medium and became non causal when set in motion.*

It seems Einstein did not have complete knowledge of wave motion, wave theory and how to solve the wave equation generally for both source and observer motion relative to the medium.

Einstein was aware that light transmission was not straight forward. But he was not aware of the detailed moving medium model, described below, or the absolute necessity of a medium needed to solve the wave equation. Otherwise, he would not have tried to ‘simplify’ the situation by removing the medium altogether. This created his homogeneous, ether-less aspect, whereby one energy/information motional transform was attempted to fit all situations. Therefore, Einstein’s concept of relativity cannot satisfy the wave equation or sustain the multi moving media model. The ether-less theory causes the inability to distinguish between stationary and constantly moving systems and between source and observer motion. It is non casual and non observable, the imaginary observed event can occur before the cause (source event), and the source and observer paths are indefinable. Whereas, the medium based LT and the new theory are a solution of the wave equation and their moving systems definable in the medium. They predict observations of systems in motion (are causal) and predict no paradoxes. *The new theory simply extends the LT for both source and observer motion, relative to the medium, giving for the first time the complete solution for EM systems in motion.*

There does not seem much point in questioning the existence of medium. It has a measured permeability μ (an electrical density or mass term), and a permittivity ϵ (an electrical elastic or springiness term), enabling EM waves to ‘bounce’ through the medium (propagate). For optical media, μ varies little, it’s ϵ that determines the distinguishing light properties. A diamond has a large ϵ varying widely with frequency, giving it its spectacular sparkle. *Whereas a vacuum, with a small but finite μ and ϵ is therefore not empty space.* Away from gravitational bodies, the cosmic microwave background (CMB), discovered by Penzias and Wilson [19], radiates uniformly in all directions throughout the universe, implying that the EM medium is at rest in space. The medium has been confirmed through the Cosmic Background Explorer COBE [20]. Here it is shown that motion relative to the CMB increases the EM radiation (energy) collection, similar to trawling fish nets catch more fish than stationary ones. If the stationary medium, containing the propagating energy radiating uniformly, did not exist, then there would be no energy increase in the direction of motion.

4. MOVING MEDIUM MODEL

It is a fact that the event time is invariant and the propagation time is symmetrical and invariant in and against the direction of the Earth’s motion through the universe, as established in the MMX. Einstein’s ether-less model cannot explain the invariant MMX. The propagation time variance and asymmetry, predicted by the LT in a moving frame, relative to stationary medium, does not explain the MMX either. The only remaining explanation seems to be that the propagation medium moves with the Earth. Further support is provided through the symmetrical propagation time across the Solar System, moving with its galaxy through space at high speed. This is according to the Mars-Viking Lander project, Reasenburg et al. [21], where the communication time in each direction between Mars and Earth was found to be symmetrical. If the propagation medium had been at rest in the universe, there would have been a large asymmetry in the communication time, caused through the motion of the Solar System through the medium. *These propagation medium symmetries around the Earth and Sun imply that the propagation medium must move with large gravitational systems.* Data listed in reference [8], confirms that the medium moves (is entrained) with large gravitational bodies, as Stokes [22] suggested.

Bradley’s stellar aberration, where star light appears to be deflected by a small angle through observer (telescope) motion, is unaffected by the medium moving with the Earth. This non effect was used in an attempt to discredit Stokes entrainment model. However, this lack of effect is now explained quite naturally through the unity refractive index of a vacuum. According to Fresnel [16], there is no convection or deflection at a moving interface for this special case of refractive index $n = 1$. Light propagation cannot be altered by the universal medium changing velocity or direction at its boundary. Although the medium (vacuum) has finite permeability μ and permittivity ϵ it has no matter, no moving atomic structure to deflect the light, as explained in detail in [23]. *Thus a moving gravitational body can entrain a propagation medium, but the entrained medium cannot convect light.* Therefore, with no rigid (atomic) structure, a vacuum appears to behave as a fluid medium, on average at rest in space. It is attracted to and forms medium gradients around moving gravitational bodies, suns, planets, etc., including the Earth.

Measured properties support this model. The medium in the near field appears to rotate with the Earth’s surface relative to a ‘stationary’ medium surrounding and orbiting with the Earth. This model is supported by: Sagnac [13], he showed the medium is ‘stationary’ on the Earth’s surface.

Michelson and Gale [24], supported a surface medium rotating relative to the surrounding ‘stationary’ medium. Hafele and Keating [25], confirmed the surrounding ‘stationary’ medium orbits with the Earth. Saburi et al. [26] and GPS [14], established the Earth and satellites rotate relative to the surrounding ‘stationary’ medium. Reasenburg et al. [21], showed the Earth’s surrounding ‘stationary’ medium moves relative to a ‘stationary’ medium surrounding the Sun and Solar System. The Solar System medium, in turn, moves as a whole through the medium at ‘rest’ in the universe, as illustrated in Figure 1. Here W , is the effective extent of the Earth’s gravitational field of dominance (GFOD) in the presence of the Sun. It is estimated to be about 50 Earth radii, using Newton’s inverse square law. And w is the effective width of the Earth’s surface boundary layer, it is less than one Earth radius using Schwartzschild [17] inverse law. Its measureable effect appears to be less than 10 km, according to Hafele and Keating. Further details are given in references [8, 23, 27], but more details are required.

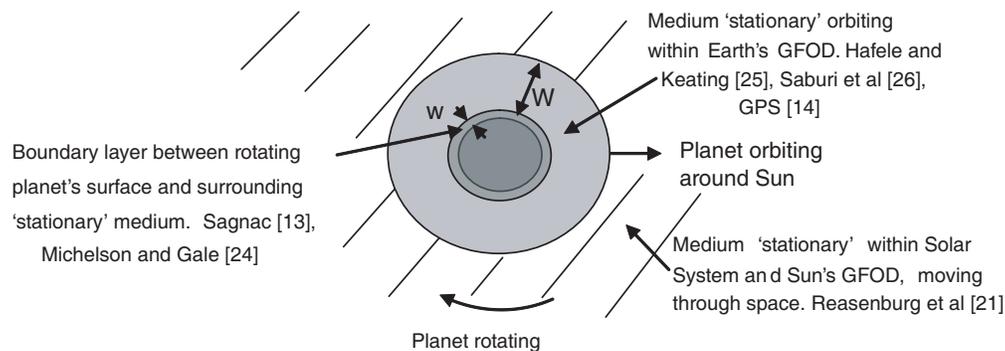


Figure 1: Propagation medium velocity profile around an orbiting planet.

5. CONCLUSION

In spite of Einstein’s claims to the contrary, it appears that the propagation medium for the transmission of EM waves (light) exists, and that Einstein’s SR, as a result, is incomplete. A new complete medium based theory, describing both source and observer motion, accounts for the ambiguities and paradoxes present in Einstein’s SR. The new theory predicts those properties of Einstein’s SR that have been measured and removes those that have not. It predicts further properties that have been measured that Einstein’s SR cannot predict. The conclusions from the 5 papers considered in this series suggest an electrical theory of everything.

REFERENCES

1. Einstein, A., “On the electrodynamics of moving bodies,” *Annalen der Physik*, Vol. 17, 891–921, 1905.
2. Minkowski, H., “Address delivered at the 80th assembly of German natural scientists and physicians,” *Cologne*, September 21, 1908; *Acad. Science Amsterdam*, Vol. 1, 427–442, 1899.
3. Lorentz, H. A., “Simplified theory of electrical and optical phenomena in moving systems,” *Proc. Acad. Science Amsterdam*, Vol. 1, 427–442, 1899.
4. Wright, S. E., “Electromagnetic sources and observers in motion IV — New motional optics,” *PIERS Proceedings*, Marrakesh, Morocco, March 20–23, 2011.
5. Isaac, N., *Philosophiae Naturalis Principia Mathematica*, 1687.
6. Maxwell, J. C., “A dynamical theory of the electromagnetic field,” *Philosophical Transactions of the Royal Society of London*, Vol. 155, 459–512, 1865.
7. Michelson, A. A. and E. W. Morley, “On the relative motion of the Earth and the luminiferous ether,” *American Journal of Science*, Vol. 34, 203, 333–345, 1887.
8. 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.
9. 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.

10. Wright, S. E., “Electromagnetic sources and observers in motion III — Derivation and solution of the electromagnetic motional wave equation,” *PIERS Proceedings*, 1151–1155, Cambridge, US, July 5–8, 2010.
11. Galilei, G., *Dialogue Concerning the Two Chief World Systems* using the Example of a ship travelling at constant velocity, 1632.
12. Doppler, C., “On the coloured light of the binary refracted stars and other celestial bodies — Attempt of a more general theory including Bradley’s theorem as an integral,” Monograph, 1842.
13. 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.
14. Logsdon, T., “The NAVSTAR Global positioning system,” *GPS*, Van Nostrand, Reinhold, 1992.
15. Bradley, J., *Third Astronomer Royal*, Reprint from the *Philosophical Transactions of Stellar Aberration: On the Motion of the Fixed Stars*, 1725.
16. Fresnel, A. J., “Lettre d’Augustin Fresnel a Francios Arago sur l’influence du mouvement terrestre dans quelque phenomenes d’optique,” *Annales de chimie el de physique*, Vol. 9, 57–66, 1818.
17. Schwarzschild, K., “Über das gravitationsfeld eines massenpunktes nach der einstein’schen theorie,” *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, Vol. 1, 189–196, 1916.
18. Poincaré, H., “La théorie de Lorentz et le principe de réaction,” *Archives Néerlandaises des Sciences Exactes et Naturelles*, Vol. 5, 253–278. 1900.
19. 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.
20. COBE Cosmic Background Explorer, NASA Goddard Space Flight Center, 1992.
21. 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.
22. Stokes, G. G., “On the aberration of light,” *Philosophical Magazine*, Vol. 27, 9–15, 1845.
23. Wright, S. E., *Problems with Einstein’s Relativity*, Trafford Publishing, 2010.
24. 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.
25. Hafele J. and R. Keating, “Around the world atomic clocks. Predicted relativistic time gains,” *Science* Vol. 177, No. 4044, 166–168, 1972.
26. Saburi, Y., M. Yamamoto, and K. Harada, “High precision time comparison via satellite and observed discrepancy of synchronization,” *IEEE Transactions on Instrumentation and Measurement*, Vol. 25, 473–477, 1976.

Efficient Use of “White Spaces” in the UHF Band (470–862 MHz) Employing Genetic Algorithms

Nikolaos C. Kapsalis and Panayotis G. Cottis

Division of Information Transmission Systems and Material Technology
School of Electrical and Computer Engineering, National Technical University of Athens, Greece

Abstract— The present work proposes a Genetic Algorithm based channel allocation scheme that allocates the Digital Dividend spectrum (790–862 MHz) to the cells of a mobile network taking into account the interference caused by neighboring TV broadcasters. The channel allocation is optimal in the sense that it maximizes frequency utilization and at the same time assures an acceptable level of fairness among the network cells.

1. INTRODUCTION

The present work follows a Genetic Algorithm (GA) approach to implement [1–3] dynamic channel allocation to cellular networks destined to operate in the 790–862 MHz frequency band or in white spaces dispersed over the Digital Dividend. The resource allocation scheme proposed takes into account the interference caused by neighboring broadcasters covering parts of the geographical area served by the cellular networks.

The GA approach has been widely used to implement adaptive resource allocation in wireless broadband systems. GAs are employed to dynamically assign frequencies in GSM networks [4], or subchannels in the framework of OFDM modulation [5]. Based on the extremely high computational power now available, GAs have become more than a theoretical tool and have given solutions to many practical problems [6–8]. Currently, GAs are used in a variety of optimization problems including the design of large scale networks [9] and frequency allocation [10], particularly dynamic channel allocation in the framework of mobile computing [11]. In the present work, the GA approach is used to formulate a dynamic channel allocation scheme for cellular networks taking into account interference from neighboring TV broadcasters. The proposed scheme maximizes the aggregate spectrum utilization whereas at the same time it guarantees: (i) that all the channels used are not interfered by neighboring TV broadcasters and (ii) a specified level of acceptable fairness among the network cells. The rest of the paper is organized as follows: In Section 2, the channel allocation problem under consideration is formulated. In Section 3, the GA approach is employed to implement optimal channel allocation both in the presence of interference and in uncommon situations. In Section 4, numerical results are presented demonstrating the feasibility of the proposed resource allocation scheme and its effectiveness in achieving the required trade-off between optimal frequency utilization and specified level of fairness.

2. EMPLOYING GENETIC ALGORITHMS FOR CHANNEL ALLOCATION TO CELLULAR NETWORKS SUBJECTED TO INTERFERENCE FROM BROADCASTERS

To eliminate any intracell interference in a cellular network, neighboring cells are not allowed to use the same channels. It is assumed that the transmit power and the distance between the Base Stations (BS) are determined so that no interference exists between non neighboring cells. Anyway, the proposed channel allocation scheme is adaptable to any other relevant assumptions.

2.1. Genetic Algorithms. A Short Overview

Genetic Algorithms are stochastic iterative methods that model biological phenomena such as genetic inheritance. They use a vocabulary borrowed from genetics. A population consists of individuals which represent potential solutions to a problem. An evolutionary process running on a population of individuals is a search through a space of potential solutions. To deal with a particular optimization problem, a GA must specify the following:

1. A genetic representation of candidate solutions to the problem.
2. A way to create an initial population of individuals.
3. A metric quantifying the solutions with regard to their fitness.
4. Genetic operators that determine the formation of children.
5. Appropriate values for the various parameters used by the GA.

After the initialization phase, the fitness of each individual is evaluated, new individuals are selected based on the current level of their fitness and individuals belonging to the population are transformed via the operations of mutation and crossover. After a certain number of updating the generations and if no further improvement in the aggregate fitness is observed, the resulting final population exhibits the highest aggregate fitness and constitutes the optimal solution.

To maximize spectrum utilization the channels available must be appropriately reused. Frequency reuse is achieved by assigning the channels pattern (A, B, C, D) to blocks of four cells as shown in Fig. 1. Hence,

$$a_{mn} = a_{m(n-2)} = a_{m(n+2)} = a_{(m-2)n} = a_{(m+2)n} \tag{1}$$

The proposed GA scheme starts by creating a random population of individuals which represent allocations of the available channels $i, i = 1, 2, \dots, F$, to the network cells $n, n = 1, 2, \dots, N$. Any possible channel allocation to the non interfered cells of the cellular network may be appropriately represented by a $(N \times F)$ matrix \mathbf{C} denoting how the channels available are allocated to the network cells taking into account both II and BI. The elements of \mathbf{C} ,

$$c_{ni} = a_{nm} \cdot b_{mi}, \quad m = 1, 2, \dots, N \tag{2}$$

are either 1, if channel i is allocated to cell n , or 0, otherwise.

The proposed GA scheme aims at maximizing spectrum utilization by appropriately allocating the available channels to the network cells. Since the cellular network is divided into four-cell blocks and a fair channel allocation is required, the maximum number of channels that can be assigned to a cell is equal to $F/4$. In every iteration of the scheme, the aggregate fitness is determined as a figure of merit of the solution, i.e., the current channel allocation plan. In the course of appropriately defining the aggregate fitness, vector \mathbf{E} is introduced to quantify any possible solution that satisfies the restrictions imposed by II and BI. Its elements

$$e_n = \sum_{i=1}^F c_{ni}, \quad n = 1, 2, \dots, N \tag{3}$$

denote the number of channels allocated to cell n . If $r_n, n = 1, 2, \dots, N$, is the number of channels required, the respective individual fitness is defined by

$$i_n = [1 + (r_n - e_n)]^{-1} \leq 1, \quad n = 1, 2, \dots, N \tag{4}$$

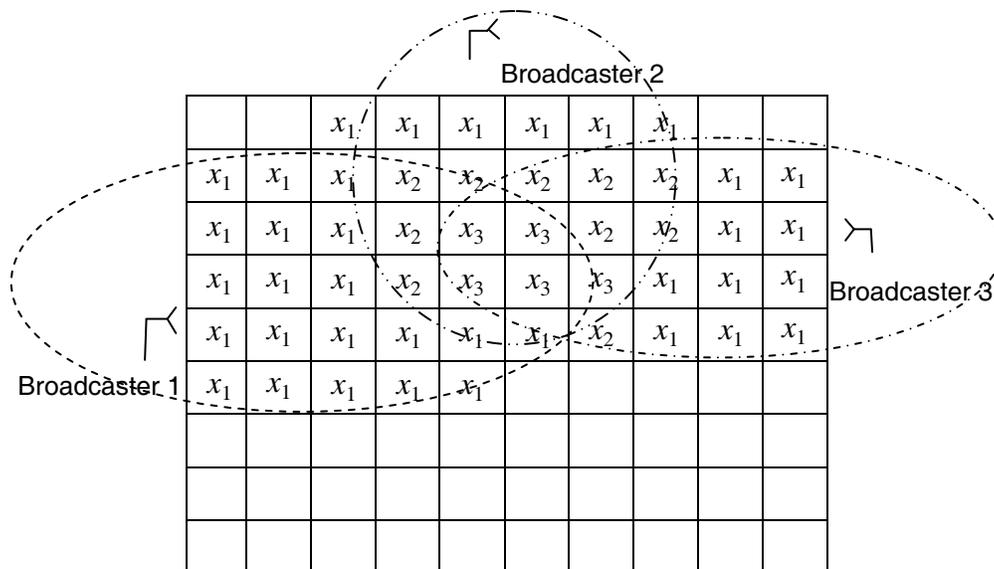


Figure 1: Cellular network subjected to BI.

where $r_n \geq e_n$. An individual is completely fit if $i_n = 1$, i.e., $e_n = r_n$. Then, the aggregate fitness of the population is defined by

$$I = \left[1 + \sum_{n=1}^N (r_n - e_n) \right]^{-1} \leq 1 \quad (5)$$

The aggregate fitness takes its maximum value which is equal to 1, if all the network cells may use as many channels as they require. The proposed GA scheme updates \mathbf{C} until the aggregate fitness is maximized, that is, the final population becomes the optimal channel allocation to the cellular network. The final values c_{ni} of the elements of \mathbf{C} are: 1, if channel i , $i = 1, 2, \dots, F$, is allocated to cell n , $n = 1, 2, \dots, C$, or 0, if channel i is not allocated to cell n .

To take into account fairness, Jain's fairness index [12] is introduced to quantify how fairly the candidate solutions allocate channels to the network cells. In this course,

$$JFI = \left[\sum_{n=1}^N e_n \right]^2 / N \cdot \sum_{n=1}^N e_n^2 \quad (6)$$

In the absence of BI, the number of channels allocated to all the network cells is $F/4$ and $JFI = 1$, which corresponds to a totally fair channel allocation.

3. NUMERICAL RESULTS AND DISCUSSION

In the simulations performed, a hypothetical cellular network consisting of $N = 100$ rectangular cells has been considered. Up to three interfering broadcasters have been assumed for the simulation of a hypothetical cellular network subjected to BI. The simulations were performed for $F = 20$ and 28 and for population sizes $P = 20, 30$ and $P = 40$.

A number of simulations have been performed concerning cases where the proposed GA based scheme is asked to allocate channels optimally and fairly in the presence of BI. For the simulations, BI was assumed to affect up to 3 out of $F/4$ frequency channels, and 10, 20, 30 or 40 network cells, respectively.

In Fig. 1, the number of usable channels has been plotted with respect to the extent of BI as quantified by the number of interfered channels, for a hypothetical cellular network consisting of $N = 100$ cells employing $F = 20$ channels for various values of P . From Fig. 1, it is apparent that

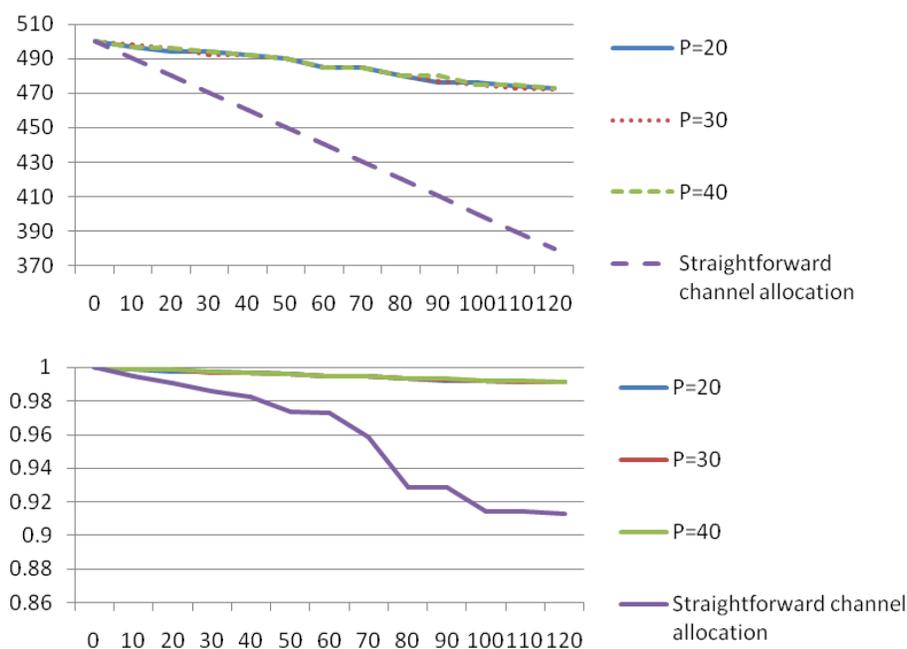


Figure 2: Total number of channels allocated and respective JFI.

the proposed GA based scheme achieves a significantly better spectrum utilization compared to the straightforward channel allocation option of excluding the interfered channels from operation.

In Fig. 2, JFI has been plotted with respect to the number of interfered channels for the same hypothetical cellular network ($N = 100$, $F = 20$) and various values of P . It is easily observed that the minimum fairness threshold set at $JFI = 0.95$ is always exceeded and that a slightly better fairness is achieved employing higher values of P . On the other hand though, when the straightforward option of excluding the interfered channels is preferred, fairness becomes significantly worse constituting an additional drawback of the straightforward option.

The above dynamic channel allocation is done in less than 15 s employing a laptop with an Intel Centrino processor with 2.5 GB RAM using Windows Vista. When the proposed GA based scheme was executed by a Desktop PC Intel Quad Core 2 with 4 GB RAM using Windows 7, the execution time was significantly reduced to approximately 2 s.

4. CONCLUSIONS

The proposed scheme takes into account both intracell interference and interference due to neighboring broadcasters. It follows a Genetic Algorithm approach to allocate the available channels to the network cells aiming at maximizing the aggregate spectrum utilization and at the same time assuring a specified level of fairness among the network cells. The simulations concerning the various scenarios examined demonstrate how the proposed GA based channel allocation scheme achieves a very high spectrum utilization ($> 95\%$). At the same time, a high fairness level is assured (> 0.95).

REFERENCES

1. Michalewicz, Z., *Genetic Algorithms + Data Structures = Evolution Programs*, 3rd Edition, Springer-Verlag, 1999.
2. Alex, F. and B. Donald, *Computer Models in Genetics*, McGraw-Hill, New York, 1970.
3. Melanie, M., *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA, 1996.
4. Pinagapany, S., A. V. Kulkarni, and D. Y. Patil, "Solving channel allocation problem in cellular radio networks using genetic algorithm", *3rd International Conference on Communication Systems Software and Middleware and Workshops, COMSWARE 2008*, Jan. 6–10, 2008.
5. Chen, J., S. Olafsson, and X. Gu, "A biologically inspired dynamic channel allocation technique in 802.11 WLANs with multiple access points," *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2007*, Sept. 3–7, 2007.
6. Murray, K. and D. Pesch, "Adaptive radio resource management for GSM using neural networks and genetic algorithms," Adaptive Wireless Systems Group, Department of Electronic Engineering Cork Institute of Technology, Cork, Ireland, 2001.
7. Cha, S.-H. and C. C. Tappert, "A genetic algorithm for constructing compact binary decision trees," *Journal of Pattern Recognition Research*, 2009.
8. Fogel, D. B., *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, IEEE Press, Piscataway, NJ, third edition, 2006.
9. Quintero, A. and S. Pierre, "On the design of large-scale UMTS mobile networks using hybrid genetic algorithms," *IEEE Transactions on Vehicular Technology*, Vol. 57, No. 4, 2007.
10. Idoumghar, A., M. Alabau, and R. Schott, "New hybrid genetic algorithms for the frequency assignment problem," *IEEE Trans. Broadcasting*, Vol. 48, No. 1, Mar. 2002.
11. Khanbary, L. M. O. and D. P. Vidyarthi, "A GA-based effective fault-tolerant model for channel allocation in mobile computing," *IEEE Transactions on Vehicular Technology*, Vol. 57, No. 3, May 2008.
12. Jain, R. K., D. M. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," *Digital Equipment Corporation, Technical Report DEC-TR-301*, Maynard, Massachusetts, 1984.

A Planar Parabolic Patch Antenna for UWB Applications

Mohamed Hayouni¹, Nabil Dakhli¹, Fethi Choubani¹, Tan Hoa Vuong², and Jacques David²

¹6'tel Research Unit, Higher School of Communications of Tunis, Sup'Com
University 7th November at Carthage, Tunisia

²High School of Electronic, Electro Technical, Hydraulic and Telecommunications of Toulouse, France

Abstract— For the purpose of wideband operations, partial grounded substrate patch antenna is investigated. The antenna consisted of a parabolic planar monopole printed on a thin dielectric FR4_epoxy substrate of $\epsilon_r = 4.4$ permittivity and is fed through a partial grounded substrate microstrip line. Various prototypes with different parabolic contours equations and the same length are analysed using Ansoft HFSS electromagnetic simulator based on the finite element method and measured using a vector network analyser Anritsu 37369C (40MHz to 40 GHz) where the calibration plane is the SMA connector jack used to connect the antenna. The best simulated and measured results were obtained for a parabolic planar patch monopole antenna defined by the geometrical equation $y = 0.044x^2$. These results concord and proved the ultra wide band behaviour of this design. Nevertheless, the bandwidth is ranging from 3.5 GHz to 16 GHz for a $VSWR < 2$. It includes the wireless local area (WLAN) band such as IEEE802.11, wireless personnel area network (WPAN) such as Bluetooth and it contains the two ranges of the ISM (Industrial, Scientifics, and Medical band) bands defined by the ITU-R (International Telecommunication Union-Radio communication sector) which are [2.4 GHz, 2.5 GHz] and [5.725 GHz, 5.875 GHz]. The proposed UWB antenna can be also used in the biomedical systems. Simulated results in terms of electrical parameters and radiation pattern show that this parabolic planar antenna is characterized by an omnidirectional radiation pattern.

1. INTRODUCTION

The design of a small ultra-wide band (UWB) antenna printed on short range radio devices plays a significant role in modern wireless communications. It is a carrier less short range communications technologies which transmits the information in the form of very short pulses. This aim encourages the research community to design ultra-wide band antennas where size, cost, weight, performance, ease of installation, and aerodynamic profile are constraints, like microstrip and printed slot antennas.

The variety in design that is possible with microstrip antennas probably exceeds that of any other type of antenna element. Microstrip antennas where size, weight cost better performance, compatibility with microwave and millimetre wave integrated circuits “MMICs”, robustness, ability to conform to planar and non-planar surfaces, etc. are required. Despite the advancements in this field, microstrip antennas are still limited by the inherent narrow bandwidth, low gain (6 dB), the most of the patch microstrip antennas radiate in the half plane, support exclusively lower power (100 W) and radiation losses by surface waves. For this reason, many works have been devoted to enhancing the bandwidth of microstrip patch antennas which can grow up to 70% by using multilayer substrate and the gain can grow up to 30% by using microstrip patch antennas array. Notably, other broad banding techniques include: thicker substrate, shaping the patch antennas [1], stacking [2], aperture coupling [2], slot compact planar design [3] and parasitic patch either in another layer (stacked geometry) or in the same layer (coplanar geometry) [4]. Microstrip technology using these techniques meets the demands of today's mobile communication operators.

The UWB technique presented in this paper is the use of a non uniform planar patch antenna printed on the FR4_epoxy substrate backed by a partial grounded substrate and fed through microstrip line in order to get different electric lengths which excite different close resonance frequencies and reduce the VSWR of the bandwidth. It is based on parabolic geometrical profile. Several prototypes with different equations are simulated in order to study the correlation parameter between the parabolic equation and the impedance matching.

2. UWB ANTENNA DESIGN

The antenna, as depicted by Figure 1, consisted of parabolic shaped planar patch geometry. It is printed on the FR4_epoxy substrate of 4.4 dielectric permittivity and 0.02 tangent losses. The substrate layer is of L_{sub} by W_{sub} dimensions and a thick of 1.6 mm. It is fed through a 50 Ω impedance characteristic feed line. A partial ground plane is glued at the bottom. Various antennas

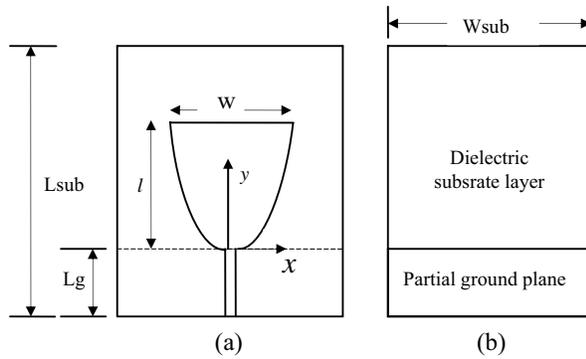


Figure 1: Antenna geometry design: Top view (a), bottom view (b).



Figure 2: Photograph of the realised antenna.

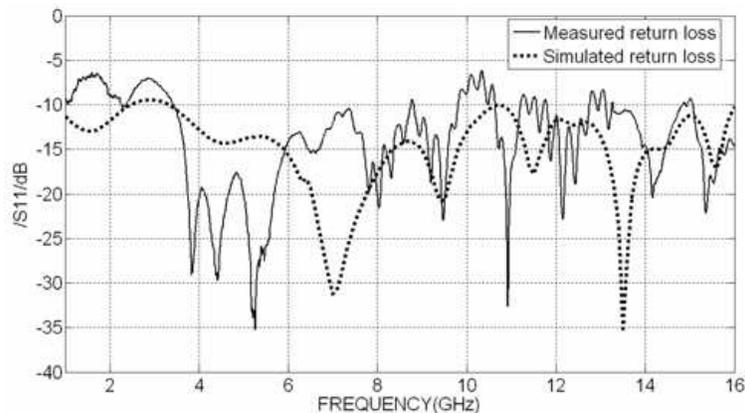


Figure 3: Measured and simulated return loss of the parabolic planar antenna.

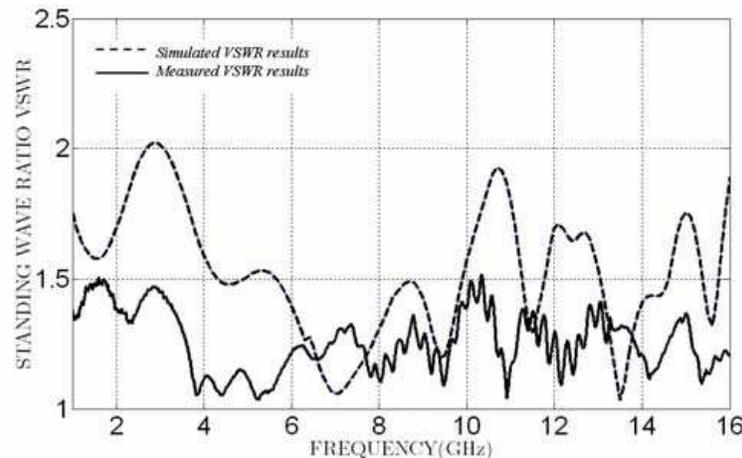


Figure 4: Measured and simulated return loss of the parabolic planar patch antenna.

of $y = ax^2$ kind equation was simulated using Ansoft HFSS electromagnetic simulator based on the finite element method in order to study its fundamental parameters such as impedance matching, gain, directivity and radiation pattern.

3. MEASURED AND SIMULATED RESULTS

3.1. S Parameters

Several simulations were performed using Ansoft HFSS electromagnetic commercial simulator which is based on the finite element method in order to study the correlation parameter between the parabolic curve equation and the antenna parameters. The best one curve equation corresponds to $y = 0.044x^2$; the antenna is characterized by $l = 80$ mm length and $w = 69$ mm width dimensions,

the substrate layer is of $L_{sub} = 90$ mm length and $W_{sub} = 85$ mm width and $h = 1.6$ mm thick dimensions and it is backed by a partial ground plane of 11.5 mm by 85 mm dimensions. Nevertheless, the visualization of the coefficient return loss as function of frequencies shows an ultra wide band behaviour, so the bandwidth of the simulated antenna is ranging from 3.5 GHz to 16 GHz for $|S_{11}| < -10$ dB.

The considered antenna is fabricated as depicted by Figure 2 and measured using a vector network analyser Anritsu 37369C (40 MHz to 40 GHz) where the calibration plane is the SMA connector jack used to connect the antenna. The measured results validate successfully those simulated in terms of return loss coefficient. Indeed, as showed by Figure 3, the measured return loss S_{11} reasonably agrees with the simulated results. A 3.5 – 16 GHz frequency range below -10 dB of VSWR is obtained as depicted by Figure 4.

3.2. Radiation Pattern

The simulated results at 5 GHz and 9.5 GHz frequencies in terms of radiation pattern as depicted by Figures 5 and 6 respectively, show that the pattern is directional in the azimuth plane [$f(\Phi), \theta = \pi/2$] and nondirectional in the elevation plane [$f(\theta), \Phi = \pi/2$], so it is an omnidirectional pattern; since it has an essentially nondirectional pattern in elevation plane and an essentially directional pattern in any orthogonal plane (in this case in the azimuthal plane).

The Tables 1 and 3 expose the antenna parameters simulated results of the parabolic planar antenna while the Tables 2 and 4 expand its maximum data at 5 GHz and 9.5 GHz frequencies

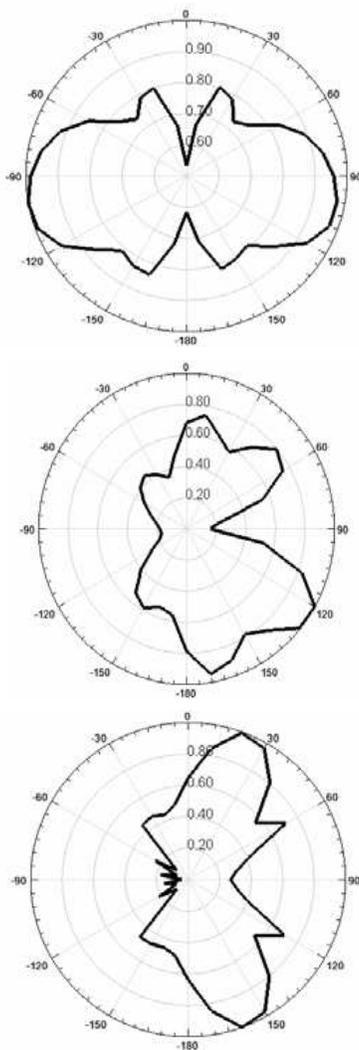


Figure 5: Simulated normalized E field radiation patterns at 5 GHz: (a) XY plane; (b) XZ plane; (c) YZ plane.

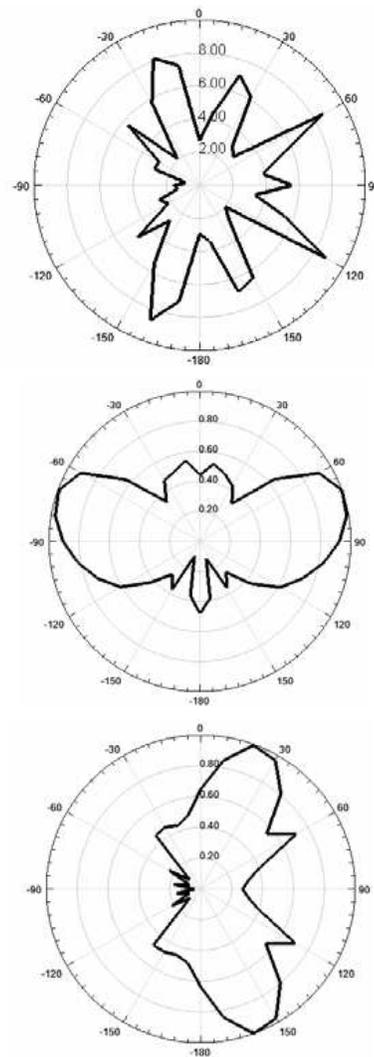


Figure 6: Simulated normalized E field radiation patterns at 9.5 GHz: (a) XY plane; (b) XZ plane; (c) YZ plane.

Table 1: Antenna parameters at 5 GHz frequency.

| Antenna parameter | Value |
|----------------------|--------------|
| Max U | 0.11209 W/Sr |
| Peak directivity | 1.6115 |
| Peak gain | 1.4657 |
| Peak realized gain | 1.4086 |
| Radiated power | 0.87408 W |
| Accepted power | 0.96105 W |
| Incident power | 1 W |
| Radiation efficiency | 0.9095 |
| Front to back ratio | 21.684 |

Table 2: Maximum field datas at 5 GHz frequency.

| E field | Value | At phi (°) | At theta (°) |
|--------------|----------|------------|--------------|
| Total | 9.1932 V | 140 | 100 |
| X | 7.1866 V | 290 | 270 |
| Y | 7.5805 V | 330 | 260 |
| Z | 3.9751 V | 90 | 120 |
| Phi | 9.1376 V | 320 | 270 |
| Theta | 4.59 V | 270 | 240 |

Table 3: Antenna Parameters at 9.5 GHz frequency.

| Antenna parameter | Value |
|----------------------|--------------|
| Max U | 0.21378 W/Sr |
| Peak directivity | 3.443 |
| Peak gain | 2.7186 |
| Peak realized gain | 2.6865 |
| Radiated power | 0.78026 W |
| Accepted power | 0.98818 W |
| Incident power | 1 W |
| Radiation efficiency | 0.78959 |
| Front to back ratio | 5.732 |

Table 4: Maximum field datas at 9.5 GHz frequency.

| E field | Value | At phi (°) | At theta (°) |
|--------------|----------|------------|--------------|
| Total | 12.696 V | 340 | 280 |
| X | 7.489 V | 60 | 90 |
| Y | 11.521 V | 160 | 90 |
| Z | 6.1848 V | 360 | 60 |
| Phi | 12.291 V | 340 | 280 |
| Theta | 7.1416 V | 360 | 60 |

respectively. Nevertheless, the peak gain exceeds 3.4 at 9.5 GHz

4. CONCLUSION

Both measured and simulated results proved successfully that this planar patch parabolic antenna with a geometrical equation of $y = 0.044x^2$, a partial grounded FR4 epoxy substrate layer and fed through a 50Ω impedance characteristic microstrip line enhances the bandwidth and reduces the impedance variation. It can be a suitable candidate for UWB applications since its bandwidth is ranging from 3.5 GHz to 16 GHz. Its radiation pattern is omnidirectional in the elevation plane and the peak gain can exceed 3.4 at some resonance frequencies such as 9.5 GHz. The bandwidth includes wireless local area (WLAN) band such as IEEE802.11, wireless personnel area network (WPAN) such as Bluetooth and it contains the two ranges of the ISM (Industrial, Scientifics, and Medical band) bands defined by the ITU-R (International Telecommunication Union-Radio communication sector) witch are [2.4 GHz, 2.5 GHz] and [5.725 GHz, 5.875 GHz]. The proposed UWB antenna can be also used in the biomedical systems.

REFERENCES

1. Choi, S., H., J. K. Park, S. K. Kim, and J. Y. Park, "A new ultra-wide band antenna for uwb application," *Microwave and Optical Technologies Letters*, Vol. 40, No. 5, March 5, 2004.
2. Eldek, A. A., "A broadband double dipole antenna with triangle and rhombus shapes and stable end-fire radiation patterns for phased array antenna systems," *Microwave Journal*, Jackson State University, October 2006.
3. Sldovieri, F., G. Prisco, A. Brancaccio, and G. Leone, "Characterization of ultra-wideband bow-tie antennas for ground penetrating radar systems," *Microwave Journal*, Vol. 49, No. 8, Pharmaceutical News Index page 186, August 2006.
4. Jang, Y.-W., "Characteristics of a wide-band, high-element characteristics of a wide-band, high gain, eight-element slot antenna for Pcs, Imt-2000 and Wll-band," *Microwave Journal*, Keukdong College, Chungcheongbuk Do, Korea, June 2004.

A Novel Compact Ultra-wideband Rectangular Shaped Antenna

Mohamed Hayouni¹, Fethi Choubani¹, Mohsen Denden¹,
Tan Hoa Vuong², and Jacques David²

¹6'tel Research Unit, Higher School of Communications of Tunis, Sup'Com
University 7th November at Carthage, Tunisia

²ENSEEIH, France

Abstract— In this paper, novel Compact Ultra Wideband (UWB) antennas are proposed. It consisted of rectangular shaped patch of $Lx = 15$ mm by $Ly = 14.5$ mm dimensions with two circled concave or convex profiles into its two corners. All simulated and measured prototypes are printed on a partial grounded FR4.epoxy substrate of 30 mm by 35 mm dimensions. Both of these inserted geometrical shapes provide different electric lengths with smoothly variations witch are adjusted to enhance the impedance bandwidth and reduce the VSWR between the main resonance frequencies of the rectangular patch. The excitation is launched through a 50 Ω microstrip feed line. Design and performances are analysed using Ansoft HFSS electromagnetic simulator based on the finite element method and measured using a vector network analyser Anritsu 37369C (40 MHz to 40 GHz) where the calibration plane is the SMA connector jack used to connect the antenna. Various prototypes are simulated and measured in order to view the correlation parameter between the impedance bandwidth and the circled convex or concave radius. The best antenna simulated and measured results correspond to a prototype with two convex circled forms of 3 mm radius. Nevertheless, the bandwidth is ranging from about 3.55 GHz to 4.6 GHz and from 7.4 GHz to 11.7 GHz for $S_{11} < -10$ dB.

1. INTRODUCTION

In recent applications of RF communications and remote sensing at radio, microwave and terahertz frequencies, the design of a small ultra-wide band (UWB) antenna printed on short range radio devices plays a significant role in modern wireless communications. It is a carrier less short range communications technologies witch transmits the information in the form of very short pulses. This aim encourages the research community to design ultra-wide band antennas where size, cost, weight, performance, ease of installation, and aerodynamic profile are constrains, like microstrip and printed slot antennas.

The variety in design that is possible with microstrip antennas probably exceeds that of any other type of antenna element. Microstrip antennas where size, weight cost better performance, compatibility with microwave and millimetre wave integrated circuits (MMICs), robustness, ability to conform to planar and non-planar surfaces, etc. are required. Despite the advancements in this field, microstrip antennas are still limited by the inherent narrow bandwidth, low gain (6 dB), the most of the patch microstrip antennas radiate in the half plane, support exclusively lower power (100 W) and radiation losses by surfaces waves. For this reason, an enormous amount of research literature is available on broad banding techniques for microstrip patch antennas witch can grow up to 70% by using multilayer substrate and the gain can grow up to 30% by using microstrip patch antennas array. Notably, other broad banding techniques include: thicker substrate, shaping the patch antennas [1], stacking [2], aperture coupling [2], slot compact planar design [3] and parasitic patch either in another layer (staked geometry) or in the same layer (coplanar geometry) [4]. Microstrip technology using these techniques meets the demands of today's mobile communication operators.

Seok H. Choi has proved in [1] using a rectangular shaped partial grounded substrate with two inserted steps of 9*1 mm and 12*1.5 mm dimensions and rectangular slot ultra wideband behaviour. In order to more enhance the bandwidth, reduce the VSWR between the main resonance frequencies of the rectangular planar antenna and create a smoothly variation of the return loss coefficient, we have proposed a non uniform radiator whose profile is characterized by two convex and concave circled curves into its two corners. Various prototypes have been simulated using the commercial electromagnetic simulator Ansoft HFSS witch is based on the finite element method to view the correlation parameter between the radius of each circled curve and the return loss coefficient without forgetting the radiation parameters such as the gain, directivity, radiation pattern. A comparison with measured results in terms of impedance bandwidth is also established for the best UWB simulated prototype in order to view the concordance degree between the simulated and measured results.

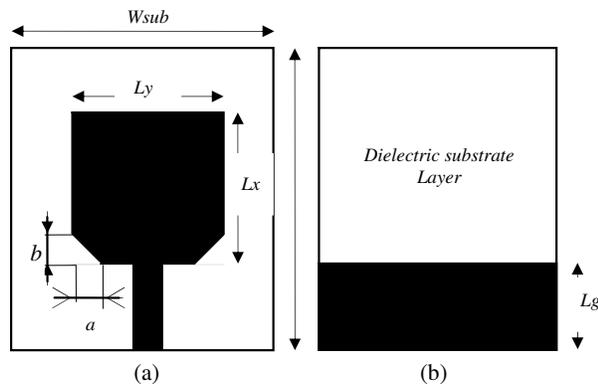


Figure 1: Geometry of the proposed UWB antenna with two triangular removed shapes, (a) front view, (b) back view.

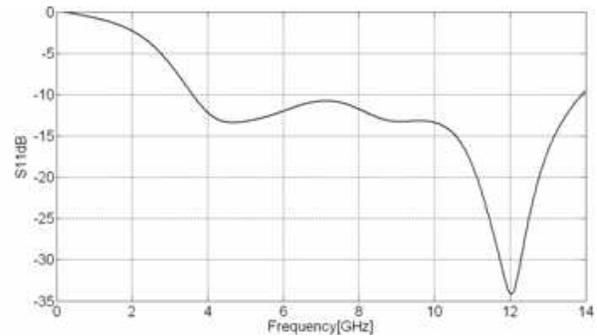


Figure 2: Return loss coefficient $/S_{11}/$ dB of the rectangular patch antenna with two removed triangular shapes.

2. ANTENNA GEOMETRY AND SIMULATED RESULTS

2.1. Antenna Design

The antenna showed by Figure 1 consists of a rectangular patch printed on an FR4 epoxy dielectric substrate with a permittivity constant of 4.4, a tang losses of 0.018 and a thick of 1.6 mm, it is fed through a 50Ω microstrip line partial grounded substrate. Two triangular removed forms defined by $a = b = 3$ mm dimensions as depicted by Figure 1 are truncated from two corners of the rectangular patch in order to reduce the VSWR between the resonance frequencies. The dimensions of the proposed antenna are summarized in the Table 1 below.

Table 1: Antenna dimensions.

| | |
|--|---------|
| Lx (length of the rectangular patch) | 15 mm |
| Ly (Width of the patch) | 14.5 mm |
| $Lsub$ (length of the substrate layer) | 35 mm |
| $Wsub$ (width of the substrate layer) | 30 mm |
| Lg (width of the ground plane) | 11.5 mm |
| a (length of the removed triangular conductor sheet) | 3 mm |
| b (width of the removed triangular conductor sheet) | 3 mm |

2.2. S Parameters

Several simulations using Ansoft HFSS commercial software were performed in order to study the impedance bandwidth of the antenna. A commonly used definition of impedance bandwidth is the frequency range over the voltage standing wave ratio (VSWR) is less than 2 : 1. Based on this impedance bandwidth, simulated results depicted by Figure 2 and Figure 3 in terms of return loss coefficient and VSWR respectively proved that this antenna can operate from 3.5 GHz to 14 GHz for a $VSWR < 2$.

3. PLANAR NON-UNIFORM UWB RADIATORS

3.1. Antenna Design

Several electromagnetic simulations have been realised in order to more reduce the return loss coefficient at the bandwidth frequency of this antenna. For this aim, we have essayed smoothly variation profiles between the two points A and B such as concave and convex circled curves as depicted by Figure 4. Several prototypes were simulated and analysed using Ansoft HFSS electromagnetic software in order to view the correlation parameter between each circled curve radius and the impedance bandwidth without forgetting the exam of radiation pattern, gain and directivity.

3.2. S Parameters Simulated Results

The behaviour of the return loss for different radius of the removed concave forms is showed by Figure 5. It proves that if the radius R increases the resonance frequency around 12 GHz moved forward at the left.

3.3. Antenna with Two Convex Circled Forms

In order to create a another smoothly variations of the electrical length between the two points A and B we have inserted convex circled shapes into the two corners of the rectangular patch. Compared to the return loss coefficient of the previous antennas with concave circled shapes, the $/S_{11}/(\text{dB})$ depicted by Figure 6 shows that with this geometrical form the return loss coefficient can be decreased around under about 8 GHz. Indeed, a $/S_{11}/\text{dB} < -15 \text{ dB}$ is achieved for the frequency band ranging from 8 GHz to 12 GHz.

3.3.1. Simulated and Measured Results of a Realised Antenna with Two Convex Circled Forms

A realised rectangular patch with two convex circled curves of $R = 3 \text{ mm}$ radius as depicted by Figure 7 was measured using an Agilent Performance Vector Network analyser [reference de l'analyseur] (10 MHz to GHz). The calibration plane is the SMA connector jack used to connect the antenna. Figure 8 shows the measured and simulated return loss coefficient using Ansoft HFSS concurs from 3.55 GHz to 4.6 GHz and from 7.4 GHz to 11.7 GHz for $S_{11} < -10 \text{ dB}$.

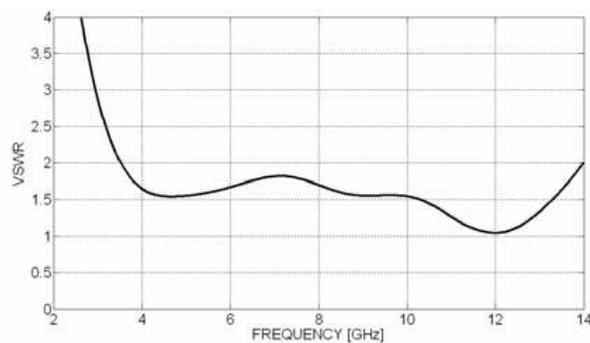


Figure 3: VSWR of the rectangular patch antenna with two triangular removed shapes.

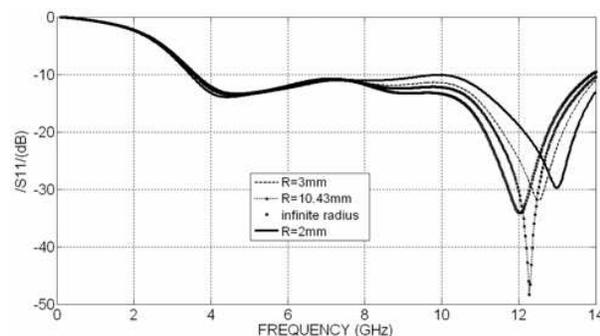


Figure 5: Behaviour of the return loss coefficient as function of the radius of the two concave circled profiles.

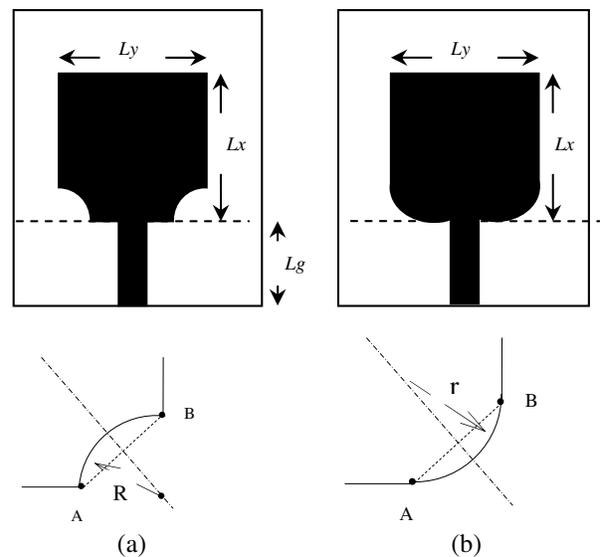


Figure 4: Geometries of the proposed UWB antennas, (a) antenna with concave circled curve, (b) antenna with convex circled curve.

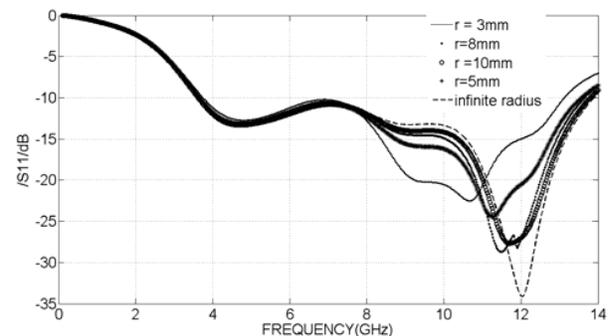


Figure 6: Behaviour of the return loss coefficient as function of the radius of the two convex profiles.

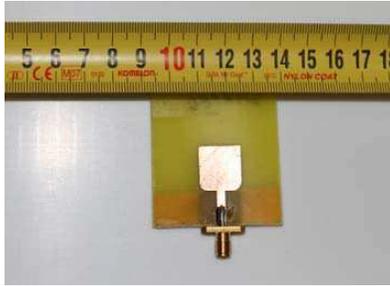


Figure 7: Photograph of the realised rectangular patch with two inserted convex circled curves of $R = 3$ mm radius.

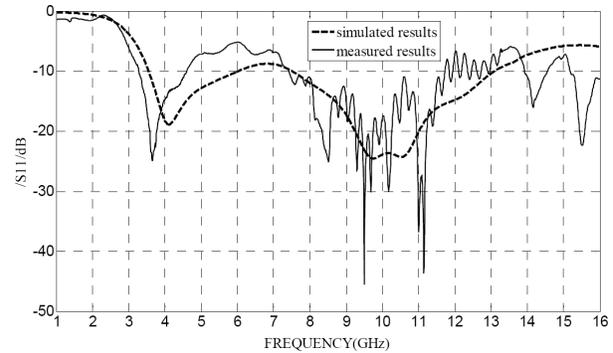


Figure 8: Measured and simulated return loss of an antenna with two convex circled curves of 3 mm radius.

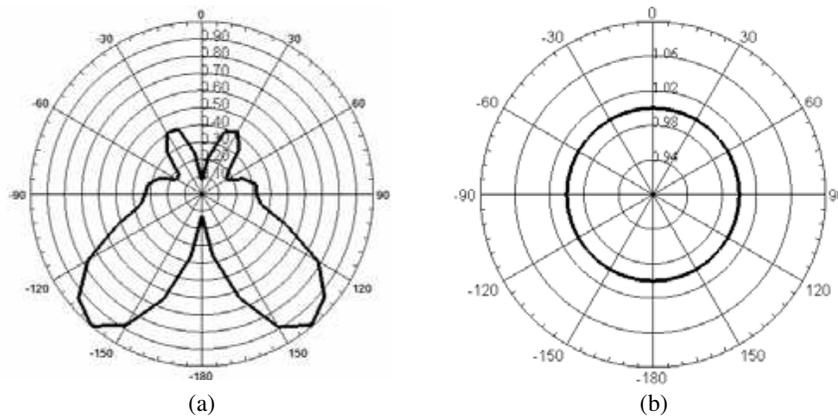


Figure 9: Simulated radiation pattern of an antenna with two convex circled curves of $R = 3$ mm radius at 9.5 GHz. (a) Theta cuts for 90° . (b) Theta cuts for 0° .

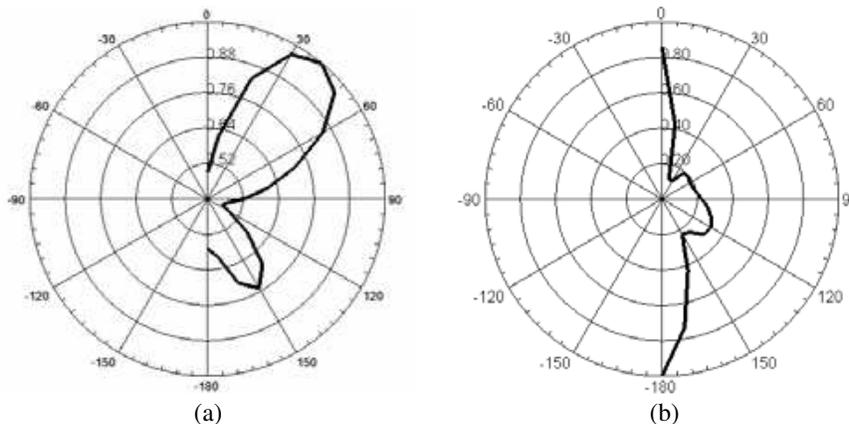


Figure 10: Simulated radiation pattern of an antenna with two convex circled curves of $R = 3$ mm radius at 9.5 GHz. (a) Phi cuts for 90° . (b) Phi cuts for 0° .

3.3.2. Radiation Pattern and Electrical Parameters

Figures 9 and 10 show the simulated conical (theta) and planar (phi) cuts normalized radiation patterns at 9.5 GHz. It is apparent that the proposed antenna has maxima for $\varphi = [110^\circ, 165^\circ]$ and for $\varphi = [205^\circ, 255^\circ]$, nulls at $\varphi = 0^\circ$ and $\varphi = 180^\circ$. The radiation pattern is omnidirectional at theta = 0° .

Tables 2 and 3 summarizes the main electrical parameters of this antenna at 9.5 GHz frequency.

Table 2: Antenna parameters at 9.5 GHz frequency. Table 3: Maximum field datas at 9.5 GHz frequency.

| Antenna parameter | Value |
|----------------------|--------------|
| Max U | 0.25651 W/Sr |
| Peak directivity | 3.733 |
| Peak gain | 3.242 |
| Peak realized gain | 3.2234 |
| Radiated power | 0.86349 W |
| Accepted power | 0.99428 W |
| Incident power | 1 W |
| Radiation efficiency | 0.86846 |

| E field | Value | At phi (°) | At theta (°) |
|--------------|----------|------------|--------------|
| Total | 13.907 V | 140 | 80 |
| X | 10.04 V | 130 | 80 |
| Y | 10.649 V | 140 | 70 |
| Z | 6.1258 V | 290 | 50 |
| Phi | 13.848 V | 140 | 80 |
| Theta | 8.408 V | 280 | 40 |

4. CONCLUSION

In order to get a smoothly variation of the electrical length and reduce the impedance bandwidth between the resonance frequencies of a rectangular partial grounded substrate planar patch antenna, some geometrical forms have been inserted into two corners witch are triangular, concave and convex circled shapes. Ultra wide band behaviour is achieved making these prototypes suitable for mobile and wireless network applications. Nevertheless, measured and simulated results of an antenna with two convex circled forms of 3 mm radius concords successively and showed that this compact radiator has a bandwidth that is ranging from 3.55 GHz to 4.6 GHz and from 7.4 GHz to 11.7 GHz for $S_{11} < -10$ dB. The simulated electrical parameters results showed also that this antenna is characterized by a peak directivity of 3.6657 at 9.5 GHz frequency. Numerical analysis using a CFDTD code will be presented in the final version of this paper.

REFERENCES

1. Choi, S. H., J. K. Park, S. K. Kim, and J. Y. Park, "A new ultra-wide band antenna for UWB application," *Microwave and Optical Technologies Letters*, Vol. 40, No. 5, March 5, 2004.
2. Sun, D. and L. You, "A broadband impedance matching method for proximity-coupled microstrip antenna," *IEEE Transaction on Antenna and Propagation*, Vol. 58, No. 4, April 2010.
3. Sim, C.-Y.-D., W.-T. Chung, and C.-H. Lee, "Compact slot antenna for UWB applications," *Antenna and Wireless Propagation Letters*, Vol. 9, 2010.
4. Eldek, A. A., "A broadband double dipole antenna with triangle and rhombus shapes and stable end-fire radiation patterns for phased array antenna systems," *Microwave Journal*, October 2006.
5. Sldovieri, F., G. Prisco, A. Brancaccio, and G. Leone, "Characterization of ultra-wideband bow-tie antennas for ground penetrating radar systems," *Microwave Journal*, Vol. 49, No. 8, 186, August 2006.
6. Jang, Y.-W., "Characteristics of a wide-band, high-element characteristics of a wide-band, high gain, eight-element slot antenna for Pcs, Imt-2000 and Wll-band," *Microwave Journal*, June 2004.

A Novel Printed Circular Antenna for Ultra Wideband Applications

Mohamed Hayouni¹, Mohsen Denden¹, Fethi Choubani¹,
Tan Hoa Vuong², and Jacques David²

¹6'tel Research Unit, Higher School of Communications of Tunis, Sup'Com
University 7th November at Carthage, Tunisia

²ENSEEIH, France

Abstract— For the purpose of wideband operations, a new half Hollow cylindrical monopole antenna for ultra wideband applications is investigated. It consisted of a half-disc monopole antenna with inserted rectangular step printed on a half hollow cylindrical substrate shape antenna and it is characterised by a small size of 30 mm × 35 mm, a simple geometry and it is fed through a 50 Ω microstrip feed line. Simulations results using electromagnetic simulator such as Ansoft HFSS proved the efficiency of this antenna in terms of bandwidth, indeed it is ranging from 4 to 6 GHz for a VSWR < 2 that include wireless local area (WLAN) band such as IEEE802.11 and wireless personnel area network (WPAN) such as Bluetooth, the lower range of the Federal Communications Commission band and ISM (Industrial, Scientifics, and Medical) band. Some techniques are investigated in order to have a notched band at this operating frequency range. Compared to the planar geometry, half Hollow cylindrical monopole antenna reduce the space occupied by the planar antenna to about 36, 28%.

1. INTRODUCTION

There is a requirement for low profile and lightweight antennas for microwave communications equipment. According to these requirements, a microstrip antenna is best suited. One of its main advantages is that they can be conformed to arbitrary curved surfaces. Indeed, conformal microstrip structures such as transmission lines, antennas and antenna arrays have witnessed a huge amount of research work in recent years owing to their interesting features for several applications, in particular to meet the increasing demand for smaller and more portable microwaves, for aircraft and telecommunication equipments. However microstrip antennas are limited in terms of bandwidth witch is approximately 10 to 20 percent. In the last decade, many researchers have studied bandwidth widening techniques for microstrip antennas to become suitable for UWB systems, so Wireless Personal Area Network (WPAN) and Wireless Body Area Network (WBAN) are seen as one of the major fields where such UWB characteristics can be potentially exploited. Some of these broad banding techniques include appending parasitic radiator [1–3], impedance matching network [4], L-probe feeding [5], shaping the patch antennas, using a partial backed grounded substrate and inserting some steps to adjuster the main resonance frequencies of the antenna in order to reduce the impedance matching [6]. So planar monopole antennas of various configurations have been proposed and offer different attractive features such as the possibility to operate in ultra-wide-bandwidth.

In this paper, we propose firstly a planar partial grounded substrate patch antenna witch consists of a compact half disc with two steps, secondly this planar antenna is transformed to a novel compact Half Hollow cylindrical monopole antenna for UWB systems by saving the same effective dimensions of the planar case. The prototypes are simulated using commercial electromagnetic simulators such as Ansoft HFSS witch is based on the finite element method.

2. PLANAR ANTENNA DESIGN AND RESULTS

2.1. Planar Antenna Design

The proposed planar antenna as illustrated by Figure 1(a) consisted of a half-disc of 10 mm radius associated with a rectangular patch of 10.4 mm by 1 mm dimensions printed on a partial grounded PTFE dielectric substrate layer of $\epsilon_r = 2.05$ permittivity, $h = 1.5$ mm thick and of 46 mm by 32 mm dimensions. At the bottom of the substrate, a partial ground plane of 36 mm by 18 mm is glued as depicted by Figure 1. The excitation is launched through a 50 Ω microstrip feed line.

2.2. Simulated Results

2.2.1. S Parameters

Simulated results using Ansoft HFSS witch is based on the finite element method proves as depicted by Figure 3 the UWB behaviour of the planar partial grounded substrate. Indeed it shows that the

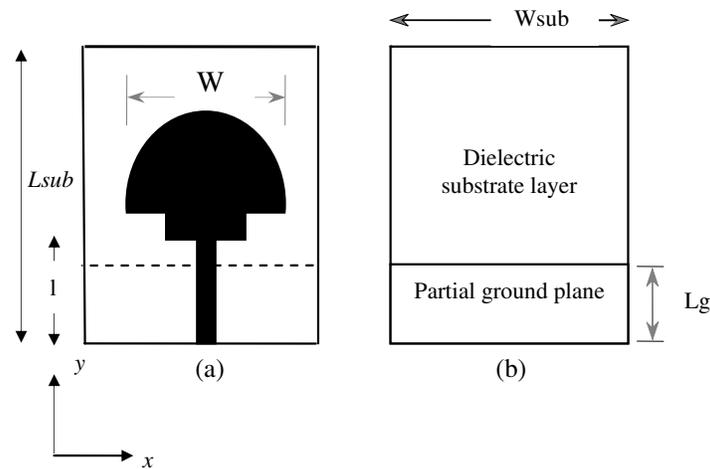


Figure 1: Antenna geometry design, (a) top view of the design, (b) back view.

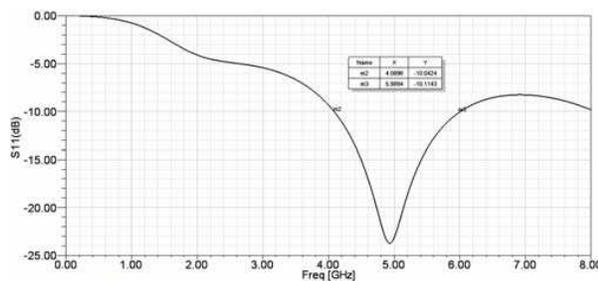


Figure 2: Simulated return loss for the proposed antenna.

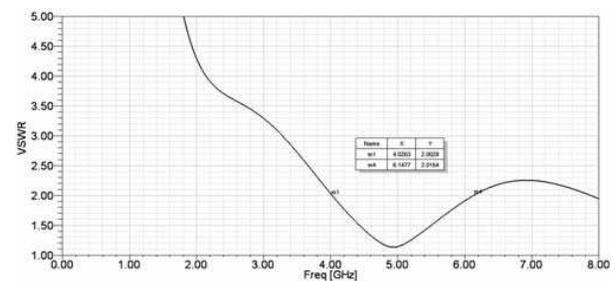


Figure 3: Simulated VSWR for the proposed antenna.

Table 1: Antenna parameters at 5 GHz.

| Quantity | value |
|------------------|-------------|
| Max U | 0.13033 W/S |
| Peak directivity | 1.6251 |

Table 2: Maximum field data at 5 GHz.

| E field | value | φ | θ |
|-------------|----------|-----------|----------|
| E_{total} | 9.913 V | 250 | 170° |
| E_x | 4.842 V | 230° | 90° |
| E_y | 9.806 V | 220° | 170° |
| E_z | 5.2887 V | 270° | 50° |
| E_{phi} | 9.7406 V | 360° | 180° |
| E_{theta} | 9.9058 V | 270° | 170° |

bandwidth of the planar antenna is $f_h - f_l \approx 2 \text{ GHz} \geq 0.25 f_c$ for a VSWR < 2 with $f_l \approx 4 \text{ GHz}$, $f_h \approx 6 \text{ GHz}$, and $f_c = \frac{f_h + f_l}{2} \approx 5 \text{ GHz}$, which proves an ultra wide band of the planar antenna.

2.2.2. Electric Field

Figure 4 represents the xy -plane, yz -plane and xz -plane normalized E field pattern in polar form at 5 GHz. The planar antenna is characterized by $HP = 120^\circ$ at xy -plane. Table 1 and Table 2 summarize the antenna parameters and the maximum E total field at 5 GHz frequency.

3. HALF HOLLOW CYLINDRICAL ANTENNA

3.1. Antenna Design

In recent years, microstrip patch antennas have received much attention, due to their interesting characteristic such as simple structure, low cost, being mechanically robust, simple to manufacture, and capability of integration with a printed communication system's circuitry. In the last few years, circular monopole antennas have been studied extensively for UWB communications systems, they present some attractive features, e.g., easy fabrication, Omni directional radiation patterns. Due to these advantages, we have essayed to design and simulate different half hollow cylindrical microstrip antennas in order to characterize these antennas in terms of bandwidth, radiation pattern, gain and

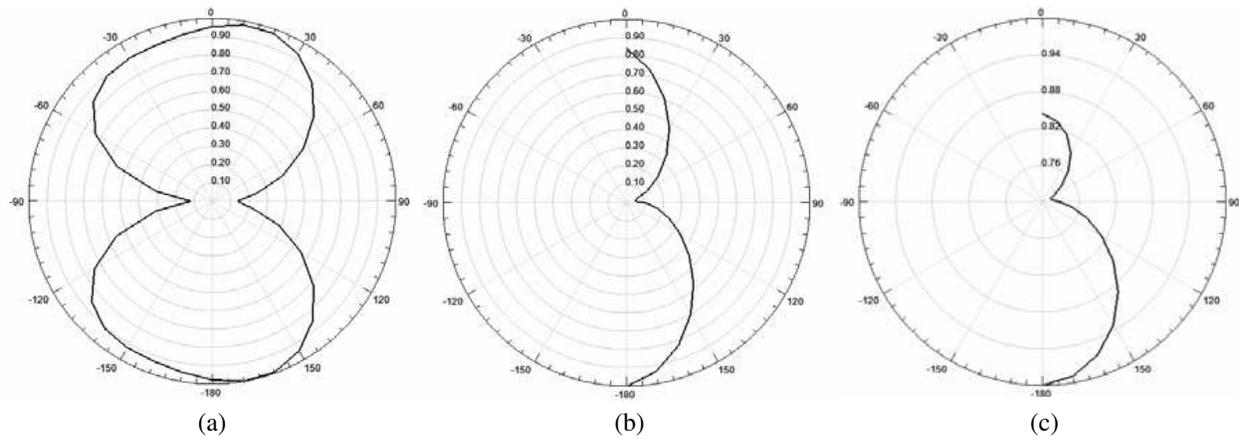


Figure 4: Radiation pattern of the UWB planar antenna at 5 GHz. (a) E -total (φ); theta = $90^\circ \rightarrow XY$ plane, (b) E -total (θ); phi = $90^\circ \rightarrow YZ$ plane, (c) E -total (θ); phi = $0^\circ \rightarrow XZ$ plane.

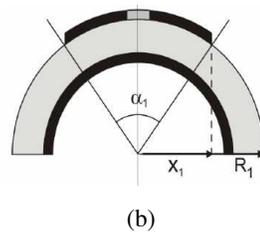
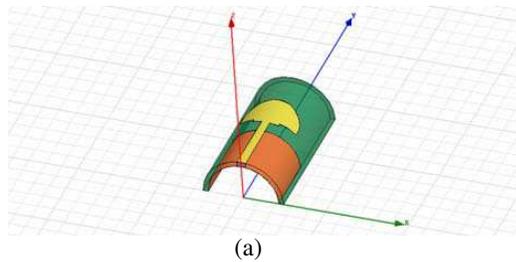


Figure 5: The proposed hollow half cylindrical antenna design, (a) perspective drawing and (b) side view.

Figure 6: Photograph of the realised half hollow cylindrical UWB antenna.

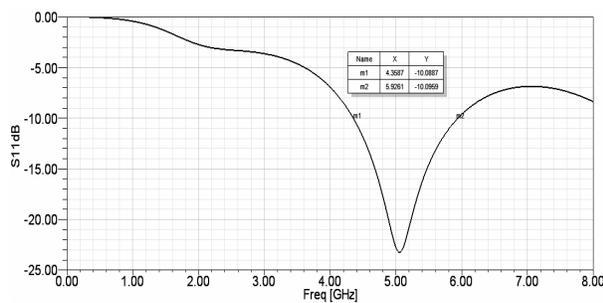


Figure 7: Simulated return loss coefficient of the proposed hollow cylindrical antenna.

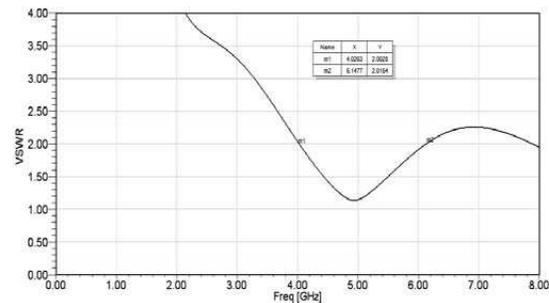


Figure 8: Simulated VSWR of the proposed hollow cylindrical antenna.

efficiency. Figure 5 presented a half hollow cylindrical antenna which is printed on a half cylindrical PTFE dielectric substrate with the same length and width of the planar substrate layer case. The curved antenna is patched on this substrate by maintaining the same dimensions contour as the planar case. Getting the previous dimensions, we have determined the different parameters required to design the half hollow cylindrical antenna such as R_1 and the angles that intercept each arc of circle. Indeed $R_1 = W_{sub}/\pi$ and $\alpha_1 = W/R_1$.

3.2. Simulated Results

Simulated results using electromagnetic Ansoft HFSS simulator shows ultra wide band behaviour of the proposed antenna. Indeed, Figure 8 shows that the bandwidth of the planar antenna is $f_h - f_l = 1.5674 \text{ GHz} \geq 0.25f_c$ for a $\text{VSWR} < 2$, where $f_l = 4.3587 \text{ GHz}$, $f_h = 5.9261 \text{ GHz}$ and $f_c = \frac{f_l + f_h}{2} = 5.14 \text{ GHz}$ which proves ultra wide band behaviour of the half hollow cylindrical antenna. Figure 9 shows the radiation pattern in XZ and YZ planes at 5 GHz frequency while Table 3 and Table 4 summarize the half hollow cylindrical antenna parameters and the maximum

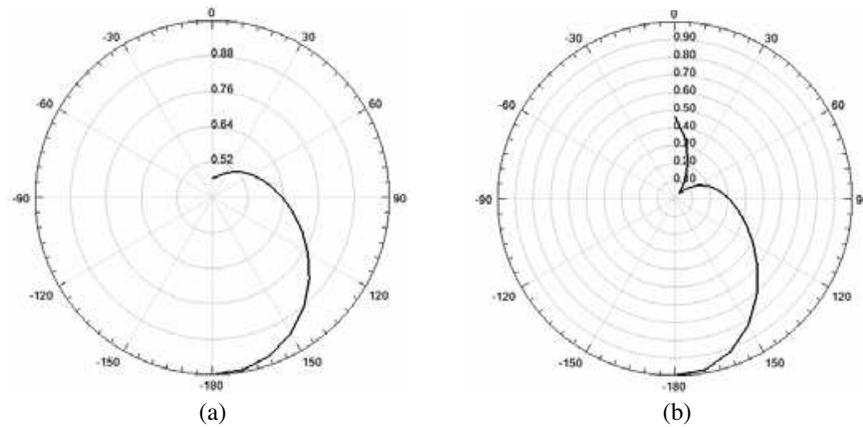


Figure 9: Radiation pattern of the UWB half hollow cylindrical antenna at 5 GHz. (a) $E_{tot}(\theta)$; $\phi = 0^\circ$, (b) $E_{tot}(\theta)$; $\phi = 90^\circ$.

Table 3: Half hollow cylindrical antenna parameters at 5 GHz.

| Quantity | value |
|------------------|-------------|
| Max U | 0.23178 W/S |
| Peak directivity | 2.9073 |

Table 4: Maximum field data at 5 GHz.

| E field | value | φ | θ |
|-------------|----------|-------------|-------------|
| E_{total} | 13.22 V | 290° | 180° |
| E_X | 5.5063 V | 310° | 100° |
| E_Y | 13.22 V | 290° | 180° |
| E_Z | 6.5754 V | 270° | 50° |
| E_{phi} | 13.22 V | 360° | 180° |
| E_{theta} | 13.22 V | 270° | 180° |

E total field at 5 GHz frequency.

4. CONCLUSION

Both planar partial grounded substrate antenna and half Hollow cylindrical monopole antenna have been simulated using Ansoft HFSS electromagnetic simulator. Simulated results in terms of return loss coefficient and radiation pattern show ultra wide band behaviour. Nevertheless, the bandwidth of the first planar antenna is ranging from 4 GHz to 6 GHz while whose of the half hollow cylindrical antenna is ranging from 4.3587 GHz to 5.9261 GHz for $S_{11} < -10$ dB. Compared to the monopole planar antenna, the geometry of the half hollow cylindrical antenna has reduced the space occupied by the antenna of about 36, 28%.

A comparison between measured results of the realized antenna as depicted by Figure 5 with simulated results will be established in the final version of this paper.

REFERENCES

1. Katehi, P. B., N. G. Alexopoulos, and I. Y. Hsia, "A bandwidth enhancement method for microstrip antennas," *IEEE Transaction on Antennas and Propagation*, Vol. 35, No. 1, 5–12, Jan. 1987.
2. Luk, K. M., K. F. Lee, and Y. L. Chow, "Proximity-coupled stacked circular-disc microstrip antennas with slots," *Electron. Lett.*, Vol. 34, No. 5, 419–420, Mar. 1998.
3. Rowe, W. S. T. and R. B. Waterhouse, "Investigation into the performance of proximity coupled stacked patches," *IEEE Transaction on Antennas and Propagation*, Vol. 54, No. 6, 1693–1698, Jun. 2006.
4. Waterhouse, R. B., *Microstrip Patch Antennas: A Designer's Guide*, Kluwer Academic, Boston, MA, 2003.
5. Luk, K. M., C. L. Chow, and K. F. Lee, "Broadband microstrip antenna," *Electron. Lett.*, Vol. 34, No. 15, 1442–1443, Jul. 1998.
6. Choi, S. H., J. K. Park, S. K. Ki, and J. Y. Park, "A new ultra-wideband antenna for UWB applications," *Microwave and Optical Technologies Letters*, Vol. 40, No. 5, Mar. 5, 2004.

Circular Patch Antenna Directivity Enhanced by Left-handed Material Cavity

Mondher Labidi, Nabil Dakhli, Jamel Belhadj Tahar, and Fethi Choubani

6thTel Research Unit, Higher School of Communications of Tunis, Sup'Com

University 7th November at Carthage, Tunisia

Abstract— In this paper, we study the influence of metamaterials on the performances of a circular patch antenna. For this purpose, metamaterials are based on Artificial Magnetic Conductor (AMC) and on inductive and capacitive planar structures. These structures allow to obtain a High Impedance Surfaces (HIS) and Partially Reflective Surfaces (PRS). Application of these structures allows to improve the performance of antenna, increase the gain and make it more directive.

1. INTRODUCTION

During the past few year, many investigation have been proposed for designing patch antennas. Feresidis [1] and Zhou [2] showed that the half wavelength restriction of a Fabry-Perot cavity antenna can be reduced to respectively a quarter wavelength and a tenth (10th) wavelength by using a novel type of metamaterial-based resonant cavity in order to design compact directive electromagnetic sources based on a single radiating antenna.

In the present work, we study the properties and characterization of an optimized cavity of a circular patch antenna at about 10 GHz. The cavity is composed of a substrate, a high impedance surface (HIS) and a partially reflective surface (PRS). The high impedance surface (HIS) [3, 4] consists of periodic metallic patterns printed on a dielectric substrate; they have very interesting electromagnetic properties, because they allow the propagation of electromagnetic waves along the surface only for certain frequency bands.

Moreover, the HIS has very important property of resonant impedance surface. This resonance leads to the zero crossing of the phase coefficient reflection. The HISs creates a current in phase with the surface source. Therefore, these surfaces behave as an artificial magnetic conductors (AMC). The HIS has been designed especially for the suppression of TM_0 mode dominant for the printed antennas. Indeed, the proposed design has a band gap for waves propagating laterally in TM mode and TE.

Furthermore, in the same band the surface waves are totally reflected. In this study, we use an artificial magnetic conductor (AMC) surface with reflection phase zero instead of a perfect electric conductor (PEC) surface and a partially reflective surface (PRS) with a frequency dependent reflection phase. In addition, we utilize Left Handed Materials (LHM) composed of HIS-AMC and PRS-AMC in the near environment of a circular patch antenna in order to enhance its performances. Using LHM technique, the gain and directivity of the patch antenna are greatly improved and the results obtained from the numerical simulations using Ansoft's HFSS show an enhancement of ~ 28 dB for the gain and a better directivity of the antenna in order of 30° instead of 85.5° .

This paper is organized as follows, Section 1 presents an introduction to the application of Fabry-Perot Cavity to the antenna, in Section 1, in Section 2 we introduce the model of cavity in use. Moreover, the characteristics optimization of circular patch antenna is presented in Section 3 and the conclusion is drawn in Section 4.

2. CONCEPTION OF FABRY-PEROT CAVITY

In 2005, Feresidis proposed for the realization of Fabry-Perot cavity, a totally reflective surface having the characteristic of reflecting waves with zero phase at its resonance frequency [1]. The cavity studied later by the team Feresidis was fueled by a patch antenna (microstrip patch antenna). But this structure was more compact than most any voids created by Von Trentini. Indeed, with the phase reflection near to zero, the thickness of the cavity was reduced by a factor of two. The subwavelength resonant cavity proposed in this paper, is optimized to operate in X-band [8.2 GHz; 12.4 GHz]. It consists of an Artificial Magnetic Conductor (AMC) composed of a PEC surface and a Partially Reflective Surface (PRS) [5]. The latter consists of a capacitive metallic square patch milled on one face and an inductive metallic mesh on the other. This cavity is optimized to obtain

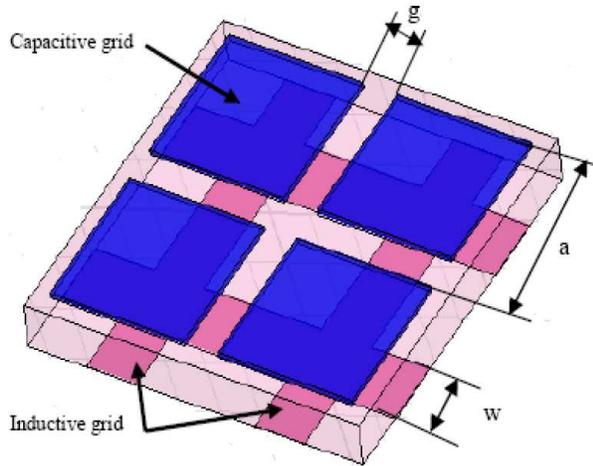


Figure 1: Metasurface reflector used as PRS with $a = 4$ mm; $w = 1.2$ mm.

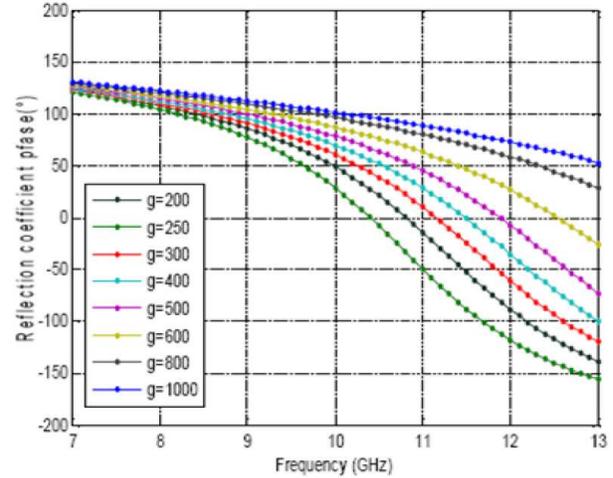


Figure 2: Reflection coefficient phase (deg) of PRS with different gap spacing “ g ” (g is in micron (μm)).

a phase shift of the reflection coefficient between -180° and 180° . As shown in Fig. 1, the PRS is placed on Rogers R04232 substrate of 0.81 mm thick ($\epsilon_r = 3.2$ and $\tan\delta = 0.0018$).

The width of the mesh (w) is equals to 1.2 mm, the width of the square patch (a) is = 3.8 mm. Both (the mesh and patch) have a periodicity of 4 mm. The gap spacing between two patches is g . To study the effect of variation of this gap width g between the patches on the reflection coefficient phase, several simulations were made on a unit cell with $a = 4$ mm, $w = 1.2$ mm. To study the effect of variation of this gap width g between the patches on the reflection coefficient phase, several simulations were made on a unit cell with $a = 4$ mm, $w = 1.2$ mm.

The AMC behavior extends over a wide frequency range (30%) in which the phase remains between -90° and 90° . It accompanies the behavior of the proposed structure of a HIS in this band. Simulation results are illustrated in Figs. 2 and 3. Fig. 2 shows the variation of the reflection coefficient phase ($^\circ$), Fig. 3 depicts the variation of the transmission coefficient phase ($^\circ$). As shown in these figures we can note that the variation of the gap spacing (g) causes a variation in the resonance frequency, increasing the gap width causes a decrease in the value of the capacitance between two cells.

3. STUDY OF THE PATCH ANTENNA

In this section, we first study the circular patch antenna only, then, the second part is devoted to describe the performances metamaterial based on Fabry-Perot resonance cavity of the antenna. The antenna used in this study is a circular patch antenna designed on an FR4-epoxy dielectric substrate with a thickness of 1.2 mm and has a permittivity of 4.4 and $\tan\delta = 0.0197$. The optimization of the patch antenna structure is deemed necessary in the simulation process to ensure that the result gain is optimum within the required frequency band. Thus the radius r of the patch antenna is chosen to operate in X-band. Knowing that the resonant frequency of patch antenna (f_r) is given as:

$$f_r = \frac{x'_{nm} c_0}{2\pi r_e \sqrt{\epsilon_r}} \quad (1)$$

where x'_{nm} is the n th non-zero root of the derivation of the Bessel function of m th order ($k_{11} = 1.84118$), c is the speed of light in vacuum, ϵ_r is the relative permittivity and r_e represent the effective radius of the patch antenna. It is given by the following relation:

$$r_e = r \left[1 + \frac{2h}{\pi r \epsilon_r \mu_r} \left(\ln \frac{\pi r}{2h} + 1.7726 \right) \right]^{1/2} \quad (2)$$

where r is the physical radius of the patch antenna and h is the thickness of the dielectric substrate.

For a desired working frequency in a fundamental mode, the calculated radius is chosen to be 4 mm. Fig. 3 shows the structure of the circular patch antenna. The simulation shows that the results antenna has a return loss of -16 dB at 10.86 GHz as shown in Fig. 4. After we have sized

antenna printed in the previous paragraph, we study in this part of the resonant cavity based on metamaterials, which will be fed by this antenna. We use the HIS and the PRS sized and characterized previously. The Fabry-Perot cavity consists of a reflecting plane above which is filed a partially reflective surface. The total structure is excited by a printed antenna inserted at the plane reflector as shown in Fig. 3. The previously circular patch antenna is designed to operate at 11 GHz. It is used as an excitation source of the cavity. It is surrounded by an array of periodic cells HIS-AMC to form one of two planes of the cavity mirrors. It's very interest to study the effect of SHI-CMA on different characteristics of the antenna.

In this section, we discuss simulation results of the cavity antenna composed of PEC and a PRS metasurface with a metamaterial PRS is also composed of an inductive and capacitive grid. A HIS-ACM and a PRS (capacitive grid) consist of a periodic array of copper with a periodicity and a width are 4 and 1.2 mm respectively. The gap spacing between two cells is $g = 400 \mu\text{m}$, optimized to operate in X-band and to provide a sufficiently high reflectance near to 10 GHz. In this paper, several position of the distance h between the patch antenna and the PRS, h presents a cavity thickness and it's given by the following relation [6]:

$$h = (\phi_{PRS} + \phi_r) \frac{\lambda}{4\pi} - (n \times e) \pm N \frac{\lambda}{2} \quad (3)$$

where: ϕ_{PRS} and ϕ_r are the reflection phase of the PRS reflector and the feeding antenna's ground plane respectively, are the index refractive and a thickness of the substrate respectively and N is

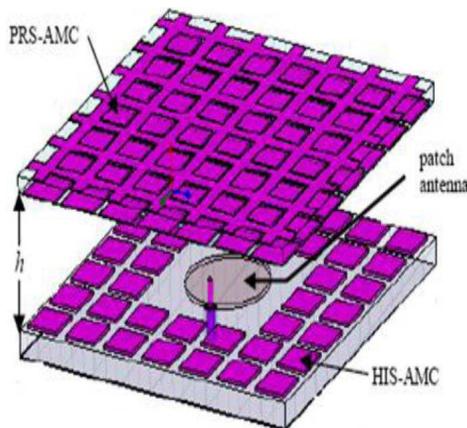


Figure 3: Optimized cavity composed of circular patch antenna, PEC, HIS, and composed metamaterial showing capacitive and inductive grid.

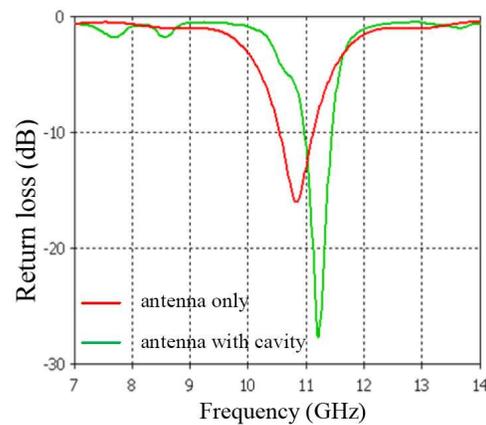


Figure 4: Return loss to the antenna only and with cavity ($a = 4$, $w = 1.2 \text{ mm}$, $g = 400 \mu\text{m}$ and $h = 1 \text{ mm}$).

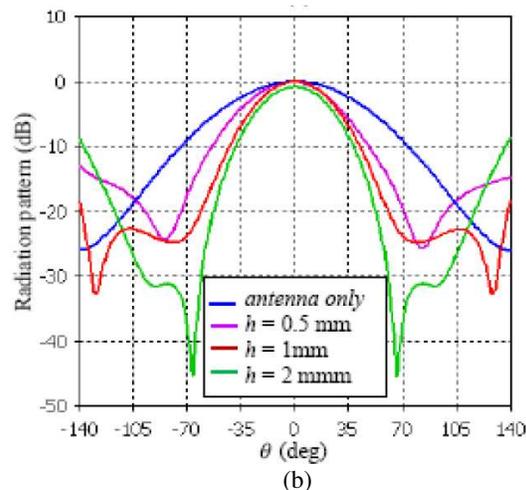
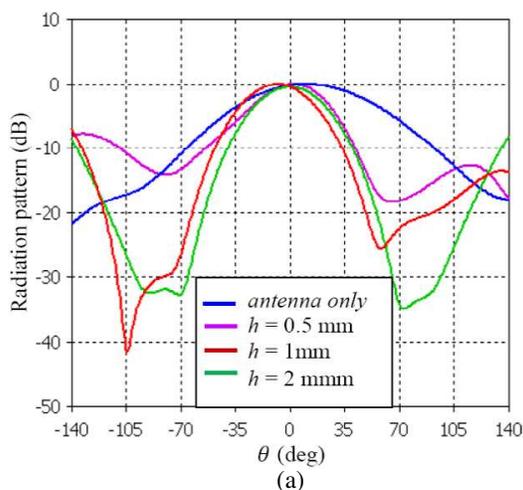


Figure 5: Radiation patterns in the (a) E -plane and (b) H -plane to the cavity.

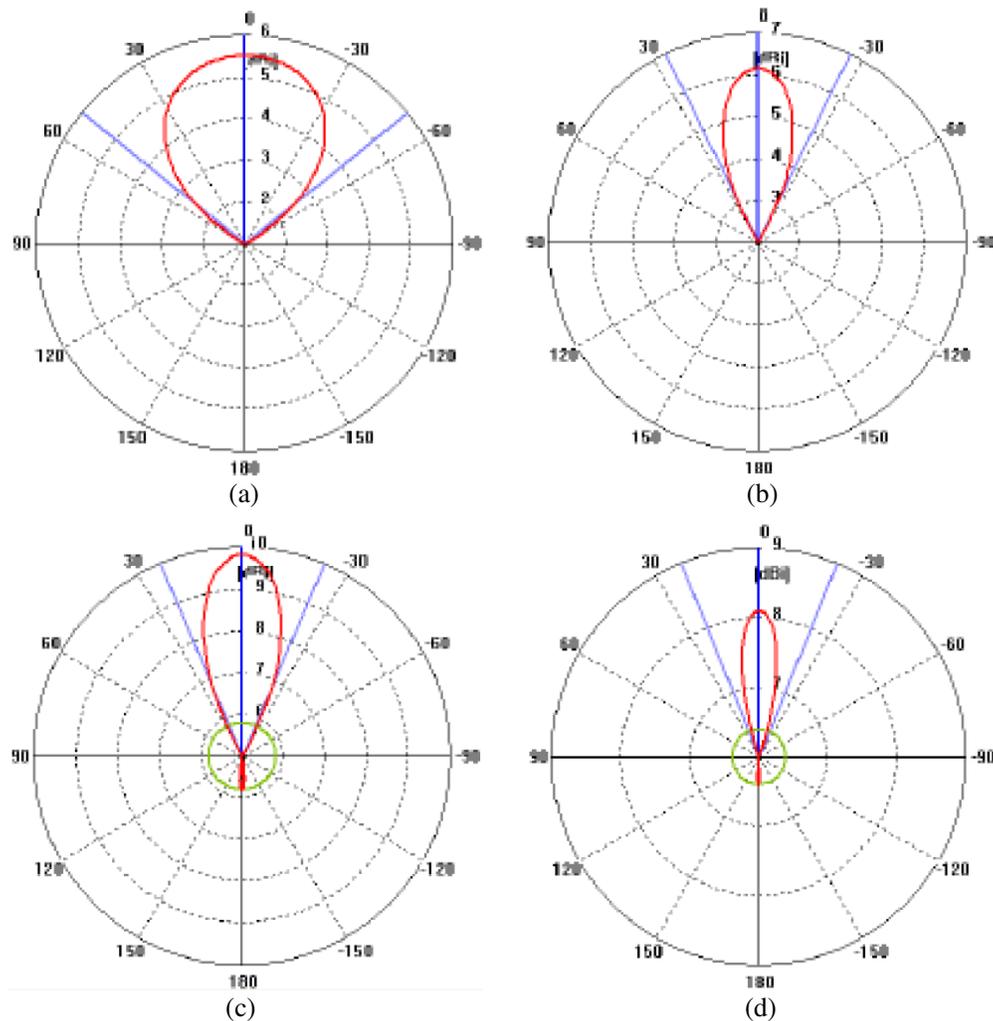


Figure 6: Gain patterns in the (a), (b), (c), (d) E -plane for the simulated antenna alone simulated gain patterns for the cavity antenna with several thickness.

an integer qualifying the electromagnetic mode of the cavity. Simulation results of the radiation patterns of the reconfigurable cavity antenna are presented in Fig. 5 and Fig. 6. Figs. 5 (a) and (b) show both the E -plane ($\phi = 90^\circ$) and H -plane ($\phi = 90^\circ$) radiation patterns and Fig. 6 presents a radiation gain in the E -plane. With several cavity thicknesses $h = 0.5, 1$ and 2 mm, we denote that the simulated cavity antenna presents respectively a gain of 6.7 dBi, 9.8 dBi and 8.1 dBi instead 5.5 dBi for the antenna alone. Also we denote at the best cavity antenna directivity ($h = 1$ mm), transmission coefficient of the simulated antenna is -26.33 dB at 11.18 GHz instead of -16 dB for the antenna alone, the frequency shift observed between the two return loss is due to the frequency varying effective parameters of metamaterial based resonant cavity. Also we notice that the beam is normal on the plane of the antenna. Moreover, there is no deflection observed as shown in Fig. 6. Furthermore, the aperture at -3 dB for the simulated cavity antenna is decreased from 102° to the antenna alone to 47° for cavity thickness of $h = 0.5$ mm, 39° for $h = 1$ mm and 44° with $h = 2$ mm. Hence, we conclude that the antenna is more directive.

4. CONCLUSION

The improvement of antenna characteristics using surfacebased planar artificial magnetic conductors is shown in this paper. A Fabry-Perot cavity is designed from these metamaterials. The performance of a circular patch antenna in terms of directivity has been increased. To conclude, more significant benefits are obtained after introducing the metamaterial based AMC to the antenna domain. They concern the gain, directivity and compact design of these structures. A demonstration of some of these advantages has been presented in this paper.

REFERENCES

1. Feresidis, A. P., G. Goussetis, S. Wang, and J. C. Vardaxoglou, “Artificial magnetic conductor surfaces and their application to low-profile highgain planar antennas,” *IEEE Trans. Antennas Propag.*, 209–215, 2005.
2. Zhou, L., H. Li, Y. Qin, Z. Wei, and C. T. Chan, “Directive emissions from subwavelength metamaterial-based cavities,” *Appl. Phys.*, 2005.
3. Broas, R. F. J., D. Sievenpiper, E. Yablonovitch, “A high-impedance ground plane applied to a cell phone hand set geometry,” *IEEE Trans. Microw. Theory Tech.*, 1262–1265, 2001.
4. Sievenpiper, D., L. Zhang, R. F. J. Broas, N. G. Alexópoulos, and E. Yablonovitch, “High-impedance electromagnetic surfaces with a forbidden frequency band,” *IEEE Trans. Microw. Theory Tech.*, 2059–2074, 1999.
5. Ourir, A., A. de Lustrac, and J.-M. Lourtioz, “Optimization of metamaterial based subwavelength cavities for ultracompact directive antennas,” *Microwave Opt. Technol. Lett.*, 2573–2577, 2006.
6. Ourir, A., A. de Lustrac, and J.-M. Lourtioz, “All-metamaterial-based subwavelength cavities ($\lambda/60$) for ultrathin directive antennas,” *Appl. Phys. Lett.*, Vol. 88, 2006.

Coupled Non Uniform Transmission Lines: Modeling and Crosstalk Performances

Mnaouer Kachout, Jamel Bel Hadj Tahar, and Fethi Choubani

6^oTel Research Unit, Higher School of Communications of Tunis, Sup'Com
University of the 7th November at Carthage, Ghazala Technopark, Ariana 2083, Tunisia

Abstract— In this paper, we propose a rigorous method to evaluate the crosstalk performance for coupled non uniform transmission lines. In this method, the width of the non uniform line is subdivided into a large number of rigorously coupled narrow lines. So, non uniform lines can be considered as a coupled multi-conductor transmission line. An electric equivalent model is developed to calculate the crosstalk between coupled non uniform transmission lines. Determination of the per-unit length capacitance and inductance parameters of the model is introduced. Finally, we demonstrate that coupled non uniform lines give the best performance in term of crosstalk compared to coupled microstrip lines.

1. INTRODUCTION

Single and coupled non uniform transmission lines (NTLs) are widely used in RF and microwave circuits as resonators [1], impedance matching [1, 2], delay equalizers [3], filters [4], wave shaping [5], analog signal processing [6], VLSI interconnect [7] and etc.. MTL is a system consisting of N conductors and a conducting shield or common return conductor, and serving to transmit signals via guided electromagnetic waves — mostly of transverse electromagnetic (TEM) or quasi-TEM type [8]. Such a system is further called $N + 1$ MTL. From telecommunication cables (of the lengths of meters or kilometers) through interconnects of equipment modules (of the lengths of centimeters) down to the micrometer scale for micrometer integrated circuit (MMIC) interconnects, all can be treated as MTL.

Modern trends in circuit designs such as operating at higher frequencies, lowering threshold voltages, and shrinking device geometries have made accurate prediction of electromagnetic (EM) compatibility an indispensable component in the design cycle. Susceptibility to electromagnetic interference (EMI) can severely degrade the signal integrity of the system. One of the main sources for the EMI is the coupling between the electrical interconnects, which serve as antennas at high frequencies. Many researches about MTL are conducted, such as multi-pin connector characteristic, crosstalk prediction of cable bundle, field analysis of digital bus.

There are several classical methods for the analysis of such structures in the literature. Almost all of these methods can be categorized into two methods, the transmission line method [9, 10] and full-wave analysis [11, 12]. The transmission line method is a simple model that cannot consider the current distribution in the width of the strip, because the whole width of the strip is considered as one line.

In this paper, a rigorous method is proposed and detailed to calculate the crosstalk for coupled non uniform transmission lines. We used the structure sketched in Fig. 1.

This paper is structured as follows: first, we propose an electric equivalent model, in Section 2. The per-unit length capacitance and inductance parameters needed for proposed model have been extracted in Section 3. Finally, to evaluate the crosstalk performance we make comparison between coupled microstrip line structures and coupled non uniform transmission lines.

2. MODELING OF NON UNIFORM TRANSMISSION LINES

2.1. Coupled Non Uniform Transmission Lines

A non uniform three-conductor transmission lines structure is sketched in Fig. 1.

A voltage source $V_S(t)$, with internal resistance R_S , is connected to a load R_L via both a generator conductor and reference conductor. A receptor circuit shares the same reference conductor and connects two terminations R_{NE} and R_{FE} by a receptor conductor.

We subdivide this structure into 5 equal parts (P_1, P_2, \dots, P_5), each part have the same line length of. In all these parts, conductors are assumed to be uniform. In this case, non uniform lines can be considered as a coupled multi-conductor transmission line.

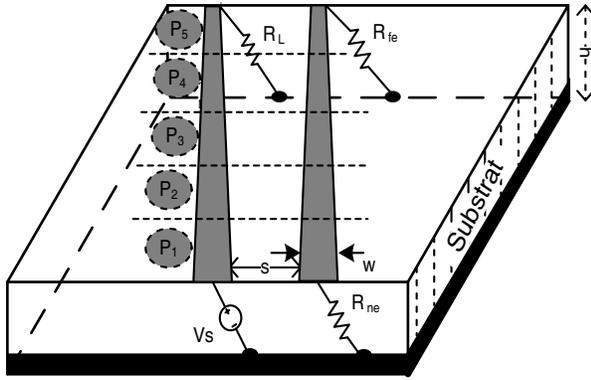


Figure 1: Coupled non uniform transmission lines.

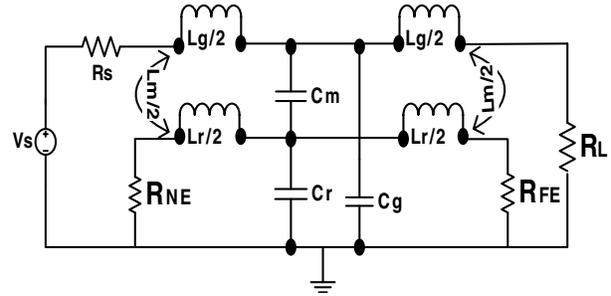


Figure 2: T-model.

2.2. Electric Equivalent Model

In this section, we develop T-electric equivalent model for each part of the presented structure. Fig. 2 shows the proposed T-model, where, $L_m = l_m * l_w$ represents the mutual inductance between conductors, $L_g = l_g * l_w$ is the self-inductance of generator conductor, $L_r = l_r * l_w$ is the self-inductance of the receptor conductor. Where l_w , l_m , l_g and l_r denote the conductor length, the per-unit length mutual inductance between generator and receptor conductors, the per-unit length inductance of the generator conductor, and the per-unit length inductance of the receptor conductor, respectively. $C_m = c_m * l_w$ is the mutual capacitance between conductors, $C_r = c_r * l_w$ is the capacitance of receptor conductor, $C_g = c_g * l_w$ is the capacitance of generator conductor. Where c_m , c_r and c_g denote the per-unit length mutual capacitance between two conductors, the per-unit length capacitance of the receptor conductor, and the per-unit length capacitance of the generator conductor, respectively.

3. CROSSTALK CALCULATION

In order to evaluate the crosstalk between non uniform conductors, we deal first with various per-unit-length parameters. In principle, the method of moments is a common and widespread technique. In order to illustrate this method, let us reconsider the parallel-plate capacitor problem. We assume that the charge distribution over each plate is uniform, that is, does not vary over the plates. In reality, the charge distribution will peak at the edges. To model this, in Fig. 3 we break each plate into small rectangular areas Δ_{si} , and assume the charge over each subarea as being constant with an unknown level, α_i . The total charge on each plate having been divided into N subareas is:

$$Q \cong \sum_{i=1}^N \alpha_i \Delta S_i \quad (1)$$

The heart of this method is to determine the total voltage of each subarea as the sum of the contributions from the charges on each subarea. Hence the total voltage of a subarea is the sum of the contributions from all the charges of all the subareas (including the subarea under consideration):

$$V_j = K_{j1}\alpha_1 + \dots + K_{jN+1}\alpha_{N+1} + \dots + K_{j2N}\alpha_{2N} \quad (2)$$

Each term K_{ji} represents as basic subproblem relating the voltage of a subarea V_j to the charge amplitude on another subarea.

$$K_{ji} = V_j / \alpha_j \quad (3)$$

Because of symmetry (both plates are identical), we can assign the voltage of the top plate (with respect to infinity) as $+V$ and the voltage of the bottom plate (with respect to infinity) as $-V$. The voltage between the two plates is then $2V$, so that the capacitance is

$$C = \frac{Q}{2V} \quad (4)$$

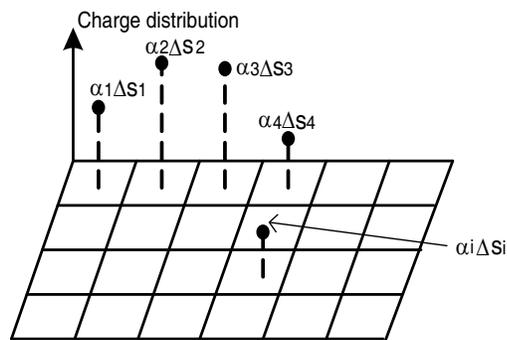


Figure 3: Approximating the charge distribution on the plates of parallel-plate capacitor.

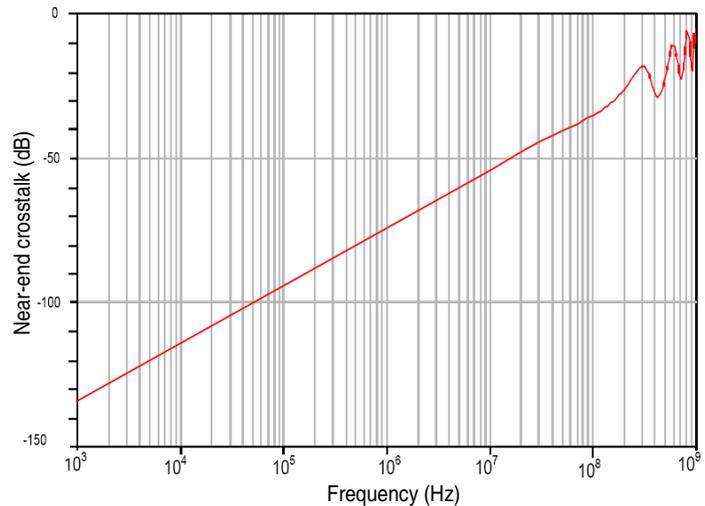


Figure 4: Near-end crosstalk versus frequency for coupled non uniform transmission line.

Grouping (2) for all subareas gives a matrix equation to be solved (which is the final result for all such MoM schemes):

$$\begin{bmatrix} K_{11} & \cdots & K_{1(2N)} \\ \vdots & \ddots & \vdots \\ K_{(2N)1} & \cdots & K_{(2N)(2N)} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_{2N} \end{bmatrix} = \begin{bmatrix} +V \\ \vdots \\ -V \end{bmatrix} \quad (5)$$

We have assigned all subareas on the top plate to have voltages of $+V$ and all subareas on the bottom plate have voltages of $-V$. Once (5) is solved for all the α_i charge distribution coefficients, the total charge on each plate can be determined from (1) and the total capacitance can be determined from (4).

In our case, we consider non uniform transmission lines structure shown in Fig. 1, where, $h = 47$ mils, and $\epsilon_r = 4, 7$ (glass epoxy). Conductors are assumed to be immersed in an inhomogeneous medium.

The per-unit length capacitance parameter matrix is:

$$C = \begin{bmatrix} C_g + C_m & -C_m \\ -C_m & C_r + C_m \end{bmatrix} \quad (6)$$

The per-unit length inductance parameter matrix is:

$$L = \begin{bmatrix} l_r & l_m \\ l_m & l_g \end{bmatrix} \quad (7)$$

In the configuration presented in Fig. 1 we find that $l_r = l_g$ and $C_r = C_g$.

For the above mentioned values, the per-unit length inductance and capacitance parameters for each part of the structure are presented in Table 1, where w is the line width and s is the separation distance between non uniform conductors.

These parameters can now be used to simulate the mentioned electrical equivalent model; the model is implemented in Advanced design system (ADS) of Agilent. Fig. 4 describes the near-end crosstalk variation versus frequency for non uniform transmission lines.

Simulation results show that the coupling increases gradually versus frequency. Maximum crosstalk is achieved for a frequency of 500 MHz.

4. NON UNIFORM TRANSMISSION LINES VERSUS MICROSTRIP LINE

To evaluate the crosstalk performance for non uniform conductors, in this section, we make comparison between uniform and non uniform conductors in term of near-end crosstalk. We consider

Table 1: Inductance and capacitance per-unit length parameters.

| | W (mils) | S (mils) | l_m (nH/m) | l_r (μ H/m) | C_m (pF/m) | C_r (nF/m) |
|----------------|----------|----------|--------------|--------------------|--------------|--------------|
| P ₁ | 225 | 100 | 19.05 | 0.20 | 6.10 | 0.19 |
| P ₂ | 200 | 150 | 14.77 | 0.23 | 3.20 | 0.18 |
| P ₃ | 175 | 200 | 12 | 0.24 | 1.86 | 0.16 |
| P ₄ | 150 | 250 | 10.07 | 0.26 | 1.15 | 0.14 |
| P ₅ | 125 | 300 | 8.65 | 0.29 | 0.73 | 0.13 |

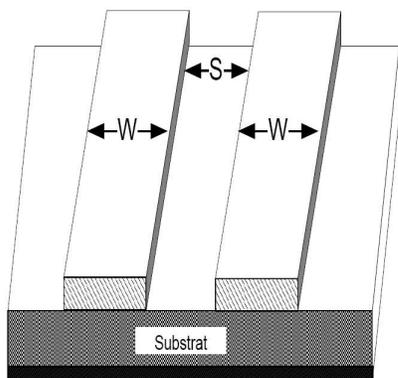


Figure 5: Coupled microstrip line.

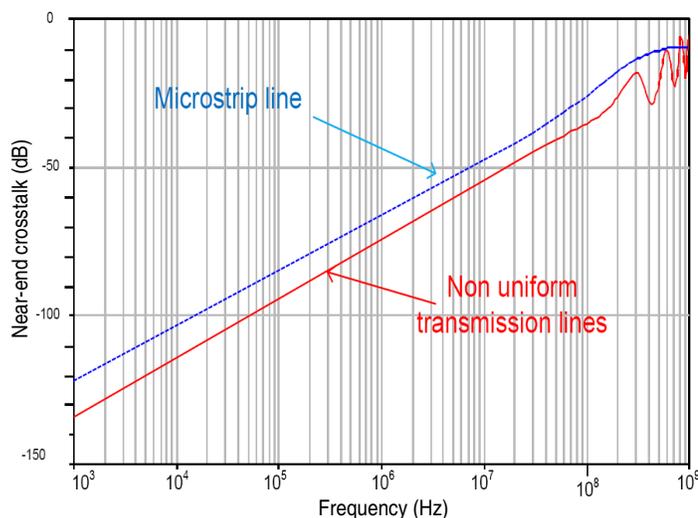


Figure 6: Near-end crosstalk comparison between coupled non uniform transmission lines and microstrip line.

coupled microstrip line having mainly the same demission as the non uniform conductors. Coupled microstrip line configuration is shown in Fig. 5, where the microstrip width is $w = 100$ mils and the distance separation $s = 100$ mils. Comparison results are presented in Fig. 6.

Comparison results show that non uniform transmission line have the best performance in term of crosstalk compared to microstrip line with same dimension. We demonstrate that we can reduce crosstalk more than 10 dB if we use non uniform transmission lines instead of microstrip line.

5. CONCLUSION

A rigorous method is presented and detailed to calculate the crosstalk for coupled non uniform transmission lines. We proposed an electric equivalent model and its capacitance and inductance per-unit length parameters are introduced. This method is much simpler and faster than other method. Also, to evaluate the crosstalk performance, we compared non uniform transmission lines versus uniform microstrip line. Obtained result shows that we can reduce crosstalk more than 10 dB if we use non uniform transmission lines instead of microstrip line.

REFERENCES

1. Collin, R. E., *Foundations for Microwave Engineering*, McGraw-Hill, New York, 1996.
2. Gholam-Tamimi, M. and M. Khalaj-Amirhosseini, "Conical coupled stripline transmission lines," *IEEE Symp. Mic. Ant. Prop. EMC, MAPE*, 1084–1087, Beijing, China, Aug. 8–12, 2005.
3. Tang, C. C. H., "Delay equalization by tapered cutoff waveguides," *IEEE Trans. Mic. Theory and Tech.*, 608–615, Nov. 1964.
4. Roberts, P. P. and G. E. Town, "Design of microwave filters by inverse scattering," *IEEE Trans. Mic. Theory and Tech.*, 739–743, Apr. 1995.
5. Burkhart, S. C. and R. B. Wilcox, "Arbitrary pulse shape synthesis via nonuniform transmission lines," *IEEE Trans. Mic. Theory and Tech.*, 1514–1518, Oct. 1990.

6. Hayden, L. A. and V. K. Tripathi, “Nonuniform coupled microstrip transversal filters for analog signal processing,” *IEEE Trans. Mic. Theory and Tech.*, 47–53, Jan. 1991.
7. Dhaene, T., L. Martens, and D. D. Zutter, “Transient simulation of arbitrary nonuniform interconnection structures characterized by scattering parameters,” *IEEE Trans. Circuits and Sys. I*, 928–937, Nov. 1992.
8. Khalaj Amirhosseini, M., “Analysis of coupled or single nonuniform transmission lines using step-by-step numerical integration,” *Progress In Electromagnetics Research*, Vol. 58, 187–198, 2006.
9. Collin, R. E., *Foundations of Microwave Engineering*, McGraw-Hill, 1992.
10. Pozar, D. M., *Microwave Engineering*, Addison Wesley, 1990.
11. Laroussi, R. and G. I. Costache, “Finite-element method applied to EMC problems [PCB environment],” *IEEE Trans. Electromagnetic Compatibility*, 178–184, May 1993.
12. Dhaene, T., L. Martens, and D. De Zutter, “Transient simulation of arbitrary nonuniform interconnection structures characterized by scattering parameters,” *IEEE Trans. Circuits Systems I: Fundamental Theory and Appl.*, 928–937, Nov. 1992.

Design of Non Uniform Meander Line Antennas for Passive RFID Tags in the UHF Band

Karim Bentaher¹, Fethi Choubani¹, Tan-Hôa Vuong², and Jacques David²

¹6'tel Research Unit, High School of Communications, Sup'Com
University 7 Novembre of Carthage, Tunisia

²ENSEEIH, University of Toulouse, France

Abstract— Development of far-field passive tags occurred during the first half of the twentieth century thanks to technological advances in the field of low power consumption integrated circuits. Passive tags are essentially composed of a chip and an antenna laying on a substrate. They are actually seen as the “future wireless barcodes”. But further research aiming at lowering manufacturing cost is still needed. This makes dipole based antennas so popular commercially, and an attractive choice for passive tags designers. In the UHF band, the dipole length (~ 16.5 cm) is above standard label dimensions (10×5 , 10×10 , 10×15 cm). This size constraint is solved using different miniaturization techniques. The most known method is meandering. In this paper, we study the effect of line width on non-uniform meandered antenna with inductively coupled feed. The performance is explored in terms of power transfer efficiency (τ) and tag free space read range (d_{\max}).

1. INTRODUCTION

The main reason for keeping RFID system design as simple as possible is to keep costs down. To make RFID tags attractive to retailers, they should be as cheap as possible. Admitting that the cost of the embedded integrated circuit (IC) has reached reasonable limit, the cost of a RFID tag is primarily set by the antenna design. Furthermore, in a printing process, the cost of the antenna is proportional to the amount of conductive ink used to print the layout. That's why passive UHF RFID antennas are mostly based on a printed dipole. This is the basic structure, but modifications are then made to resolve essentially two issues: power transfer, and size reduction. The first problem is resolved by a proper matching between the chip and the antenna. Good matching means maximum of power transfer and is essentially realized by loop-like structures [1–3]. We study here the inductively coupled feed variant, which is mostly used for chips with high impedance angle.

The second issue comes from a size constraint. In deed, most RFID tags are designed to fit within 101.6 mm (4 inches) long labels. But a resonant dipole in the UHF band is about 150 mm. So we have to shrink the size of the dipole. Meandering is the most used solution because it didn't use an extra conductive material [4, 5].

2. IMPEDANCE MATCHING

Obviously, a resonant meandered dipole can't be the solution to efficiently power up the tag chip. This is due to the capacitive reactance presented by the integrated circuit, which must be compensated by the antenna input impedance to assure a conjugate match needed for an efficient power transfer. The proposed solution here is an inductive coupling between a small loop and the meandered dipole acting as a radiating body (Figure 1).

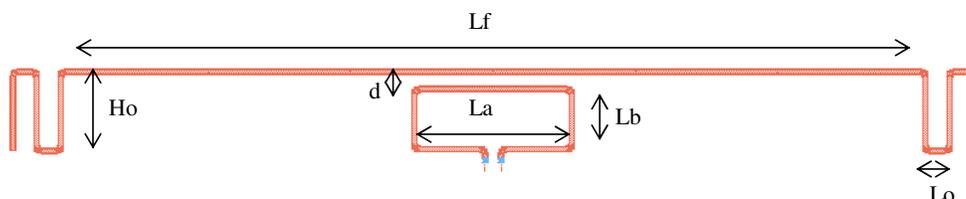


Figure 1: Inductively coupled non-uniform meander line structure.

The theory governing this coupling is already presented in [1]. The inductive coupling is modelled by a transformer. The input impedance seen from the loop's terminals is given by:

$$Z_a = R_a + jX_a = Z_{loop} + \frac{(2\pi f M)^2}{Z_{rb}} \quad (1)$$

where : Z_{loop} and Z_{rb} are the impedances of the radiating body and the feed loop, M is the mutual inductance characterizing the coupling between them.

At resonance, we have:

$$R_{a,0} = R_a(f = f_0) = \frac{(2\pi f_0 M)^2}{R_{rb,0}} \quad (2)$$

$$X_{a,0} = X_a(f = f_0) = 2\pi f_0 L_{loop}$$

We see that the antenna reactance depends only on L_{loop} , and the total input resistance is only related to the mutual inductance M .

The design procedure is composed of the following steps:

- Calculate the loop dimensions that gives the inductive reactance XL ,
- Assuming that R_{rb} is constant over the frequency range and equal to R_{rb} at resonance, and given the loop dimensions, we must adjust the mutual inductance M to a value approaching $(R_c \cdot R_{rb})^{1/2} / \omega_0$.

This procedure was applied to match two inductively coupled meandered antennas of different line widths (0.5 mm and 1 mm) to an RFID tag chip with $Z_c = 6.7 - j197$ @915 MHz (ST XRA00). Like [1], we match to $2R_c$ to have wideband characteristics over the FCC frequency range beginning from 902 MHz and ending at 928 MHz. The simulated results (obtained using ADS Momentum) are shown in Figure 2.

Detailed dimensions of both antennas are given in Table 1.

Both antennas are meandered within a 90 mm by 10 mm bounding box, and the conductor is assumed to be perfect with zero thickness. The permittivity of the substrate is $\epsilon_r = 4.34$ and its thickness is 1.6 mm.

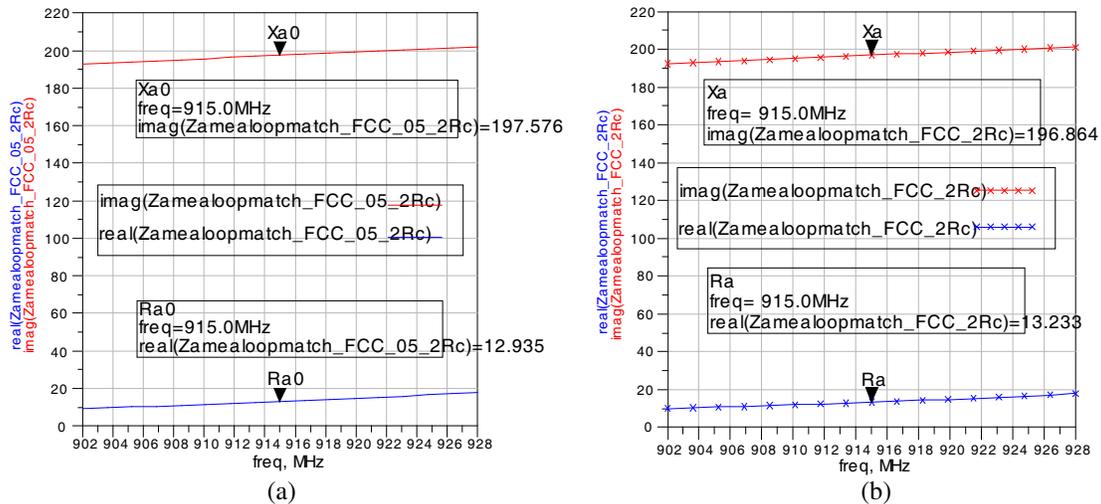


Figure 2: Impedance matching with $Z_c = 6.7 - j197$ @915 MHz for (a) line width $w = 0.5$ mm, (b) $w = 1$ mm.

Table 1: Geometric dimensions of the final layout designs giving a proper matching at $f_o = 915$ MHz.

| Antenna | Width (mm) | L_b (mm) | L_a (mm) | d (mm) | L_{dip} (mm) | L_f (mm) | H_o (mm) | L_o (mm) | Total conductor surface (mm ²) |
|---------|------------|------------|------------|----------|----------------|------------|------------|------------|--|
| #1 | 0.5 | 5.6 | 14.0 | 0.7 | 134.2 | 81 | 7.37 | 2.25 | 76.9 |
| #2 | 1.0 | 5.7 | 17.8 | 1.3 | 145.0 | 81 | 9.17 | 2.25 | 192 |

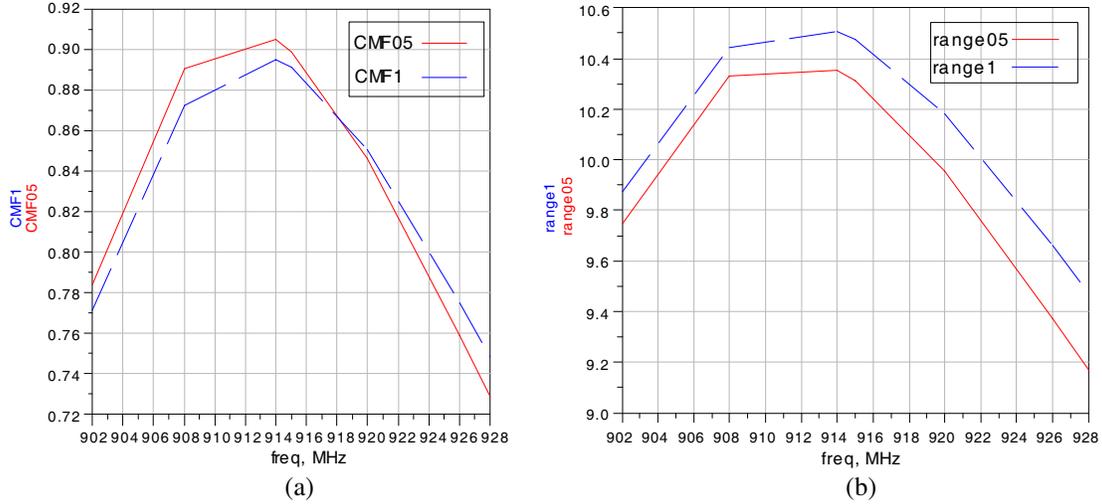


Figure 3: (a) Conjugate match factor of antennas presented in Table 1 with Z_c , (b) free space read range.

Table 2: Mean free space read range calculated for antenna 1 and 2.

| Antenna | Mean read range (m) |
|---------|---------------------|
| #1 | 9.89 |
| #2 | 10.08 |

3. POWER TRANSFER EFFICIENCY AND READ RANGE

The most used parameter to measure the quality of the impedance match between the antenna and the chip is the power transfer coefficient or the Conjugate Match Factor in some software manuals [7]. If CMF equals 1, then the antenna and the chip are perfectly matched, and maximum power is transferred between them.

$$CMF = \tau = 4 \cdot R_a \cdot R_c / |Z_a + Z_c|^2 \quad (3)$$

Figure 3(a) illustrates the conjugate match factor of both antennas over the FCC frequency range (902–928 MHz).

Assuming polarization matching between reader and tag antennas, the read range beyond which the tag can not be activated, is given by [6]:

$$d_{\max}(\theta, \phi) = \frac{c}{4\pi f} \sqrt{\frac{EIRP_R}{P_{chip}} \tau G_{tag}(\theta, \phi)} \quad (4)$$

where

- c = speed of light ($3 \cdot 10^8$ m/s),
- f = frequency,
- $EIRP$ = Effective Isotropic Radiated Power,
- P_{chip} = Power threshold to activate the tag chip,
- τ = power transfer coefficient, and
- G_{tag} = gain of the tag antenna.

The free space read range is calculated for an EIRP of 36 dBm and a P_{chip} of -14 dBm. Simulated results are presented in Figure 3(b). The mean read range for the two structures are given in Table 2.

We see that by using thinner line width we can obtain nearly 40% of conductor amount reduction at the cost of 2% of mean read range degradation.

4. CONCLUSION

Two non-uniform meandered dipoles have been designed and properly matched to an integrated chip, over the FCC frequency range (902–928 MHz) using an inductively coupled feed structure. Two values of the width line were tested (0.5 and 1 mm). Both values give acceptable (free space) read ranges. However, with the thinner line, we can reduce the amount of used conductor with a negligible degradation of the tag performance. This fact can be very useful for printed antennas using ink jet process.

REFERENCES

1. Son, H. W. and C. S. Tyo, “Design of RFID tag antennas using an inductively coupled feed,” *Electronics Letters*, Vol. 41, No. 18, 994–996, September 2005.
2. Choi, W., H. W. Son, C. Shin, J.-H. Bae, and G. Choi, “RFID tag antenna with a meandered dipole and inductively coupled feed,” *IEEE Antennas and Propagation Society, International Symposium*, 619–622, Albuquerque, NM, July 2006.
3. Michishita, N. and Y. Yamada, “A novel impedance matching structure for a dielectric loaded 0.05 wavelength small meander line antenna,” *IEEE Antennas and Propagation Society, International Symposium*, 1347–1350, Albuquerque, NM, July 2006.
4. Marrocco, G., “Gain-optimized self-resonant meander line antennas for RFID applications,” *IEEE Antennas and Wireless Propagation Letters*, Vol. 2, 302–305, 2003.
5. Toccafondi, A. and P. Braconi, “Compact load-bars meander line antenna for UHF RFID transponder,” *European Conference on Antennas and Propagation*, 804, Nice, France, November 2006.
6. Rao, K. V. S., P. V. Nikitin, and S. F. Lam, “Antenna design for UHF RFID tags: A review and a practical application,” *IEEE Transaction on Antennas and Propagation*, Vol. 53, No. 12, 3870–3876, December 2005.
7. IE3D User’s Manual, Release 14.1, May 2008, Zeland Software, Inc.

Design of Composite Electromagnetic Wave Absorber Made of Fine Spherical Metal Particles Dispersed in Polystyrene Resin

Y. Guan, K. Sakai, Y. Sato, and S. Yoshikado

Department of Electronics, Doshisha University, Japan

Abstract— For the purpose of designing electromagnetic wave absorbers which can operate at frequencies above 1 GHz, the ability to control the relative complex permeability μ_r^* and the relative complex permittivity ε_r^* for the composite made of fine metal particles dispersed in polystyrene resin was investigated. Method to calculate μ_r^* for sample made of dispersed metal particles having the distribution of the size is proposed. In order to clarify the effect of particle shape on μ_r^* and ε_r^* , samples made of dispersed non-spherical aluminium particles in polystyrene resin were prepared and the values of μ_r^* and ε_r^* were compared. Moreover, in order to clarify the effect of particle size on μ_r^* and ε_r^* , samples made of dispersed spherical particles of copper with the size of several micro meters were prepared. The calculated value of μ_r^* using the proposed formula considering the size distribution was more in agreement with the measure value compared with that calculated using formula applicable to single particle size. Proposed formula was also applicable to calculation of μ_r^* with multiple kinds of average particle size. Moreover, the calculated value of μ_r^* was qualitatively in agreement with the measured value for the samples with different shape or size. The measured values of μ_r^* for the composite made of the mixture of copper particles with two kinds of average particle size were in agreement with the calculated values. Thus, it was found that the values of μ_r^* and absorption characteristics can be controlled by the particle size distribution, the particle shape, or mixing ratio of particles with different average particle size.

1. INTRODUCTION

Telecommunication devices such as wireless local area network (LAN) and electronic toll collection system (ETC), use electromagnetic waves with frequencies higher than 1 GHz and frequencies used by these devices will shift to the high frequency range above 10 GHz in the future. For this reason, the development of an electromagnetic wave absorber suitable for these frequency bands is required [1–4]. To design such an absorber, control of the frequency dependences of the relative complex permeability μ_r^* and the relative complex permittivity ε_r^* is important for a metal-backed single layer electromagnetic absorber because the absorption of electromagnetic waves is determined by μ_r^* and ε_r^* . In particular, μ_r' , the real part of μ_r^* , must be less than unity to satisfy the non-reflective condition for incident electromagnetic wave at frequencies above several GHz. In the previous study, we have reported that the values of μ_r' for the composite made of aluminium particles dispersed in polystyrene resin became less than unity owing to the magnetic moment generated by an eddy current [3]. Thus, the values of μ_r^* and ε_r^* could satisfy the non-reflective condition at frequencies above several GHz. Moreover, magnetic loss arises by the eddy current loss and the composite made of metal particles dispersed in polystyrene which satisfies the non-reflective condition can absorb electromagnetic wave power in the high frequency range. In addition, it was also found from theoretical calculation that the frequency where μ_r^* begins to decrease shifts to the high frequency side and the improvement of absorption characteristics is expected in the high frequency range [2, 4]. Also, theoretical calculation suggested that the composite made of metal particles with two kinds of particle size has the different frequency dependences of μ_r^* compared with that with single particle size. However, the shape of the particle used for measurement is non-spherical, particle size had distribution, and the theoretical calculation is applicable only to single particle size. Moreover, the increase of the magnitude of ε_r' with increasing the amount of metal particles deteriorate the absorption characteristic of the electromagnetic wave absorbers [4]. The real part ε_r' of ε_r^* is determined by the capacitance between particles. It is speculated that the capacitance is affected by the contact between metal particles and is determined by the shape of a particle.

In this study, a method to calculate μ_r^* for sample dispersed metal particles having the distribution of the size for the composite is proposed. In order to clarify the effect of particle shape on μ_r^* and ε_r^* , samples dispersed non-spherical aluminium particles or spherical copper and polystyrene resin are prepared and the values of μ_r^* and ε_r^* were compared. Moreover, in order to clarify the effect of particle size on μ_r^* and ε_r^* , samples dispersed spherical particles of copper with the size

of several micro meters were prepared. The frequency dependences of μ_r^* , ε_r^* and the absorption characteristics of the composite were investigated in the frequency range from 1 to 40 GHz. Furthermore, the composite made of the mixture of copper particles with two kinds of average particle size was also evaluated to control μ_r^* and absorption characteristics more flexibly.

2. EXPERIMENTS

The metal particles used in this study were commercially available non-spherical aluminum particles and spherical copper particles. The average particle size (diameters) of non-spherical aluminum was approximately 8 μm (maker specification). The average particle size of spherical copper was approximately 1.5, 2.5, 5 or 10 μm (maker specification). The particle size was measured using a particle-size distribution meter (Microtrac 3300EX II). Chips of polystyrene resin were dissolved in acetone and the particles were mixed in until they were uniformly dispersed within the resin. The volume mixture ratio of aluminum particle was 16.4, 33.8 or 50.0 vol%. The volume mixture ratio of each copper particle was 50.0 vol% to investigate the effect of particle size. The mixture was then heated to melt the polystyrene resin and was hot-pressed at a pressure of 5 MPa to form a pellet. This was allowed to cool naturally to room temperature and was processed into a toroidal-core shape (outer diameter of approximately 7 mm inner diameter of approximately 3.04 mm) for use in a 7 mm coaxial line in the frequency range 1 to 12.4 GHz, or into a rectangular shape (P-band: 12.4–18 GHz, 15.80 mm \times 7.90 mm, K-band: 18–26.5 GHz, 10.67 mm \times 4.32 mm, R-band: 26.5–40 GHz, 7.11 mm \times 3.56 mm) for use in a waveguide. The sample was mounted inside the coaxial line or waveguide using silver past to ensure that no gap existed between the sample and the walls of the line/waveguide. The complex scattering matrix elements S_{11}^* (reflection coefficient) and S_{21}^* (transmission coefficient) for the TEM mode (coaxial line) or TE₁₀ mode (rectangular waveguide) were measured using a vector network analyzer (Agilent Technology, 8722ES) by the full-two-port method in the frequency range from 1 to 40 GHz. The values of μ_r^* ($\mu_r^* = \mu_r' - j\mu_r''$, $j = \sqrt{-1}$) and ε_r^* ($\varepsilon_r^* = \varepsilon_r' - j\varepsilon_r''$) were calculated from the data of both S_{11}^* and S_{21}^* . The complex reflection coefficient Γ^* for a metal backed single layer absorber was then determined from the values of μ_r^* and ε_r^* . The return loss R for each sample thickness was calculated from Γ^* using the relation $R = 20 \log_{10} |\Gamma^*|$. R was calculated at 0.1 mm intervals in the sample thickness range 0.1 to 30 mm.

3. THEORETICAL CALCULATION OF RELATIVE COMPLEX PERMEABILITY CONSIDERING THE SIZE DISTRIBUTION OF METAL PARTICLE

We have previously reported that μ_r' and μ_r'' for a composite composed of metal particles with the cylindrical shape of $2a$ in diameter and $2a$ in height dispersed in polystyrene resin are given as follows [4].

$$1 - \mu_r' = N\pi a A e^{-\frac{\omega}{\delta}} [\delta (2\delta^2 - 2a\delta + a^2) (e^{\frac{\omega}{\delta}} - 1) - a\delta(2\delta - a)] = N f(a) \quad (1)$$

$$\mu_r'' \approx \mu_r' \frac{P}{\frac{\omega}{2} \mu_r' \mu_0 \iiint H^2 dv} \quad (2)$$

Here, if $2a$ is constant, the number N of cylindrical particles per unit volume of the composite is given by

$$N = V/2\pi a^3. \quad (3)$$

Here, V is the volume mixture ratio of the particles in the composite, ω is the angular frequency of the electromagnetic wave and δ is the skin depth for an electromagnetic wave and is given by

$$\delta = \sqrt{\frac{2}{\omega \sigma \mu_0 \mu_{Mr}'}}. \quad (4)$$

Here, σ is the electrical conductivity of the metal, μ_0 is the permeability in free space, μ_{Mr}' is the real part of the relative complex permeability of the metal (μ_{Mr}' is almost 1) and H is the magnetic field in the cylinder. The integral in the nominator and the denominator of Equation (2) is given

as follow, respectively

$$P = N \frac{\pi a A^2 e^{-2\frac{a}{\delta}}}{\sigma} \left[-\frac{\delta^2}{4} (e^{2\frac{a}{\delta}} - 1) + \frac{a\delta}{2} e^{2\frac{a}{\delta}} \right] H_0^2 = Ng(a) H_0^2, \quad (5)$$

$$\begin{aligned} \frac{\omega}{2} \mu'_{dr} \mu_0 \iiint H^2 dv &= \frac{\omega}{2} \mu'_{dr} \mu_0 (1 - V) H_0^2 + N \pi a \omega \mu'_{dr} \mu_0 \left[a^2 + 4A\delta \left(\frac{1}{2} a^2 - \delta a + \delta^2 - \delta^2 e^{-\frac{a}{\delta}} \right) \right. \\ &\quad \left. + A^2 \delta^2 \left(a^2 - 3\delta a + \frac{7}{2} \delta^2 - 4\delta^2 e^{-\frac{a}{\delta}} + \frac{1}{2} e^{-2\frac{a}{\delta}} \delta^2 \right) \right] H_0^2 = \frac{\omega}{2} \mu_0 (1 - V) H_0^2 + Nh(a) H_0^2 \quad (6) \end{aligned}$$

Here, H_0 is the external magnetic field and A is given by

$$A = -\frac{(\omega \mu_0 \sigma)^2 \delta \left(\frac{1}{2} a^2 + \delta^2 - \delta^2 e^{-\frac{a}{\delta}} - a\delta \right)}{1 + \left(\frac{\omega \mu_0 \sigma}{a} \right)^2 \delta^2 \left[\left(\frac{1}{2} a^2 + \delta^2 - \delta^2 e^{-\frac{a}{\delta}} - a\delta \right) \right]^2}. \quad (7)$$

Here, to clarify the effect of magnetic moment generated by eddy current, the real part of μ_r^* is shown as $1 - \mu'_r$. It is found from above equations that the frequency where the values of μ'_r and μ''_r vary increases with decreasing the particle size of metal. This is because the skin depth δ is the function of frequency as given by Equation (4). Therefore, the frequency f' where the radius of the particle is equal to δ is given by

$$f' = \frac{1}{a^2 \pi \sigma \mu'_r \mu_0}. \quad (8)$$

It is found from Equation (7) that f' is inversely proportional to a and the frequency where the effect of magnetic moment generated by the eddy current is apparent can be shifted to a high frequency range using small sized particle.

However, the measured values of $1 - \mu'_r$ and μ''_r did not precisely agreed with the values calculated using Equations (1) and (2). The main factor for the disagreement might be the size distribution of metal particles. Therefore, the values of $1 - \mu'_r$ and μ''_r were calculated by taking the particle size distribution into consideration. The values of real and imaginary parts of the relative complex permeability μ'_{dr} and μ''_{dr} obtained by taking the particle size distribution into consideration are given by

$$1 - \mu'_{dr} = \int_0^\infty \rho(a) f(a) da. \quad (9)$$

$$\mu''_{dr} \approx \mu'_{dr} \frac{\int_0^\infty \rho(a) g(a) da}{\frac{\omega}{2} \mu'_{dr} \mu_0 (1 - V) + \int_0^\infty \rho(a) h(a) da}. \quad (10)$$

Here, $f(a)$, $g(a)$ and $h(a)$ are given by Equations (1), (5) and (6). The $\rho(a)$ denotes the distribution function of a and

$$\int_0^\infty \rho(a) da = N, \quad (11)$$

$$\int_0^\infty 2\pi a^3 \rho(a) da = V. \quad (12)$$

The $\rho(a)$ is given approximately by

$$\rho(a) = \sum_i \rho_{0i} \exp \left[-\left(\frac{a - a_{0i}}{w_i} \right)^2 \right]. \quad (13)$$

The values of ρ_{0i} , a_{0i} , and w_i are determined by the least square fitting using $\rho(a)$ given by Equation (13) for measured distribution of a' (radius of a spherical particle, $a' = \sqrt[3]{3/2}a$) assuming that the volume of a cylindrical particle is the same as that of a spherical particle. In the theoretical calculation considering the particle size distribution, the distribution of a' for all metal particles was measured. Figure 1 shows the measured distribution of a' for copper with the average diameter of 2.5 μm (maker specification) and the fitted line using $\rho(a)$ given by Equation (13).

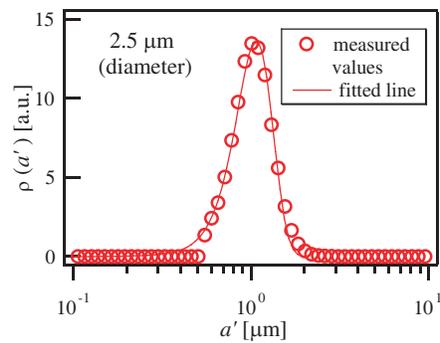


Figure 1: Measured distribution of a' for copper with the average diameter of $2.5 \mu\text{m}$ and the fitted line using $\rho(a)$ given by Equation (13).

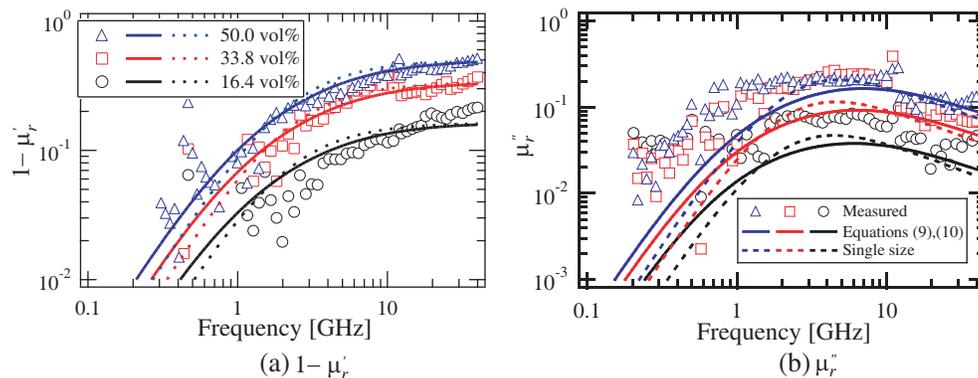


Figure 2: Frequency dependences of $1 - \mu_r'$ and μ_r'' for the composite made of aluminum particles with average sizes of $8 \mu\text{m}$. The volume mixture ratio of the aluminum was 16.4 vol%, 33.8 vol% or 50.0 vol%. Dashed lines show the values calculated using Equations (1) and (2). Solid lines show the values calculated using Equations (9) and (10).

4. RESULTS AND DISCUSSION

4.1. Frequency Dependence of $1 - \mu_r'$ and μ_r'' for a Composite Composed of Aluminum Particles

Figure 2 and shows the measured values of $1 - \mu_r'$ and μ_r'' for the composites containing non-spherical aluminum particles with diameters of approximately $8 \mu\text{m}$ in [4]. The volume mixture ratio was 16.4, 33.8 or 50.0 vol%. Dashed lines show the values calculated using Equations (1) and (2). Solid lines show the values calculated using Equations (9) and (10) considering the size distribution of aluminum particles. When the particle size distribution is considered, the solid lines showing the frequency dependences of $1 - \mu_r'$ and μ_r'' tend to approach to measured value. This result indicates that the frequency dependences of $1 - \mu_r'$ and μ_r'' are determined by not only the size but also the size distribution of metal particles and that the values of $1 - \mu_r'$ and μ_r'' can be controlled by not only the volume mixture ratio and particle size but also the particle size distribution of metal particles.

4.2. Frequency Dependence of $1 - \mu_r'$ and μ_r'' for a Composite Composed of Spherical Copper Particles with the Size of Several Micro Meter

To develop an absorber suitable for a high frequency range, the copper particles with the size of several micro meters were investigated, because good absorptions was expected by the theoretical calculation using Equations (9) and (10) when the size of metal particles decreases. Figure 3 shows the measured and calculated values of $1 - \mu_r'$ and μ_r'' for the composites containing copper particles with size of 1.5, 2.5, 5 or $10 \mu\text{m}$. The volume mixture ratio was 50.0 vol%. Plots show the measured values and solid lines shows the values calculated using Equations (9) and (10) considering the particle size distribution. Frequency dependence of $1 - \mu_r'$ quantitatively agreed with the calculated curves. However, the values of μ_r'' did not agreed with the calculated values quantitatively. One reason for this is considered to be the aggregation of copper particle in the process of mixing copper particle and polystyrene resin. In addition, the oxidation of copper is also considered because the conductivity of copper changes when the copper is oxidized.

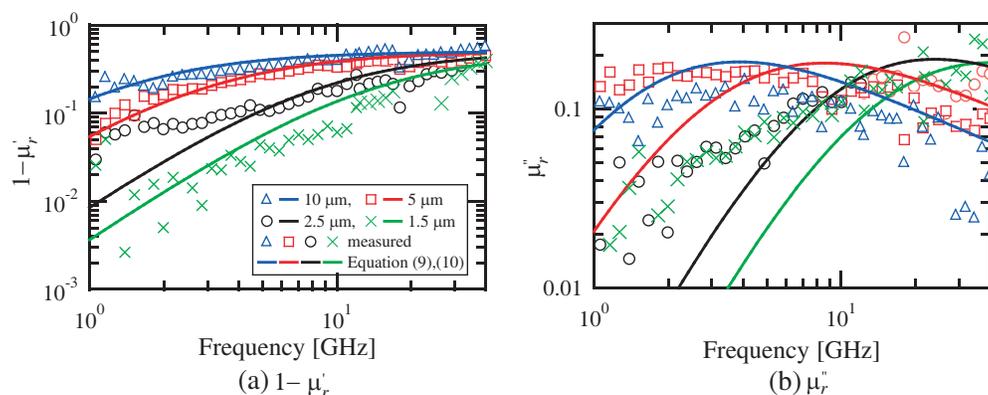


Figure 3: Frequency dependence of $1 - \mu_r'$ and μ_r'' for composites made of polystyrene resin and copper particles with size of 1.5, 2.5, 5, or 10 μm . The copper volume mixture ratio is 50.0 vol%. Lines show values calculated using Equations (9) and (10).

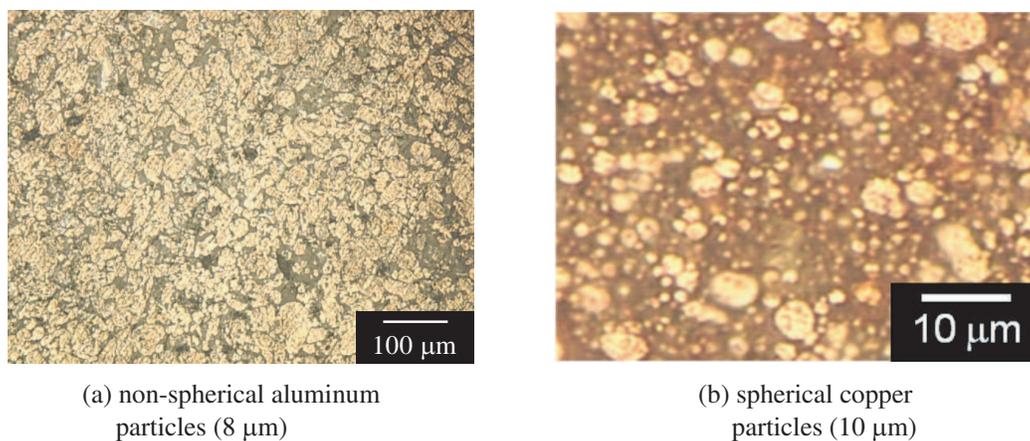


Figure 4: The microphotograph of the sample surface for the composite containing non-spherical aluminum particles with size of 8 μm and spherical copper particles with size of 10 μm . The volume mixture ratio is 50.0 vol%.

When the particle size decreased, the frequency where the frequency dependence of $1 - \mu_r'$ bent and μ_r'' became the maximum shifted to the high frequency side. This quantitative agreement is obtained by considering the particle size distribution in the theoretical calculation. In all particle size cases, the shape of the frequency dependence of μ_r'' similar to that obtained in the theoretical calculation using Equation (10).

From above results, the qualitative agreement with the measured and calculated values in both cases of aluminum and copper particles indicates that the theoretical calculation of μ_r' and μ_r'' are applicable to any metal particles with the different size of shape as shown in Figure 4 without magnetism simply substituting the value of σ for metal particle into Equations (9) and (10).

4.3. The ϵ_r' , ϵ_r'' for Composites Composed of Non-spherical Aluminum of Spherical Copper Particles

Figure 5 shows the frequency dependences of ϵ_r' for composites made of non-spherical aluminum particles of 8 μm and spherical copper particles of 10 μm . The measured values of ϵ_r' were found to be independent of frequency. Magnitude of ϵ_r' for composite made of non-spherical aluminum particles was larger by approximately 1.5 times than that made of spherical copper particles. The reason for the difference of ϵ_r' is considered to be as follows. The ϵ_r' is determined by the capacitance between particles. When non-spherical particles are dispersed in polystyrene resin, the contact of each particle tends to be plane contact as shown in Figure 4(a). Meanwhile, the contact of each spherical particle becomes point contact because of its particle shape as shown in Figure 4(b). Therefore, it is found that non-spherical particles are more likely to contact with each other compared with the spherical particles and as a result, the magnitude of ϵ_r' increases.

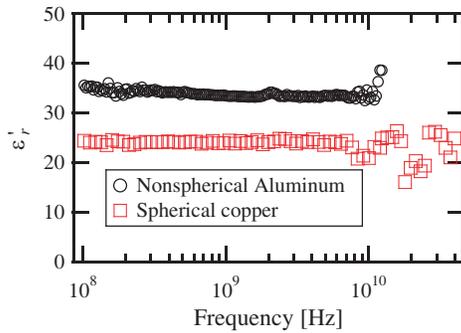


Figure 5: Frequency dependences of ε'_r for composites made of non-spherical aluminum particles of $8\ \mu\text{m}$ and spherical copper particles of $10\ \mu\text{m}$. The volume mixture ratio of metal is 50.0 vol%.

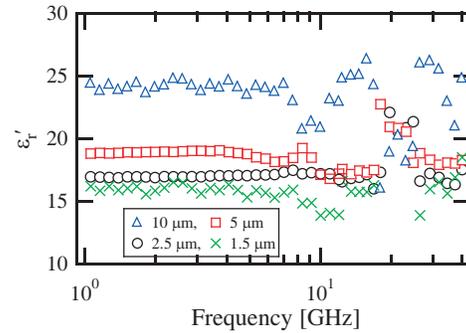
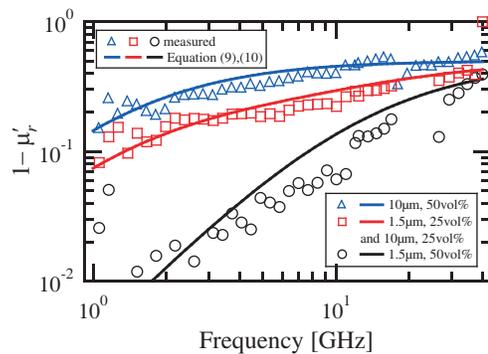
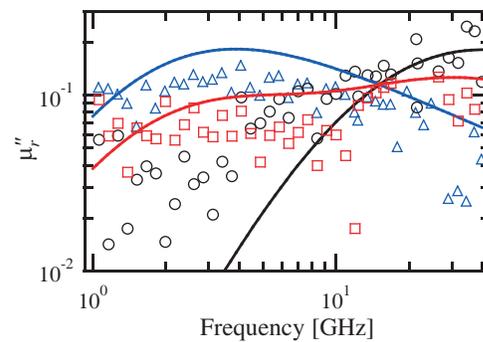


Figure 6: Frequency dependences of ε'_r for the composites containing copper particles with sizes of 1.5, 2.5, 5 and $10\ \mu\text{m}$. The copper volume mixture ratio is 50.0 vol%.



(a) $1 - \mu'_r$



(b) μ''_r

Figure 7: Frequency dependences of $1 - \mu'_r$ and μ''_r for the composite made of copper particles with two kinds size of 1.5 and $10\ \mu\text{m}$. The total volume mixture ratio of copper is 50 vol%. Solid lines show the values calculated using Equations (9) and (10).

Figure 6 shows the frequency dependences of ε'_r for the composites containing copper particles with sizes of 1.5, 2.5, 5 and $10\ \mu\text{m}$. The measured values of ε'_r were found to be independent of frequency and decreased with decreasing the particle size although these four composites were the same volume mixture ratio of 50.0 vol%. The values of both ε'_r and ε''_r decreased with decreasing the particle size. The reason for the difference is not known at the present stage.

4.4. Frequency Dependence of $1 - \mu'_r$ and μ''_r for a Composite Made of Copper Particles with Two Kinds of Sizes

To enlarge the control range of $1 - \mu'_r$ and μ''_r , composites made of copper particles with two kinds of size, such as 1.5 and $10\ \mu\text{m}$, was prepared changing the mixture ratio of each particle so that the total volume mixture rate is 50.0 vol%. Figure 7 shows the frequency dependences of $1 - \mu'_r$ and μ''_r for the composites. Plots are measure values and solid lines are values calculated using Equations (9), (10), and (13). The frequency dependence of both $1 - \mu'_r$ and μ''_r for the composite made of copper particles of 25 vol% for 1.5 μm and 25 vol% for 10 μm in the low frequency range was similar to that for composite made of copper particles of 50.0 vol% for 1.5 μm and in the high frequency range was similar to that made of copper particles of 50.0 vol% for 10 μm . In addition, these frequency dependences for the composite made of copper particles of 1.5 and $10\ \mu\text{m}$ qualitatively agreed with the calculated lines considering the distribution of two kinds of particle size. From this result, it is concluded that the frequency dependences of μ'_r and μ''_r can be changed by adjusting the particle size distribution of metal particle dispersed in resin.

The return loss of this composite was less than $-20\ \text{dB}$ in the frequency range 28 to 40 GHz. Although, the return loss of less than $-20\ \text{dB}$ could be obtained in the composite made of copper particles with single size ($2.5\ \mu\text{m}$), the frequency range where the return loss is less than $-20\ \text{dB}$ became broader when the composite made of copper particles with two kinds of size was used. Therefore, it was found that the composite made of metal particles with two or more kinds of size

contributed to the improvement of absorption characteristics.

5. CONCLUSIONS

The frequency dependence of μ_r^* calculated considering the size distribution of metal particles qualitatively agreed with the measured values for the composites made of non-spherical or spherical metal. In addition, the values of ε_r' and ε_r'' decreased when the spherical particles were used or the size became smaller. Furthermore, the frequency dependences of μ_r^* for the composite made of copper particles with two kinds of particle sizes could be explained by the qualitative theory considering the size distribution of metal particles well. Therefore, it is found that the composite made of metal particles with two or more kinds of size had different absorption characteristics from that for single particle size distribution and absorption characteristics of for a metal-backed single layer electromagnetic absorber wave absorber made of such composite can be improved

ACKNOWLEDGMENT

This work was supported by the Japan Society for the Promotion of Science (JSPS) and the RCAST of Doshisha University.

REFERENCES

1. Liu, J. R., M. Itoh, T. Horikawa, E. Taguchi, H. Mori, and K. Machida, "Iron based carbon nanocomposites for electromagnetic wave absorber with wide bandwidth in GHz range," *Appl. Phys. A*, Vol. 82, 509–513, 2006.
2. Yamane, T., A. Nishikata, and Y. Shimizu, "Resonance suppression of a spherical electromagnetic shielding enclosure by conductive dielectrics," *IEEE Trans. on Electromagnetic Compatibility*, Vol. 42, 441–448, 2000.
3. Wada, Y., N. Asano, K. Sakai, and S. Yoshikado, "Preparation and evaluation of composite electromagnetic wave absorbers made of fine aluminum particles dispersed in polystyrene medium," *PIERS Online*, Vol. 4, No. 8, 838–845, 2008.
4. Sakai K., Y. Wada, Y. Sato, and S. Yoshikado, "Design of composite electromagnetic wave absorber made of fine aluminum particles dispersed in polystyrene resin by controlling permeability," *PIERS Online*, Vol. 5, No. 7, 663–670, 2009.

Identifying EMC Interference Sources of a Microwave Transmission Module in Order to Locate Them

P. Descamps¹, G. Ngamani-Njomkoue², D. Pasquet¹, C. Tolant²,
D. Lesénéchal¹, and P. Eudeline¹

¹LaMIPS, Laboratoire Commun CRISMAT, UMR 6508 CNRS
6 boulevard Maréchal Juin, CAEN, France

²Thales Air Systems SA, Technical Unit Radio Frequency (TU-RF)
Technology & Innovation (REIRI-Y), Z.I. du Mont Jarret, Ymare 76520, France

Abstract— In the design of power microwave transmitter modules, many signals of different types coexist in a small space. It is possible that module dysfunctions occur due to the presence of electromagnetic disturbance type. It's the reason why it is important to understand and quantify these disturbances and in what form they spread in the structure. To do this we have studied a microwave simplified demonstrator placed in a shielded box by reducing the complex electromagnetic study of the complete transmission module taking into account a maximum of electromagnetic effects with and without shield. The electromagnetic (EMC) measurements of the demonstrator were conducted in a GTEM cell to isolate the demonstrator from all external interference and to compare to the 3D electromagnetic simulations, allowing to locate parts of the sensitive circuit to disturbances and to determine the sources of disturbances of the module. The simulation of the demonstrator using electromagnetic simulator allowed us to obtain a real internal electromagnetic mapping of the demonstrator. This mapping was used to characterize the electromagnetic behavior of the overall structure issue and identify sources of disturbance. This study should be complemented by measurements to scan the surface of the circuit using appropriate probes and serve to validate the method.

1. INTRODUCTION

The designers of the Transmission/Receivers (T/R) modules give particular attention to the problems of electromagnetic compatibility (EMC) of their products. Usually, RF power transmitters are electromagnetically shielded by metal cavities in order to be immune to external disturbances but also to avoid disturbing from close circuits. In radar applications, the output power of such transmitters can exceed hundreds of kilowatts which make very high radiated power emissions. Indeed, microwave components such as transistors find themselves in a severe EM environment that can cause failure mechanisms of their structure. Perfect knowledge of that EMC environment is then essential to carry out reliability studies. This work aims to provide a solution to the characterization of radiated emissions of microwave structures found in metallic housing, i.e., cavities.

Transmitter modules are built in a confined area containing several radiating elements. They may interact each other and it is often difficult to evaluate their mutual influence by conventional simulations. The modules are usually large and contain many small elements leading to a huge memory consuming electromagnetic simulation. The electromagnetic fields distribution can be evaluated by direct field measurement with the open module or by internal detection devices placed at strategic points. But this does not give the actual distribution when the module is closed or when the real devices are included. In particular, many box modes appear in the close module but could not exist when open. We have tried to propose a non-invasive approach using external EMC measurements in order to detect any anomaly that deviates from a standard behavior.

We have defined a simplified demonstrator in which the electromagnetic study is simple. In particular, the parasitic resonance modes are easy to calculate with an electromagnetic simulator. The EMC measurements are performed in a GTEM cell. The frequencies where electric field orientation perturbations appear are compared with resonance frequencies expected by the e-m simulation. Hence, we can obtain a comparison between the EMC measurement singularities and the parasitic phenomena in the module.

2. DEFINITION OF THE DEMONSTRATOR

The demonstrator is composed of few passive devices printed on a PCB (i.e., FR4 substrate) included into a box. The PCB layout is shown in Fig. 1. It consists of a microstrip line, a coupler and two patch antennas.

The microstrip line is a 150 mm-long 50Ω transmission line. The coupler has a 20 dB coupling at 3 GHz and has not been optimized for the directivity. He has been placed in order to discriminate the propagation direction of possible parasitic waves along the microstrip line. The antennas have been calculated for a 5 GHz resonance frequency and they are placed in perpendicular positions. In fact, one is magnetically coupled to the line and the other is electrically coupled. These devices also generate higher modes inside the demonstrator in order not to have the only quasi-TEM mode, but also higher modes.

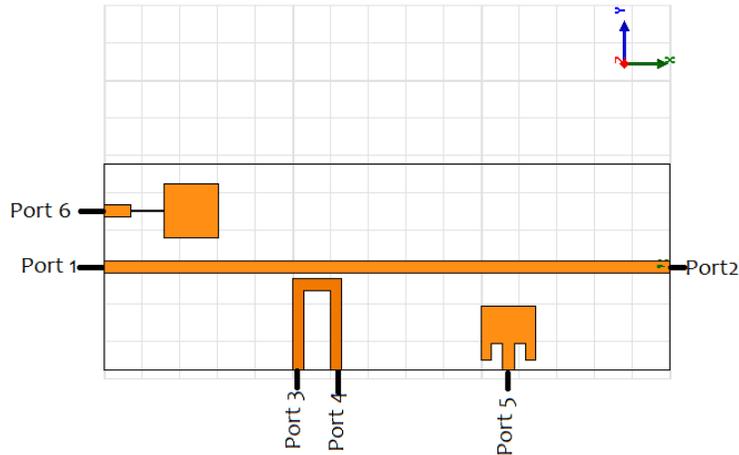


Figure 1: Layout of the PCB.

Table 1: Frequencies of the resonance modes in GHz (1st line calculated for the empty box, 2nd line simulated with substrate and without printed metal, 3rd line measured on S_{41}).

| Mode | TE ₁₀₁ | TE ₁₀₂ | TE ₁₀₃ | TE ₀₁₁ | TE ₀₁₂ | TM ₁₁₀ | TE ₁₀₄ | TM ₁₁₁ | TE ₁₁₁ |
|------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| Calculated | 2.90 | 3.38 | 4.05 | 4.12 | 4.47 | 4.84 | 4.84 | 4.94 | 4.94 |
| Simulated | 2.85 | 3.32 | 3.97 | 4.12 | 4.46 | 4.74 | 4.83 | 4.88 | 4.93 |
| Measured | 2.85 | 3.32 | 3.85 | 3.95 | 4.33 | 4.67 | 4.72 | 4.8 | 4.83 |

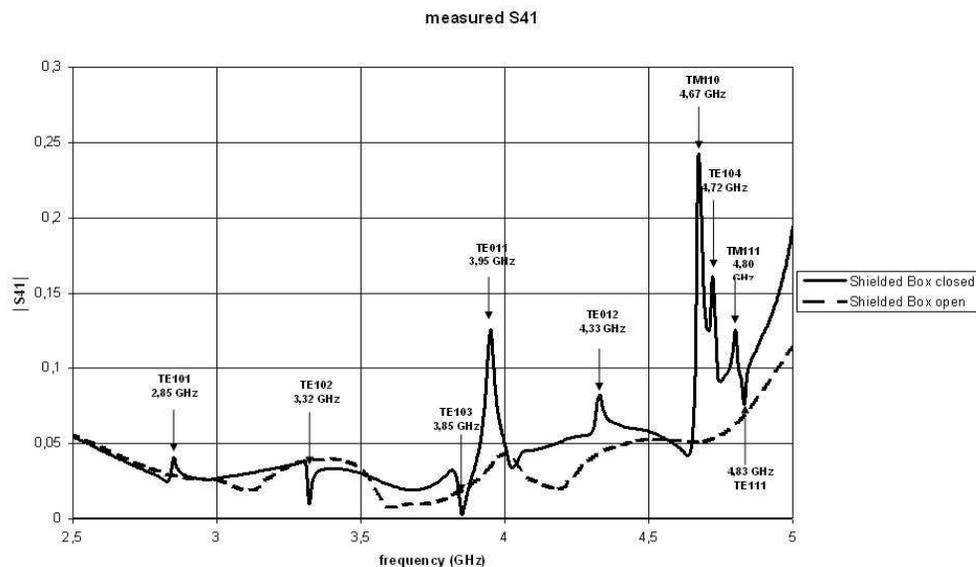


Figure 2: Measured S_{41} parameter of the demonstrator.

3. ELECTROMAGNETIC SIMULATION AND MEASUREMENT OF THE DEMONSTRATOR

In a first step, the resonance modes have been calculated for the empty and shielded box in order to identify the mode indices. Then, the resonance frequencies have been calculated by a 3D e-m simulator including the PCB substrate but without printed metal. The results are presented in Table 1.

The complete demonstrator has then been simulated and measured as a 6-port with the closed box. All the S -parameters between SMA ports were extracted with and without shielding box (open and close). The parasitic modes appear as perturbations (see Fig. 2 for the measured S_{41}) between the line and the coupled access of the coupler. The identified resonance frequencies are given in the third line of Table 1. We can notice in this figure that the measured resonance frequencies are very close to the calculated and simulated resonance modes of the cavity. These frequencies appear only when the cavity is closed as shown in Fig. 2.

4. GTEM CELL MEASUREMENTS

The closed demonstrator has been included into a GTEM cell. The input port 1 has been fed with a 20 dBm RF signal. The amplitude of the electric field has been measured by the septum in the three XYZ directions as defined in Fig. 6 [1].

Figure 4 shows the electric field components in a small bandwidth. From this result that we can observe that the E_y component of the electric field is lower than the other two.

The orientation of the electric field in xOz plane shown in Fig. 4 is more significant (0° corresponds to Oz axis) than in the other planes. Peaks appear for the resonance frequencies. They are identified in Fig. 5.

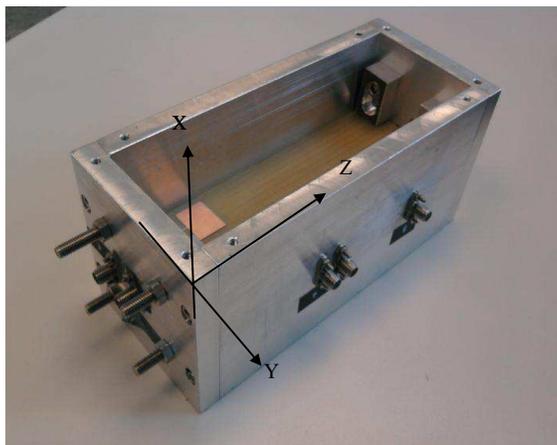


Figure 3: Definition of XYZ axis.

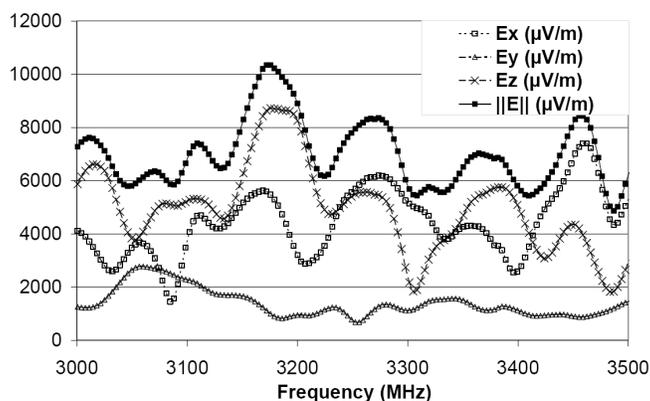


Figure 4: Electric field components measured with GTEM cell.

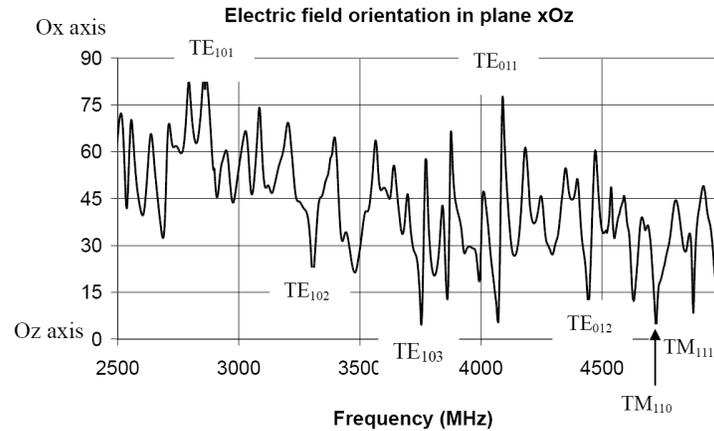


Figure 5: Orientation of the electric field in xOz plane (0° corresponds to Oz , 90° corresponds to Ox).

5. CONCLUSION

We have demonstrated that it is possible to observe the parasitic resonances from the measurement of the electric field outside the closed box. Once the frequencies have been identified, it is possible to know which resonance is involved. As the field patterns of all the modes are known, it is thus possible to know what the field repartition inside the box is and to know where it is adequate not to put components that are liable to radiate. These results show that we can better understand the interaction between the cavity and high frequency circuits. In the case of a closed box, the electromagnetic environment is disturbed by the presence of the cover. When the parasitic resonance frequencies are known, it will be possible to identify the parasitic resonance frequencies excited by the transmitter module.

ACKNOWLEDGMENT

This work has been supported by MOVEO cluster in the project AUDACE.

REFERENCES

1. Klingler, L., V. Deniau, S. Egot, B. Demoulin, and T. Sarkar, "Measuring radiation of small electronic equipment in three-dimensional TEM cells," *Electrical Performance of Electronic Packaging*, 21–24, 2001.
2. Fernández-López, P., C. Arcambal, D. Baudry, S. Verdeyme, and B. Mazari, "Radiation modeling and electromagnetic simulation of an active circuit," *7th Workshop on Electromagnetic Compatibility of Integrated Circuits (EMC Compo 09)*, paper no. 58, Toulouse, France, November 17–19, 2009.
3. Jansen, R. H. and L. Wiemer, "Full-wave theory based development of mm-wave circuit models for microstrip open end, gap, step, bend and tee," *IEEE MTT-S Digest*, 779–782, 1989.

Low Frequency Monopole-like Small Metamaterial Antenna

Nabil Dakhli¹, Mohamed Hayouni¹, Fethi Choubani¹, and Jacques David²

¹Research Unit of Telecommunication Systems (6'Tel) at Sup'COM, Tunisia

²ENSEEIH, France

Abstract— In this paper; we propose a monopole-like small antenna resonating at low frequency. The main radiating elements of the antenna are the two vias which are connected in series to an inductor at each edge of a transmission line in between. The inductance values are tuned in order to calibrate the flowing currents in the posts to be in phase so that the radiated fields are constructive in the far field region. Full-wave simulations were carried in order to study the properties of the antenna at 434 MHz. The antenna offers a fractional bandwidth of 1.08%, a radiation efficiency exceeding 50%, and a monopole-like radiation pattern.

1. INTRODUCTION

Metamaterials are artificial structures and represent a new breakthrough in electromagnetic engineering. They exhibit unique electromagnetic properties such as negative electric permittivity negative magnetic permeability and infinite wavelength propagation at specific non-zero frequency. A class of metamaterials is engineered by an artificial transmission line (TL) or LC loaded transmission line [1, 2] called in literature, CRLH (Composite Right/Left Handed) or DNG (Double Negative) structures.

In this paper, we present a low frequency monopole-like small resonant antenna working at 434 MHz based on ENG (Epsilon Negative) TL, implemented in microstrip technology. The antenna was designed to incur an insertion phase of 0° at the zeroth-order mode, which allows the two inductive vias (posts) to be fed in phase and act as the main radiating elements, the flowing currents in each of them are equal in magnitude and phase which produce a monopole-like radiation pattern.

2. THEORY OF LC LOADED TRANSMISSION LINE

The equivalent electric models of a lossless DNG and ENG unit cells are given by Figure 1.

The DNG unit cell can be represented by a host transmission line with characteristic impedance Z_0 and physical length $d < \frac{\lambda}{4}$ (homogeneity condition) loaded with a series capacitance C_0 and shunt inductance L_0 . For the ENG unit cell, there's no series capacitance.

By applying Bloch-Floquet theory to the DNG and ENG unit cells, we obtain the following dispersion relations [2, 5]

$$\begin{aligned}\beta_{DNG} &\approx \pm\omega\sqrt{\left[L - \frac{1}{\omega^2 C_0 d}\right] \left[C - \frac{1}{\omega^2 L_0 d}\right]} \\ \beta_{ENG} &\approx \pm\omega\sqrt{L \left[C - \frac{1}{\omega^2 L_0 d}\right]}\end{aligned}\quad (1)$$

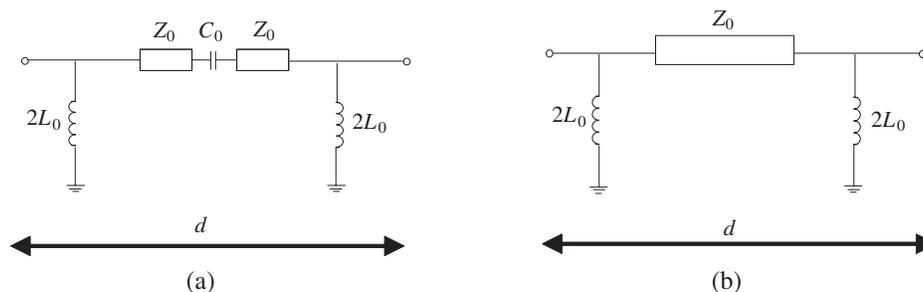


Figure 1: Electric model of (a) DNG TL and (b) ENG TL unit cells.

where L and C are the equivalent distributed inductance and capacitance, respectively of a short section of conventional transmission line of length d .

For a short-ended CRLH transmission line, the resonant condition of $\beta_m l = m\pi$ must be satisfied, where $m = (0, \pm 1, \pm 2, \pm 3, \dots, \pm(N-1))$ is the resonance mode for the DNG TL and $m = (0, +1, +2, +3, \dots, +(N-1))$ for the ENG TL, with $l = N \times d$ the total physical length of the transmission line [2]. In the case where $m = 0$ (zeroth-order resonance mode), an infinite guided wavelength is supported which corresponds to 0° phase shift.

$$\phi = -\beta d = -\frac{2\pi}{\lambda_g} d \quad (2)$$

The exact determination of the Eigen-frequencies for which $\beta = 0$ can be solved by studying the dispersion relations given by (1). The zeroth-order resonance frequency depends only on L_0/C (shunt-tank resonant circuit) and L/C_0 (series resonant circuit) parameters for the ENG TL, DNG TL respectively and not on the physical length as in pure right-handed TL, as a result; the number N of CRLH unit cells can be set arbitrarily.

3. ANTENNA DESIGN

A monopole-like small antenna composed of one ENG TL unit cell, consists of two posts closely located and connected each to an edge of a transmission line by series inductors (L_1, L_2). The posts can radiate in phase by simply tuning the different values of inductance of the series inductors so that the antenna can operate at only one Eigen-frequency ($N = 1$) corresponding to $\beta = 0 \rightarrow \phi = 0^\circ$. In the other hand, the input impedance Z_{in} of the radiator can be matched to 50Ω , the physical size of the transmission line is kept unchanged [3, 4]. The zeroth-order frequency $f_0 = \frac{1}{\sqrt{L_0 C_d}}$ is lowered by increasing L_1 and L_2 , such that $L_1 = L_0$ and $L_1 = \frac{1}{\alpha} L_2$ with α a scaling factor [5]. Each via can be modeled as a radiation resistance R_r and an internal inductor $L_v \ll L_1, L_2$ as shown in Figure 2.

It was shown that the antenna's input resistance $\Re(Z_{in})$ is expressed by the following equation [5, 7]

$$\Re(Z_{in}) = n^2 R_r \left(1 + \frac{1}{\alpha^2} \right) \quad (3)$$

$$\text{and } R_r = 160\pi^2 \left(\frac{h}{\lambda} \right)^2 \quad (4)$$

where n and h represent the number and the height of the posts in the antenna respectively.

Thus an appropriate chosen scaling factor α provides an easy way to adjust the radiator's input impedance.

The ZOR antenna was designed to operate at 434 MHz using microstrip technology on FR4 epoxy substrate with dielectric constant $\epsilon_r = 4.4$, footprint of $L_{sub} \times W_{sub} = 34.6 \text{ mm} \times 3 \text{ mm}$ and thickness $h_d = 1.52 \text{ mm}$, while a ground plane of dimension $100 \text{ mm} \times 100 \text{ mm}$ is separated from the substrate by 50 mm an air layer height, corresponding to $\lambda_0/14$. The etched transmission line has a size of $L_{TL} \times W_{TL} = 24.5 \text{ mm} \times 3 \text{ mm}$. The diameters of the posts are equal to 0.9 mm. Top and side views of the proposed antenna are shown in Figure 3.

The radiation quality factor of the antenna is related to its volume by,

$$Q_r = \frac{\omega_0 W_T}{P_{rad}} \propto \frac{1}{V} \quad (5)$$

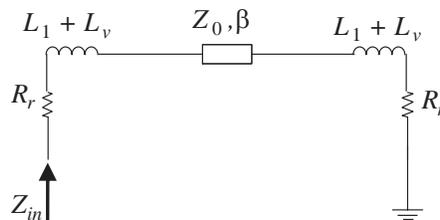


Figure 2: Electric equivalent circuit of the ENG TL.

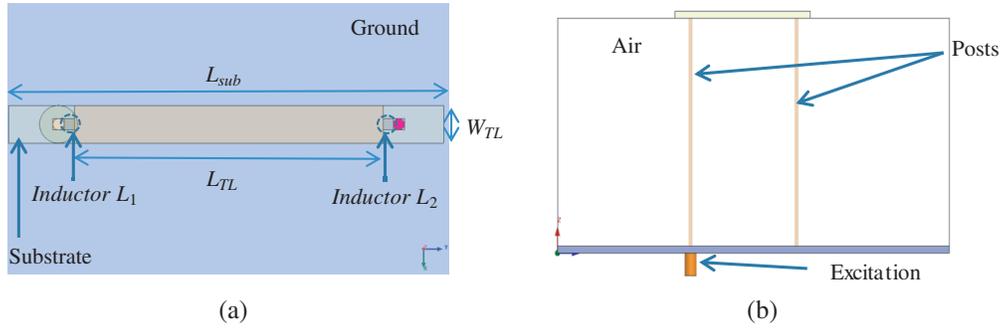
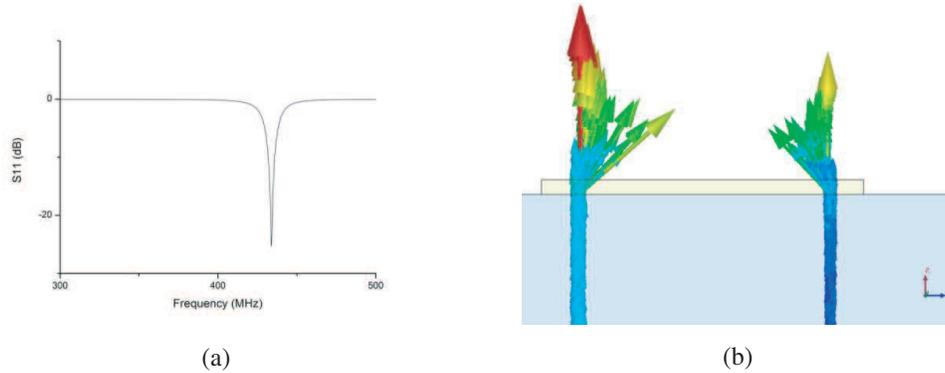


Figure 3: (a) Top and (b) side view of the ZOR antenna layout.


 Figure 4: (a) Return loss S_{11} , (b) flowing currents in the two posts of the ZOR antenna.

with ω_0 the zeorth-order frequency, W_T the total stored energy, P_{rad} the radiated energy and V the antenna volume. The quality factor Q of this radiator is calculated by the following expression [8]:

$$Q(\omega) \approx \frac{S - 1}{FBW\sqrt{S}} \quad (6)$$

where $FBW = \frac{f_+ - f_-}{f_0}$ is the fractional matched bandwidth band $S = VSWR(f_+) = VSWR(f_-)$, S is chosen to be equal to 2.

The antenna has $\frac{\lambda_0}{230} \times \frac{\lambda_0}{25} \times \frac{\lambda_0}{14}$ volume; the radius of an imaginary sphere encompassing the maximum dimension of this antenna was approximately $a = 51.92$ mm. Using the criterion of $a = \frac{\lambda}{4\pi}$, this structure would be considered electrically small at the operating frequency. For the excitation, a coaxial cable was used with inner and outer radii fixed to 0.45 mm and 1.05 mm respectively. The chosen values of the inductances are $L_1 = \frac{L_2}{1.416} = 300$ nH and $L_2 = 425$ nH for which the antenna is matched to 50Ω and the currents flowing in vias are in phase.

4. SIMULATION RESULTS

A full-wave electromagnetic simulation carried using Ansoft HFSS gives the return loss coefficient $S(1, 1)$, displayed in Figure 4(a). The zeorth-order resonance mode is located at $f_0 = 433.9$ MHz with corresponding $S_{11} = -25.39$ dB. The simulation suggests that the antenna's fractional bandwidth of operation is in the order of $FBW \approx 1.08\%$ and thus $Q \approx \frac{1}{FBW\sqrt{2}} = 65.47$.

In Figure 4(b), the currents that are supported by the two radiating posts are shown; these currents are in phase and thus maximize radiation in far field region. The input impedance of the radiator is $Z_{in} = 45.04 - j1.39$. The radiation resistance of one post can be deduced either from Equation (3) or Equation (4), from the first, $R_r \approx 7.51 \Omega$, where $n = 2$ and $\alpha = 1.416$, from the latter $R_r = 8.25 \Omega$, where $h = 50 \Omega$ and $\lambda = 691.4$ mm, it is observed that the two calculated values of R_r are slightly different. The radiation efficiency was enhanced to approximately 70.75% at this frequency. A maximum simulated gain and directivity obtained at f_0 were of values 1.339 dBi and 2.842 dB respectively. The radiation patterns E and H shown in Figure 5(a) and Figure 5(b), respectively, reveal the expected monopolar radiation pattern at f_0 .

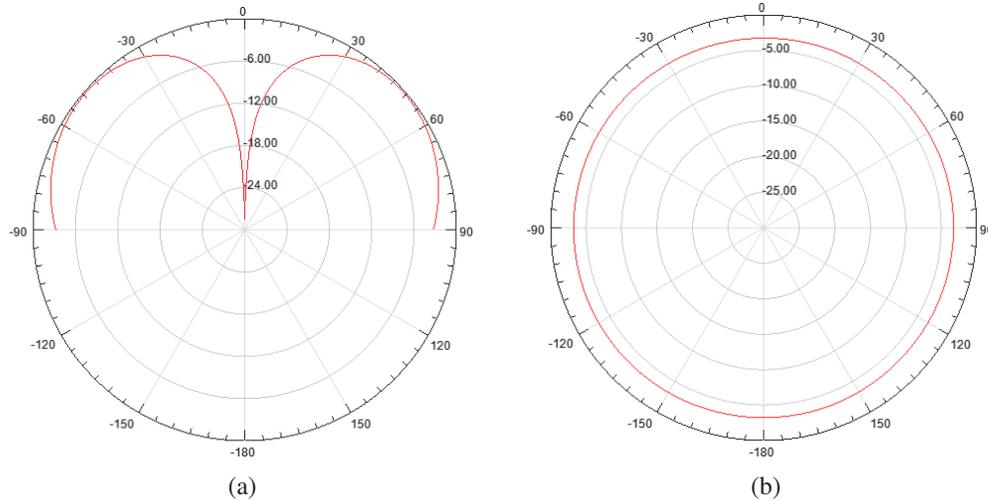


Figure 5: Radiation patterns (a) E and (b) H .

5. CONCLUSION

An electrically small monopole-like zeroth-order antenna has been presented and demonstrated to exhibit higher radiation efficiency and acceptable gain at the fundamental resonant mode than other metamaterial antennas. Generation of a vertical uniform electric field by the antenna is confirmed at 434 MHz.

REFERENCES

1. Lai, A., K. Leong, and T. Itoh, "Infinite wavelength resonant antennas with monopolar radiation pattern based on periodic structures," *IEEE Trans. Antenna Propag.*, Vol. 55, No. 3, 868–875, March 2007.
2. Caloz, C. and T. Itoh, *Electromagnetic Metamaterials: Transmission Line Theory and Microwave Applications*, John Wiley & Sons, Press, 2006.
3. Lee, D. H., A. Chauraya, J. C. Vardaxoglou, and W. S. Park, "Low frequency tunable metamaterial small antenna structure," *The Second European Conference on Antennas and Propagation, EuCAP 2007*, 375–379, November 2007.
4. Kokkinos, T., A. P. Feresidis, and J. C. Vardaxoglou, "A low-profile monopole-like small antenna with embedded metamaterial spiral-based matching network," *The Second European Conference on Antennas and Propagation, EuCAP 2007*, 261–265, November 2007.
5. Qureshi, F., M. A. Antoniadis, and G. V. Eleftheriades, "A compact and low-profile metamaterial ring antenna with vertical polarization," *IEEE Antennas and Wireless Propagation Letters*, Vol. 4, 333–336, September 2005.
6. Best, S. R., "The performance properties of electrically small resonant multiple-arm folded wire antennas," *IEEE Antennas and Propagation Magazine*, Vol. 47, No. 4, 13–27, February 2006.
7. Antoniadis, M. A. and G. V. Eleftheriades, "A folded-monopole model for electrically small NRI-TL metamaterial antennas," *IEEE Antennas and Wireless Propagation Letters*, Vol. 4, 333–336, September 2005.
8. Yaghjian, A. D. and S. R. Best, "Impedance, bandwidth, and Q of antennas," *IEEE Trans. Antenna Propag.*, Vol. 7, 425–428, November 2008.

Optimization of a Patch Antenna Performances Using a Left Handed Metamaterial

Akram Boubakri and Jamel Bel Hadj Tahar
 Graduate School of Communications of Tunis, Sup'Com
 University of 7 November of Carthage
 Route de Raoued, Cité Elghazela, Ariana 2083, Tunisia

Abstract— Among many solution's adopted to improve antenna performances, using a left handed metamaterial in it's near environment can lead first to achieve an amelioration in gain, directivity and size and second to avoid drawbacks shown in other solution's. In this example, the use of left handed metamaterial above a circular patch antenna operating at 10 GHz, increase the gain of 5.2 dB and offer a better directivity.

1. INTRODUCTION

After the first experimental verification of left handed metamaterials (LHM) [1, 3],... research into those structures has exponentially grown. In particular, LH metamaterials has made it possible to realize novel microwave devices such as negative refractive lens and small resonant antenna etc.

However use of an array of several antenna, or placing a high value permittivity or permeability structure above the patch antenna, are solutions which offer enhancement of antenna properties. LH metamaterials give the opportunity to improve antenna properties with avoiding many draw back in other solutions like costs, coupling in antenna array...

To our knowledge there are two way of using LHM in antenna conception, either to build it with LHM [4, 5] or to place this last at the near environment of the structure which is the solution proposed by Nakano et al. [6].

Here, the work is based on the use of a LHM in the near environment of a circular patch antenna in order to enhance its performances. With the LHM technique, the gain and directivity of the patch antenna are greatly improved and the results obtained using Ansoft's HFSS simulations show an improvement of 5.2 dB for the gain and a better directivity of the antenna.

2. PATCH ANTENNA PARAMETERS

The antenna used for this study is a circular patch designed on a dielectric board which has a relative permittivity of 2.5 and permeability of 1. The radius r of the patch antenna is about 12 mm and the working frequency is about 10 GHz. The simulated S_{11} parameter of the antenna is given on Fig. 1(a) and the 3-D radiation pattern is shown in Fig. 1(b). As it's shown, the circular patch antenna has a gain of 6.1 dB and a return loss of about -11 dB.

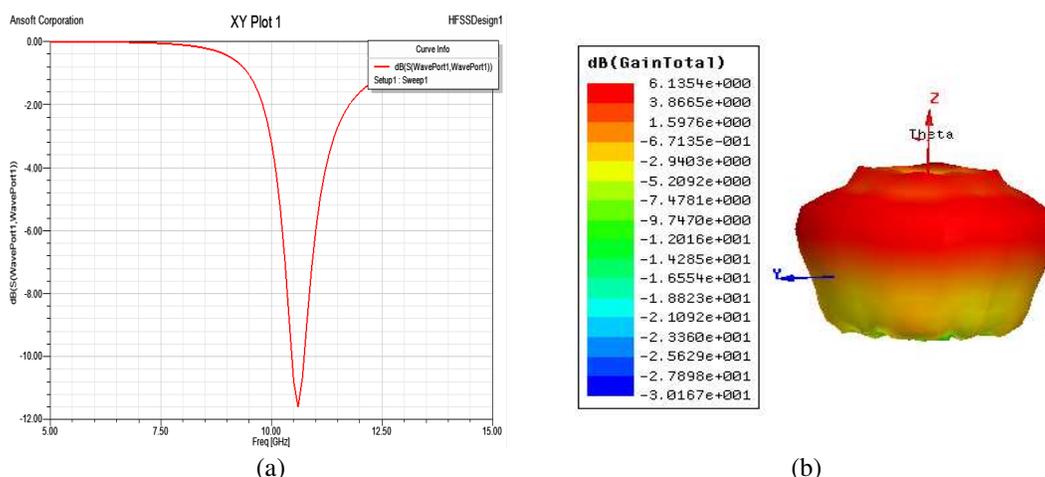


Figure 1: (a) S_{11} parameter of circular patch antenna. (b) 3D radiation pattern.

3. PATCH ANTENNA BASED ON LHM

The LHM used is characterized by an effective permittivity and permeability of -2.5 . Such a medium can be achieved by constructing a periodic array composed of SRRs and metallic wires [2].

Placing the LHM in the near environment of the patch antenna studied before, has shown its influence on the antenna performances.

The antenna with the LHM present's a gain of 11.3 dBi and an aperture at -3 dB of 60° , compared to the antenna without LHM, this aperture has been decreased and the backward radiation has been weakened. Besides the distance had been varied to determinate its influence on antenna

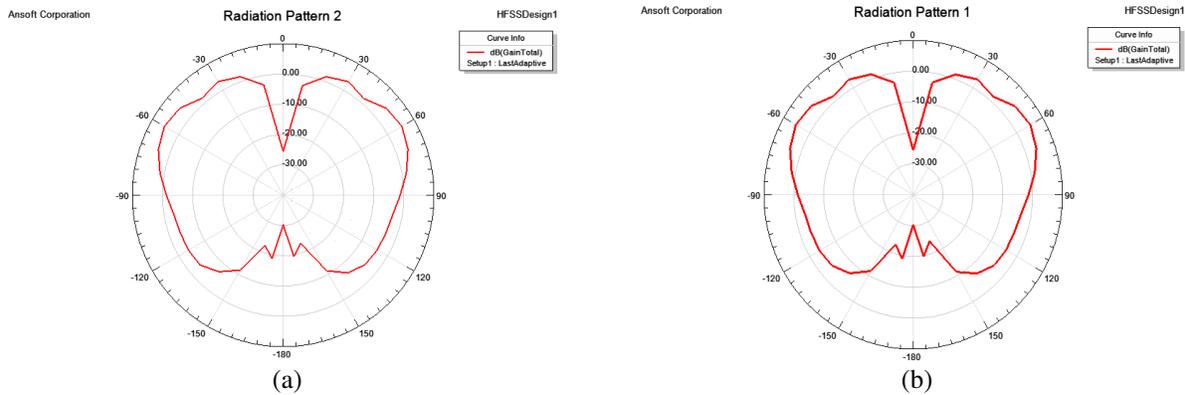


Figure 2: (a) Radiation pattern E plane. (b) Radiation pattern H plane.

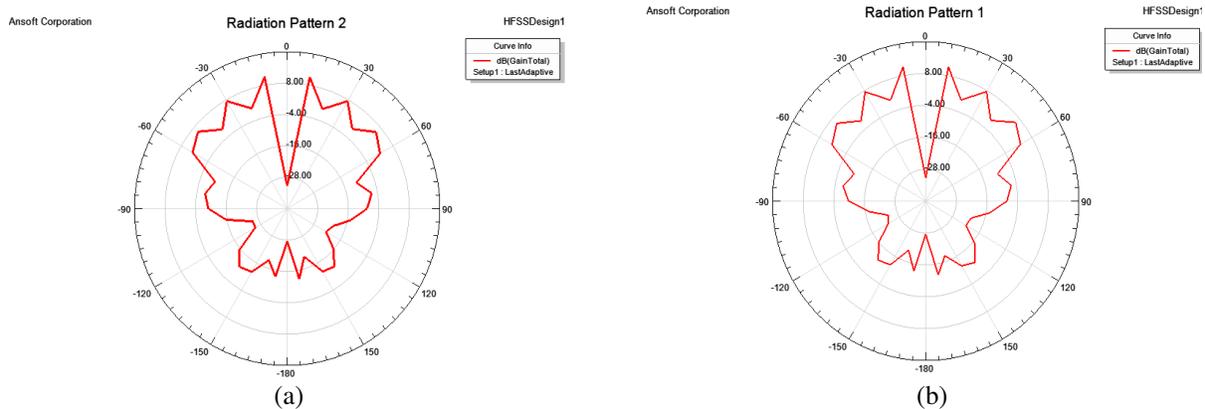


Figure 3: (a) Radiation pattern H plane. (b) Radiation pattern E plane.

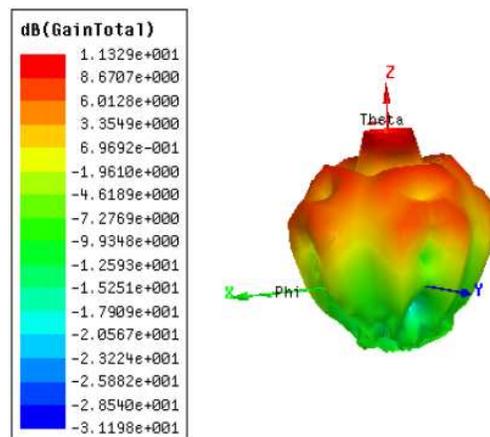


Figure 4: 3D radiation pattern of antenna with LHM above.

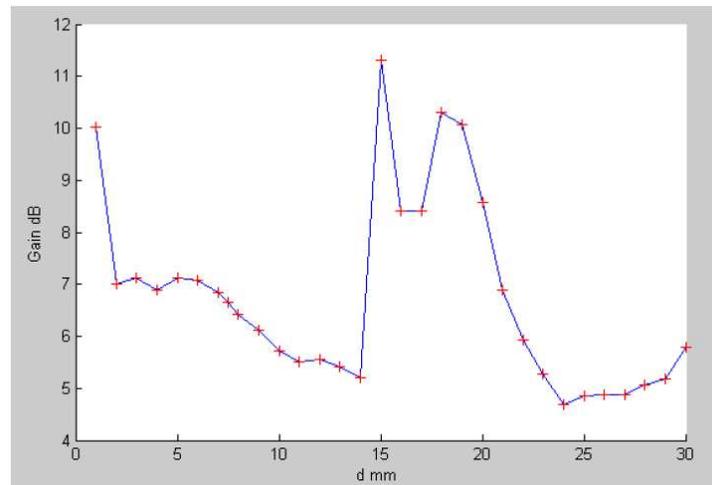


Figure 5: Gain versus distance between antenna and LHM.

gain. The optimal value (a gain of 11.3 dB) was obtained for $d = \lambda/2$ (15 mm).

4. CONCLUSION

In this work dealing with patch antenna performance's amelioration, we show that an improvement of the gain by 5.2 dB is achieved and a more directional antenna is obtained by the use of a LHM in its near environment. In addition we realized a miniaturization on antenna size by replacing the gain given by an array of three antenna at least by just one antenna with LHM and also we have shown the relation gain versus distance d (between antenna and LHM) and how the gain depend on d , the optimal value was for $d = \lambda/2$.

REFERENCES

1. Smith, D. R. and N. Kroll, "Negative index refraction in left handed metamaterials," *Physical Review Letter*, Vol. 85, No. 14, October 2000.
2. Smith, D. R., et al., "Composite medium with simultaneously negative permeability and permittivity," *Physical Review Letters*, Vol. 84, No. 18, 4184–4187, May 2000.
3. Pendry, J. B., "Negative refraction makes a perfect lens," *Physical Review Letter*, Vol. 85, No. 18, 3966–3969, October 2000.
4. Wu, B.-I., W. Wang, J. Pacheco, X. Chen, J. Lu, T. M. Grzegorzczuk, J. A. Kong, P. Kao, P. A. Theophilakes, and M. J. Hogan, "Anisotropic metamaterials as antenna substrate to enhance directivity," *Microwave Optical Technology Letters*, Vol. 48, No. 4, 681–683, April 2006.
5. Wu, B.-I., W. Wang, J. Pacheco, X. Chen, T. M. Grzegorzczuk, and J. A. Kong, "A study of using metamaterials as antenna substrate to enhance gain," *Progress In Electromagnetics Research*, Vol. 51, 295–328, 2005.
6. Nakano, H., M. Ikeda, K. Hitosugi, and J. Yamauchi, "A spiral antenna sandwiched by dielectric layers," *IEEE Trans. Antennas Propag.*, Vol. 52, No. 6, January 2004.

Global Maps of TEC and Conditions of Radio Wave Propagation in the Mediterranean Area

O. A. Maltseva, N. S. Mozhaeva, and G. M. Glebova
Institute of Physics, Southern Federal University, Russia

Abstract— On the example of the Mediterranean area it is showed that the joint use of global maps of the total electron content and data of ionospheric vertical sounding can be useful for: 1) the choice of maps, the most appropriate experimental values of the critical frequency foF2, 2) filling in missing data of foF2, 3) determination of foF2 in zone with a radius greater than the radius of spatial correlation of foF2. Such use allows to investigate and more accurately predict the conditions of the propagation of radio waves in the ionosphere.

1. INTRODUCTION

Conditions of radio wave propagation in the ionosphere are closely related to the critical frequency foF2 (the maximum electron density NmF2). In recent years, the ionospheric total electron content TEC is increasingly being used for assessing the state of the environment. The transition from ground-based monitoring of the ionosphere by vertical sounding (NmF2) to satellite remote sensing (TEC) needs knowledge of the coefficient of proportionality τ ($TEC = \tau * NmF2$) and the absolute values of TEC to determine NmF2. Study of the behavior of τ was conducted in many papers on specific GPS stations for the European region, including the Mediterranean area (e.g., [1]). In the given paper, it is proposed to use the global maps of TEC to analyze the behavior of τ and assess its capacity to determine NmF2 (foF2). The aim of this work is to show that the hyperbolic regression dependence τ (NmF2) allows to: 1) select a map, the most appropriate experimental values of foF2, 2) fill in missing data foF2, 3) determine foF2 in the zone with a radius exceeding the radius of spatial correlation of foF2. These aspects define the relevant sections of the article. Discussion of results and additional possibilities to use τ is given in the final section. According to the project MIRTO [2], created to provide ionospheric data for this area, the boundaries of the region are $-10 \div 30^\circ N$ in longitude and $10 \div 40^\circ$ in latitude. Main stations are ElArenosillo, Ebre, Rome, therefore, in this paper the connection between the global maps and propagation conditions is studied, mainly on the example of these stations.

2. CHOISE OF MAPS

Global maps of total electron content (TEC) are a huge array of data provided by various organizations: Jet Propulsion Laboratory (JPL), The University of Berne (CODE), Universitat Politecnica de Catalunya (UPC), European Space Agency (ESA). These are available on the Web site <ftp://cddis.gsfc.nasa.gov/gps/products/ionex> as IONEX files and cover a period longer than the solar activity cycle. As is known, the values of various maps differ from each other (e.g., [3]) and from the values obtained at individual stations (e.g., [4]) due to various methods of determining the TEC. Differences can reach a factor of 1.5–2. In this paper, the choice of maps is associated with the assessment of the possibility to determine the critical frequency using the array of global TEC maps, since knowledge of these frequencies and M3000 factors largely determine the propagation of radio waves, particularly in the HF band. Fig. 1 shows examples of the differences of TEC for various maps from [3] for st. Kiruna (25.04.2002) and from calculations for st. Ebre (20.03.2001). Notations for JPL (◆), CODE (■), UPC (▲), ESA (●) are similar. The left panel of Fig. 1 contains also curves for RAL (dashed line), the IRI model (pluses). The choice of map is performed by using an equivalent slab thickness of the ionosphere τ , which is the coefficient of proportionality between the TEC and NmF2 ($TEC = \tau * NmF2$). Since there is an inverse proportionality between τ and NmF2 ($\tau = TEC/NmF2$), natural to use a hyperbolic dependence τ (hyp) = $b_0 + b_1/Nm$ to build a regression relationship, in which $Nm = foF2 * foF2$ and which is defined for each map. Mean absolute deviation $|\Delta foF2| = |foF2(obs) - foF2(hyp)|$ of critical frequencies corresponding to the hyperbolic regression from the experimental ones is used to identify the best maps. Fig. 2 shows deviations for April, 2001 and 2007 for stations Ebre and Rome.

The difference between the maps is determined by scattering τ from the regression curve. Fig. 3 shows the corresponding values of τ together with their regression curves.

Some statistical results are given in Table 1 for stations Ebre and Rome and 2001 (first 8 lines), 2007 (second 8 lines). The smallest values are indicated by bold type.

It is seen that in this area the foF2 values closest to the experimental ones are provided by the regression curves for the map JPL, so the remaining results are presented for calculations using the TEC (JPL).

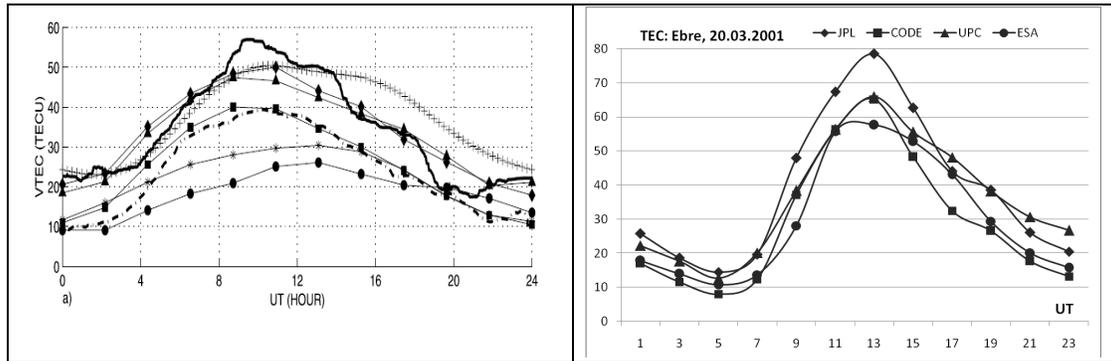


Figure 1: Examples of differences between TEC maps from [3] and additional calculations.

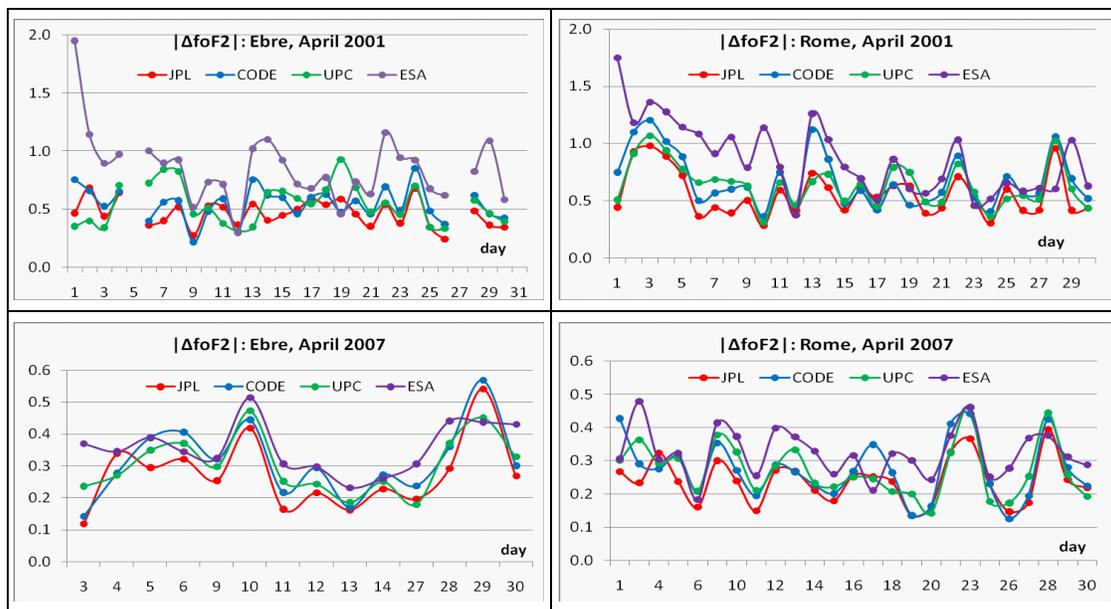


Figure 2: Illustrations for the choice of maps.

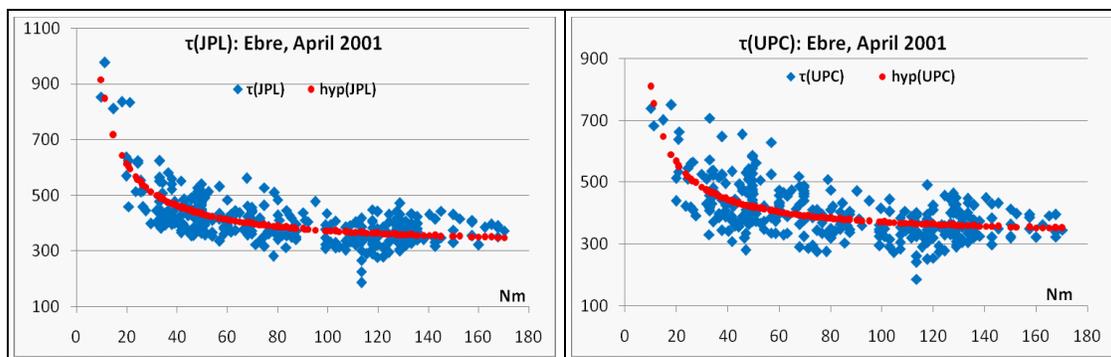


Figure 3: Examples of TEC amount for two maps together with the regression curves.

3. FILLING THE GAPS WITH AN EQUIVALENT SLAB THICKNESS OF THE IONOSPHERE

The filling in of missing data is necessary for many tasks, dealing with a large number of measurements. In the most cases, filling for any parameter is fulfilled on the basis of statistical relations obtained for this parameter. In given case, another parameter TEC and thickness τ are used to fill in missing values. Having the experimental value of the TEC for a specific point in time and hyperbolic regression line constructed for the month (or 27 days), we can determine foF2. For the data in Table 1, the corresponding average absolute deviations $|\Delta\text{foF2}|$ are shown in Table 2, together with $|\Delta\text{foF2}|$ for the IRI model, which is often used to fill the gaps. Overall, the results are in 1.5–2 times better than ones obtained by using the IRI model.

Table 1: Statistics illustrating the choice of map.

| | ΔfoF2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|----------|---------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| E | JPL | 0.28 | 0.37 | 0.49 | 0.46 | 0.46 | 0.45 | 0.60 | 0.43 | - | 0.80 | 0.67 | 0.57 |
| b | CODE | 0.33 | 0.45 | 0.57 | 0.55 | 0.52 | 0.55 | 0.67 | 0.54 | - | 0.86 | 0.72 | 0.64 |
| r | UPC | 0.39 | 0.40 | 0.62 | 0.54 | 0.48 | 0.47 | 0.60 | 0.45 | - | 0.89 | 0.89 | 0.70 |
| e | ESA | 0.38 | 0.56 | 0.74 | 0.85 | 0.56 | 0.58 | 0.58 | 0.62 | - | 0.90 | 0.87 | 0.79 |
| R | JPL | 0.40 | 0.43 | 0.61 | 0.56 | 0.46 | 0.43 | 0.37 | 0.36 | - | - | 0.65 | 0.50 |
| o | CODE | 0.48 | 0.51 | 0.73 | 0.68 | 0.47 | 0.53 | 0.47 | 0.45 | - | - | 0.70 | 0.54 |
| m | UPC | 0.52 | 0.52 | 0.73 | 0.64 | 0.46 | 0.46 | 0.42 | 0.45 | - | - | 0.98 | 0.52 |
| e | ESA | 0.45 | 0.57 | 0.79 | 0.87 | 0.61 | 0.61 | 0.58 | 0.58 | - | - | 0.78 | 0.76 |
| E | JPL | - | - | 0.23 | 0.26 | 0.43 | 0.51 | 0.48 | 0.41 | 0.21 | 0.20 | 0.18 | 0.19 |
| b | CODE | - | - | 0.29 | 0.31 | 0.50 | 0.58 | 0.55 | 0.47 | 0.25 | 0.25 | 0.21 | 0.25 |
| r | UPC | - | - | 0.25 | 0.28 | 0.49 | 0.51 | 0.50 | 0.44 | 0.26 | 0.27 | 0.21 | 0.23 |
| e | ESA | - | - | 0.29 | 0.34 | 0.55 | 0.60 | 0.55 | 0.47 | 0.28 | 0.29 | 0.28 | 0.22 |
| R | JPL | - | - | 0.18 | 0.25 | 0.39 | 0.51 | 0.55 | 0.37 | 0.21 | 0.22 | 0.21 | 0.23 |
| o | CODE | - | - | 0.22 | 0.29 | 0.45 | 0.55 | 0.60 | 0.41 | 0.24 | 0.25 | 0.23 | 0.28 |
| m | UPC | - | - | 0.24 | 0.30 | 0.51 | 0.53 | 0.58 | 0.43 | 0.26 | 0.29 | 0.24 | 0.24 |
| e | ESA | - | - | 0.28 | 0.34 | 0.54 | 0.62 | 0.64 | 0.47 | 0.29 | 0.31 | 0.25 | 0.25 |

Table 2: Statistics illustrating filling the gaps of foF2.

| ΔfoF2 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| rec(E) | 0.41 | 0.54 | 0.57 | 0.55 | 0.45 | 0.41 | 0.81 | 1.02 | - | 0.91 | 0.93 | 0.94 |
| IRI(E) | 0.78 | 0.85 | 0.83 | 0.91 | 0.88 | 0.78 | | 1.02 | - | 0.84 | 1.06 | 0.94 |
| rec(R) | 0.37 | 0.39 | 0.67 | 0.51 | 0.46 | 0.34 | 0.39 | 0.40 | - | - | 0.63 | 0.62 |
| IRI(R) | 0.92 | 0.82 | 0.96 | 0.92 | 0.81 | 0.68 | 0.87 | 1.20 | - | - | 1.04 | 0.85 |
| rec(E) | - | - | 0.24 | 0.33 | 0.54 | 0.69 | 0.64 | 0.56 | 0.33 | 0.35 | 0.35 | 0.38 |
| IRI(E) | - | - | 0.57 | 0.53 | 0.82 | 0.90 | 0.85 | 0.72 | 0.49 | 0.53 | 0.40 | 0.48 |
| rec(R) | - | - | 0.21 | 0.47 | 0.59 | 0.73 | 0.78 | 0.51 | 0.34 | 0.33 | 0.41 | 0.58 |
| IRI(R) | - | - | 0.47 | 0.47 | 0.72 | 0.79 | 0.88 | 0.62 | 0.43 | 0.48 | 0.37 | 0.45 |

4. FILLING IN THE GAPS USING TEC OF TWO STATIONS

Filling the foF2 gaps using data of the single station is a special case of the more common definition of foF2 of the single station according to data of another station, which can be used to predict the critical frequencies in the area of spatial correlation of foF2. The results of the reconstruction method used in this paper are compared with results of other methods, based on use of: a) linear regression (“lin” notation in Fig. 4, [5]), b) a special formula (denoted by “Gul” in Fig. 4, [6]), c) the IRI model (“IRI” notation in Fig. 4, [7]), d) the original IRI model (“mod” notation in Fig. 4 (left panel)). Methods from [5, 6] require knowledge of the median foF2 of the second station.

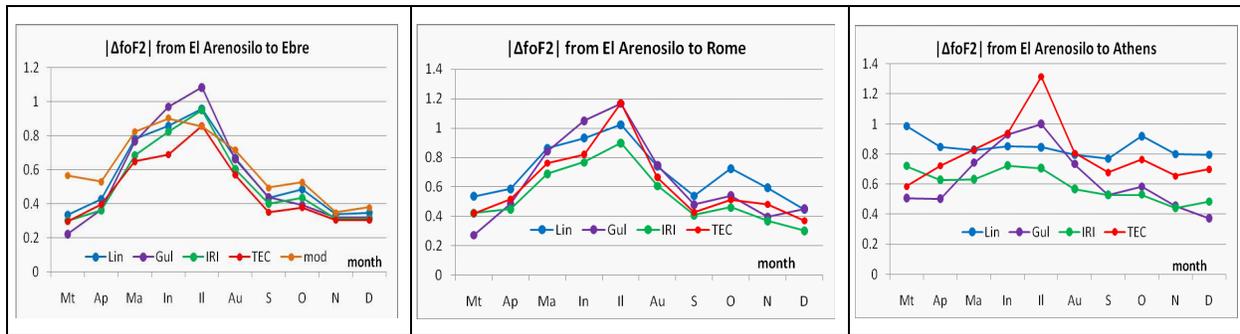
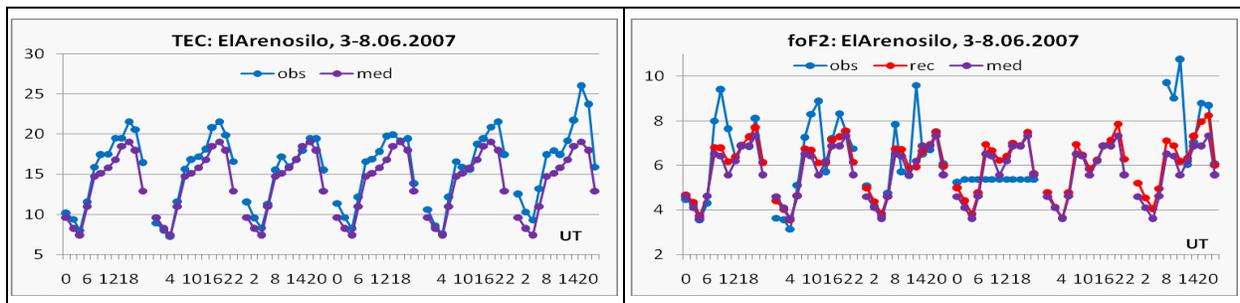

 Figure 4: Differences $|\Delta foF2|$ at various distances from the st. El Arenosilo.


Figure 5: Example of foF2 reconstruction for cases of presence of the Es layer.

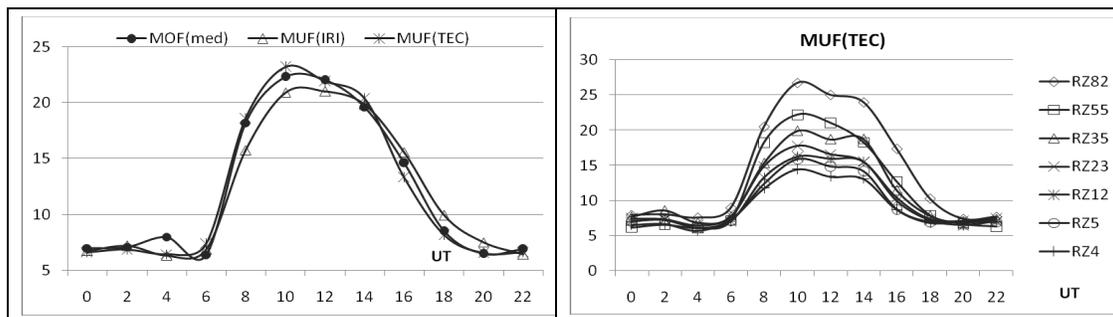


Figure 6: Example of MUF calculations using the TEC values.

Method from [7] and method used here do not use the values of foF2 of the second station. The results for the chain of stations are presented in Fig. 4 in the form of differences $|\Delta foF2|$ at various distances from the st. El Arenosilo.

Method from [7] defines the distance (R_m), for which the foF2 value reconstructed becomes equal to the model ones. At distances less than R_m , the use of the TEC gives results better than any other methods. At distances greater than R_m , all the methods give results worse than the model. In the example of Fig. 4, the first case corresponds to the left panel, the second case — The other two. These findings have direct relevance to the problem of assimilation of TEC data.

5. CONCLUSION AND DISCUSSION OF RESULTS

It is shown that the joint use of global maps of the total electron content and data of ionospheric vertical sounding can be useful for: 1) the choice of maps, the most appropriate experimental values of critical frequency foF2, 2) filling in missing data of foF2, 3) determination of foF2 in the zone with a radius greater than radius of spatial correlation of foF2. A few more aspects may be noted when such use would be helpful. In the first, the given region is characterized by increased activity of the Es layer [8]. In this case, the definition of foF2 is difficult. Use of the TEC allows to determine foF2 with sufficient accuracy (Fig. 5), but in this case, it is necessary to use only the median τ (med) value. In the second, it is necessary to evaluate the possibility of using the TEC for investigation and real-time determination of the propagation conditions on the HF

path. Unfortunately, the authors do not have data for the MUF in this region. The only indirect example is MUF for December 2003 for Inskip-Rome path [9]. Getting a good agreement between experimental and calculated using the TEC values of MUF (left panel of Fig. 6), we determined the MUF for a given path for the remaining years (2002–2008), which exhibit a linear dependence on solar activity index RZ12 (right panel of Fig. 6). Because it is expected that the maximum RZ12 in the 24 cycle does not exceed the value ~ 58 [10], the use of TEC can simplify prediction of HF propagation conditions. Finally, it is necessary to point out a couple of discussion questions. In the first, how to improve the definition of τ for the disturbed conditions? It was seemed to be attractive an idea to build a regression of the difference $\Delta\tau = \tau(\text{obs}) - \tau(\text{med})$ from any index of geomagnetic activity, e.g., Dst, as its prediction is rather realistic [11]. Attempt was made to use such a relationship, but positive results were sometimes obtained only for very strong disturbances. Typical is the absence of a certain relationship. In the second, what is percentage of synchronic cases of disturbances of TEC and NmF2? Despite the high correlation coefficient between the TEC and NmF2, simple visualization of the behavior of the TEC and NmF2, as well as the correlation coefficients between δTEC and δNmF2 suggest that this percentage is not great. Despite the huge amount of works, defining and explaining the various delays in the response of the magnetosphere and ionosphere to the disturbance, e.g., [12], to predict this delay is not yet possible.

REFERENCES

1. Kouris, S. S., K. V. Polimeris, L. R. Cander, and L. Ciruolo, "Solar and latitude dependence of TEC and SLAB thickness," *J. Atmos. Solar-Terr. Physics*, Vol. 70, 1351–1365, 2008.
2. Alfonsi, L., C. N. Mitchell, V. Romano, and P. Spalla, "MIRTO: a prototype for real-time ionospheric imaging over the Mediterranean area," *Annals of Geophysics*, Vol. 50, No. 3, 447, 2007.
3. Arican, F., C. B. Erol, and O. Arican, "Regularized estimation of vertical total electron content from global positioning system data," *J. Geophys. Res.*, Vol. 108, NA12, doi: 10.1029/2002JA009605, 2003.
4. Orus, R., L. R. Cander, and M. Hernandez-Pajares, "Testing regional vertical total electron content maps over Europe during the 17–21 January 2005 sudden space weather event," *Radio Sci.*, Vol. 42, RS3004, doi: 10.1029/2006RS003515, 2007.
5. Cander, L. R., "Spatial correlation of foF2 and vTEC under quiet and disturbed ionospheric conditions: A case study," *Acta Geophysica*, Vol. 55, No. 3, 410–423, 2007.
6. Gulyaeva T. L., I. Stanislawski, and M. Tomasik, "Ionospheric weather: Cloning missed foF2 observations for derivation of variability index," *Annals of Geophysics*, Vol. 26, 315–321, 2008.
7. Maltseva, O. A., V. T., Rodionova, and A. S. Schlupkin, "Using the total electron content for current diagnosis of ionospheric channel," *Geomagn. and Aeronomie*, Vol. 45, No. 4, 480–486, 2005.
8. Pietrella, M. and C. Bianchi, "Occurrence of sporadic-E layer over the ionospheric station of Rome: Analysis of data for thirty-two years," *Adv. Space Res.*, Vol. 44, 72–81, 2009.
9. Zolesi, B., G. Fontana, L. Perrone, M. Pietrella, V. Romano, G. Tutone, A. Belehaki, I. Tsagouri, S. S. Kouris, F. Vallianatos, J. P. Makris, and M. J. Angling, "A new campaign for oblique-incidence ionospheric sounding over Europe and its data application," *J. Atmos. Solar-Terr. Physics*, Vol. 70, doi:10.1016/j.jastp.2007.02.015, 2008.
10. Kane, R. P., "Size of coming solar cycle 24 based on Ohl's precursor method, final estimate," *Annals of Geophysics*, Vol. 28, 1463–1466, 2010.
11. Amata, E., G. Pallochia, G. Consolini, M. F. Marcucci, and I. Bertello, "Comparison between three algorithms for Dst predictions over the 2003–2005 period," *J. Atmos. Solar-Terr. Physics*, Vol. 70, No. 2–4, 496–502, 2008.
12. Liu, J., B. Zhao, and L. Liu, "Time delay and duration of ionospheric total electron content responses to geomagnetic disturbances," *Annals of Geophysics*, Vol. 28, 795–805, 2010.

Analysing the Attenuation at Mobile Phone Bands Provided by Vegetation Supported by Lattice Structures

Paula Gómez, Iñigo Cuiñas, and Ana V. Alejos

Dept. Teoría do Sinal e Comunicacóns, Universidade de Vigo, Spain

Abstract— The use of lines of trees or shrubs as barriers to reduce the radio electric pollution at mobile frequency bands has demonstrated to provide additional attenuation up to 20 dB at 2100 MHz. An improvement to common vegetation fences is provided in this work: the using of lattice structures along with lighter vegetation specimens supplies higher attenuations which are barely influenced by the wind, reducing the scattering but maintaining its aesthetic and electromagnetic conditions. An extensive measurement campaign has been carried out with several barrier layouts. Results show the influence of the material and lattice patterns.

1. INTRODUCTION

Population in many countries is afraid of the presence of radio electric antennas transmitting in the surroundings of their houses or workplaces. Although most of the medical research indicates that there is no risk for the human health exposed to controlled electromagnetic fields [1], under some restrictions defined by the International Commission for Non-Ionizing Radiation Protection (ICNIRP) [2]; the World Health Organization (WHO) [3] also recognizes a lack in the research in some fields: long-term exposure, or children and elderly exposure. Trying to mitigate the human contact to electromagnetic fields, a proposal involving the construction of vegetation walls to provide shadowed areas has been previously done [4]. This current contribution goes beyond: the wall or barrier could be constructed by means of a lattice structure (as the supports of climbing plants in parks and gardens) holding some foliated plants or bushes.

Section 2 explains the details of the measurement campaign, including vegetation species and lattice structure, materials and patterns. A comparison of the results obtained for different lattice structure and materials is exposed in Section 3. Finally, a resume of the main conclusions obtained is exposed in Section 4.

2. MEASUREMENT DESCRIPTION

A measurement campaign has been performed in frequency bands assigned to mobile phone services (GSM and UMTS). The layout has been deployed within an anechoic chamber with both transmitter and receiver at fixed locations. The vegetation specie chosen was *Ficus benjamina*. The transmitter consisted of a signal generator transmitting continuous waves at each of the frequencies (900, 1800, and 2100 MHz), and the receiver was built around a spectrum analyzer. Both ends used log-periodic antennas.

Up to 10.000 samples have been taken in each measurement event, distributed in 20 traces of 500 values each one of them. Six lattice supports made of different materials were selected, to be measured under two orthogonal linear polarizations, in order to test the behavior of different materials and geometries. Six barrier configurations including lattice structures and line of shrubs were considered, which are depicted at Figure 1. A reference is also measured, in free space conditions, with no barrier between transmitter and receiver.

The barrier was placed 1.45 m from transmitter antenna and 2.15 from receiver antenna. Such distances verify far field conditions at the considered frequencies [5]. Both antennas have been installed a height of 1.70 m, pointed to the centre of the lattice structure. Structures made by iron wires (one sample), wood planks (two samples, with different design), resin (also two samples), and PVC plastic (one sample) have been used. The main differences among the structures, as well the own material, are the dimensions of the lattice: how large the holes are, and how width the solid parts are. Figure 2 defines the lattice parameters. All the lattices used had similar patterns, with different dimensions, which are summarized in Table 1. These structure configurations were measured both in calm and under windy conditions. Three different wind speeds have been tested: 1.9, 3.0, and 4.7 m/s. The wind generator was placed in two different positions: nearby the transmitter and on the side of the receiver.

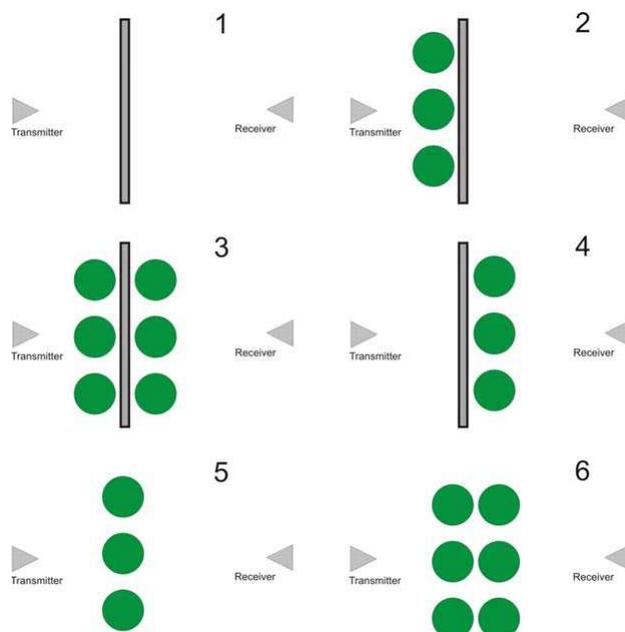


Figure 1: Barrier configurations.

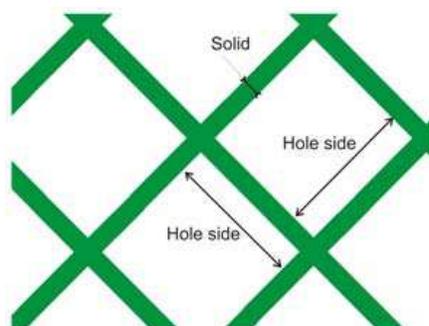


Figure 2: Definition of lattice structure parameters.

Table 1: Lattice supporting structure dimensions.

| Material | Structure | | | Lattice | |
|----------|------------|-----------|------------|------------|----------------|
| | Height (m) | Width (m) | Depth (cm) | Solid (cm) | Hole side (cm) |
| Iron | 1.8 | 1.8 | 0.5 | 0.5 | 9.0 |
| Wood 1 | 1.8 | 1.8 | 1.0 | 3.0 | 12.0 |
| Wood 2 | 1.8 | 1.8 | 0.5 | 6.0 | 7.0 |
| Resin 1 | 1.0 | 2.0 | 0.5 | 3.5 | 6.5 |
| Resin 2 | 1.0 | 2.0 | 0.5 | 2.5 | 2.7 |
| PVC | 1.0 | 2.0 | 0.5 | 1.5 | 2.5 |

3. EXPERIMENTAL RESULTS AND ANALYSIS

Figure 3 shows the increasing of the additional attenuation when including a lattice structure in the radio path by making a comparison between the attenuation provided by configurations 2, 4 and 5. While configuration 2 and 4 have a lattice between the antenna and the barrier, configuration number 5 only have the barrier itself, with no lattice structure supporting it. Analyzing the results for several frequencies and polarizations, it can be seen that, among this three configurations, configuration 5 is which provides the lowest attenuations in all cases, so the using of a lattice structure to obtain a surplus attenuation is hereby demonstrated.

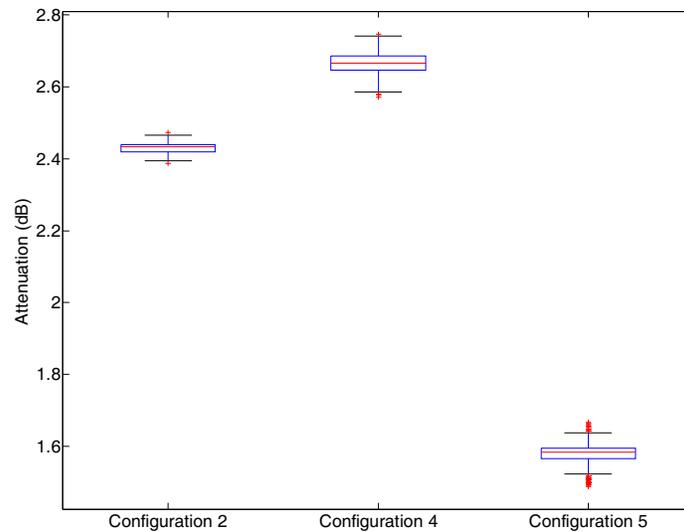


Figure 3: Attenuation for 2100 MHz, resin 1 lattice and horizontal polarisation.

Table 2: Attenuations provided by lattices and vegetation.

| frequency | | 900 MHz | | 1800 MHz | | 2100 MHz | |
|--------------|---------------|----------|------------|----------|------------|----------|------------|
| polarisation | | vertical | horizontal | vertical | horizontal | vertical | horizontal |
| barrier | configuration | | | | | | |
| Iron | 1 | -0.18 | 3.54 | 1.73 | 0.24 | 0.15 | 0.58 |
| | 2 | 0.39 | 5.88 | 10.72 | 3.39 | 5.15 | 9.83 |
| | 3 | 4.36 | 7.45 | 10.26 | 7.76 | 6.98 | 16.31 |
| | 4 | 1.23 | 5.72 | 9.54 | 3.23 | 7.56 | 11.17 |
| Wood 1 | 1 | -1.41 | -0.05 | 2.10 | -0.71 | -0.91 | 5.03 |
| | 2 | 1.26 | 1.80 | 11.60 | 1.06 | 5.52 | 4.27 |
| | 3 | 6.24 | 6.17 | 18.64 | 2.89 | 9.46 | 3.91 |
| | 4 | 3.99 | 1.85 | 6.82 | 2.36 | 2.75 | 6.20 |
| Wood 2 | 1 | -1.02 | 0.09 | 2.18 | -0.83 | -0.91 | 3.14 |
| | 2 | 0.97 | 1.60 | 5.03 | 2.44 | 13.69 | 6.01 |
| | 3 | 4.52 | 5.11 | 6.00 | 7.58 | 12.06 | 7.81 |
| | 4 | 1.99 | 3.22 | 8.30 | 1.88 | 11.53 | 4.70 |
| Resin 1 | 1 | -0.21 | -0.02 | -0.61 | 0.02 | 0.36 | -0.06 |
| | 2 | 2.57 | 1.88 | 5.57 | 2.25 | 3.46 | 2.17 |
| | 3 | 7.90 | 4.25 | 19.74 | 3.63 | 13.15 | 3.37 |
| | 4 | 4.29 | 2.69 | 8.69 | 2.00 | 8.28 | 2.77 |
| Resin 2 | 1 | -0.62 | 0.06 | -0.38 | 0.26 | 1.60 | -0.05 |
| | 2 | 4.35 | 1.79 | 3.91 | 2.06 | 3.37 | 2.90 |
| | 3 | 25.18 | 4.64 | 17.94 | 2.20 | 18.06 | 2.94 |
| | 4 | 13.70 | 2.24 | 10.60 | 0.83 | 10.62 | 1.21 |
| PVC | 1 | -0.97 | 0.22 | -1.19 | -0.26 | 1.86 | 0.91 |
| | 2 | 3.10 | 1.88 | 4.61 | 1.26 | 13.00 | 0.73 |
| | 3 | 13.31 | 6.04 | 21.50 | 1.62 | 16.62 | 1.08 |
| | 4 | 5.13 | 2.91 | 6.38 | -0.20 | 19.47 | 0.93 |
| No lattice | 5 | 2.05 | 1.56 | 4.08 | 1.38 | 4.82 | 1.61 |
| | 6 | 10.79 | 5.01 | 4.13 | 2.57 | 8.03 | 5.38 |

Median attenuation values obtained for each frequency, polarization and barrier configuration are shown in Table 2 for every lattice materials. Attenuation has been computed by comparing the median received powers when the line of sight is obstructed by the barriers with the median received power in free space conditions. Negative values in few cells indicate a certain power gain, probably explained by diffractions at the borders of the structures or even at the holes. These negative values only appear when no shrub is in the composition of the barrier: when the vegetation is present, the scattering mechanism, which is induced by the leaves, dominates the possible diffraction.

As the results show in Table 2, polarization is rather affected by lattice pattern than the material itself. Tighter lattice patterns provide more attenuation than looser ones. In general, vertical polarization provides deeper attenuation levels. This effect can also be found when measuring vegetation barriers with no lattice supporting structures [4], and could be because of the own geometry of the trees, which are vertical elements themselves, as well as their leaves.

The influence of the wind seems to be negligible, in opposition to reported in [6] for shrubs with no lattice support. Data resulted to be very homogeneous even at maximum speeds. Results could be explained because of the lattice structure, which is a homogeneous obstacle that does not vary in time or with the wind. This is a desirable effect when designing electromagnetic fences, because it adds stability to the attenuation provided by the vegetation barrier and makes the barrier gain independence of the wind conditions.

4. CONCLUSIONS

The validity of using combined barriers of shrubs and lattice supports to reduce the electromagnetic pollution in sensitive areas has been tested by means of experimental research. The lattice supporting structure improves the rigidity of the barrier, as well as it enhances the attenuation the vegetation induces. The best lattice supporting structure, in terms of strongest attenuation induced, appears to be the one that presents the highest conductivity or even the largest surface to the wave propagating. A single shrub barrier combined with a lattice has a better behavior than the barrier or the lattice by themselves. However, the largest attenuation levels are obtained when using a double shrub barrier at both sides of the lattice structure. Also, although vegetation is influenced by the wind, adding a lattice removes most of the random behavior that wind introduces in the system, making the barrier almost independent of the wind.

Depending on the lattice pattern and its dimensions, vertical and horizontal polarizations are differently affected. For the patterns used in this work, vertical polarization suffers bigger attenuations than horizontal for almost all the lattices and materials. The material influence is also remarkable. Some materials, as iron or some resins, attenuate more than lattice woods or PVC, depending on the frequency or polarization. However, there are materials introducing the same attenuation levels independently of these factors, and thus they are suitable to be used in all type of electromagnetic fences, no matter the radiofrequency system and frequency. The lattice pattern is also an important factor, providing better attenuations when the pattern is tighter than the open ones.

ACKNOWLEDGMENT

This work has been supported by Xunta de Galicia, Spain, under project InCiTe 08MRU045322PR.

REFERENCES

1. Kabat, G. C., *Hyping Health Risks*, Columbia University Press, 2008.
2. ICNIRP, "Guidelines for limiting exposure to time-varying electric, magnetic and electromagnetic fields (up to 300 GHz)," *Health Physics*, Vol. 74, No. 4, 494–522, 1998.
3. "Establishing a dialogue on risks from electromagnetic fields," International EMF Project, World Health Organization.
4. Cuiñas, I., A. V. Alejos, and M. G. Sánchez, "Vegetal barriers for minimizing electromagnetic pollution at cellular phone bands," *IEE Electronics Letters*, Vol. 41, No. 6, 340–341, 2005.
5. Balanis, C. A., *Antenna Theory: Analysis and Design*, 2nd Edition, John Wiley & Sons, 2005.
6. Gómez, P., I. Cuiñas, A. V. Alejos, M. G. Sánchez, and R. F. S. Caldeirinha, "Shrub-blown time variability in attenuation and scattering at cellular frequencies," *IET Microwaves, Antennas and Propagation*, Vol. 4, No. 4, 526–542, 2010.

Microstrip Antenna for Microwave Imaging Application

S. Adnan, R. A. Abd-Alhameed, H. I. Hraga, I. T. E. Elfergani, J. M. Noras, and R. Halliwell
 Mobile and Satellite Communications Research Centre, University of Bradford
 Bradford, West Yorkshire, BD7 1DP, UK

Abstract— A compact microstrip antenna design to be used in breast cancer detection is presented. The antenna consists of a radiating patch mounted on two vertical plates, fed by coaxial cable. A study is carried out on different parameters of the antenna. Simulation results show that the antenna possesses a wide bandwidth and this is confirmed experimentally. In experiments, a homogeneous dielectric box, having similar properties to human tissue is used to study the interaction of the antenna with tissue. Even without added matching medium or lumped loads there is good matching when the antenna is in contact with the tissue. Finally a two-element antenna array is investigated numerically, with promising results.

1. INTRODUCTION

Breast cancer is the most common form of cancer found in women; early detection is the best factor in long time survival. Currently the most popular screening method is X-ray mammography, namely X-ray imaging of the compressed breast. According to reports [1], serious limitations of X-ray mammography provide clear motivation for the development of a new imaging tool to assist in early detection. For a long time, microwave engineers have tried to implement non-ionizing electromagnetic waves to detect cancer in the human body. Over the past several years significant progress has been made in using microwaves for breast cancer detection. Microwave imaging can be defined as seeing the internal structure of an object by illuminating the object with low power electromagnetic wave at microwave frequencies. In the microwave frequency range, passive, hybrid and active approaches to breast cancer detection are being researched [2]. Such methods would be beneficial to patients because both ionizing radiation and breast compression are avoided.

The active method using microwaves radiated into the breast to detect the presence of tumours is the most widely adopted method, and this may be sub-divided into tomography and radar-based approaches [3]. In tomography a single transmitter radiates into the breast while a number of antenna placed around the breast receive scattered waves. The procedure is repeated for various positions of the transmitter. Depending on the acquisition of data these data will be processed to produce a 2-dimensional or 3-dimensional image of the breast [4].

For radar-based microwave imaging, a short pulse is transmitted from a single ultra-wideband (UWB) antenna into the breast and any back-scattering is detected by the same antenna. This process will be repeated for different locations around the breast. The presence of a tumour produces strong scattering, and such a response can be interpreted to estimate the location of the tumour. The travel times of signals at various locations are recorded and computed [3]. As with any radar-based system, this system does not require complex image reconstruction algorithms, and hence offers more detailed information than the tomography microwave imaging method.

The antenna which is used as transmitter and receiver to transmit UWB signals must permit a high level of resolution. This requirement thus limits the class of antenna that can be utilized. Large fractional bandwidth, low side-lobes and low mutual coupling (i.e., when two antennas are used for system operation) are the characteristic factors which must also be considered. The existing antennas used for breast cancer detection require lumped loads as a trade-off for wider bandwidth [5, 6]. Most metallic-type sensors must be immersed in a matching medium with permittivity similar to that of breast tissue so that the reflection between the free space and the breast surface can be minimised [7–12]. However in reality it is difficult to implement the matching medium to surround the patient breast. This paper presents a compact microstrip antenna design for the RF detection system of cancer cells inside the human breast tissue. The operation of two elements including the breast model is also demonstrated.

2. ANTENNA GEOMETRY

Figure 1 shows the overall dimensions of the antenna. The antenna consists of a radiating patch mounted on two vertical plates placed on a ground plane of dimension $L = W = 30$ mm and 0.5 mm thickness. A $50\ \Omega$ coaxial probe excites the antenna, and design parameters are obtained

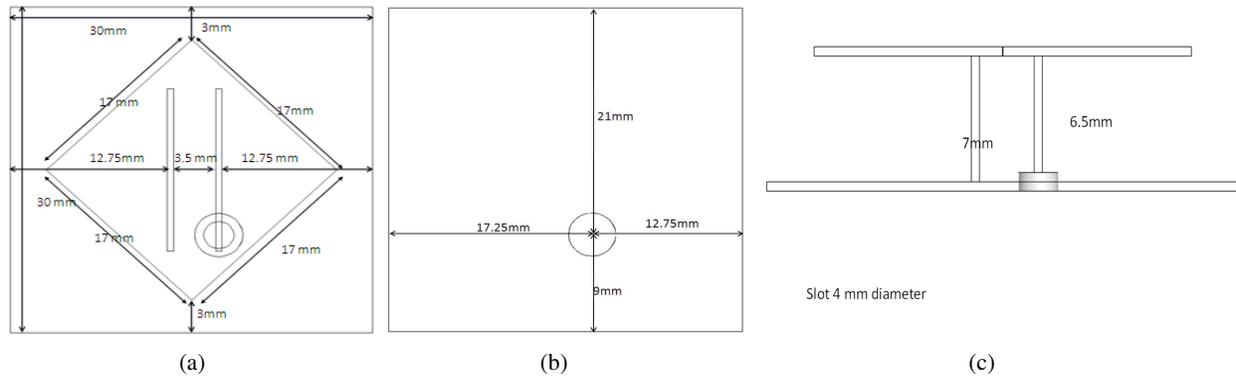


Figure 1: Geometry of the Antenna. (a) Top view. (b) Ground plane. (c) Side view.

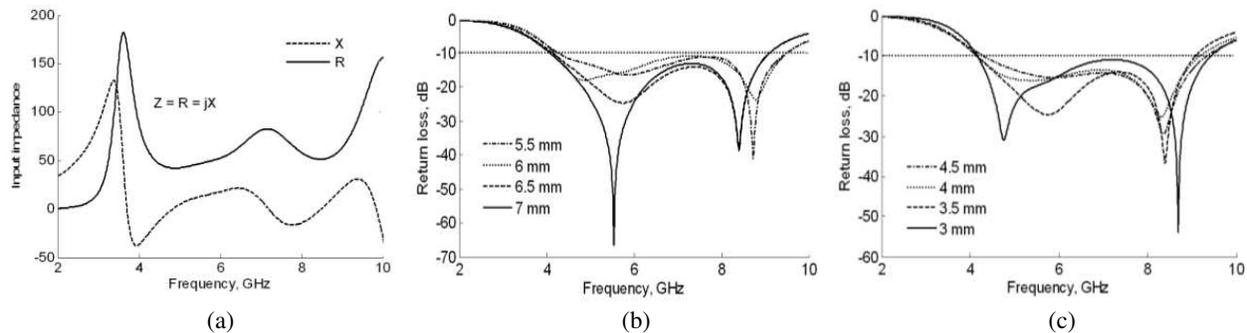


Figure 2: (a) Input impedance. (b) Simulated reflection coefficient of the antenna for different heights. (c) Simulated reflection coefficient for the gap between two vertical plates.

by performing parametric studies with simulation software HFSS [13]. Figure 2(a) shows the input impedance of the proposed antenna, where the real part of the impedance is close to 50Ω over the targeted bandwidth (i.e., 4 to 8 GHz). Figures 2(b) and 2(c) show the reflection coefficients of the antenna for different heights of the radiating patch from the ground plane and varying gap between the two vertical plates. These show that the reflection coefficient improves with increasing height of the antenna and a narrowing of the gap between the vertical plates.

3. RESULTS AND DISCUSSION

The return loss of this fabricated antenna, see Figures 3(a) and 3(b), is measured using an HP8510 network analyzer. The simulated and the measured return loss of the antenna are compared in Figure 3(c), indicating good agreement. The measured and simulated frequency range covered between 4 GHz and 9 GHz, with $S_{11} < -10$ dB, corresponds to a relative bandwidth of 77%.

The antenna is designed to be placed directly on the breast, so the interaction between the antenna and tissue is investigated. A homogenous box with similar properties to breast tissue ($\epsilon_r = 10$ and $\sigma = 0.4$ S/M) [14] is modelled. Figure 4(a) shows the input impedance of the antenna while in contact with the tissue, both before and after optimisation. The results show good matching after optimisation when the sensor touches the tissue, so that more power is penetrating into the tissue. This is obtained without adding any matching materials or lumped loads. After optimisation the new height of the antenna from the ground plane is 4 mm and the gap between the two plates is 8 mm respectively. The remaining dimensions of the antenna remain the same. The configuration of a two-element antenna array in the presence of tissue is illustrated in Figure 4(b). The two antennas are aligned face to face with 50 mm separation. Figure 4(c) shows the return loss (S_{11}) and transmission coefficients (S_{21}) of the antenna elements, indicating wideband characteristics for the impedance and transmission.

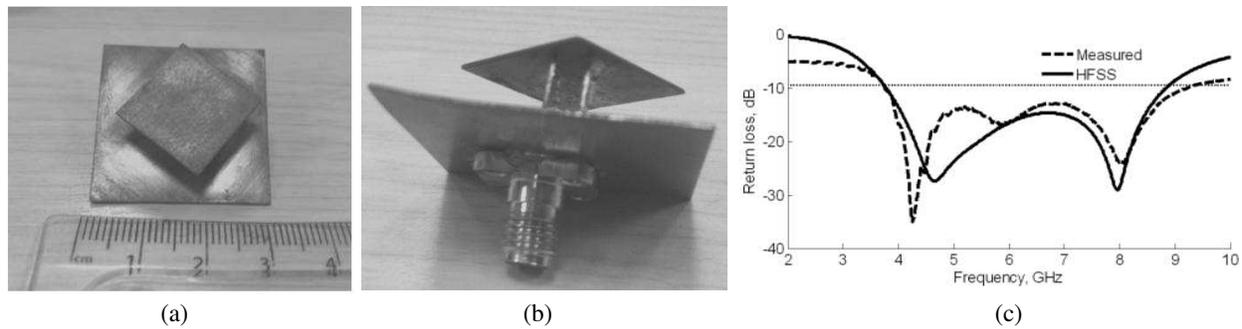


Figure 3: (a) and (b) Antenna prototype. (c) The input return loss of the proposed antenna measured and simulated.

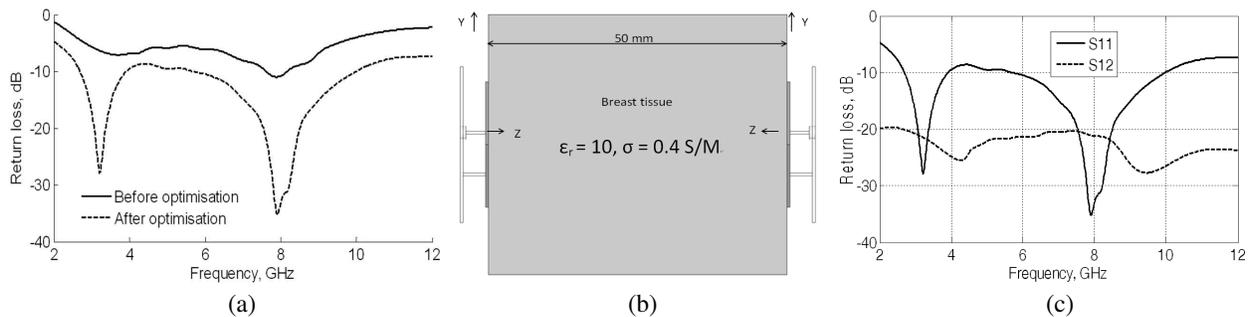


Figure 4: (a) Simulated reflection coefficients of the sensor in contact with breast tissue before and after optimisation. (b) Configurations of two-element arrays. (c) Return loss S_{11} and transfer function S_{21} .

4. CONCLUSION

In this paper a new compact UWB antenna design has been presented. The effect of the various antenna parameters on the bandwidth and the resonance characteristics were discussed. The operating bandwidth of the antenna achieved at a minimum workable return loss of 10 dB was 4 to 9 GHz. The measurement results show good agreement with simulations. A homogenous box having similar properties to human breast tissue was also used to check the interaction of the antenna with the tissue. The antenna showed good matching without the need to add any lumped loads or introduce dielectric matching media.

ACKNOWLEDGMENT

The authors would like to thanks Mr Dhirajlal Chavda for his help in the fabrication of the antenna.

REFERENCES

1. Kopans, D. B., R. H. Moore, K. A. McCarthy, D. A. Hall, C. A. Hulka, G. J. Whitman, P. J. Slanetz, and E. F. Halpern, "Positive predictive value of breast biopsy performed as a result of mammography: there is no abrupt change at age 50 years," 1996.
2. Fear, E. C., P. M. Meaney, and M. A. Stuchly, "Microwaves for breast cancer detection," *IEEE Potentials*, 12–18, Feb./Mar. 2003.
3. Fear, E. C., "Microwave imaging of the breast," *Technology in Cancer Research & Treatment*, Vol. 4, 69–82, Feb. 2005.
4. Meaney, P. M., M. W. Fanning, D. Li, S. P. Poplack, and K. D. Paulsen, "A clinical prototype for active microwave imaging of the breast," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 48, 1841–1853, 2000.
5. Xu, L., S. C. Hagness, K. Choim, and D. W. Vanderweide, "Numerical and experimental investigation of an ultrawideband ridged pyramidal horn antenna with curved launching plane for pulse radiation," *Antennas Wirel. Propag. Lett.*, Vol. 2, 259–262, 2003.
6. Kanj, H. and M. Popovic, "Miniaturized microstrip-fed dark eyes antenna for near-field microwave sensing," *Antennas Wirel. Propag. Lett.*, Vol. 4, 397–401, 2005.

7. Craddock, I. J., M. Klemm, J. Leendertz, A. W. Preece, and R. Benjamin, “An Improved hemispherical antenna array design for breast imaging,” *The Second European Conference Antennas and Propagation EuCAP*, Nov. 2007.
8. Hagness, S. C., “FDTD analysis of a pulsed microwave confocal system for breast cancer detection,” *19th International Conference Proc*, Vol. 2506–2508, Oct. 1997.
9. Fear, E. C., J. Sill, and M. A. Stuchly, “Experimental feasibility study of confocal microwave imaging for breast tumor detection,” *IEEE Transactions on Microwave Theory and Techniques*, Vol. 51, 887–892, 2003.
10. Fear, E. C., S. C. Hagness, P. M. Meaney, M. Okoniewsk, and M. A. Stuchly, “Enhancing breast tumor detection with near-field imaging,” *IEEE Microwave Magazine*, 48–56, Mar. 2002.
11. Shannon, C. J., E. C. Fear, and M. Okoniewski, “Dielectric-filled slotline bowtie antenna for breast cancer detection,” *Electronics Letters*, Vol. 41, Mar. 2005.
12. Hossain, I., S. Noghianian, and S. Pistorius, “A diamond shaped small planar ultra wide band (UWB) antenna for microwave imaging purpose,” *IEEE Antennas and Propagation Society International Symposium*, 5713–5716, 2007.
13. Ansoft High Frequency Structure Simulator Corporation, V11, 2007.
14. Bindu, G., S. J. Abraham, A. Lonappan, V. Thomas, C. K. Aanandan, and K. T. Mathew, “Active microwave imaging for breast cancer,” *Progress In Electromagnetics Research*, Vol. 58, 149–169, 2006.

Propagation Characteristics of 24 GHz Frequency Band for Automotive Collision Avoidance Radar

Deock-Ho Ha¹, Yeon-Wook Choe², Jee-Youl Ryu¹, and Sung-Un Kim¹

¹Department of Telecommunication Engineering, Pukyong National University, Busan, Korea

²Department of Control and Measurement Engineering, Pukyong National University, Busan, Korea

Abstract— In this paper, we describe the propagation characteristic of 24GHz frequency band used for automotive collision avoidance radar (ACAR). To investigate the radio propagation characteristic of ACAR signals, we used 24 GHz UMCW (Un-Modulated Continuous Wave) signals using by several polarized antenna combinations. To detect the accurate message data for safety guiding ACAR, we analyzed and investigated the optimum receiving condition from the short range UMCW signal. From the measurement analysis, it can be clearly seen that the short range UMCW signal can be encountered to multipath interference due to scattering rays from the ground plane and adjacent moving vehicles. In addition, it was also found that the received signal strength of circularly co-polarized antenna shows the most strong envelopes and best fading reduction effect.

1. INTRODUCTION AND MEASUREMENT RESULTS

We investigate which is the best propagation condition to reduce the multipath fading when the combinations of transmitting and receiving polarized waves are adopted. In the narrowband signal measurement, the signal generator transmits the 24 GHz UMCW radio frequency band. And also, we examined the broadband signal which show the characteristic of amplitude fluctuation within broad bandwidth which the transmitting frequency is sweeping from 23.9 GHz to 24.1 GHz. In order to obtain maximum power transfer between two antennas, we used several polarized antenna which generate the vertical, horizontal, and circular polarized waves.

2. OUTLINE OF UMCW MEASUREMENTS

To investigate the radio propagation characteristic of ACAR signals, we used 24 GHz UMCW (Un-Modulated Continuous Wave) signals using by several polarized antenna combinations in the short range radio propagation environments. The UMCW signal measurements are conducted in a conference room which size is 11.11 m×7.99 m, height is 2.7 m and two sides are windows and the other sides are walls. A measuring distance is about 7 m. In the moving measurement system as shown in Fig. 1, we used vertical, horizontal polarized antenna and omni-directional circularly polarized antenna as the transmitting and receiving antenna. Circularly polarized antenna was combined by vertical and horizontal antenna composed of patched micro-strip antenna, which antenna can be used as a vertical antenna and a horizontal antenna, respectively. In the broadband signal measurement as shown in Fig. 2, we fixed receiver in each measuring point, and then measured the fluctuation characteristics of signal strength within 200 MHz bandwidth. In this case, spectrum analyzer received the signal strength of UMCW signal which transmitted the center frequency 24 GHz from sweep generator with the transmitting power 0 dBm.

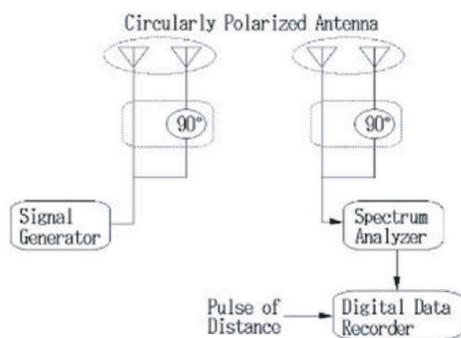


Figure 1: Moving measurement system.

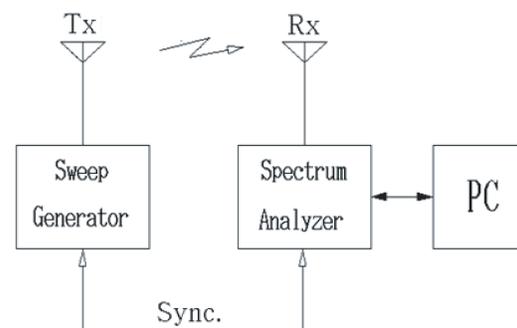


Figure 2: Broadband measurement system.

3. FADING REDUCTION EFFECTS

For a more detailed review, we investigated the fading reduction effect by polarized antenna of every kind. Fig. 3 is an example of the narrowband signal strength for the C-C (circular co-polarization), H-H (horizontal-horizontal), V-V (vertical-vertical), and C-X (circular cross-polarization) of transmitting-receiving antenna combinations, respectively. Fig. 3 shows the results of fading reduction effects for the polarization combinations for the case of LOS propagation environment. Fig. 4 and Fig. 5 show the broadband signal envelopes for the case of LOS and NLOS environment, which signal envelopes are received by right handed circular polarized antenna combination (RHCP-RHCP: C-C), left handed circular polarized antenna combination (LHCP-LHCP: C-C), LHCP-RHCP combination (C-X) and RHCP-LHCP combination (C-X), respectively. Fig. 6 shows the broadband signal envelopes for the case of NLOS conditions existing together with some moving obstacles between the transmitting antenna and the receiving antenna. From the measurement analysis, it was found that the circular co-polarization combination shows better fading reduction effect than the cases of circular cross-polarization, vertical polarization and horizontal polarization combination, respectively. And also, for the broadband signal envelopes received by frequency sweeping, it was found that the circular co-polarization combination shows better fading reduction result than the cases of circular cross-polarization condition.

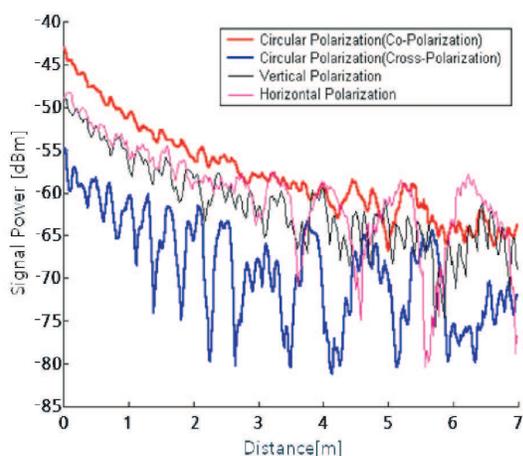


Figure 3: Fading signal envelopes of each polarized radio waves.

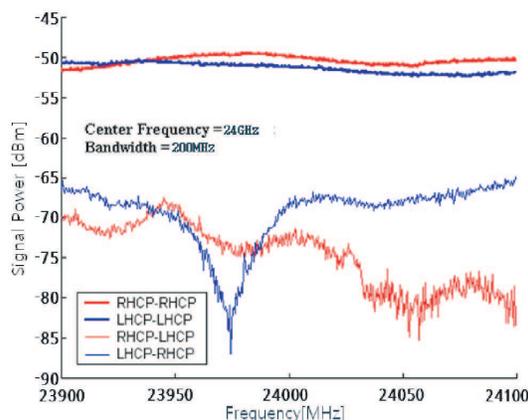


Figure 4: Broadband signal experiment results for circular polarization combination of LOS condition.

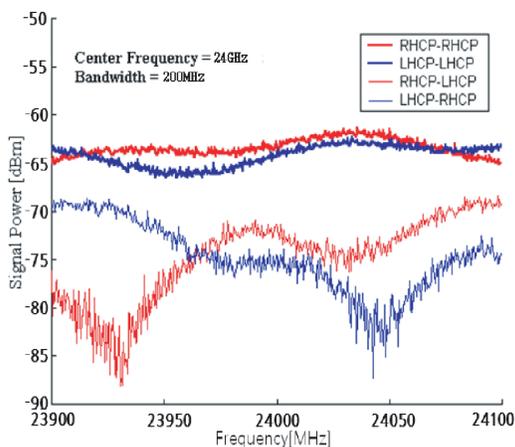


Figure 5: Broadband signal experiment results for circular polarization combination of NLOS condition.

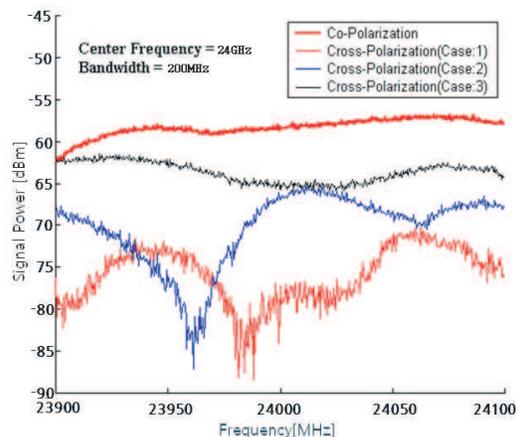


Figure 6: Broadband signal experiment results for circular polarization combination of obstacle condition.

4. CONCLUSIONS

To investigate the radio propagation characteristic of ACAR signals, we transmitted 24 GHz UMCW (Un-Modulated Continuous Wave) signals using by several polarized antenna combinations. Especially, in order to optimally reduce the fading signals received by a vehicle in motion, circular polarized antenna combination methods are adopted, and then we examined the narrow band propagation characteristic of 24 GHz frequency band used for automotive collision avoidance radar (ACAR). From the experimental analysis, it was found that the circular co-polarized antenna combination method could reduce the multipath fading effectively in short range propagation environments. And also, it was found the circular co-polarized antenna combination method could improve the broadband signal envelope to the more flat characteristic. In conclusion, it can be inferred the effective fading reduction method in short range propagation environment for the ACAR signals is to create the co-polarized antenna combination when a circularly polarized antenna is used transmitting end. It will be necessary to compose the real experimental measurement system like a real road condition and also to make sure diversity antenna reception effect for the ACAR signals. This will be done in further research.

ACKNOWLEDGMENT

Following are results of a study on the “Human Resource Development Center for Economic Region Leading Industry” Project, supported by the Ministry of Education, Science & Technology (MEST) and the National Research Foundation of Korea (NRF).

Swept Versus Real-time Spectrum Analyzer Ability to Accurately Assess Electromagnetic Exposure due to Wireless Communications Signals in the Environment: An Analysis

Paul Bechet and Simona Miclaus

Land Forces Academy, Sibiu, Romania

Abstract— Present study aimed to identify optimal settings of a swept spectrum analyzer (SSA) used to measure the electric field (E -field) level in the vicinity of WLAN sources by comparison to the results obtained by a real-time spectrum analyzer (RTSA), which is highly advanced but prohibitive for many metrological labs. Measurement sessions investigated both stable WLAN signals produced in the lab in ideal conditions by a vector signal generator and realistic WLAN signals produced by an access point, in various operational conditions, with the scope of protocol improvement. The major problem is that WLAN signals are pulsed, short duration and wide-banded. One of the key aspects when using a classic SSA concerns the establishment of its proper settings, main of them being the sweep time, which must be chosen according to the active durations of the signal. Analysis of settings influence upon E -field precision assessment is made. Obtained results show that SSA may be used for accurate exposure assessment of stochastic-like signals, but an adequate procedure should be applied.

1. INTRODUCTION

Broadband wireless local networks (WLAN) have become widespread in the last years, as a variant of fast and efficient implementation of data communication services, especially in indoor transmission environment. There is a need to increase precision of field level assessment (for human exposure purposes) of stochastic signals for data communication like those used in WLANs [1–4]. The variant of WLAN communication standard chosen in present study is IEEE 802.11g. It operates in the 2.4 GHz band, allowing high data rates — 54 Mbps, while the maximum radiated power by the router’s antenna is of 100 mW. WLAN signals are pulsed, burst like, of short duration [5], characteristics that impose specific requirements of the field-level measurement systems and proper measurement protocol. Presently, a standardized procedure for exposure assessment due to such signals lacks.

The paper is focused on analysis of a swept spectrum analyzer (SSA) capability to accurately measure exposure levels of WLAN signals, by comparison to a vector signal analyzer (real time spectrum analyzer — RTSA).

A SSA sweeps a narrowband filter across a range of frequencies (SPAN), sequentially measuring one frequency at a time. Unfortunately, sweeping the input works well for stable signals but will not accurately represent signals that change during the sweep. The RTSA simulate a parallel bank of filters overcoming swept limitations. If the input is a transient signal, the entire signal event is captured — by using Fast Fourier Transform (FFT) and the “instantaneous” complex spectra can be displayed in real time. It has been shown that there are cases when a SSA can be appropriate for WLAN field-level assessment, but a set of optimal settings must be chosen [1]. The procedure recommended in [1] indicates the clear influence of sweep time (SWT) setting on the measurements results accuracy. By using recommendations given in [1] as a starting point, our paper further investigates improvements to be settled when choosing a SSA for WLAN signal measurements by comparing electric field (E -field) level values measured by the SSA with the ones measured by the RTSA (considered as reference). SSA capability and its proper use in various conditions are identified for this one, in order to get reliable results. Present study considered two experimental scenarios. In the first one, a stationary, perfect reproducible and ideal WLAN signal is provided by a vector signal generator, and in the second one, a real WLAN signal is provided by an IEEE 802.11g access point (AP), used in two operating modes: idle mode and broadcasting mode. E -field levels are measured in each scenario, by comparison, by the two spectrum analyzers and results are compared for precision estimation.

2. MATERIALS AND METHOD

The two spectrum analyzer models were: a RTSA model Keithley 2820 and a SSA model FSH6 Signal Analyzer by Rohde & Schwarz. For emission of the stationary, reproducible (non-stochastic),

ideal WLAN signal, a SM 300 RF Signal Generator by Rohde & Schwarz was used (frequency range 100 kHz–3 GHz), coupled to the Dual Channel Arbitrary/Function Generator AM 300 by Rohde & Schwarz, which formed the signal under test in the base-band by using the WLAN Software package WinIQSim. The antenna used for emission was a ridged horn antenna model BBHA 9120A by Schwarzbeck Mess Elektronik. For emission of the real WLAN signal, a Trendnet TEW-432BRP wireless broadband router/AP was used and a workstation and a PC connected to the AP through UTP cable were enabled in infrastructure architecture. At the reception, Precision Conical Dipole Antenna PCD8250 by ARC Seibersdorf (antenna factor 42.7 dB/m@ 2.42 GHz) was connected via a RF cable (−2.6 dB cable loss) to both spectrum analyzers consecutively for measurements. The antenna was positioned in three orthogonal directions so as a finally isotropic response to be achieved for the measured total E -field level (root mean square — RMS value). The experimental set-up is presented in Fig. 1(a). Significant settings at the reception are presented below.

Taking into account the stochastic nature of the WLAN signal it is necessary to use a RMS or power-average detector to avoid underestimating of the field level [1, 3]. To correctly measure the power of the signal, the resolution bandwidth (RBW) has to be larger than frequency separation, Δf , between two consecutive display points. At the same time, RBW requires a value that allows minimizing adjacent channel interference. In [6] it is recommended that RBW to be 1%–3% of channel bandwidth, so as to obtain a sufficient attenuation of adjacent channel interference. It can be concluded that the frequency SPAN and RBW are selected as follows:

$$\Delta f = \frac{SPAN}{M} < RBW < 0.03 \cdot Channel\ BW \quad (1)$$

In case of WLAN signals, a channel of 20 MHz needs a 600 kHz RBW, to distinguish between different channels and minimize overlap if adjacent channels are present. To reduce the effects of noise on the display trace, spectrum analyzers include a variable video filter with a bandwidth (VBW) that is usually selected to be $(3 \div 10) \times RBW$ for pulsed signals. In case of RTSA analyzer, it is important to consider the real time bandwidth (RTBW) performance indicator, which is the maximum frequency SPAN that can be continuously processed without missing any event in the input signal. In present case the maximum SPAN must be:

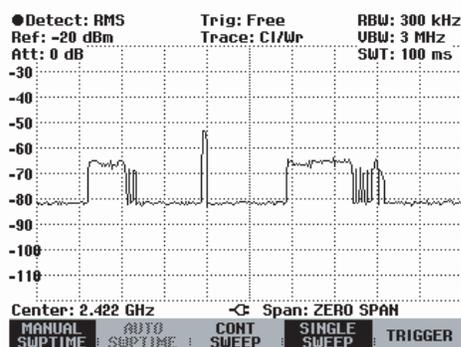
$$SPAN_{max} < RTBW \quad (2)$$

If one uses a RMS detector, than setting a long SWT has the effect of underestimating the power level because the analyzer takes samples both in the periods when the signal is active and when it is not active. For a short SWT an overestimate of the field is expected. The optimal value of the SWT (SWT_{optim}) may be obtained taking into account the period when the signal is active (t_{active}) and the response time and display time of the instrument [1]. In relationship (3) below, M represents the number of display points of the analyzer:

$$SWT_{optim} = t_{active} \cdot M \quad (3)$$



(a)



(b)

Figure 1: (a) Experimental set-up. (b) Display capture of the FSH6 analyzer in span zero indicating active time in broadcasting mode.

For the FSH6 analyzer $M = 301$. Duration of t_{active} depends on traffic conditions and on the communication standard features. They impose the physical frame duration (that is the same as t_{active}) [5]:

$$\text{Physical frame duration} = (\text{Preamble} + \text{Header}) \text{ duration} + \frac{\text{length decoded packet}}{\text{data rate}} + \frac{\text{length Frame Check Sequence}}{\text{data rate}} \quad (4)$$

The determination of the signal duty cycle, T , needs measurement in Zero Span mode, with settings: $RBW = 300$ kHz; $VBW = 3$ MHz; detector = RMS ; $SWT =$ sufficient to distinguish between individual packets and to assure a correct estimate of T . Duty cycle T is determined either by direct measurement or by calculation, like:

$$T = \frac{t_{active}}{t_{total}} \cdot 100\% \quad (5)$$

The RMS level of the E -field measured by the PCD8220 antenna is calculated from its three orthogonal components by relationship:

$$E_{rms}^{active} = \sqrt{E_{rmsOx}^2 + E_{rmsOy}^2 + E_{rmsOz}^2} \quad (6)$$

The total average E -field level is then corrected by taking into consideration the duty cycle T :

$$E_{total}^{cor} = E_{rms}^{active} \cdot \sqrt{T} \quad (7)$$

3. EXPERIMENTAL CASES, RESULTS AND DISCUSSIONS

The results of measurements in stable signal conditions — when WLAN signal is provided by a vector signal generator, and in realistic emission conditions — when signal is provided by an IEEE 802.11g AP, are presented in Table 1. The table indicates, by comparison, the E -field level measured by the two analyzer types — the RTSA (Keithley 2820) and the SSA (FSH6, Rohde & Schwarz).

R & S SM300 signal generator was set on Channel 3 of the IEEE 802.11g standard (2.422 GHz) with a power = 0 dBm for the carrier frequency. IQ signal characteristics, obtained by using the R & S AM300 generator and the WinIQSim software, were: modulation PSDU 64 QAM (54 Mbps), data length = 4095 bytes, idle time = 100 μ s. Antenna heights, for both emission and reception antennas was 125 cm above the ground, and the distance between the antennas was 100 cm. Trendnet TEW-432BRP AP was also set to emit on Channel 3 (2.422 GHz) at a constant data rate of 54 Mbps and positioned on a tripod at a height of 125 cm above the ground at a distance of 30 cm from the PCD 8250 receiving antenna. Both idle and traffic modes of the router were enabled and the E -field measured for the exposure assessment. A SPAN of 30 MHz was chosen for both analyzers in order to respect the real time bandwidth (RTBW) of the RTSA. The duty cycle correction (T) of the total average E -field was determined by direct measurement of the signal active duration (t_{active}).

Table 1: Measurement results: $RBW = 300$ kHz, $VBW = 3$ MHz, $SPAN = 30$ MHz, RMS detector, Max-Hold Trace.

| Scenario | RTSA Keithley 2820 | | | | SSA FSH6 | | | | | | | |
|-----------------------|--------------------|-------------------|-------------------|--------------------|----------------------------|---------------------|------------|-------------------|-------------------|-------------------|---------------------|--------------------|
| | E_{Ox} [V/m] | E_{Oy} [V/m] | E_{Oz} [V/m] | E_{tot} [V/m] | t_{active} [μ s] | SWT_{opt} [ms] | T [%] | E_{Ox} [V/m] | E_{Oy} [V/m] | E_{Oz} [V/m] | E_{meas} [V/m] | E_{tot} [V/m] |
| WLAN-signal generator | 0.13 | 0.14 | 0.08 | 0.21 | 630 | 190 | 86 | 0.14 | 0.16 | 0.09 | 0.24 | 0.22 |
| AP-idle mode | 0.14 | 0.15 | 0.24 | 0.31 | 900 | 2.7 | 0.9 | 0.75 | 0.5 | 0.38 | 0.97 | 0.29 |
| AP-traffic mode | 0.27 | 0.18 | 0.11 | 0.34 | 30 | 9 | 8 | 0.76 | 0.45 | 0.46 | 0.99 | 0.28 |

In case of stable WLAN signal, very good agreement between RTSA and SSA measurements was obtained (Table 1, first line). Due to the high value of T , the SWT setting of the SSA is not critical, so the comparative results obtained by the two analyzers remain very similar even for SWT values very different from the optimal value.

For realistic WLAN signal, small and variable values of T were obtained (lines 2 and 3 in Table 1) and the SWT parameter, according to the signal active duration (t_{active}), becomes a critical parameter. In idle mode, due to limited FSH6 analyzer performance, SWT has been set to 20 ms (minimal possible), which led to an underestimation of the E -field by the SSA. In broadcasting mode (traffic), establishing the optimal value of the SWT of the SSA becomes highly important in achieving accurate measurements. Otherwise, significant differences between RTSA and SSA results of the field level are obtained. A correct estimation of t_{active} is a very difficult task in realistic, broadcasting conditions due to the stochastic nature of the signal. Therefore exposure assessment should take into account the “worst case”, which corresponds to the minimal SWT. Even in this case it is very hard to establish the appropriate SWT value, and for this a detailed statistical analysis is indicated. For example, in our experimental case three more experimental cases were chosen: $SWT_1 = 20$ ms; $SWT_2 = 50$ ms; $SWT_3 = 300$ ms. With these, the relative variations of the total E -field value measured by SSA as referenced to the RTSA values, were: $(\Delta E/E_{RTSA})_1 = 17\%$, $(\Delta E/E_{RTSA})_2 = 7\%$, $(\Delta E/E_{RTSA})_3 = 40\%$. The second case here indicates that a $SWT = 50$ ms gives a more accurate value of the exposure than the SWT_{optim} , which took into account “the worst case”. The relative variations of the field level indicated above demonstrate that a SSA has clear limitations in making accurate estimations when measuring variable active-duration of WLAN signals.

4. CONCLUSIONS

Main problem in accurately assess the electromagnetic exposure due to WLAN signals by a SSA is connected to their stochastic nature. The study concerned measurements on one single channel of the IEEE 802.11g band due to the restricted RTBW of the RTSA (used as a reference). Comparative results of E -field level measured by a SSA and a RTSA showed that a SSA can assess sufficiently accurate field level if some settings are carefully chosen (Span, RBW, VBW, detector type, trace mode, SWT and number of sweeps). The critical settings that influence precision are: minimal RBW, maximal Span, minimal SWT and minimal number of display points. The most important setting is however the SWT, which must be equal to the duration of the active WLAN signal. A SSA having a $SWT < 1$ ms is recommended for all cases. For analyzers having a small number of display points ($M < 300$) it is difficult to ensure a span enough large to cover the whole IEEE 802.11 spectrum (13 channels), so that such analyzers are not suitable for an accurate assessment.

Comparative results obtained by using the two types of analyzers clearly showed that relative inaccuracies may reach as much as 40%, depending on the SWT value setting in relation to the duration of the active signal (very variable in traffic conditions). For the ideal case, i.e., when a laboratory generator gives the WLAN signal, evaluation of active time is not hard to make. However for real WLAN signals and in broadcasting mode, when active time is random, correct measurements by using a SSA are dependent on averaging technique. For indoor propagation in relatively stable conditions, a reasonable averaging is possible, so as accurate results to be obtained. Further research concerns involvement of advanced statistical techniques that will ensure a more reliable evaluation of active time in variable traffic conditions.

ACKNOWLEDGMENT

This work was developed with the support of “IDEI” Project 367/2007 granted to Land Forces Academy, Sibiu, in year 2010 by Romanian Ministry of Education and Research.

REFERENCES

1. Verloock, L., W. Joseph, G. Vermeeren, and L. Martens, “Procedure for assessment of general public exposure from WLAN in offices and in wireless sensor network testbed,” *Health Physics*, Vol. 98, 628–638, April 2010.
2. Bornkessel, C. and M. Wuschek, “Exposure measurements in different WLAN-scenarios,” *International Conference and COST 281 Workshop on Emerging EMF Technologies, Potential Sensitive Groups and Health*, Graz, April 20–21, 2006.
3. Schmid, G., P. Preiner, D. Lager, R. Uberbacher, and R. Georg, “Exposure of the general

- public due to wireless LAN applications in public places,” *Radiat Protect Dosim*, Vol. 124, 48–52, 2007.
4. Miclaus, S., P. Bechet, I. Bouleanu, and R. Helbet, “Radiofrequency field distribution assessment in indoor areas covered by wireless local area networks,” *Advances in Electrical and Computer Engineering*, No. 1, 2009, ISSN 1582-7445.
 5. IEEE 802.11g, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Part 11, IEEE, 1999.
 6. Myhr, J., “Measurement method for exposure to electromagnetic field strength from WLAN systems,” Chalmers University of Technology, Goteborg, Sweden, 2004.

Frequency Tuned Planar Inverted F Antenna with L Shaped Slit Design for Wide Frequency Range

I. T. E. Elfergani¹, Abubakar Sadiq Hussaini^{1,2}, R. A. Abd-Alhameed¹, C. H. See¹,
M. M. Abusitta¹, H. I. Hraga¹, A. G. Alhaddad¹, and Jonathan Rodriguez²

¹Mobile and Satellite Communications Research Centre, University of Bradford
Bradford, West Yorkshire, BD7 1DP, UK

²Instituto de Telecomunicações, Aveiro, Portugal

Abstract— A frequency tuned antenna has been designed to meet the coverage requirements of the DCS, PCS, UMTS and WLAN bands. The antenna consists of a main patch, and a planar inverted L (PIL) slot. The radiator patch is fed, and shorted, using simple feed lines with broadband characteristics. The handset represents the finite ground plane, and a varactor diode is mounted across the middle of the slot for tuning purposes. Initial tuning was obtained by placing lumped capacitors, instead of the varactor, over the radiator. Good agreement is obtained between the predicted and measured input return loss, gain and radiation pattern over the tuned frequency range.

1. INTRODUCTION

PIFA designs tend to be low profile, light weight and efficient space filling structures, and as such, are particularly attractive for handset and terminal applications. Conventional PIFA designs have constrained bandwidth; however, it is possible to realize novel structures which are electronically tunable over most of the wireless communication bands. Tunable multi-functional handset modules employ the same basic design aims to provide favorable trade-offs in terms of volume, weight and performance. Many interesting PIFA, and more general tuned printed antennas, have been proposed [1]. Various switching technologies, such as RF switches, MEMs switches, PIN diodes and varactor diodes have been used in reconfigurable antenna designs [2–4]. The varactor diodes in particular seem to offer a rich possibility for future designs over a wide frequency range, due to their excellent DC voltage controlled reactance property.

The increasing trend for slimmer, highly integrated mobile handsets and terminals, operating over several distinct and over-lapping frequency bands, has created a need for reconfigurable (tuned) antenna designs [5]. However, this trend towards slimmer, more space efficient integrated designs, poses a significant problem for the antenna designer. The trade-off between small antenna size, adequate bandwidth, and high efficiency, creates a difficult competing multi-objective optimization, which is also subject to fundamental physical constraints. To some extent this coverage problem can be solved by using multiple antennas, but the design of such an antenna module would add mutual coupling effects to an already substantial list of physical constraints.

Planar inverted F antennas (PIFA) may be modified to operate over multiple frequency bands by using, e.g., parasitic elements, and/or by the formation of several surface current paths through the introduction of slots in the radiator element. Recently, PIFA structures providing comprehensive coverage of the frequency ranges of six operating standards have been presented [6]. A common approach to tuning the operating frequency of a PIFA is to reconfigure a short circuit connection with an external tuning circuit. This has been used in single and dual band designs [7, 8]. Other reported tuning methods include the use of adjustable reactive components between the PIFA patch and a ground plane [9, 10], and switched tuning stubs, which have been applied in both single and dual band designs [3, 11].

2. ANTENNA DESIGN CONCEPT

The design optimization for this study requires an antenna structure compatible with contemporary handset chassis dimensions. The structural and lumped element parameters were simulated using the general electromagnetic analysis packages HFSS [12]. The optimization was tracked in terms of the antenna return loss, radiation patterns and power gain. Initial frequency tuning is tested and evaluated by placing lumped capacitors in the range of 0.5 to 3.3 pF along the slot of the radiator. This process is equivalent to replace the varactor operation. The design was further constrained by the impedance matching requirements covering a frequency range comprising the DCS, PCS, UMTS and W-LAN bands.

The prototype radiator patch is suspended over a $50\text{ mm} \times 80\text{ mm}$ (W_{board} , L_{board}) ground plane with a $2\text{ mm} \times 7\text{ mm}$ shorting pin, and the antenna is fed by means of a vertical plate, of maximum height 6.5 mm and width 2 mm as shown in Fig. 1. It is also connected to the feeding probe through the slot in the ground plane for which the effective substrate is considered air. The rectangular patch dimensions, (L_{ant} , W_{ant}), are $14\text{ mm} \times 50\text{ mm}$. The L-slot has a uniform width (S_s) of 1 mm , and the lateral dimensions, (L_s , W_s , and W_{gap}), are 8 mm , 17.5 mm , and 15 mm .

3. IMPACT OF THE LOADING ON THE ANTENNA

The impact of the loading capacitance on overall antenna performance, including the return loss, gain and radiation pattern needs to be investigated both in simulation, and through performance assessments on the physical prototype. For comparison purposes in simulation, we require a fully characterized reference antenna which omits the varactor; this structure resonates at 2.47 GHz (see Fig. 2). A loading capacitor with values selected from $\{0.5\text{--}3.3\}\text{ pF}$ is placed over the slot radiator in a fixed location, making the antenna resonate over wide range $1.75 \leq f_0 \leq 2.4\text{ GHz}$. In the first instance the tuning is investigated via a lumped (ceramic) capacitor, instead of a varactor, therefore the DC bias and circuit parasitic effects are excluded. The effects of lumped capacitors have been tested on the working prototype, in order to find out how the loaded capacitors could affect the antenna performance; and the optimal location on the slot ‘arm’. The working prototype in Fig. 3 was tested without the bias circuit. The loading capacitor was varied over $[0.5, 3.3]\text{ pF}$, which was found to be sufficient to control the antenna response.

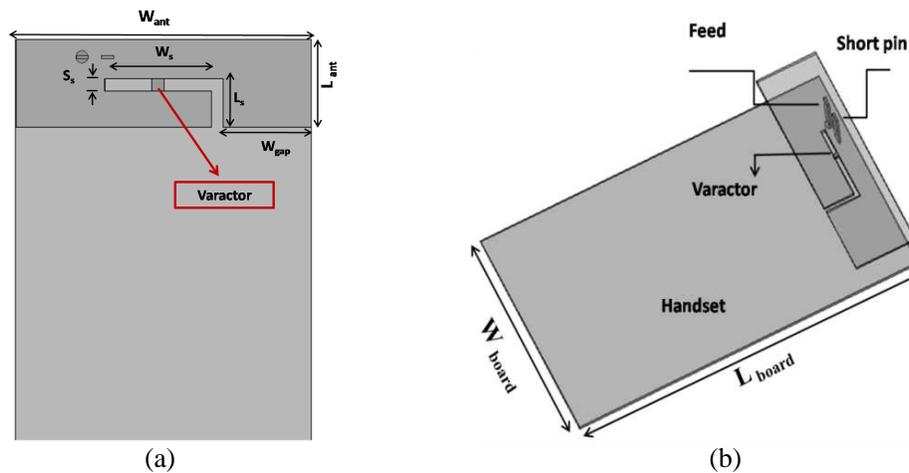


Figure 1: Basic antenna structure. (a) Top view, (b) 3D.

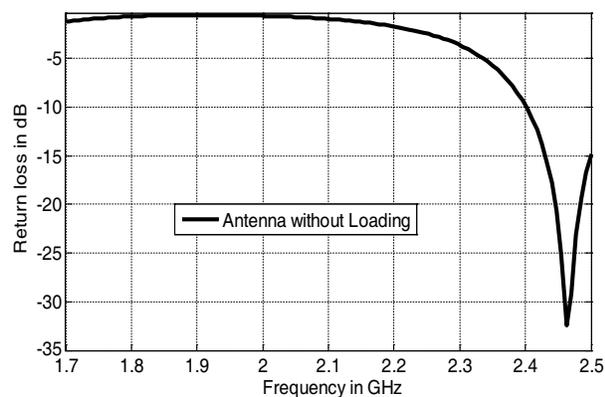


Figure 2: Return loss of antenna prototype, without loading capacitor present.



Figure 3: The prototype (loaded) antenna structure.

4. SIMULATED AND MEASURED RESULT DISCUSSIONS

This tuned antenna design relies on the introduction of the L-shaped slot on the radiator arm, and a varactor diode is attached to a location over the slot to achieve the required tuning. As the capacitance is varied, tuning is exhibited over a wide frequency band. The first resonance occurs at 2400 MHz, and then shifts, to the lower frequency band of 1750 MHz. However, the resonant frequencies are not only be controlled by changing the value of applied capacitance of the varactor, but also by its location along the slot.

Figure 4 illustrates the simulated and the measured return loss, the simulated was generated by using HFSS software package while the measured was obtained using a HP 8510C VNA. Five frequencies were selected: 1750 MHz, 1850 MHz, 2040 MHz, 2200 MHz, and 2400 MHz, respectively. The results show a satisfactory agreement between the results achieved by both simulation and measurement. It can be seen that the antenna is capable of covering the frequency ranges of digital cellular system (DCS: 1710–1880 MHz), personal communication service (PCS: 1850–1990 MHz), universal mobile telecommunications system (UMTS: 1900–2200 MHz), wireless LAN (WLAN: 2400 MHz).

Figure 5 shows the measured and simulated gains of the proposed antenna across the 2400 MHz, 2200 MHz, 2040 MHz, 1850 MHz and 1750 MHz bands. It can be seen that the maximum measured

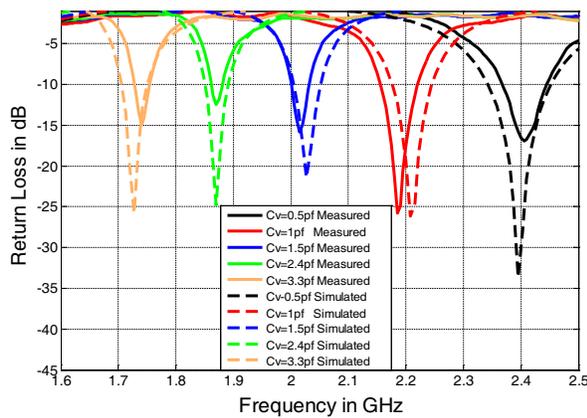


Figure 4: Measured and simulated return loss.

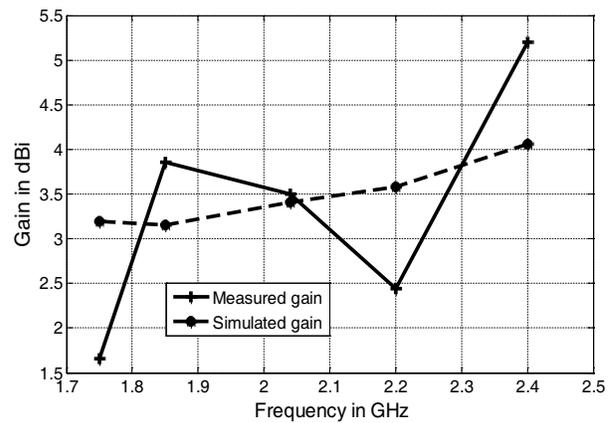
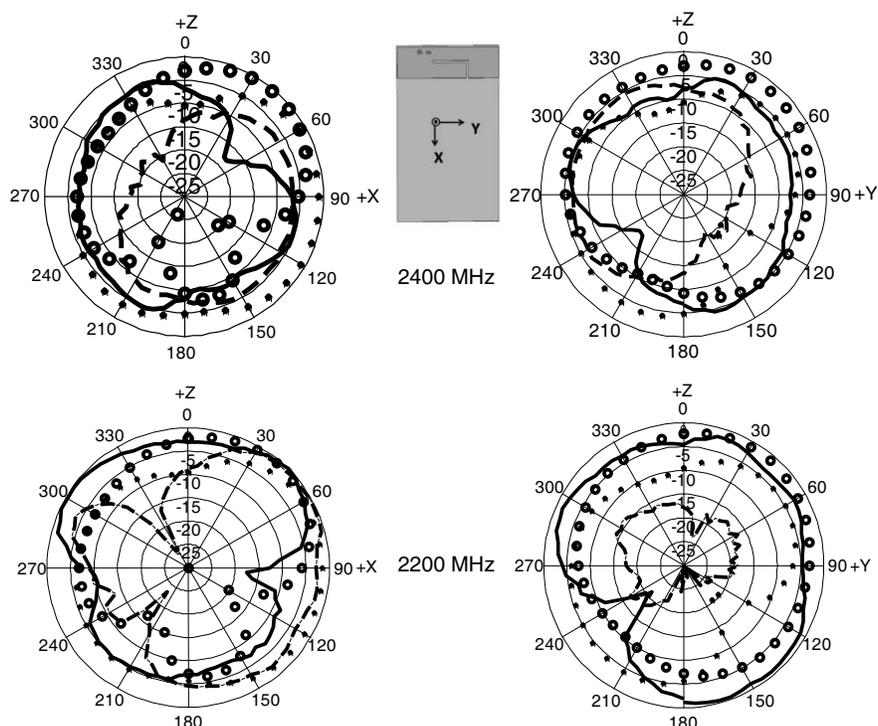


Figure 5: Simulated and measured antenna gains.



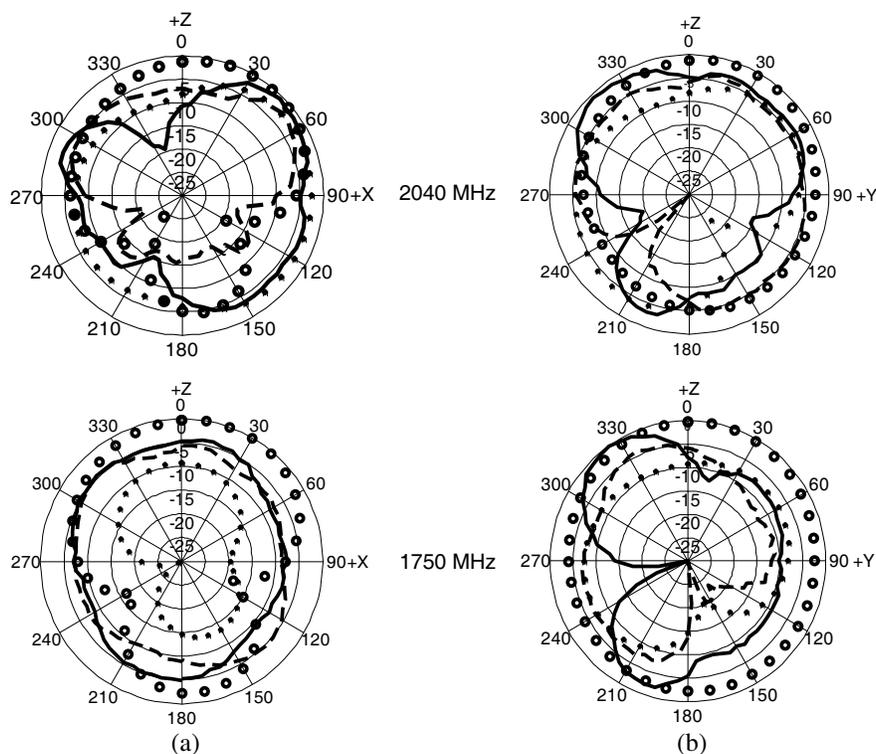


Figure 6: Measured and simulated normalised radiation field patterns of the proposed antenna for various operating frequencies. (a) xz plane, (b) yz plane. ‘o’: Simulated co-polar, ‘*’: Simulated cross-polar, ‘—’: Measured co-polar, ‘---’: Measured cross-polar.

gains values for the mentioned five bands were 1.6 dB_i , 3.8 dB_i , 3.2 dB_i , 2.5 dB_i , and 5.2 dB_i , respectively. The maximum simulated gains over the five bands were 3.3 dB_i , 3.4 dB_i , 3.5 dB_i , 3.55 dB_i , and 4.1 dB_i , respectively.

The simulated and measured radiation patterns are quite similar, and are shown in Fig. 6. The slight variations in these field patterns may be attributed in inaccuracies introduced during fabrication, and construction, the patterns are essentially omni-directional. The far field measurements were performed in an anechoic chamber, the simulated patterns were generated using HFSS. Two pattern cuts were selected for three operating frequencies covering the whole designated bandwidth; radiation patterns in the xz -plane and yz -plane at 2400 MHz, 2200 MHz 2040 MHz, and 1750 MHz were measured.

5. CONCLUSIONS

In this paper, a PIFA-tuned antenna is proposed and validated through practical measurement. By carefully selecting different sets of optimal structural and tuning parameters, basically, this antenna has wide frequency range operation that covers DCS, PCS, UMTS, and WLAN that makes it suitable for multifunctional mobile handsets or terminals. By attaching a varactor diode on the L-slot, the proposed antenna achieved more wideband frequency coverage. In addition, the antenna size was small enough to be fit within the mobile handsets. The proposed frequency-tuning concept can be extended to more complex PIFA structures as well as other antenna types. Instead of the varactor diode, RF diode switch, or RF MEMS components may be used to further improve the electrical performance and functionality of the antenna.

REFERENCES

1. Wu, J. F., C. J. Panagamuwa, P. McEvoy, J. C. Vardaxoglou, and O. A. Saraereh, “Switching a dual band PIFA to operate in four bands,” *IEEE Antennas and Propagation Society International Symposium*, 2675–2678, 2006.
2. Chiu, C. W. and F. L. Lin, “Compact dual-band PIFA with multi-resonators,” *Electronics Letters*, Vol. 38, No. 12, 538–540, Jun. 2002.

3. Behdad, N. and K. Sarabandi, “A varactor-tuned dual-band slot antenna,” *IEEE Transactions on Antennas and Propagation*, Vol. 54, No. 2, Part 1, 401–408, Feb. 2006.
4. Panaia, P., C. Luxey, G. Jacquemod, R. Staraj, G. Kossiavas, L. Dussopt, F. Vacherand, and C. Billard, “MEMS-based reconfigurable antennas,” *IEEE Int. Symp. on Industrial Electronics*, Vol. 1, 175–179, 2004.
5. Lindberg, P., “Wide-band active and passive antenna solutions for hand held terminals,” Ph.D. Dissertation, Uppsala University, Sweden, Feb. 2007.
6. Park, H., K. Chung, and J. Choi, “Design of a planar inverted-F antenna with very wide impedance bandwidth,” *IEEE Microw. Wireless Comp. Lett.*, Vol. 16, No. 3, 113–115, Mar. 2006.
7. Louhos, J. P. and I. Pankinaho, “Electrical tuning of integrated mobile phone antennas,” *Proc. Antenna Applications Symp.*, 69–97, Monticello, IL, 1999.
8. Karmakar, N. C., “Shorting strap tunable stacked patch PIFA,” *IEEE Transactions on Antennas and Propagation*, Vol. 52, No. 11, 2877–2883, Nov. 2004.
9. Panayi, P. K., M. O. Al-Nuaimi, and I. P. Ivrisimtzis, “Tuning techniques for planar inverted-F antenna,” *IEEE Electron. Lett.*, Vol. 37, No. 16, 1003–1004, Aug. 2001.
10. Karmakar, N. C., “Shorting strap tunable stacked patch PIFA,” *IEEE Transactions on Antennas and Propagation*, Vol. 52, No. 11, 2877–2883, Nov. 2004.
11. Oh, S. K., H. S. Yoon, and S. O. Park, “A PIFA-type varactor-tunable slim antenna with a pill patch feed for multiband applications,” *IEEE Antennas and Wireless Propagation Letters*, Vol. 6, 103–105, 2007.
12. Ansoft High Frequency Structure Simulator v10 Users Guide, CA, USA.

Beam Steering of Time Modulated Antenna Arrays Using Particle Swarm Optimization

M. M. Abusitta¹, R. A. Abd-Alhameed¹, I. T. E. Elfergani¹, A. D. Adebola¹, and P. S. Excell²

¹Mobile and Satellite Communications Research Centre, University of Bradford, UK

²Centre for Applied Internet Research, Glyndŵr University, Wrexham, Wales, UK

Abstract— In this paper, a simple switching process is employed to steer the beam of a vertically polarised circular antenna array. This is a simple method, in which the difference resulting from the induced currents when the radiating/loaded element is connected/disconnected from the ground plane. A time modulated switching process is applied through particle swarm optimisation.

1. INTRODUCTION

The current growth in the wireless communications industry has brought about the need for greater performance through improved capacity, data rates and reduce interference [1, 2]. The application of beam forming techniques on multiple-antenna arrays has increased in importance in the commercial wireless sector [3–6]. There are two generic approaches to beam-forming: dynamic phased array beam-forming, and adaptive beam-forming. In dynamic phased array beam-forming, the gain is maximised towards an intended user or target. Adaptive beam-forming places the nulls of an array pattern towards interfering or redundant signals, in addition to directing the maximum beam of the array towards the intended user (signal) [7]. Although adaptive beam-forming may be the preferred method, there are issues in the multi-criteria optimisation required to steer the main-lobe of the array in the desired direction(s), in a dynamically changing signal environment. This optimisation will operate through the amplitude and phase of each element, either continuously, or in discrete steps. The performance of the array is affected principally by its geometry and radiation pattern.

Most previous research on beam steering has been for linear [9], circular or planar arrays [2–8, 10], although it may be applied to other geometries. The uniform linear array provides the simplest array geometry, and the array processing is straight forward, however, it does not provide full view in the azimuthal range because of its use of edge elements. Even though this limitation can be overcome by use of several uniform arrays placed in a triangular or rectangular shape, this solution intrinsically increases costs and processing overheads. Beams formed by uniform linear arrays tend to broaden in a significant manner when the array is steered away from its bore sight. Uniform rectangular arrays, which have no omnidirectional elements are also unable to provide full azimuthal coverage. Uniform circular arrays on the other hand, provides full view of the azimuthal range, and depending on the element radiation pattern, can also provide a certain degree of source elevation information. By breaking down the array excitation into series of symmetrical spatial components, circular arrays have been shown to have a high potential of overcoming the effects of mutual coupling in antenna arrays.

Phase shifting is the most popular method of realizing beam steering. The feed to each element is varied such that the received or transmitted signals from all the elements are in phase in a given direction. Phase shifting may be realised using ferrite or electronic phase shifters at RF (or IF), or by using digital signal processing at the baseband.

Beam steering may also be realised through a time modulated array in which RF switches are used to control the ON and OFF time of the radiator elements to obtain a time average radiation pattern at the fundamental frequency, and with low side lobes. This method is constrained by the generation of spurious signals which occur as multiples of the fundamental frequency, which contributes to energy losses, as well as interference. This constraint may be turned into an advantage if through a simultaneous scan operation, based on time modulation, the unwanted sideband beams are used to point in another direction. This method may be seen as preferable in broadband operations as the shape of the main beam will remain constant over a broad bandwidth.

The optimisation techniques utilise random transition rules rather than deterministic ones, they are not limited by local minima, and can be efficient in large scale optimisation problems. Prominent optimisation algorithms that have been applied in previous research on beam steering are genetic algorithm, ant colony optimisation and PSO. Among these, PSO is our algorithm of choice as it is

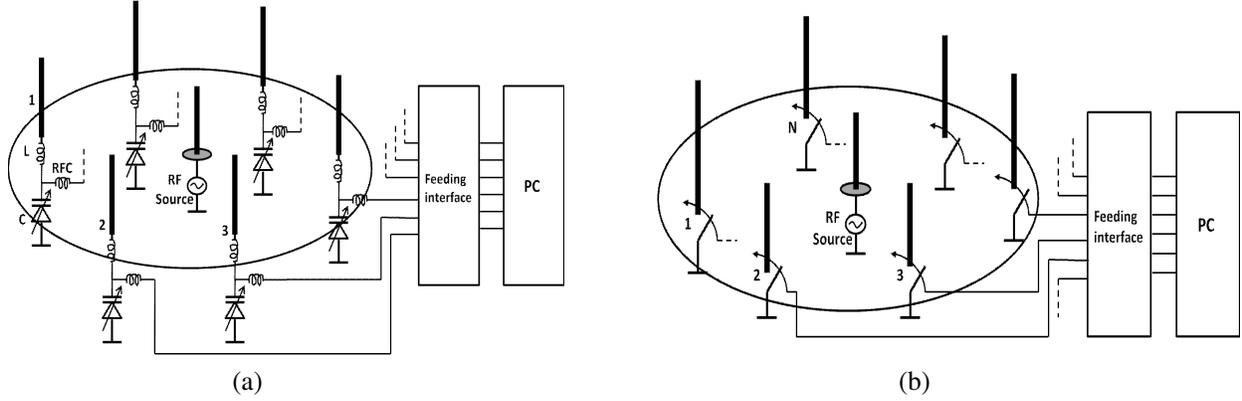


Figure 1: Beam steering ring antenna array; (a) Using capacitive loaded elements, (b) Using simple time switches.

simpler to implement, uses fewer parameters, and has the ability to perform local search, control convergence, and is insensitive to scaling of design variables [11].

This paper describes a simple beam steering exercise for a circular array, using time modulated elements, through switching the conducting current from each radiator to the ground plane. The differenced values of the currents for switch states are used to fully generate the required radiation pattern, with beam steering control facilitated by particle swarm optimisation (PSO). The present work demonstrates the method by simply connecting one element at a short time of one harmonic cycle, and leaving the others disconnected to ground plane. Several results are presented to validate the concept of this method.

2. ANTENNA MODEL AND ANALYSIS

The vertically polarised antenna geometry is given in Figure 1. The field pattern can be given by,

$$E = E_m + \sum_{i=1}^N (E_i) \quad (1)$$

For the horizontal polarisation case, then $E_m = 0$. Taking the central element as the reference element, and assuming all the radiators are quarter-wavelength, then Eq. (1) can be approximated by the following form:

$$E = I_m \left(\hat{d}_m \cdot \hat{u}_\theta \hat{a}_\theta + \hat{d}_m \cdot \hat{u}_\phi \hat{a}_\phi \right) + \sum_{i=1}^N \left\{ I_i \left[\begin{array}{l} \hat{d}_{ei} \cdot \hat{u}_\theta \hat{a}_\theta + \\ \hat{d}_{ei} \cdot \hat{u}_\phi \hat{a}_\phi \end{array} \right] e^{-jk\bar{P}_i \cdot \hat{r}} \right\} \quad (2)$$

where d_m and d_e are unit vectors describing the orientations of the main and outer elements. These may be simply defined as,

$$\hat{d} = \sin(\theta_d) \cos(\phi_d) \hat{a}_x + \sin(\theta_d) \sin(\phi_d) \hat{a}_y + \cos(\theta_d) \hat{a}_z \quad (3)$$

θ_d , and ϕ_d are the zenith and azimuth. I_m and I_i are the current maxima of the main elements, and the i th outer elements, respectively. The i th \hat{r} , \hat{u}_θ and \hat{u}_ϕ ; position vector of the i th radiator is denoted by p_i , with reference to Figure 1, this vector can be written as,

$$\bar{p}_i = r_a \cos \phi_i \hat{a}_x + r_a \sin \phi_i \hat{a}_y \quad (4)$$

The unit vectors (\hat{r} , \hat{u}_θ and \hat{u}_ϕ) are given by the following:

$$\hat{r} = \sin \theta \cos \phi \hat{a}_x + \sin \theta \sin \phi \hat{a}_y + \cos \theta \hat{a}_z \quad (5)$$

$$\hat{u}_\theta = \cos \theta \cos \phi \hat{a}_x + \cos \theta \sin \phi \hat{a}_y + \sin \theta \hat{a}_z \quad (6)$$

$$\hat{u}_\phi = -\sin \phi \hat{a}_x + \cos \phi \hat{a}_y \quad (7)$$

Using the time switching process on these elements, the fields may still be expressed by Eq. (2). It should be noted that,

$$\frac{|E_{\text{ion}}(\theta)|}{\max |E_{\text{ion}}(\theta)|} \approx \frac{|E_{\text{ioff}}(\theta)|}{\max |E_{\text{ioff}}(\theta)|} \quad (8)$$

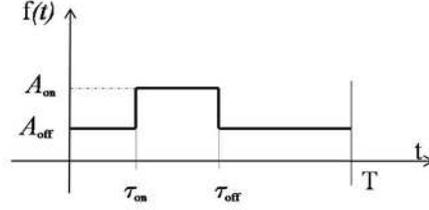


Figure 2: The variation of current over one time cycle.

In other words, the normalized fields due to ON and OFF states on the i th vertically polarised radiator are equivalent. These assumptions are based on the separation distance, d_s , for each radiator from the ground plane to be $d_s \ll \lambda$. The induced current on the i th radiator for ON and OFF shown in Figure 2 are expressed by A_{on} and A_{off} , respectively. The limited time constraint is applied on i th radiator as $|\tau_{\text{off}} - \tau_{\text{on}}| < T$.

The Fourier series coefficient for this pulse shape, can be given (using [2]):

$$f_{mi} = A_{\text{off}}T \sin c(0.5\pi mT)e^{-j0.5\pi mT} + (A_{\text{ion}} - A_{\text{ioff}}) \frac{\sin(\pi m(\tau_{\text{ioff}} - \tau_{\text{ion}}))}{\pi m} e^{-j\pi m(\tau_{\text{on}} + \tau_{\text{off}})} \quad (9)$$

Using the results in [2], the time switching can be found subject to the weighted induced currents (w_i) assumed over the antenna structure. This can be put into similar format as following:

$$\frac{\sin(\pi m(\tau_{\text{ioff}} - \tau_{\text{ion}}))}{\pi m} e^{-j\pi m(\tau_{\text{on}} + \tau_{\text{off}})} = \frac{w_i - A_{\text{off}}T \sin c(0.5\pi mT)e^{-j0.5\pi mT}}{(A_{\text{ion}} - A_{\text{ioff}})} = w_{di} \quad (10)$$

Assuming the same procedure used in [2], the switch-on and switch-off of the i th element can be given as follows:

$$\tau_{\text{ioff}} = \frac{1}{2\pi m} \left(\frac{\gamma_i}{\pi m} + \frac{1}{\pi m} \sin^{-1} \pi m |w_{di}| \right) \quad (11)$$

$$\tau_{\text{ion}} = \frac{1}{2\pi m} \left(\frac{\gamma_i}{\pi m} - \frac{1}{\pi m} \sin^{-1} \pi m |w_{di}| \right) \quad (12)$$

The induced currents A_i^{on} and A_i^{off} are known for all cases that can be considered within the antenna operation. It should also be noted, for a given baseband bandwidth B_a , the time period T might be predicted by $T > 1/B_a$. Thus, the i th element times, i.e., switch-on and switch-off can be easily determined.

3. LOCUS VARIATION OF THE WEIGHTING COEFFICIENTS

The boundary variations of the weighted coefficients can be estimated subject to Eqs. (11) and (12). This can be explained in following example when m equals 1. It should be noted that the coefficient w_{di} can be given by the following inequality:

$$\left| \frac{w_i - A_{\text{off}} \sin(0.5\pi) e^{-j0.5\pi}}{(A_{\text{ion}} - A_{\text{ioff}})} \right| \leq \frac{1}{\pi} \quad (13)$$

The above equation may be reduced further:

$$\pi \left| \frac{w_i - jA_{\text{off}}/\pi/2}{(A_{\text{ion}} - A_{\text{ioff}})} \right| \leq 1 \quad (14)$$

Substituting $w_i = x + jy$ into Eq. (14),

$$\left(x - \frac{\pi}{2} a_{if} \right)^2 + \left(y + \frac{\pi}{2} a_{rf} \right)^2 \leq \frac{1}{\pi^2} |A_{\text{ion}} - A_{\text{ioff}}|^2 \quad (15)$$

where a_{rf} and a_{if} are the real and imaginary parts of the complex current in the off state. Thus, the locus variation can be simplified using polar co-ordinates as follows:

$$x = a + r_d \cos \phi \quad y = b + r_d \sin \phi \quad (16)$$

where:

$$r_d = \frac{1}{\pi} |A_{\text{ion}} - A_{\text{ioff}}| \quad a = \frac{2}{\pi} a_{if} \quad b = -\frac{2}{\pi} a_{rf}.$$

4. SIMULATION RESULTS

Two examples are considered for this method. For each case a six element ring array of radius 0.375 wavelength is used. For the first example all the elements were fed through switching process, in which the mutual coupling is ignored, and no feed element is considered. Figure 3 shows the radiation pattern for four selected steering angles, and their time switching sequence is presented in Figure 4.

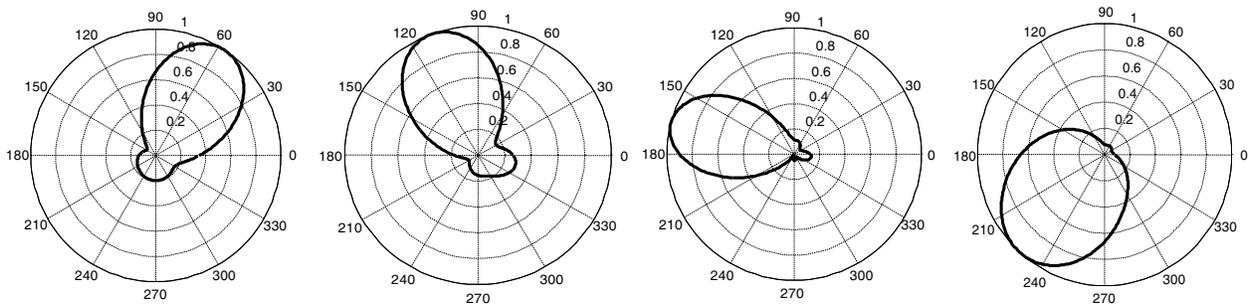


Figure 3: The radiation pattern of the beam steering angles 57°, 110°, 165°, 228°.

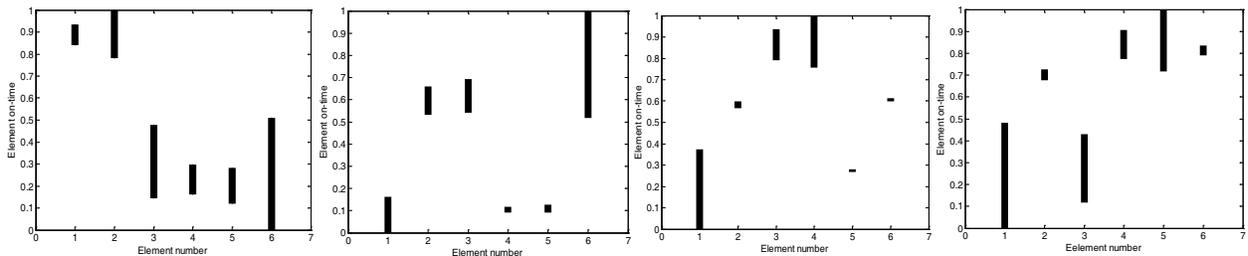


Figure 4: Normalized time-sequence corresponding to the antenna ring array steering angles in Figure 3.

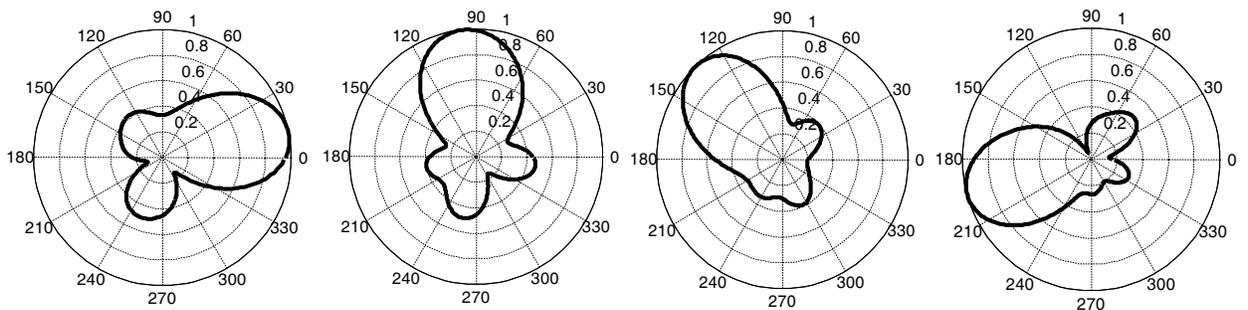


Figure 5: The radiation pattern of the beam steering angles 10°, 100°, 130°, 200°.

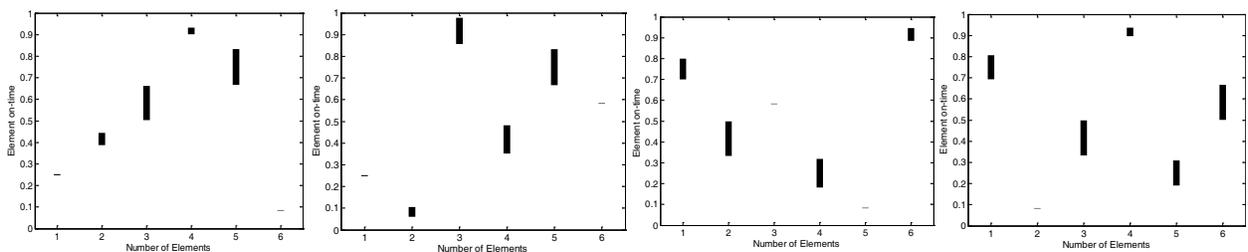


Figure 6: Normalized time-sequence corresponding to the antenna ring array steering angles in Figure 5.

Clearly, the required overlap switching for these elements is a very difficult task when mutual coupling is present.

In the second example, mutual coupling is included, and the feed line at the centre is applied. The induced currents were computed, when one element is on, and others are off. These are recorded as input data to the PSO. The radiation pattern requested for a few particular directions and their associated time sequences are presented in Figures 5 and 6. The overlap switching process is totally eliminated.

5. CONCLUSION

A simple procedure of beam steered of ring antenna array using time modulated switching process has been presented. The method is quite simple and it could be a good candidate to replace the loaded reactive steered antenna array.

ACKNOWLEDGMENT

The authors would like to thank their colleague Mr. M. B. Child for helpful comments and assistance.

REFERENCES

1. Li, G., S. Yang, Y. Chan, and Z.-P. Nie, "A novel electronic beam steering technique in time modulated antenna arrays," *Progress In Electromagnetics Research*, Vol. 97, 391–405, 2009.
2. Tong, Y. and A. Tennant, "Simultaneous control of sidelobe level and harmonic beam steering in time-modulated linear arrays," *Electronics Letters*, Vol. 46, No. 3, 201–202, 2010.
3. Tennant, A. and B. Chambers, "A two-element time-modulated array with direction-finding properties," *IEEE Antennas Wirel. Propag. Lett.*, Vol. 6, 64–65, 2007.
4. Tennant, A., "Experimental two-element time-modulated direction finding array," *IEEE Trans. Antennas and Propagation*, Vol. 58, No. 3, 986–988, 2010.
5. Bregains, J. C., et al., "Signal radiation and power losses of time-modulated arrays," *IEEE Trans. Antennas and Propagation*, Vol. 56, No. 6, 1799–1804, 2008.
6. Sugiura, S. and H. Iizuka, "Reactively steered ring antenna array for automotive application," *IEEE Trans. Antennas and Propagation*, Vol. 55, No. 7, 1902–1908, 2007.
7. Ioannides, P. and C. A. Balanis, "Uniform circular arrays for smart antennas," *IEEE Antennas and Propagation Magazine*, Vol. 47, No. 4, 192–206, 2005.
8. Abusitta, M. M., R. A. Abd-Alhameed, D. Zhou, C. H. See, S. M. R. Jones, and P. S. Excell, "New approach for designing beam steering uniform antenna arrays using genetic algorithms," 617–620, Loughborough, UK, November 16–17, 2009.
9. Morelos-Zaragoza, R. H. and M. Ghavami, "Combined beamforming and space-time block coding for high speed wireless indoor communications," *4th Int'l Symposium on Wireless Personal Multimedia Communications*, 1427–1431, Aalborg, Denmark, 2001.
10. Tan, C. M., et al., "On the application of circular arrays in direction finding. Part II: Experimental evaluation on sage with different circular arrays," *1st Annual COST 273 Workshop*, 29–30, Finland, 2002.
11. Robinson, J. and Y. Rahmat-Samii, "Particle swarm optimization in Electromagnetics," *IEEE Trans. Antennas and Propagation*, Vol. 52, No. 2, 397–407, 2004.

The Compact Design of Dual-band and Wideband Planar Inverted F-L-antennas for WLAN and UWB Applications

H. I. Hraga, C. H. See, R. A. Abd-Alhameed, S. Adnan, I. T. E. Elfergani, and F. Elmegri
Mobile and Satellite Communications Research Centre, University of Bradford, Bradford, UK

Abstract— Two miniature low profile PIFLA antennas with a compact volume size of $30\text{ mm} \times 15\text{ mm} \times 8\text{ mm}$ has presented in this paper. By applying the magnetic wall concept a reduced size dual-band and a wideband half PIFLAs for WLAN (2.4 GHz/5.2 GHz) and UWB applications are achieved. The dual-band antenna shows a relative bandwidth of 12% and 10.2% at ISM2400 and IEEE802.11a frequency bands respectively for input return loss less than 10 dB. By carefully tuning the geometry parameters of the dual-band proposed antenna, the two resonant frequencies can be merged to form a wide bandwidth characteristic, to cover 3000 MHz to 5400 MHz bandwidth (57%) for a similar input return loss that is fully covering the lower band UWB (3.1–4.8 GHz) spectrum. The experimental and simulated return losses on a small finite ground plane of size $30\text{ mm} \times 15\text{ mm}$ show good agreement. The computed and measured radiation patterns are shown to fully characterize the performance of the proposed two antennas.

1. INTRODUCTION

UWB technology is invented for short range and higher data rate communication. The lower and upper UWB spectrums are 3100 MHz to 4000 MHz and 6000 MHz to 10600 MHz respectively. There are three overlapping target segments that could benefit from short-range wireless connections enabled by UWB: PC and peripheral devices, mobile devices, and consumer electronics. By implementing this technology, it is believed capable of delivering 1 Gbit/s data rates over cable and wireless claims to double the exiting most of wireless terminal data-rate. This makes the constant high demand for designing smaller and lighter internal antennas, which immune to damage and low SAR.

There were several antennas size reduction techniques have been proposed over last decades that investigate the use of high permittivity substrate, shorting pins, shorting walls and modifying the geometry of the internal antenna [1–3]. Recently, another size reduction technique is proposed in [4–6] using the magnetic wall concept. It was interestingly found that the performances in terms of return losses, gains, radiation efficiencies, radiation patterns of half size structures of the U-slot, E-shaped [5] and UWB microstrip patch antennas [6] are comparable to their full structures.

2. ANTENNA DESIGN CONCEPT

The PIFLA shown in Fig. 1 is quite similar to the antenna design of the work reported in [7], but the size is reduced to half using the principle of the existing magnetic wall on the antenna surface [5, 6]. The initial geometry parameters of the antenna are stated as follows: $L_1 = 18.6\text{ mm}$, $L_2 = 10\text{ mm}$, $h_1 = 8\text{ mm}$, $h_2 = 4.5\text{ mm}$, $d = 3.5\text{ mm}$ and $w = 0\text{ mm}$. It should be noted that these dimensions are the same as in [7], except the width of F-shaped, L-shaped radiator and ground plane have been cropped to half, which are 7.5 mm, 8.5 mm and 15 mm respectively. The copper metal plate thickness of the proposed antenna and the gap distance for the feed are 0.5 mm.

The operation mechanism of this PIFLA is simple. As can be portrayed in Fig. 1, this antenna is a combination of a planar F-shaped antenna with rectangular plate feed and a planar L-shaped antenna. The F-shaped antenna which has a longer electrical length is designed to control the lower resonant mode (2450 MHz), where as the L-shaped antenna is used to provide the higher resonant mode (5200 MHz).

By keeping the same geometry parameters, the return losses of the half size PIFLA are studied. As can be seen in Fig. 2, the bandwidth for lower resonant mode remains, but the bandwidth for higher resonant mode degrades considerably in which 12.2% (5 GHz to 5.65 GHz) to work in [7] and 3.8% (5.15 GHz to 5.35 GHz) to the half structure.

3. SIMULATED AND MEASURED RESULT DISCUSSIONS

The prototype antennas were characterised by RF and anechoic chamber measurements. The RF measurements were carried out using an HP-8510C VNA. The radiation patterns were measured as follows. The fixed reference antenna was a broadband horn (EMCO 3115) and the spacing with

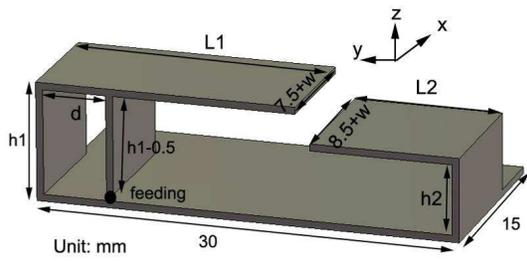


Figure 1: Geometry of the proposed miniature PIFLA.

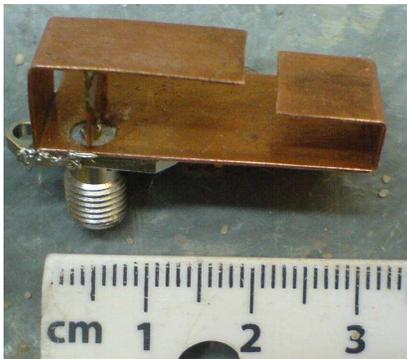


Figure 3: Practical prototype of the proposed dual-bands PIFLA.

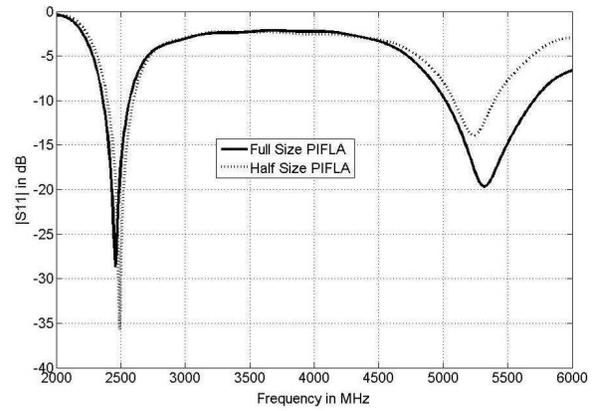


Figure 2: Simulated S_{11} (Full size PIFLA and Half size).

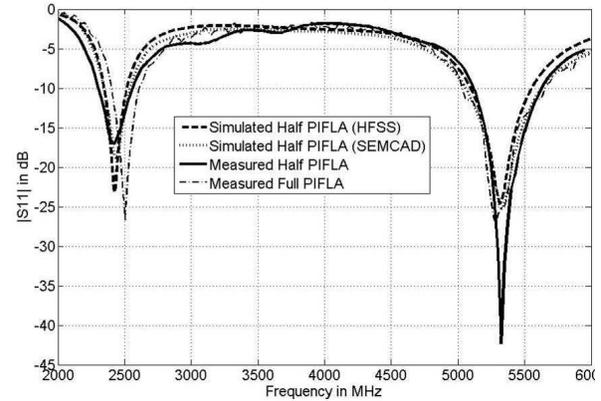


Figure 4: Measured and simulated S_{11} .

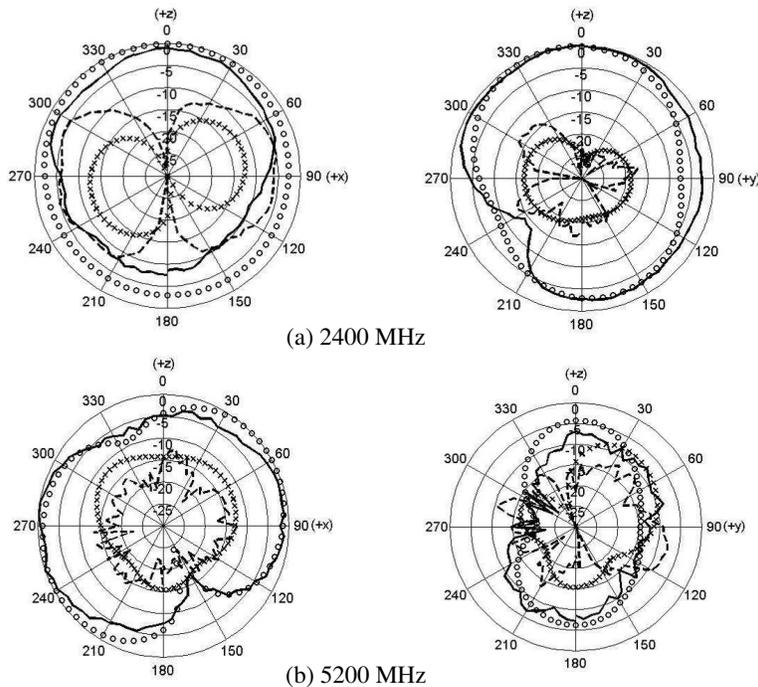


Figure 5: Simulated and measured normalised radiation patterns of the proposed dual-bands PIFLA for two planes (left: x - z plane, right: y - z plane) at (a) 2400 MHz and (b) 5200 MHz. ‘xxxx’ simulated cross-polarization; ‘oooo’ simulated co-polarization; ‘- - -’ measured cross-polarization; ‘—’ measured co-polarization.



Figure 6: Practical prototype of the proposed wideband PIFLA.

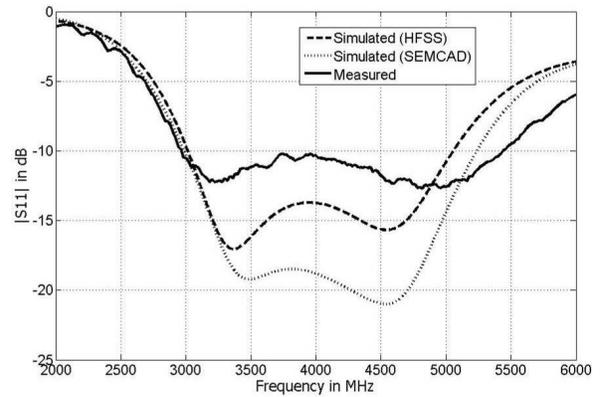


Figure 7: Measured and simulated S_{11} .

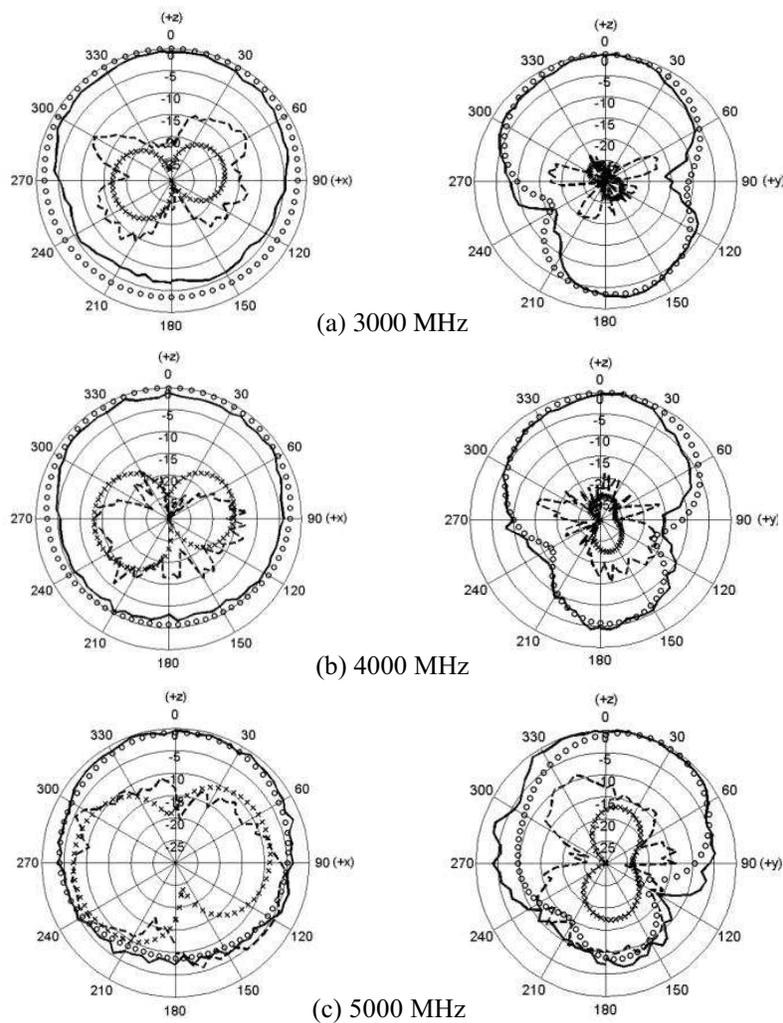


Figure 8: Simulated and measured radiation patterns of the proposed wideband PIFLA for two planes (left: x - z plane, right: y - z plane) at (a) 3000 MHz, (b) 4000 MHz and (c) 5000 MHz. ‘xxxx’ simulated cross-polarization; ‘oooo’ simulated co-polarization; ‘- - -’ measured cross-polarization; ‘—’ measured co-polarization.

the antenna under test (AUT) was fixed at 4 m. Two pattern cuts (H -plane and E -plane) were taken for each design frequency covering the target bandwidth.

3.1. Miniature Dual-band PIFLA

By scrutinizing the variations in the geometry against the impedance bandwidth, the optimal parameters can be recognized for dual-band operation. These parameters can be given as follows: $L_1 = 18.6$ mm, $L_2 = 10$ mm, $h_1 = 8$ mm, $h_2 = 5$ mm, $d = 4.5$ mm and $w = 1.5$ mm.

Figure 3 shows the first miniature dual-band PIFLA prototype. Fig. 4 illustrates the typical measured and computed antenna performance in terms of impedance bandwidth for both full-size and half-size PIFLAs by using two commercial packages (HFSS [8] and SEMCAD X [9]). The return loss is better than -10 dB, satisfying the desired IEEE802.11b/g frequency band (2400–2485 MHz) and IEEE 802.11a (5.15–5.35 GHz) bands respectively. As can be observed, simulated and measured results for half-size PIFLA were found to be in excellent agreement. The half-size and full-size PIFLA seem to show identical measured impedance bandwidth at two resonant modes.

Figures 5(a) and (b) show the simulated and measured co-polar and cross-polar radiation patterns in the x - z and y - z planes at 2400 MHz and 5200 MHz for the miniature PIFLA. The simulated and measured radiation patterns of the fabricated prototype are seen to be quite similar to each other at the two designated centre frequencies.

3.2. Wideband PIFLA

The optimised geometry parameters of the proposed PIFLA for UWB application are given as follows: $L_1 = 18.45$ mm, $L_2 = 8.5$ mm, $h_1 = 8$ mm, $h_2 = 5.5$ mm, $d = 14.5$ mm and $w = 4.5$ mm. Fig. 6 describes the experimental prototype of the PIFLA with finite ground plane of 30×15 mm, while Fig. 7 shows the typical measured and computed antenna performance in term of impedance bandwidth by using two commercial packages (HFSS [8] and SEMCAD X [9]). As can be seen, the lowest and highest frequency edges, i.e., 3000 MHz and 5400 MHz, of an input return loss ≤ -10 dB are observed. The impedance bandwidth of the proposed antenna, for -10 dB return loss, is 2400 MHz or about 57% with respect to the centre frequency at 3950 MHz (average of measured lower and higher frequencies with a -10 dB return loss), which fully covers the frequency spectrum of uplink UWB (3100 MHz to 4800 MHz). Basic agreement is achieved between the experimental and computed return loss over the desired operating frequency band.

The far-field radiation characteristics of the prototype were also investigated at 3000 MHz, 4000 MHz and 5000 MHz, as shown in Fig. 8. It is observed that the radiation patterns at different frequencies are quite similar, which is expected in a wideband antenna.

4. CONCLUSION

In this paper a simple geometry of a miniature planar inverted F-L antenna (PIFLA) has been proposed and studied experimentally and theoretically. By carefully selecting different sets of optimal geometry parameters and applying the size reduction techniques to the proposed antenna, two 50% size reduction PIFLAs have been designed and tested. By balancing the size and bandwidth constraints, these proposed antennas have a compact envelope dimension of 30 mm \times 15 mm \times 8 mm and covers the required operating frequency band for WLAN and UWB applications. These features make the proposed antenna an attractive candidate for application in a range of future mobile terminals.

REFERENCES

1. Skrivervik, A. K., J.-F. Zurcher, O. Staub, and J. R. Mosig, "PCS antenna design: The challenge of miniaturization," *IEEE Antennas and Propagation Magazine*, Vol. 43, 12–27, August 2001.
2. Lee, C. S. and K.-H. Tseng, "Size reduction of microstrip antennas," *Electron. Lett.*, Vol. 37, 1274–1275, 2001.
3. Shackelford, A. K., K. F. Lee, and K. M. Luk, "Design of small-size wideband width microstrip-patch antennas," *IEEE Antennas and Propagation Magazine*, Vol. 45, No. 1, 75–83, February 2003.
4. Deshmukh, A. A. and G. Kumar, "Half U-slot loaded rectangular microstrip antenna," *IEEE AP-S Int. Symp. USNC/CNC/URSI National Radio Science Meeting*, Vol. 2, 876–879, 2003.
5. Chair, R., C.-L. Mak, K.-F. Lee, K.-M. Luk and A. A. Kishk, "Miniature wide-band half U-slot and half E-shaped patch antennas," *IEEE Transactions on AP*, Vol. 53, No. 8, 2645–2651, August 2005.
6. Guo, L., S. Wang, X. Chen, and C. Parini, "Miniaturised antennas for UWB communications," *Proceeding of Eucap*, 3774–3778, Berlin, Germany, March 23–27, 2009.

7. See, C. H., R. A. Abd-Alhameed, D. Zhou, and P. S. Excell, “Dual-frequency Planar Inverted F-L-Antenna (PIFLA) for WLAN and short range communication systems,” *IEEE Transactions on AP*, Vol. 56, 3318–3320, October 2008
8. HFSS ver.11, Ansoft. Ltd.
9. SEMCAD X ver.14, Schmid & Partner Engineering AG, Zeughausstrasse 43, 8004, Zurich, Switzerland.

The Application in Spacecraft of High Temperature Superconducting Magnetic Energy Storage

Bo Yi¹ and Hui Huang^{1,2}

¹School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China

²State Key Laboratory of Millimeter Waves, Nanjing 210096, China

Abstract— This paper has analyzed the requirement of energy storage devices in spacecraft and introduced the present development situation of high temperature superconducting magnetic energy storage technology, then conceives an idea that applying the skill of high temperature superconducting magnetic energy storage in spacecraft to replace the existing energy storage devices in spacecraft. Also, a comprehensive comparing among the high temperature superconducting magnetic energy storage technology and other existing energy storage technologies such as flywheel energy storage technology, battery energy storage technology has been done. The conclusion that the high temperature superconducting magnetic energy storage technology has more advantages than other existing energy storage technologies in application of aerospace technology is reached.

1. THE REQUIREMENT OF ENERGY STORAGE DEVICES IN SPACECRAFT

Energy storage devices in spacecraft is used for transforming chemical energy and other types of energy into electric energy. Its main functions are below: (1) supplying electricity from spacecraft being launched to the time that solar panels started; (2) supplying electricity for spacecraft at dark period; (3) supplying auxiliary electricity when the supplying of solar battery is not sufficient at light period. Spacecrafts usually adopt battery as energy storage devices. But there are many obvious deficiencies, which will be mentioned below. Since the performance and lifetime of energy storage devices have limited the development of spacecraft for a long time, spacecraft has such aspects of requirements for energy storage devices: (1) the requirements for electrical property such as high energy, small mass and high power; (2) properly sealed in case of electrolyte evaporation for chemical battery; (3) capable of resisting for mechanical environments; (4) the requirements for heat output [1].

2. THE PRESENT DEVELOPMENT SITUATION OF HIGH TEMPERATURE SUPERCONDUCTING MAGNETIC ENERGY STORAGE TECHNOLOGY

Superconducting magnetic energy storage (SMES) is a remarkable application of superconducting magnets, especially for high temperature superconducting magnetic energy storage technology (HTSMES). The potential utilization of high-temperature superconducting materials is especially promising in order to obtain energy storage devices showing high efficiency and lifetime HTSMES device is made up of a superconducting coil, a power conditioning system and a refrigerator and a vacuum to keep the coil at low temperature. It stores electric energy as direct electric current passing through a superconducting coil. Because the superconducting material has nearly zero resistance, the electric current can almost circulate indefinitely. The energy stored in the coil can be calculated by the formula $E = 0.5LI^2$, where L is the inductance of the coil and I is the current passing through it. Storage capacity for HTSMES is quite outstanding. For SMES, it can be anything up to 20 MW. SMES devices can run for thousands of charge/discharge cycles without any degradation to the magnet, giving it a life of 20+ years. In the recent study, some scientists proposed the design of 300 MJ I-ITS superconducting magnetic energy storage system. In addition, from 2004 to 2007, a 10 MVA/20 MJ SMES prototype for a 100 MW commercial system was developed in Japan [2]. In 2010, the superconducting magnet designed with the basic requirements including a total storage energy of 2 MJ and a storage energy density of 2.73 MJ/m^3 has been achieved [3].

3. APPLYING THE SKILL OF HIGH TEMPERATURE SUPERCONDUCTING MAGNETIC ENERGY STORAGE IN SPACECRAFT

Due to the lack of air convection when spacecraft is on the orbit, the temperature of the side face to Sun can be heated to above 373 K while the temperature of the other side can be below 173 K.

The extreme temperature difference makes a perfect environment for applying of superconducting. In addition, although it costs a lot and takes a large proportion of volume and mass for a refrigerator and a vacuum to keep the coil at low temperature at superconducting storage devices, the refrigerator and devices to keep vacuum are not needed in space any more, which also makes the applying of HTS-SMES in spacecraft has more advantages.

However, the biggest problem for applying HTSMES in to spacecraft is the relatively low energy density, which is expected to solve in the inherent low temperature environment like in space.

4. A COMPREHENSIVE COMPARING AMONG THE HIGH TEMPERATURE SUPERCONDUCTING MAGNETIC ENERGY STORAGE TECHNOLOGY AND OTHER EXISTING ENERGY STORAGE TECHNOLOGIES

(1) Storage battery

Storage battery has such deficiencies: (1) the scope of its temperature admission is narrow; (2) the discharge capacity and efficiency decreased as the increasing of the number of recharge cycles; (3) the lifetime is short; (4) it takes much time for charging; (5) it may cause pollution.

(2) Flywheel energy storage

A flywheel is a mechanical device with a significant moment of inertia used as a storage device for rotational energy. The flywheel is proving to be an ideal form of energy storage on account of its high efficiency, long cycle life, wide operating temperature range, freedom from depth-of-discharge effects, and higher power and energy density — on both a mass and a volume basis [4–6]. There also some scientists proposed that the application of flywheel on satellite energy storage/attitude control system (ACS), which has advantages shown in Table 1 below [7].

Table 1: The advantages of ACS.

| the characteristic of attitude control system | advantages |
|---|---|
| long lifetime | decrease the cost of maintaining and replacing energy storage |
| large moment of control force | decrease the requirement of propulsion force |
| high efficiency | decrease the cost of energy |
| high energy density | decrease the mass of energy storage devices |

But when the ACS is operating, it can be a big disturb to the attitude of spacecraft, which will affect the normal operation of spacecraft. This is a very complicated nonlinear system, whose control is very difficult to achieve, especially for spacecraft with small mass.

(3) High temperature superconducting magnetic energy storage

HTSMES systems have the following advantages compared to other energy storage systems:

1. High cyclic efficiency — HTSMES can achieve an high overall cyclic efficiency;
2. Fast response time — HTSMES has pure electrical energy conversion, whilst other energy storage devices involve either electrical-chemical or electrical-mechanical energy conversion, which is much slower;
3. Deep discharge and recharge ability — unlike batteries, HTSMES can discharge and recharge fully for an unlimited number of times;
4. Good balance between power density and energy density — although SMES does not have the highest power density or the highest energy density, it does have a good balance between them, which is very important for aircraft systems [8].

5. CONCLUSION

As technology develops, the energy density of HTSMES can be enhanced especially for application in space environment. Compared to HTSMES, then traditional battery storage has obvious disadvantages and has no special advantages. For flywheel energy storage, its obvious advantage is that the operation of flywheel is a big disturb for the stability of spacecraft, which may make the control of the whole spacecraft too complicated. What's more, the future trend for flywheel energy storage is also adopting superconducting to increase its efficiency, which makes flywheel energy storage have the same parts of device used for keeping its soil in low temperature as the

HTSMES. On the contrary, HTSMES will not bring about a disturb. Although its energy density may be lower than flywheel storage now, its energy density will increased because there is no need for refrigerator in space And in order to ensure its operation environment, we can at least adopt hybrid system including battery and HTSMES or ACS and HTSMES, which may alleviate the disturb to an extent that can be controlled and satisfy the energy supply at the same time.

REFERENCES

1. Li, G. X., *The Introduction of Space Power System*, Chinese Astronavigation Press, 2008.
2. Katagiri, T., et al., "Field test result of 10 MVA/20 MJ SMES for load fluctuation compensation," *IEEE Trans. Appl. Supercond.*, Vol. 19, No. 3, 1993–1998, Jun. 2009.
3. Wang, Q. L., et al., "Development of large scale superconducting magnet with very small stray magnetic field for 2 MJ SMES," *IEEE Trans. Appl. Supercond.*, Vol. 20, No. 3, 1352–1355, 2010.
4. Kohari, Z. and I. Vajda, "Losses of flywheel energy storages and joint operation with solar cells," *Journal of Materials Processing Technology*, Vol. 161, No. 1–2, 62–65, 2005.
5. Long, T., W. Fred, D. Narayan, et al., "Simulation of the interaction between flywheel energy storage and battery energy storage on the international space station," *Proceedings of the Intersociety Energy Conversion Engineering Conference*, 848–854, Las Vegas, 2000.
6. Thomas, P. J., D. Decker, and V. A. K. Spector, "Spacecraft flywheel systems — benefits and issues," *Proceedings of the IEEE National Aerospace and Electronics Conference*, 589–593, Dayton, 1997.
7. Feng, X. Z. and Q. B. Wu, "The application of flywheel on satellite energy storage/attitude control system," *Control & Automation*, Vol. 20, No. 10, 29–30, 2004.
8. Yuan, W. J., et al., "Design and test of a superconducting magnetic energy storage (SMES) coil," *IEEE Trans. Appl. Supercond.*, Vol. 20, No. 3, 1379–1382, 2010.
9. Connolly, D., *A Review of Energy Storage Technologies*, University of Limerick, 2009.
10. Ali, M. H. and R. A. Dougal, "An overview of SMES applications in power and energy systems," *IEEE Transactions on Sustainable Energy*, Vol. 1, No. 1, 38–47, 2010.
11. Luo, P., et al., "The optimal design of 300 MJ I-ITS superconducting magnetic energy storage system," *Cryo. & Supercond.*, Vol. 37, No. 1, 21–24, 2009.
12. Chen, H. S., et al., "Progress in electrical energy storage system: A critical review," *Progress in Natural Science*, Vol. 19, 291–312, 2009.
13. Dai, T. Z., et al., "Design of a 10 MJ HTS superconducting magnetic energy storage magnet," *IEEE Trans. Appl. Supercond.*, Vol. 20, No. 3, 1356–1359, 2010.
14. Zhang, M., L. W. Li, T. S. Yeo, and M. S. Leong, "Scattering by a gyrotropic bianisotropic cylinder of arbitrary cross section: An analysis using generalized multipole technique — abstract," *Journal of Electromagnetic Waves and Applications*, Vol. 17, No. 7, 1049–1051, 2003.
15. Yin, W. Y., L. W. Li, and M. S. Leong, "Scattering from multiple bianisotropic cylinders and their modeling of cylindrical objects of arbitrary cross-section — abstract," *Journal of Electromagnetic Waves and Applications*, Vol. 14, No. 5, 611–612, 2000.
16. Tan, E. L. and S. Y. Tan, "Cylindrical vector wave function representations of electromagnetic fields in gyrotropic bianisotropic media," *Journal of Electromagnetic Waves and Applications*, Vol. 13, No. 11, 1461–1476, 1999.
17. Bass, F. and L. Resnick, "Spatial and temporal rotation of the polarization plane of electromagnetic waves reflected from and transmitted through a gyrotropic plate," *Journal of Electromagnetic Waves and Applications*, Vol. 17, No. 8, 1131–1137, 2003.
18. Censor, D. and M. D. Fox, "Polarimetry in the presence of various external reflection and retrodirection mirroring mechanisms, for chiral and gyrotropic media," *Journal of Electromagnetic Waves and Applications*, Vol. 11, No. 3, 297–313, 1997.
19. Huang, H., Y. Fan, B. Wu, F. Kong, and J. A. Kong, "Surface modes at the interfaces between isotropic media and uniaxial plasma," *Progress In Electromagnetics Research*, Vol. 76, 1, 2007.

Comparison Study of Eddy Current Losses of Induction Motors Fed by SPWM and SVPWM Inverters

Jingjing Han¹, Ruifang Liu¹, and Hui Huang^{1,2}

¹School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China

²State Key Laboratory of Millimeter Waves, Nanjing 210096, China

Abstract— In this paper, a comparison study of eddy current losses in three-phase induction motors fed by voltage source inverter with sinusoidal pulse width modulation (SPWM) and voltage space vector pulse width modulation (SVPWM) is presented. The simulation model of SPWM and SVPWM inverter are respectively established in MATLAB. After analyzing the harmonic spectrums of inverter output voltage, both the analytic method and FEA (finite element analysis) method are adopted to calculate the eddy current losses. The calculation results show that comparing with SPWM, SVPWM control method could decrease harmonic components of output voltage and reduce the iron losses of the motors.

1. INTRODUCTION

PWM (pulse width modulation) inverters are widely applied in speed control of motor devices. SPWM and SVPWM are two common modulation methods. Normally, there are high-order harmonics in the output voltage of PWM inverter, which increase the motor losses [1, 2]. The paper compared and analyzed the eddy current losses of induction motors fed by SPWM inverter and SVPWM inverter. After introducing the principles of SPWM and SVPWM, Matlab/simulink software was used to establish the simulation model of SPWM and SVPWM inverter respectively. The harmonic spectrums of inverter output voltage were obtained. Based on the superposition, the eddy current losses of induction motors fed by PWM inverter were calculated. The relation of the eddy current losses with the modulation ratio and carrier wave ratio were discussed. Eddy current losses discrepancy in SPWM and SVPWM was also compared. Finally, the electromagnetic field finite element model of PWM-fed induction motor was established. After time-stepping finite element calculation, the rules between the eddy current losses and control parameters in SPWM and SVPWM were obtained. The research showed that high-order harmonics had large influence to motor losses and decreased the efficiency.

2. PRINCIPLE OF SPWM AND SVPWM

An inverter bridge that consists of six switches whose state of on or off is controlled by the PWM drives is used in a most common way. The outputs are connected to each motor terminal to supply the power for motor. The structure of the voltage-source inverter (VSI) for three-phase induction motor system is shown in Fig. 1. SPWM and SVPWM schemes are used to control the switching devices to generate approximate sinusoidal signals in the stator phases. Fig. 2 shows the general principle of SPWM. An isosceles triangle carrier wave is compared with a sinusoidal modulating wave, and the points of intersection determine the switching point of power devices. The output waveform of the inverter controlled by SPWM is a series of rectangle pulses, whose amplitudes are same and the widths are different.

The switching angle control is decided by the switching frequency and modulation index, and the output fundamental wave amplitude has a linear relationship with the amplitude modulation index. And the change of the switching frequency can only change the center of the fundamental wave frequency distribution, which does little influence on the amplitude of each harmonic [3]. It can be seen that the pulses in the output waveform have a sine weighting equivalent to the reference waveform. This method was realized first with analog circuits, and it can be modeled in the MATLAB/simulink to generate a series of simulated PWM wave.

At the present time, digital implementations are preferred and SVPWM is a typical one. There are six power switches in total in the three-phase VSI and when one of the switches in the upper half inverter bridge is opened, the corresponding one in the lower half bridge will be closed. So there are actually only 8 switching modes (000, 001, ..., 111) existing in three-phase VSI. The Fig. 3 shows the corresponding eight voltage space vectors, and $U_1 \sim U_6$ are called effective vectors which have the same length but the different phases and they divide the complex plane into six

sectors I ~ VI. U_0 and U_7 are called null vectors, having the length of 0 and locating in the centre of the complex plane.

In SVPWM control technique, the reference voltage vector V_{ref} rotates in the space with an angular frequency of ω . In each sector, the two adjacent effective voltage space vectors compose the reference voltage vector V_{ref} as well as one of the two null vectors (U_0 or U_7) if necessary. These vectors are selected to equal V_{ref} by means of different operating time of various vectors. When the power switches in the inverter are driven in a sequence the different vector combinations are generated. And the inverter outputs a cycle of equivalent sinusoidal voltage. They also can be realized in the MATLAB/simulink.

3. COMPARISON OF EDDY CURRENT LOSSES WITH SUPERPOSITION METHOD

3.1. Harmonic Analysis of SPWM and SVPWM Inverters

The PWM voltage-source inverter can bring harmonic distortion. Carrier wave ratio and modulation index are two important parameters in the PWM technique. The carrier wave ratio is defined as $N = f_s/f_1$, where, f_1 is the desired fundamental frequency of the inverter and f_s is the switching frequency of the inverter. The amplitude modulation index is defined as $M = V_{con}/V_{tri}$, where, V_{con} is the peak amplitude of the control signal, and V_{tri} is the amplitude of the triangular waveform. There is a key parameter to indicate the ability the inverter output voltage and the parameter is utilization ratio of the DC link voltage. Usually it is defined as the ratio of fundamental amplitude of line voltage of inverter and DC bus voltage. There is a linear relationship between amplitude of line voltage and phase voltage, so the modulation index and DC voltage utilization ratio have a same meaning essentially.

The simulation model of SPWM and SVPWM inverters are built in MATLAB/Simulink and a FFT analyzer is used to conduct the harmonic analysis. The comparison of harmonic between SPWM and SVPWM inverter has been implemented and the simulation and FFT analysis results are obtained as in Fig. 4.

Set the two kinds of modulation methods at the same condition which is to keep $M = 0.6$ and $N = 100$, and adjust the output fundamental voltage to a same value. From the harmonic spectrum we can see that apart from the common law outsides, the SVPWM scheme has smaller harmonic wave amplitude near the hundredth order and other integral multiples of N , and less low-order harmonic wave components, which are the advantages compared to the SPWM scheme. So that

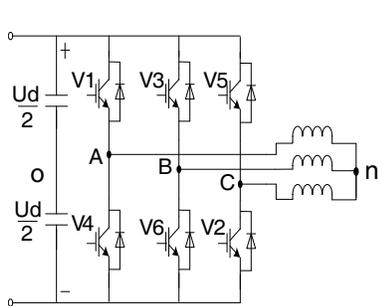


Figure 1: Structure of three-phase voltage source inverter.

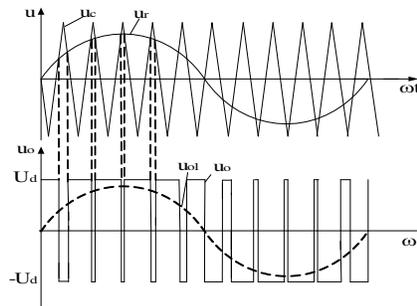


Figure 2: The modulation wave of PWM.

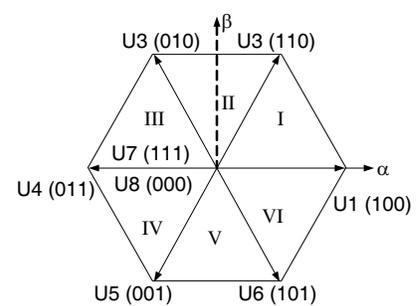


Figure 3: Space vector hexagon diagram.

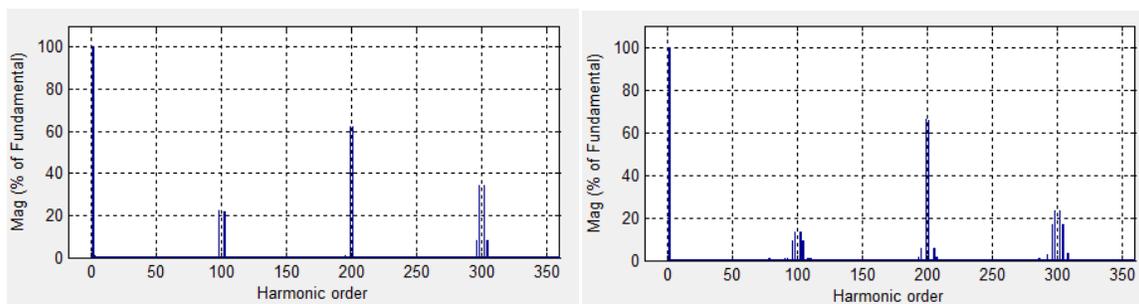


Figure 4: Comparison of harmonic between SPWM and SVPWM.

the less harmonic wave components imply the less iron losses of motor.

3.2. Eddy Current Losses Analysis of SPWM and SVPWM Schemes

The induction motor operating under non-sinusoidal supply has a more increment of losses due to the time harmonics. The analytic method based on the superposition principle is used to analyzing the influence of modulation index and switching frequency of inverters.

According to the Bertotti iron loss separate model, the conventional analysis model of losses is [4]:

$$p_{iron} = p_h + p_c + p_e = k_h f B^x + k_c f^2 B^2 + k_e f^{1.5} B^{1.5} \quad (1)$$

where the total iron losses p_{iron} is separated into three main components, the hysteresis loss p_h , the classical eddy current loss p_c , and the excess losses p_e ; k_h is the hysteresis losses coefficient; k_c is the classical eddy current losses coefficient, k_e is the excess losses coefficient. Usually, the excess losses are a very small proportion in the total losses, so that it can be ignored and has no influence on the final results. Considering the relationship between index of flux density and the losses, the iron losses increasing when PWM supplied is mainly induced by the eddy current losses increment, so we will focus on the analysis of the eddy current losses.

The output voltage of inverter contains the fundamental voltage and other high frequency harmonic voltages. Based on the superposition principle the total losses are equal to the sum of iron losses produced by fundamental voltage and all other orders harmonic voltages. So the specific expression is as following [5]:

$$p_c = \sum k_c (f_n)^2 B_n^2 = \sum k_c (f_n)^2 \left(\frac{V_n}{NSf_n} \right)^2 = \sum \frac{k_c}{(NS)^2} V_n^2 = K \sum V_n^2 \quad (2)$$

where N is the effective number of windings, S is the section of magnet core, n is the harmonic wave order and K is an assumed coefficient, and in this paper we set it as unit 1. Combing above equation and Fourier harmonic spectrum data of two PWM schemes, we can obtain the relationships between eddy current losses of and modulation index as well as carrier wave ratio of SPWM and SVPWM.

Figure 5 shows the eddy current losses comparison between SPWM and SVPWM with the change of modulation index M and carrier wave ratio N . A constant fundamental frequency of $f_1 = 50$ Hz is used. M is set as a constant value to find the influence of carrier wave ratio N on the eddy current losses, and keep the switching frequency changing from 0.5 kHz to 10 kHz. A comparison curve can be obtained. The losses have been calculated at DC bus voltage $V_d = 300$ V. We can see that only a small decrease of eddy current loss is observed when the carrier wave ratio increases. The loss is almost constant when changing carrier wave ratio N . So the influence on the eddy current losses from the N is too little to be ignored. But there is a conspicuous change of eddy current losses when change the modulation index M . The eddy current losses drop with the increasing of M , which is applied to the two different PWM schemes. The reason is that the

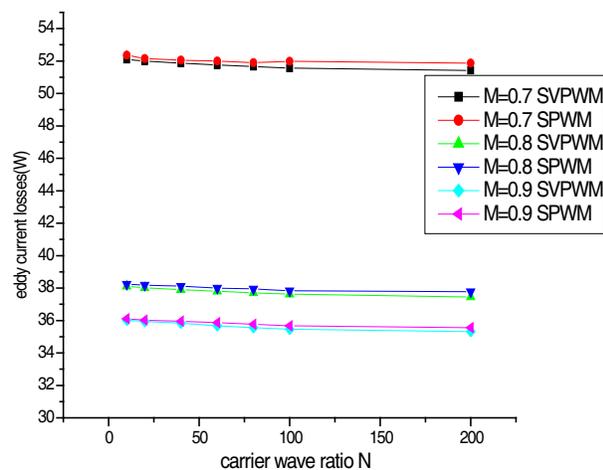


Figure 5: Effect of M and N on eddy current losses.

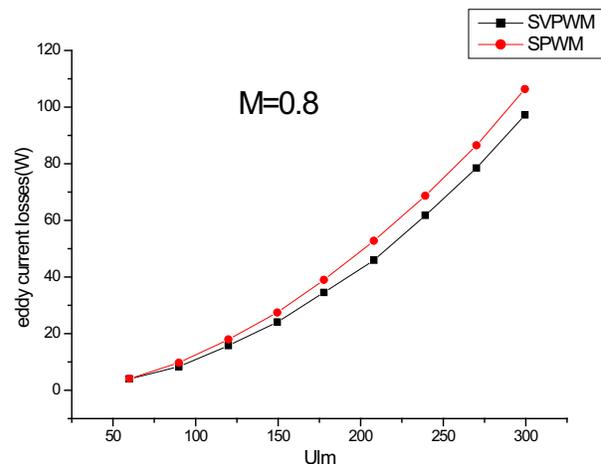


Figure 6: Losses comparison between SPWM and SVPWM at the same M .

amplitude of harmonic components is relatively lower when the modulation index is higher [5]. The iron losses will decrease with the increasing of M . Considering this case, a high modulation index should be used to reduce the iron losses when PWM supplied. In conclusion, the eddy current loss is mainly related to the modulation index M but N . Simultaneously, it can be seen that the eddy current losses under SVPWM inverter is lower than that of SPWM. Fig. 6 is also a visualized indication of variation comparison of eddy current losses at a constant modulation index when the DC bus voltage changes from 100 V to 500 V, obviously, the SVPWM scheme has less eddy current losses. All of these results indicate the SVPWM scheme has less harmonic components and iron losses and that is a tremendous advantage compared to the SPWM scheme.

4. COMPARISON OF EDDY CURRENT LOSSES WITH FEA METHOD

In this section, taking into account the nonlinear magnetic property of the iron core, the time-stepping FEA (finite element analysis) method is used to validate the results in the above calculations. An 11 kW class induction motor driven by the PWM power supply is analyzed by the 2D nonlinear time-stepping analysis. Fig. 7 shows results of the flux distribution and Fig.8 shows the curve of time-variation of the core loss at no load condition when using SVPWM scheme at given M and N .

FEA calculations are carried out under SPWM and SVPWM supply with different parameters such as N and M . In order to make the comparison conveniently, set the coefficient of eddy current losses $k_c = 1$, the coefficient of hysteresis losses $k_h = 0$ and the excess losses coefficient $k_e = 0$. In this condition, the total core loss is just the eddy current losses. Keep the fundamental frequency $f_1 = 50$ Hz, calculate the mean value of the core loss during one cycle time when the induction motor operates under steady state. Then we can obtain the iron losses when adopt different PWM schemes at a constant carrier wave ratio N and modulation index M .

Table 1 shows the comparison of eddy current losses between SPWM, SVPWM and the sine wave supply. Similarly we make the two parameters carrier wave ratio and modulation index at a constant value. Obviously it can be seen that the eddy current losses when SVPWM scheme is implemented are less than the SPWM scheme, which is accordant to the conclusions in the last chapter's analysis. The eddy current losses calculated by sine wave supply is less than the PWM supply and more than the SVPWM supply.

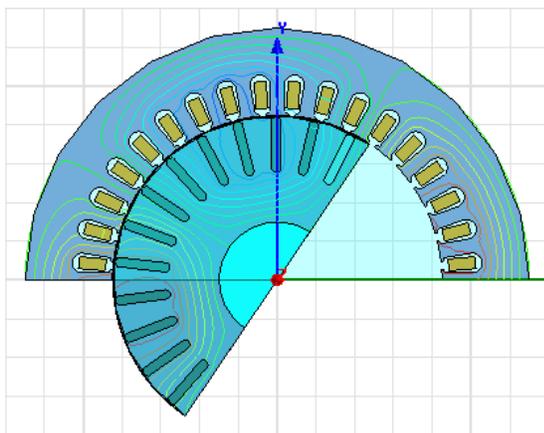


Figure 7: Flux distribution of induction motor.

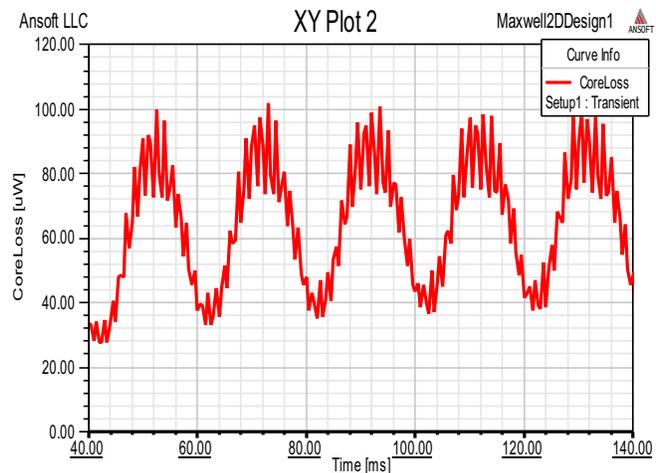


Figure 8: Time-variation of the core loss at no load condition.

Table 1: Eddy current losses comparison between different PWM schemes at constant $N = 100$ and $M = 0.9$.

| | |
|----------------------------|-----|
| P_c (W) Sine wave supply | 30 |
| P_c (W) SPWM supply | 119 |
| P_c (W) SVPWM supply | 64 |

5. CONCLUSIONS

The paper compared and analyzed the influence on the eddy current losses of induction motors fed by SPWM inverter and SVPWM inverter. The relationship of eddy current losses with the two control parameters modulation index and carrier wave ratio is studied. Not only in SPWM scheme but also in SVPWM, the analyzed and calculated results show that the modulation index plays a more important role in the eddy current losses than the carrier wave ratio. The calculation results show that comparing with SPWM, SVPWM could decrease harmonic components of output voltage and reduce the iron losses of the motors.

ACKNOWLEDGMENT

This work is sponsored by State Key Laboratory of Millimeter Waves under Contract K201012.

REFERENCES

1. Boglietti, A., P. Ferraris, M. Lazzari, and M. Pastorelli, "Influence of the inverter characteristics on the iron losses in PWM inverter fed induction motors," *IEEE Transactions on Industry Applications*, Vol. 32, No. 5, 1190–1194, Sept./Oct. 1996.
2. Yuan, D.-K., G.-Q. Xu, B. Hu, A. Xiang, and J.-S. Kang, "Research on a novel SVPWM for three-phase VSI," *IEEE Vehicle Power and Propulsion Conference (VPPC)*, Harbin, China, Sept. 3–5, 2008.
3. Zang, C., Z. Pei, J. He, T. Guo, J. Zhu, and W. Sun, "Research on the application of CPS-SPWM technology in cascaded multilevel inverter," *International Conference on Electrical Machines and Systems*, 1–4, Nov. 15–18, 2009.
4. Huang, P.-L. and Q.-S. Hu, "A novel model for predicting iron losses of electric machine fed by PWM inverter," *Proceeding of International Conference on Electrical Machines and Systems*, 1076–1079, Seoul, Korea, Oct. 8–11, 2007.
5. Liu, R., C. C. Mi, and D. W. Gao, "Modeling of Eddy-current loss of electrical machines and transformers operated by pulsewidth-modulated inverters," *IEEE Transactions on Magnetic*, Vol. 4, No. 8, 2021–2028, 2008.

Regulatory Analysis of the Intermodulation Interference between the PCS Receiver and the Low-power Radio Devices

D. O. Kim and C. Y. Kim

School of Electronics Engineering, Kyungpook National University, South Korea

Abstract— In this paper, the regulatory radiation level emanated from the low-power radio devices has been predicted. From a radio interference point of view, the personal communication services (PCS) receiver is chosen as a victim wireless device. And the low-power radio devices is offender ones. This paper primarily focuses on the intermodulation interference by unmodulated signal in adjacent channel. A theoretical analysis has been proposed using the well-known intermodulation interference theory. That method of analysis includes amplitude of interference corresponding signal-to-interference ratio (SIR) and transformation process of the electric field strength. The calculated electric field strength level becomes the quantified upper limit to avoid the intermodulation interference on the victim device, and it was $76.26 \text{ dB}\mu\text{V/m}$. To show the validity on the suggested value, the simulation was conducted by the Advanced Design System (ADS) tool for the commercial PCS receiver model. The resultant numerals show a good match between two values within less than 5%. In addition, we discussed the relation between our conclusion and the Federal Communication Commission (FCC) regulation of the low-power radio device. The outcomes of the paper may be applicable to any other wireless devices with nonlinear characteristics in many ways.

1. INTRODUCTION

Recently, the low-power radio devices, as a fast growing field, have been widely adopted in the commercial domain. But, as a variety of the low-power radio devices used unlicensed bands, there is a concern about radio interference between communication systems using a shared frequency band. With regard to this, the radiation limit of the low-power radio devices is defined by regulation like Federal Communication Commission (FCC) Part 15.209 [1]. However, the radio interference research of the low-power radio device is not yet sufficient, and also there must be more research to forecast about radio interference when the regulation is revised. So some researches about the prediction of radio interference have been reported in the published literature [2–4]. However, they have focused mainly on the co-channel interference, and give little consideration to the adjacent-channel interference as far as the immunity regulation is concerned about. Moreover, they seldom provide the academic reasoning akin to the analysis and comparison with enacted regulation. But it is important to know the scientific basis in the domain of the low-power radio devices industry.

To present the basic step between the technical approach and the current regulation the theoretical analysis has been proposed using the well-known intermodulation theory. It is noted that intermodulation is attributed to the adjacent-channel interference. Because there are many types of radio interference present, a prescribed type of radio interference should be selected by the purpose of its customers. So, in this study, the radio interference bringing about the intermodulation phenomenon has been targeted, because the intermodulation interference is commonly encountered in the receivers of almost all wireless devices exposed to the nonlinear operation. And as a target device, the PCS receiver was preferred because it is popular in a mobile communication system.

In this paper, the calculated level of electric field strength means the upper value of emission against the intermodulation interference radiated from the hostile low-power radio devices. While there are a number of interferences, what is concerned at issue here is the intermodulation interference with the unmodulated signals. With the calculated value as stated above and the analyzed interference mechanism by using the intermodulation interference, the technical basic of the regulatory radiation limit could be clearly indicated.

This paper is an extended version already reported in [5]. While the old version was concerned on CDMA (900 MHz), the new version has been devoted on PCS (1900 MHz) to the addressed question of whether the proposed analysis is equally applicable and of what becomes the new radiation limit. The first question is quite essential because the assigned frequency band dictates the detailed structure of target receiver, in here either of CDMA or PCS. On the second question, our findings are that the quantified radiation limit was $81.99 \text{ dB}\mu\text{V/m}$ for the CDMA receiver and $76.26 \text{ dB}\mu\text{V/m}$ for the PCS receiver. Although the two versions are on the ground root, the respective frequency bands pushed them away and created their own identities.

2. THEORETICAL FOUNDATION AND CALCULATION

To understand how the intermodulation interference phenomenon occurred in the PCS receiver, let's suppose a certain nonlinear system excited by two unmodulated signals of frequencies f_1 and f_2 with same amplitude A . And then the frequency spectrum of this system output can be expressed as $f_{IM} = mf_1 + nf_2$, where m and n are integers either of zero, positive or negative, and $(|m| + |n|)$ indicates the order of the IMD (Intermodulation distortion) product [6]. This indicates that two exciting signals in the nonlinear system generate multiple harmonics at the output of the system. It is assumed that the two unmodulated signals are some interference signal. This case is the intermodulation interference with the frequencies $2f_1 - f_2$ or $2f_2 - f_1$ which falls close to the frequency f_s of the desired signal. The difference frequency $2f_1 - f_2$ and $2f_2 - f_1$ are the 3rd order harmonics products (or the IMD3 products) which cause a major problem in radio reception. The ratio between the amplitudes of the output IMD3 products and the desired signal is equal to the signal-to-interference ratio (SIR) that is the key parameter deciding the existence of radio interference. In other words, the presence of the radio interference can be determined by the signal-to-interference ratio at the f_s of output of receiver. The SIR can be expressed in Equation (1) [7].

$$S/N = \frac{A_{os}}{A_{on}} = \frac{A_{IIP3}^2}{A^3} A_s \quad (1a)$$

$$A = \sqrt[3]{\frac{(A_{IIP3}^2) \cdot A_s}{(A_{os}/A_{on})}} \quad (1b)$$

where A_s and A_{os} denote the amplitudes of input and output signal. A and A_{on} are the amplitudes of the input and output interference signal respectively, while A_{IIP3} denotes the amplitude of IIP3. To decide the level of the radio interference, the reference for the signal-to-interference ratio (SIR) should be given for comparison between the reference and the device under test. In the case of the analog mobile communication, the given reference SIR is 18 dB (7.943) [8]. Therefore if A_{os}/A_{on} is under the 7.943, then the PCS receiver becomes a victim of radio interference due to the intermodulation phenomena. Hence, the amplitude A corresponding to $A_{os}/A_{on} = 7.943$ in Equation (1b) will be the amplitude of input signal to cause the radio interference.

Figure 1 shows the frequency spectrum of the intermodulation interference in the PCS receiver and its structure. Frequencies of the two interference signals are 1901.25 MHz and 1902.5 MHz with a 1.25 MHz frequency difference which is equal to a channel bandwidth of the PCS system. As shown in Figure 1, the PCS system is organized by an RF amplifier (LNA), an active mixer,

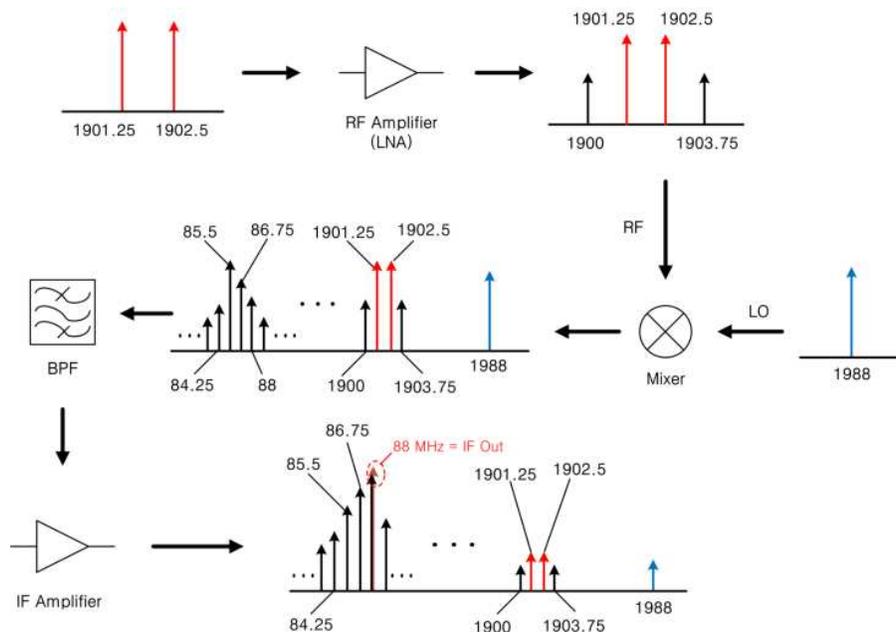


Figure 1: Frequency spectrum of the PCS receiver and its structure. (Frequency units is in MHz).

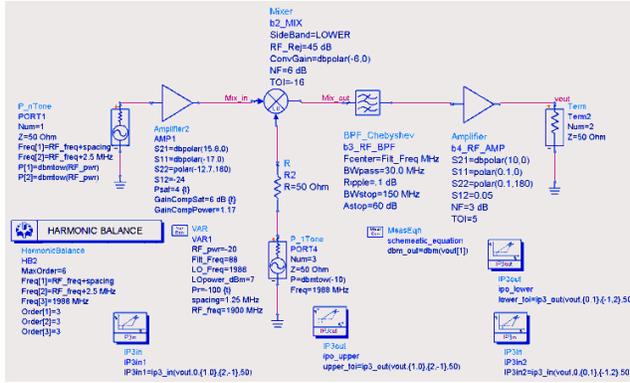


Figure 2: Circuit model of the PCS receiver using ADS simulator.

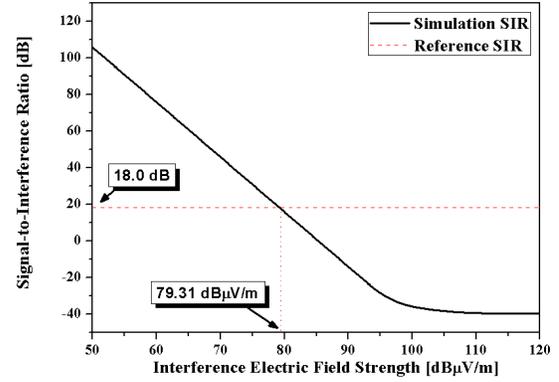


Figure 3: Simulation result of the PCS receiver on signal-to-interference ratio in term of interference field strength.

a band pass filter and an IF amplifier. And then, the building blocks of Figure 1 were modeled by using ADS (Advanced Design System) simulator of the Agilent, as shown in Figure 2. The modeling circuit is more like the structure of Figure 1. In the PCS receiver model, the same parameter values are used as those of specification on NJG1107KB2 (Low Noise Amplifier) by the New Japan Radio Co. Ltd. [9]. And the remaining parameters are built-in values by the ADS simulator run for PCS receiver [1]. The simulation shows that the power gain S_{21} (Total) is 19.8 dB, IIP3 as -40.3 dBm (corresponding to 0.002161 V), and P_{mds} as -101 dBm (corresponding to 1.99 μ V). P_{mds} is a minimum detectable signal corresponding to A_s . Substituting the above value of IIP3 and P_{mds} together with the reference SIR 18 dB into Equation (1b) produces the value A in Equation (2).

$$A = \sqrt[3]{\frac{(0.002161)^2 \times 1.99 \times 10^{-6}}{7.943}} = 105.372 \text{ } [\mu\text{V}] = 40.46 \text{ } [\text{dB}\mu\text{V}] \quad (2)$$

This voltage A appears across the antenna terminal connected to the PCS receiver. And this voltage is also determined by the incoming electric field and the antenna factor K by [11]

$$E \text{ [dB}\mu\text{V/m]} = K \text{ [dB/m]} + V \text{ [dB}\mu\text{V}] \quad (3a)$$

$$K \text{ [dB/m]} = 20 \log f \text{ [MHz]} - G \text{ [dB]} - 29.78 \quad (3b)$$

In here, G is the antenna gain and V is just equal to A . The numerical values are found to be

$$E \text{ [dB}\mu\text{V/m]} = 35.8 \text{ [dB/m]} + 40.46 \text{ [dB}\mu\text{V]} = 76.26 \text{ [dB}\mu\text{V/m]} \quad (4a)$$

$$K \text{ [dB/m]} = 20 \log(1900) \text{ [MHz]} - 0 \text{ [dBi]} - 29.78 = 35.8 \text{ [dB/m]} \quad (4b)$$

In the above calculation, the isotropic antenna is supposed, and 1900 MHz is used as the center frequency to the PCS receiver. The level of the electric field strength, 76.26 dB μ V/m shown in Equation (4a), is the emission level to deteriorate the PCS receiver operation under the intermodulation interference.

3. SIMULATION AND DISCUSSION

To validate the calculated level 76.26 dB μ V/m appeared in Equation (4a) and the above computation procedure, the simulation was conducted for the PCS receiver model by ADS seen in Figure 2. The equation of the SIR is given as

$$S/N = \frac{A_{os}}{A_{on}} = \frac{19.5 \text{ } [\mu\text{V}]}{A_{on}} \quad (5)$$

By inserting the obtained value of A_{on} from the ADS simulator to Equation (5), the signal-to-interference ratio can be plotted against the electric field strength of the interference source, as

shown in Figure 3. The numerator value $19.5 \mu\text{V}$ in Equation (5) can be converted to the power value of -81.2 dBm through Equation (6).

$$P_{os} = P_{m_{ds}} + S_{21}(\text{Total}) = -101 [\text{dBm}] + 19.8 [\text{dB}] = -81.2 [\text{dBm}] \quad (6)$$

At $\text{SIR} = 18 \text{ dB}$, the level of the interference electric field strength is $79.31 \text{ dB}\mu\text{V}/\text{m}$ as shown in Figure 3. This simulated result is in close agreement with the calculated result of Equation (4a) within only a 5% difference. In Figure 3, a tendency of decrease can be observed in SIR in terms of an increase in the electric field strength from the interference source. This tendency can also be predicted from Equation (1a). In addition, the SIR saturation region of the receiver is appeared in Figure 3. This appearance occurs due to the diminishing gain of the receiver for the overwhelming input level. The value $79.31 \text{ dB}\mu\text{V}/\text{m}$ seen in Figure 3 is the threshold level of electric field strength to provoke the intermodulation interference to the PCS receiver under the 18 dB SIR .

In the previous sections, the theoretical interference level of $76.26 \text{ dB}\mu\text{V}/\text{m}$ has been mentioned to cause the intermodulation interference on the PCS receiver. It shows a difference of 21.96 dB to the regulation value of $54 \text{ dB}\mu\text{V}/\text{m}$, which is an emission limit of the FCC Part 15.209 at 3 m distance from the intentional radiator in PCS communication band [1]. One should be reminded of the different measure of distance between FCC and this study. To make a straightforward comparison $76.26 \text{ dB}\mu\text{V}/\text{m}$ should be converted to the electric field strength at 3 m . For this, the electric field strength at 3 m is obtained to be [12],

$$E_{d=3\text{m}} = 76.26 [\text{dB}\mu\text{V}/\text{m}] + \text{Path Loss} = 76.26 [\text{dB}\mu\text{V}/\text{m}] + G_t G_r (\lambda/4\pi d)^2 = 51.38 [\text{dB}\mu\text{V}/\text{m}] \quad (7)$$

where the “Path Loss” is defined by the Friis transmission formula. And both G_t and G_r are set to unity because the FCC Part 15.204 has stipulated that the measurement antenna is of the same type and of equal or less directional gain as an antenna which is authorized with the intentional radiator antenna. Thus, the agreement between the two calculations is satisfactory within less than 3 dB . In case of the previous works about the other victim wireless device research, the corresponding value to the AMPS receiver for $76.26 \text{ dB}\mu\text{V}/\text{m}$ of Equation (4a) was $81.99 \text{ dB}\mu\text{V}/\text{m}$, contiguous to each other [5]. Now, possibly the mysterious number $54 \text{ dB}\mu\text{V}/\text{m}$ seen on FCC Part 15.209 can be understood as a consequence of intermodulation interference.

4. CONCLUSIONS

In this study, the quantitative intermodulation interference on the PCS receiver with a victim of radio interference was calculated under the worst case radio environment. An unmodulated signal was used as an interference source type for the low-power radio device. Theoretical electric field strength of the emission limit was $76.26 \text{ dB}\mu\text{V}/\text{m}$, and the simulated value was $79.31 \text{ dB}\mu\text{V}/\text{m}$. To show the adequacy of this result, the simulation was performed by using the ADS simulation tool. These calculated and simulated values remain sufficiently close to within less than 5%. In this study, we made efforts to analyze the scenario of interference between the low-power radio devices and the victim PCS receiver based on the intermodulation interference. With the analysis on the intermodulation interference occurred in the PCS receiver, we can academically understand the role of current regulation value to protect the harmful radiation from the low-power radio devices. Although considerations are done for the PCS receiver model, the proposed analysis technique might be helpful to the other type of wireless devices exposed to the radio interference.

REFERENCES

1. FCC, Code of federal regulations 47, Part 15.209, 1997.
2. Park, J. A., S. K. Park, D. H. Kim, P. D. Cho, and K. R. Cho, “Experiments on radio interference between wireless LAN and other radio devices on a 2.4 GHz ISM band,” *Proceeding of the 57th IEEE Semiannual Vehicular Technology Conference*, 1798–1801, Jeju, Korea, April 2003.
3. Ghosh, M. and V. Gaddam, “Bluetooth interference cancellation for 802.11g WLAN receivers,” *Proceeding of IEEE International Conference on Communications*, 1169–1173, Anchorage, USA, May 2003.
4. Ponnekanti, S. and S. Sali, “Non-linear interference cancellation techniques for electromagnetically dense propagation environments,” *Progress In Electromagnetic Research*, Vol. 18, 209–228, 1998.
5. Kim, C. Y. and D. O. Kim, “Prediction of the interference level from a low-power radio device provoking the intermodulation interference to the AMPS receiver,” *Progress In Electromagnetic Research*, Vol. 94, 69–81, 2009.

6. Pedro, J. C. and N. B. Carvalho, *Intermodulation Distortion in Microwave and Wireless Circuits*, Artech House, Boston, 2003.
7. Razavi, B., *RF Microelectronics*, Prentice Hall, Upper Saddle River, NJ, 1998.
8. Lee, W. C. Y., *Mobile Cellular Telecommunications*, McGraw-Hill, New York, 1995.
9. New Japan Radio Co. Ltd., *NJG1107KB2* datasheet, <http://www.jrc.co.jp/>.
10. Agilent Advanced Design System, Agilent ADS (TM) 2005A, 2005, www.agilent.com.
11. Stutzman, W. L. and G. A. Thiele, *Antenna Theory and Design*, 2nd Edition, Artech House, New York, 1997.
12. Rappaport, T. S., *Wireless Communications Principles and Practice*, Prentice-Hall, New York, 1996.

Statistical Characteristics of Region Propagation of Decametric Radiowaves in Time of Heliogeophysical Disturbances

N. P. Sergeenko and M. V. Rogova

Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio-waves Propagation
Russian Academy of Sciences, Troitsk, Russia

Abstract— The concept of use of an artificial satellite for over-the horizon information transfers was essentially added with concept about use ionosphere information transfer channels as reserve. In some cases (over-the horizon radar, positionometry of remote radio sources etc.) ionosphere channels are fundamental. In this aspect interest to properties the ionosphere channels directly connected with properties of an ionosphere has increased. However level of researches in a part of statistics of variations of critical frequency essentially did not change for last years.

The way of overcoming of limitation of applicability of Gauss distribution as model of casual variability of critical frequency of F2 layer of an ionosphere is generalized in this work. It is supposed that each sample is characterised by values selective statistical invariants, skewness A and kurtosis E . In according to the first remarkable limit and the scheme of casual pulse process use for modelling Poisson casual process which in a limiting case $A \rightarrow 0$, $E \rightarrow 0$ passes in Gauss process is proved.

The mathematical model of function of density of probability $W(x)$ is applicable in all area of experimentally observable values A and E i.e., for various heliogeophysical activity and various geographical conditions.

1. INTRODUCTION

Ionosphere plasma — statistically non-uniform environment, therefore parameters of radio systems depend essentially on characteristics of a communication channel with casually changing parameters. There are about ten theoretical and semiempirical distributions which describe experimental data on fluctuations of characteristics of decameter signals [for example, [1, 2]]. At the same time it is very important to investigate and formalize statistical properties of fluctuations of the environment in which these signals propagate. The normal distribution law $W_n(x)$, depending on two parameters—average value \bar{x} and dispersion σ^2 was the first such model of ionosphere variations of electronic concentration in F2 layer [3]. However it is possible to assert, that how real random process has been normalised, at levels small probabilities as much as big deviations from model of normal process are available. It is a probability of a big ionosphere disturbances and blackouts of reliability of work of the radio engineering systems using an ionosphere as a path of radio-waves propagation. This is related to critical frequency of F2 layer f_o as a boundary between frequency ranges of radiowaves that vertically and obliquely reflect from the ionosphere. The f_o value is proportional to maximum usable frequencies (MUFs) of radiosignals. Radiosignal frequencies, which are used to transmit information beyond the horizon due to oblique reflection from the ionospheric F region, should not exceed MUF. Correspondingly, the statistics of f_o is responsible for that:

$$\frac{\Delta f}{f} = \frac{\Delta MUF}{MUF}$$

In this connection, there is a necessity for studying of possibilities of statistical modelling set of critical frequencies of F2 layer, making more essential impact on change of conditions of propagation of decameter radio-waves, than, for example, variability of heights and semithickness of reflecting layers. This work is devoted this problem.

2. PROBLEM OF MODELLING OF RANDOM VARIATIONS $\delta F_0 F_2(T)$

Electronic concentration in F2 layer peak is characterised by considerable variability in daily, seasonal and solar cycles. As random and quasi — deterministic variabilities of properties of an ionosphere F2 layer are simultaneously. Researches of properties of files of critical frequencies of F2 layer $\{f_0\}$ have shown, that in practice division of these effects by development of an image of a “quiet” ionosphere by means of representation by a moving median in a file $\{f_0\}$ for everyone

ionosphere station of the world network is possible. Thus, the set of relative variations $\delta f_o F2$ is formed of current values $f_o(t)$ with a discreteness of one hour:

$$\delta f_o F2 = \frac{f_o(t) - f_o(t)_{mov.med.}}{f_o(t)_{mov.med.}} \quad (1)$$

Parameter $\delta f_o F2$ in practice short-term ionosphere forecasting is used as an index of ionosphere disturbance. The set of values $\{\delta f_o F2\}$ is considered as random process and is object of statistical modelling in this work. A problem of one-dimensional modelling is synthesis of function of density of probability $W(\delta f_o F2)$, calibrated parametres of set $\{\delta f_o F2\}_n$.

The first attempts of construction of statistical models of the ionosphere variability, as much as possible using the information on properties of an ionosphere, have been undertaken in [4]. In this work it is statistically authentically shown, that generally the normal law of distribution for $\delta f_o F2$ sets is inapplicable. The model with skewness A and kurtosis E has been offered for the description of statistical characteristics of process which tend to the normal distributions law at $A \rightarrow 0$ and $E \rightarrow 0$. It has been shown, that the account requirement of ionosphere disturbances in modelling leads to transition to use linearly — exponential characteristic functions. Characteristic function $f(\lambda)$ of random process $x(t)$, defined on a symmetric interval $(-\infty, +\infty)$ as the function conjugate according to Fourier to function of density of probability $W(x)$ is [5]:

$$f(\lambda) = W(x)^{i\lambda x} dx \quad (2)$$

Record (2) assumes, that out of an interval of definition of function $W(x)$ its value is equated to zero. Two kinds of the characteristic functions belonging to a class linearly-exponential functions received in works [4, 6] where as key parametres of characteristic function were used dispersion σ^2 and kurtosis E :

$$f_2(\lambda) = \left[1 + \frac{E(x)\lambda^2\sigma^2}{6} \right]^{\frac{3}{E(x)}} \quad (3)$$

$$f_2(\lambda) = \exp \left[\frac{2}{E(x)} - \sqrt{\frac{2}{E(x)} + \frac{\lambda^2\sigma^2}{E(x)}} \right] \quad (4)$$

Calibration of parametres $f(\lambda)$ in dispersion and excess terms has essential applied value as allows on these aposteriories magnitudes to synthesise functions $W(x)$ without construction of histograms. It leads to considerable reduction of expenses by processing of an experimental material and gives the proved method of extrapolation of functions $W(x)$ in the field of small values x . One of possibilities of generalisation of functions (3), (4) is connected with generalisation on a case unequal to zero of skewness of process $A(x) \neq 0$. Formally for this generalisation it is necessary to use the reception, consisting that symmetric function of density of probability $W(x)$ is multiplied on exponential multiplier $\exp(-\beta x)$. It is equivalent to introduction of new characteristic function $f_a(\lambda)$ a kind:

$$f_a(\lambda) = \frac{f_{cum}(\lambda + i\beta)}{f_{cum}(i\beta)} \quad (5)$$

Parameter β is proportional to skewness A .

However the probability density functions $W(x)$, received from characteristic functions (3)–(5), have an essential lack: they are the nonholomorphic construction in the beginning of co-ordinates. The derivative of functions suffers rupture in the beginning of co-ordinates. Therefore in work [7] attempt to carry out transition to the nonholomorphic functions is undertaken at excessively asymmetric modelling of distributions $\delta f_o F2$. In most general view this problem can be put by means of two functions $F(x)$ and $G(x)$, integrated on parts of area of change of quantity x - an interval (x_{min}, x_{max}) . It allows to present modelled function of probability $W(x)$ in the nonholomorphic construction:

$$W(x) = C \begin{cases} F(x) & x \in (x_{min}, 0) \\ G(x) & x \in (0, x_{max}) \end{cases} \quad (6)$$

The constant of calibration C easily is defined by means of a condition of standardize $\int_{x_{\min}}^{x_{\max}} W(x) \cdot dx =$

1. Functions F and G differ only in scale

$$F(x) = F\left(\frac{x}{g}\right), \quad G(x) = F\left(\frac{x}{d}\right) \quad (7)$$

At $g = d$ the symmetric case turns out.

3. MODELLING OF COMPLEX SIGNALS

Complex signals can be presented in the form of the sum:

$$x_{\Sigma} = x_N + x_d \quad (8)$$

The index Σ is chosen for total process, index N — for casual process, not necessarily normal, that is it is supposed, that the first component in the right part (8) has noisily similar character with excess $E > 0$. The second component has quasi-deterministic character. It will be assumed, that $|x_d| \geq h$. That is the second signal component in the formula (8) is limited on amplitude and exists in a band $\pm h$.

It is possible to present characteristic function of process (8) according to the formula (2):

$$f_{\Sigma}(\lambda) = \overline{\exp[i(x_N + x_d)\lambda]} \quad (9)$$

Owing to independence x_N and x_d it is possible, making use of known property of exponential function, to receive that $F(x + \gamma) = F(x) \cdot F(\gamma)$

$$f_{\Sigma}(\lambda) = f_N(\lambda) + f_d(\lambda) \quad (10)$$

One of processes in (9) — two-level when the variation $\delta f_o F2$ accepts two various values with fast or gradual transitions between levels, that corresponds to occurrence ionosphere storms in F2 layer with the sudden and gradual beginning in practice. Other process has character of pulse noise with a symmetric part of an excess and corresponds to occurrence isolated irregularities which contrast change within the large limits. Skewness of processes in both limiting cases means prevalence positive or negative irregularities.

In practice short-term ionosphere forecasting the total magnitude $\delta f_o F2$ during time ionosphere disturbances is represented in the form of the sum of regular and irregular components [8]:

$$\delta f_o F2 = R(\delta f_o F2) + IR(\delta f_o F2) \quad (11)$$

R — variation includes the most typical large-scale regular changes $\delta f_o F2$ during solar and magnetosphere storms. IR — variation is caused by occurrence of short-term disturbances and irregularities (substorms, seismic origin irregularities, man-made irregularities, etc.). Forecasting of the beginning, duration and type of ionosphere disturbance is carried in according to geomagnetic monitoring [8]. The statistical model of process (10) can be constructed by a principle (9) where characteristic function $f_d(\lambda)$ describes regular variation $R(\delta f_o F2)$, and function $f_N(\lambda)$ describes irregular variation $IR(\delta f_o F2)$.

On Figure 1(a) empirical record $\delta f_o F2(t)$ 03.09.1985 is presented by continuous thick line, a dotted line presents regular variation $R(\delta f_o F2)$ and variations $IR(\delta f_o F2)$ is presented by continuous thin line. The geomagnetic situation for these days is shown also, index Kp shows weak geomagnetic disturbance.

On Figure 1(c) statistical distributions for total $\delta f_o F2$ process (at the left), for R -variation (in the middle) and IR -variation (on the right) are presented. Here histograms are experimentally observable distributions, continuous lines — normal distributions, a dotted line — the resultant function of distribution counted on (9) for this period.

Also it is possible to present the distribution histogram of random proceses $\delta f_o F2(t)$ in the form of superposition of two shifted distributions (Figure 1(b)). Result distribution is shown by a dotted line here. Such two-topmost distributions are characteristic for $\delta f_o F2(t)$ in high latitudes. The feasibility study of phenomenological reception of approximation of two-topmost distributions also is possible on the basis of a method nonholomorphic characteristic function. It is obvious, that in

these cases it is necessary to deal with a signal of a kind (8) when the component x_N represents normal process whereas determined component x_d has more the general character:

$$W() = p\delta(h - h_1) + (1 - p)\delta(h - h_2)$$

It is obvious, that characteristic function of this process will look like

$$f(\lambda) = pe^{i\lambda x_1} + (1 - p)e^{i\lambda x_2}$$

Fourier transformation from this characteristic function will allow to receive the general function of distribution, if both components of process — normal processes:

$$W(x) = \frac{1}{2\pi\sigma} \left[pe^{\frac{(x-x_1)^2}{2\sigma^2}} + (1 - p) \cdot e^{\frac{(x-x_2)^2}{2\sigma^2}} \right].$$

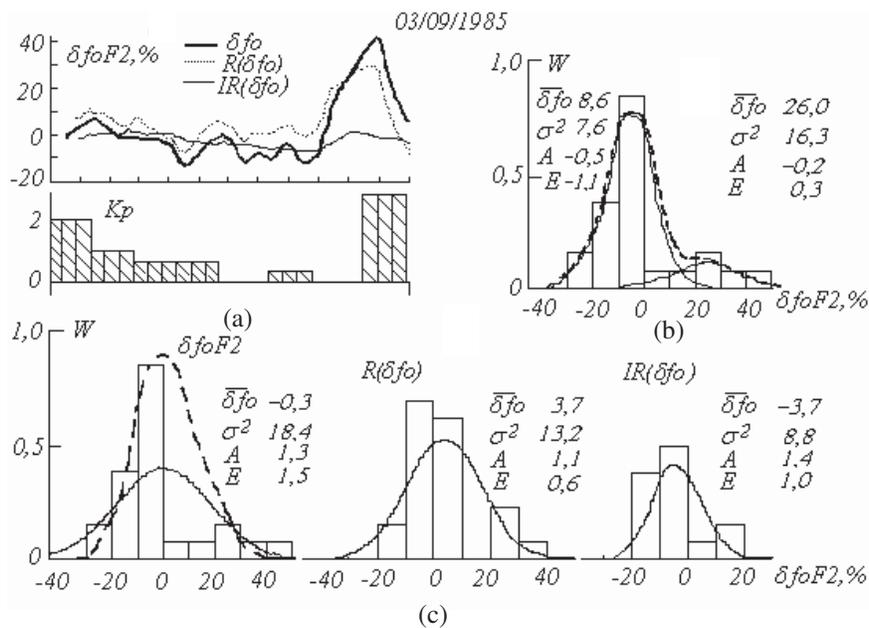


Figure 1: (a) $\delta f_o F2(t)$ and geomagnetic indexes Kp at 03.09.1985; (b) the distribution histogram of random proceses $\delta f_o F2(t)$ in the form of superposition of two shifted distributions; (c) statistical distributions for total proces (at the left), $R(\delta f_o)$ (in the middle) and $IR(\delta f_o)$ (on the right).

4. CONCLUSIONS

The theoretical-empirical approach is developed for construction of one-dimensional functions of density of probability of relative variations of critical frequency of an ionosphere. The foundation of a method is construction of analytical model of the random variable defined on all axis of ordinates and calibrated first four statistical invariant: an average, a dispersion, skewness and kurtosis. Generalisation on an asymmetric case is reached by use nonholomorphic functions. Efficiency of model is confirmed on a file of relative variations of the critical frequency received on a world network of stations of vertical sounding.

These results have applied and methodical value. $\delta f_o F2$ variations define change of the working wave bands used in work of services — the radio communication, radio navigation, a radar-location, etc. Thus a deviation of properties of distributions $W(\delta f_o F2)$ from normal law defines limiting possibilities of radio services — probabilities of erroneous signals, reliability achievements. Depending on concrete schemes of processing of signals these characteristics can be calculated on the resulted data about the $\delta f_o F2$ statistics.

REFERENCES

1. Nakagami, M., *Statistical Methods in Radio Waves Propagation*, 321, Pergamon Press, London, N.Y., 1960.

2. Siddiqui, M. M. and S. M. Ostrow, “Gram-Charlier series distribution for f_oF_2 and M (3000) F2,” *Radio Science*, Vol. 3, No. 4, 383–385, 1968.
3. Benkova, N. P. and N. I. Potapova, “Variability of ionosphere parameters,” *Proceedings of VII Scientific Conference*, Part 2, 82, Publishing by Tomsk University, 1957.
4. Vsehsvjatskaja, I. S., N. P. Sergeenko, and L. A. Yudovich, “Statistical model of geophysical processes with skewness and kurtosis of probability density functions,” *Geomagnetism and Aeronomy*, Vol. 11, No. 5, 785–789, 1971.
5. Middleton, D., *An Introduction to Statistical Communication Theory*, McGraw Hill, New York, 1960; *Sov. Radio. Moscow*, 1961.
6. Furutsu, K. and T. Ishida, “On the theory of amplitude distribution of impulsive random noise and its application to the atmosphere noise,” *J. Rad. Res. Labor*, Vol. 7, No. 32, 279–318, 1960.
7. Dzvonnkovskaya, A. L., V. A. Kuznetsov, and N. P. Sergeenko, “One dimensional statistics of the relative variations in the critical frequency of the ionospheric F2 layer at different latitudes,” *Geomagnetism and Aeronomy*, Vol. 44, No. 6, 751–755, 2004.
8. Sergeenko, N. P. and V. P. Kuleshova, “The forecasting storm from ground-based recording of the magnetic field,” *Solar-Terrestrial Predictions Proceeding of Workshop, October 16–20, 1989. Leura, Australia*, Vol. 2, 367–370, Boulder, Colorado, 1990.

Soil Parameters Retrieval Using a Neural Network Algorithm Trained by a Two Layers Multi-scale Bi-dimensional SPM Model

L. Bennaceur Farah¹, I. Hosni¹, I. R. Farah², R. Bennaceur³, and M. R. Boussema¹

¹LTSIRS, ENIT, Tunisia

²RIADI, ENSI, Tunisia

³LMPC, FST, Tunisia

Abstract— The overall objective of this paper is to retrieve soil surfaces parameters namely, roughness and soil moisture related to the dielectric constant by inverting the radar backscattered signal from natural soil surfaces. We characterize the soil surfaces and subsurfaces by a two layer geo-electrical model. The upper layer is described by its dielectrical constant, thickness, a multi-scale bi-dimensional surface roughness model by using the wavelet transform and the Mallat algorithm, and volume scattering parameters. The lower layer is described by its dielectric constant and multi-scale surface roughness. To compute surface, subsurface and volume scattering, we consider a two layers multi-scale bi-dimensional Small perturbations model. In this study, each surface of the two layers surface is considered as a band limited fractal random process corresponding to a superposition of a finite number of one dimensional Gaussian processes each one having a spatial scale. We investigated the dependence of backscattering coefficient on roughness multi-scale parameters and soil moisture parameters for different incident angles by a sensitivity analysis. This sensitivity analyses is the first step of an inversion procedure. To perform the inversion of the small perturbation multi-scale scattering model (MLS SPM) we used a multi-layer neural network (NN) architecture trained by a backpropagation learning rule. The inversion leads to satisfactory results with a relative uncertainty of 8%.

1. INTRODUCTION

The retrieval of information related to physical surface parameters is a major objective of many studies in remote sensing investigations. In that context, modeling radar backscattering through natural surfaces has become an important theme of research and active remote sensing and has shown its utility for many applications in hydrology, geology, astrophysics, etc.

Previous works [6] have developed a two layer integral equation model to simulate radar backscattering on arid subsurfaces. They used the classical statistical description of natural surfaces and characterized roughness by statistical parameters namely correlation length and standard deviation.

However, the weakness of the classical description of natural surfaces is the large spatial variability which affects the correlation function and make classical roughness parameters very variable [1, 3, 5]. In that context, many previous works have suggested that natural surfaces are better described as self affine random processes ($1/f$ processes) than as stationary processes. In previous works, we have analysed radar backscattering on multi-scale bi-dimensional surfaces [1, 2, 4]. This novel multi-scale bi-dimensional description does not depend on classical roughness parameters standard deviation and correlation length but on new parameters related to multi-scale surfaces properties.

Extracting soil moisture and roughness parameters of natural surfaces from this data has been problematic for many reasons. In fact, many previous studies have dealt with model-based retrieval algorithm and have encountered many problems like the lack of information about the characteristics of natural surface roughness as well as the range of roughness parameters to use in one hand. In another hand, the uncertainties concerning the validity of the scattering models when applied to natural roughness conditions reduces the accuracy of the retrieval procedure. In addition, the relationship between the backscattering coefficient is non linear and the problem of retrieving parameters may be ill-posed and it may be not possible to separate the contributions from different mechanisms making the retrieval of several parameters simultaneously necessary.

We propose in this paper a neural network based inversion procedure using a multi-layer neural network (NN) architecture trained by a backpropagation learning rule.

In the first section, we present the study of the impact of this multi-scale roughness description and soil moisture on radar backscattering using our two layers multi-scale bi-dimensional small perturbation model (SPM) by investigating the sensitivity of backscattering to the new surface parameters and to the dielectric constant. In the next section the neural network based inversion procedure, the results and their accuracy are presented.

2. A TWO LAYERS MULTI-SCALE BI-DIMENSIONAL SPM MODEL FOR THE STUDY OF RADAR BACKSCATTER BEHAVIOR ON SEMI-ARID SOIL SUBSURFACES

The main purpose of our study is to develop an inversion model for soil moisture and multi-scale roughness parameters retrieval over on semi-arid soil subsurfaces using remotely sensed data. The first step of this study presented in this paper, consist to model radar backscattering over a two layer geoelectrical model [6] by taking into account volume scattering. Each layer is described as a multi-scale bi dimensional surface using our multi-scale description. The radar backscattering coefficient can be expressed as the sum of a single scattering component and a multiple scattering component (1)

$$\sigma_{qp}^0 = \sigma_{qp}^S + \sigma_{qp}^M \quad (1)$$

q and p indicate the polarization state of the emitted and received wave respectively H for horizontal polarization and V for vertical polarization. In this present work, multiple scattering component is neglected.

We have considered a two layer backscattering model which can describe backscattering from Mars surfaces which are characterized by a superficial dry layer of thickness d over a second dry or wet layer of basaltic bedrock. We also take into account volume diffusion.

The total backscattering coefficient is the contribution of the two layers and a volume component (2).

$$\sigma_{qp}^0(\theta) = \sigma_{S1qp}^0(\theta) + \sigma_{V1qp}^0(\theta) + \sigma_{SS2qp}^0(\theta) \quad (2)$$

$$\text{with } \sigma_{SS2qp}^0(\theta) = \frac{\cos(\theta)}{\cos(\theta_t)} T_{12} T_{21} e^{-\frac{2\kappa_e d}{\cos(\theta_t)}} \sigma_{S2qp}^0(\theta_t) \quad (3)$$

$$\sigma_{qp}^V(\theta) = \frac{1}{2} \frac{\kappa_s}{\kappa_e} T_{12} T_{21} \left(1 - e^{-\frac{2\kappa_e d}{\cos(\theta_t)}}\right) \cos(\theta) P_{qp} \quad (4)$$

where θ is the incidence angle and θ_t is the transmission angle, P_{qp} has a value of 1.5 for the copolar case κ_s and κ_e , are the diffusion and extinction coefficients given by Fung in [3]

$$\sigma_{Siqp}^0 = \frac{k}{4} \exp(-2k^2)(\cos^2 \theta) s^2 |I_{qp}| W(-2k \sin \theta, 0) \quad (5)$$

$i = 1$ for the first layer and 2 for the second layer and $|I_{qp}|$ given by [3] and

$$W^{(n)}(-2k_x, 0) = \frac{2}{\pi} \int_0^\infty \int_0^\infty \left(\frac{r_c^i(\xi, \eta)}{r_c^i(0, 0)} \right)^n \cos(2k_x \xi) d\xi d\eta \quad (6)$$

where $W^{(n)}$ is the Fourier transform of the n th power of the multi-scale autocorrelation function given by Mattia in [5] with $n = 1$ for the SPM model [1, 2]. Previous works used classical statistical parameters namely correlation length and standard deviation in the expression of the autocorrelation function W . The principal aim of this study is to use the multi-scale surface description in the backscattering coefficient.

3. SENSITIVITY ANALYSIS

We have in a first step considered the VV polarisation and studied the sensitivity of radar backscattering and angular trends for different multi-scale roughness, for different dielectric constants of each layer. We present this sensitivity study in Figure 1 and Figure 2.

3.1. Sensitivity to Multi-scale Roughness Parameters

We have simulated the angular trends from 20 to 60 degrees of the backscattering coefficient for different roughness parameters. We kept γ_0 at 0.2 cm (Figure 1(a)) and varied the fractal parameter ν_1 from 1.3 to 2.1 in VV polarization for ten spatial scales.

Surfaces with ν between 1.7 and 2.1 can be considered as smooth where as surfaces with $\nu_1 = 1.1$ are quite rough. For all the simulations, the backscattering coefficient decreases with the incidence angle. We notice that the backscattering coefficient increases when ν increases. We also notice an increasing trend of the backscattering coefficient with the multi-scale parameter γ because the surface is rougher (Figure 1(b)).

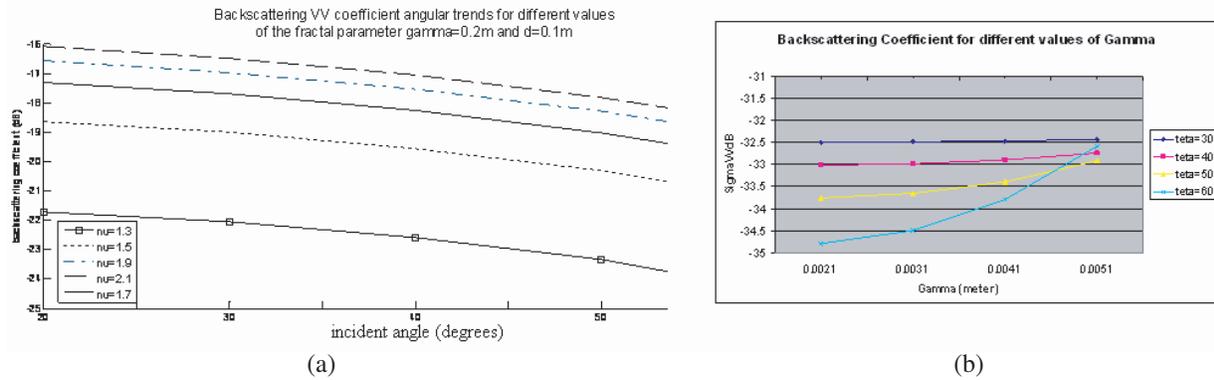


Figure 1. Backscattering coefficient dependence on fractal parameter (a) ν and (b) γ .

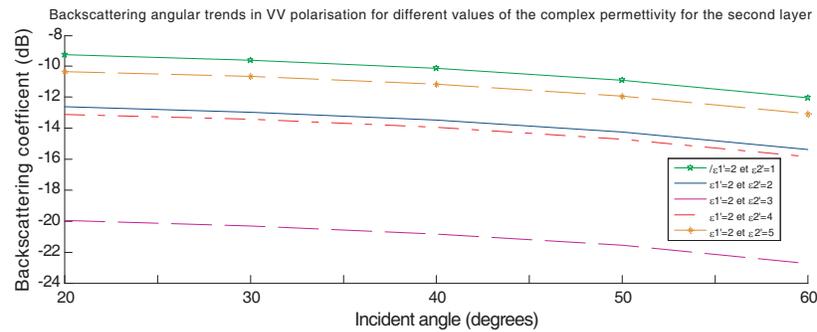


Figure 2. Backscattering angular trends in VV polarisation for different values of the complex permittivity for the second layer.

3.2. Sensitivity to Soil Moisture

Soil moisture is related to the complex dielectric constant ε . In the Figure 2, we have represented radar backscattering as angular trends for different values of the complex permittivity of the second layer. We notice that the backscattering coefficient σ_0 decreases for the low values of ε_2 and after passing by a minimum and then increases. This can be explained by the fact that when the layers are dry corresponding to a lower humidity and as a consequence a lower dielectric constant the penetration of the signal is more important and the backscattered signal is lower. As the dielectric constant increases, the surfaces and subsurface become wetter and the backscattered signal increases because the penetration is lower.

4. METHODOLOGY OF THE RETRIEVAL PROCEDURE

4.1. Inversion Procedure

In this section, an algorithm to retrieve multi-scale roughness parameters and soil moisture parameter is illustrated. The method consists of inverting the SPM direct model using a multi-layer perceptron architecture.

The inversion consists in retrieving roughness and soil moisture parameters $\gamma_1, \gamma_2, \nu_1, \nu_2, \varepsilon_1$ et ε_2 by using as input parameters the radar backscattering coefficients σ_{HH}, σ_{VV} and the incident angle θ varied from 30 to 60 degrees. The NN is trained by learning rules using the backpropagation method. Simulated data sets based on the SPM surface scattering model are used to train the neural network. Before the training of the neural network for the parameters retrieval some considerations concerning the information content of the training data need to be made. If the training data are not sensitive to some of the parameters of interest the inversion for these parameters would be ineffective [3].

4.2. Neural Network Training

The first step in the inversion procedure is the generation of a set of training patterns. In this study, a total of 320 training patterns were generated by using each of the signal models $\sigma(\Xi)$ of the SPM backscattering coefficient. The parameters of interest Ξ used to generate the training patterns were

randomly selected from within the range of parameters given by the sensitivity analysis. We have a total of 8 inputs corresponding to the backscattering coefficients σ_{HH} , σ_{VV} for 4 incident angles and 4 outputs. We have used 2 hidden layers containing 40 neurons after several tests.

4.3. Inversion Algorithm Results

To illustrate the inversion techniques described in the previous section we apply them to data simulated by the SPM. Before using the NN for the inversion we have to calculate the mean rms error of the network. We found that it converges well to a value smaller than 0.05 after 6000 iterations so that the NN is ready for the inversion procedure. Table 1 and Table 2 present respectively the original and the retrieved data for eleven sets of data from S_1 to S_{11} .

We can notice that the inversion has given quite satisfactory results as the original values were retrieved with an error of 8%.

Table 1. Original values.

| | S_1 | S_2 | S_3 | S_4 | S_5 | S_6 | S_7 | S_8 | S_9 | S_{10} | S_{11} |
|-----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|----------|----------|
| ε_1 | 10.000 | 11.000 | 12.000 | 12.000 | 14.000 | 15.000 | 12.000 | 10.000 | 15.000 | 11.000 | 12.000 |
| ε_2 | 5.00 | 5.00 | 6.00 | 8.00 | 5.00 | 6.00 | 8.00 | 5.00 | 6.00 | 5.00 | 6.00 |
| ν_1 | 1.30 | 1.50 | 1.70 | 1.30 | 1.50 | 1.50 | 1.30 | 1.30 | 1.50 | 1.50 | 1.70 |
| ν_2 | 1.30 | 1.50 | 1.70 | 1.30 | 1.50 | 1.50 | 1.30 | 1.30 | 1.50 | 1.50 | 1.70 |

Table 2. Retrieved values after the inversion by the NN.

| | S_1 | S_2 | S_3 | S_4 | S_5 | S_6 | S_7 | S_8 | S_9 | S_{10} | S_{11} |
|-----------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|----------|----------|
| ε_1 | 11.227 | 12.852 | 13.005 | 11.579 | 13.662 | 13.387 | 11.579 | 11.232 | 13.394 | 12.861 | 13.007 |
| ε_2 | 5.92 | 5.63 | 5.44 | 5.93 | 5.03 | 5.42 | 5.93 | 5.91 | 5.41 | 5.63 | 5.44 |
| ν_1 | 1.23 | 1.52 | 1.54 | 1.31 | 1.48 | 1.50 | 1.31 | 1.23 | 1.50 | 1.52 | 1.54 |
| ν_2 | 1.23 | 1.50 | 1.54 | 1.30 | 1.49 | 1.49 | 1.30 | 1.23 | 1.49 | 1.51 | 1.54 |

5. CONCLUSIONS

A sensitivity analysis of backscattering coefficient to the multi-scale roughness parameters permit us to determine the range of parameters to use in the training of the network. To perform the inversion of the small perturbation multi-scale scattering model (MLS SPM) we used a multi-layer neural network trained by a backpropagation learning rule. The inversion procedure has given quite satisfactory results with a mean error of 8%.

REFERENCES

- Davidson, M., T. Le Toan, F. Mattia, G. Satalino, T. Manninen, and M. Borgeaud, "On the characterization of agricultural soil roughness for radar sensing studies," *IEEE Trans. Geosc. Rem. Sens.*, Vol. 38, 630–640, 2000.
- Farah, L. B, I. R. Farah, R. Bennaceur, Z. Belhadj, and M. R. Boussema, "Soil multi-scale roughness parameters and soil moisture retrieval from radar backscattering using a neural network technique," *Progress In Electromagnetics Research Symposium*, Tokyo, Japan, August 2–5, 2006.
- Fung, A. K, *Microwave Scattering and Emission Models and Their Applications*, Artech House, 1994.
- Mallat, S. G, "A theory for multi-resolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, 674–693, 1989.
- Mattia, F. and T. Le Toan, "Backscattering properties of multi-scale rough surfaces," *Journal of Electromagnetic Waves and Applications*, Vol. 13, No. 4, 493–528, 1999.
- Paillou, P., Y. Lasne, E. Heggy, J. M. Malezieux, and G. Ruffie, "A study of P-band synthetic aperture radar applicability and performance for Mars exploration: Imaging subsurface geology and detecting shallow moisture," *Journal of Geophysical Research*, Vol. 111, 2006.

Accuracy of Wind Field Deduced from Envisat WSM SAR Images along the Range

P. Trivero, W. Biamino, M. Borasi, and M. Cavagnero

Dipartimento di Scienze dell’Ambiente e della Vita
Università del Piemonte Orientale “Amedeo Avogadro”, Italy

Abstract— Since several years the Synthetic Aperture Radar (SAR) is used as a reliable tool for sea surface observation. SAR provides information about the sea surface roughness generated by wind. In order to extract wind field from SAR images the direction and the intensity of the wind vector must be considered separately. Particularly the well-established methods for determining the wind intensity required the wind direction as input. We analysed the Envisat WSM images: the wind direction has been deduced applying the Continuous Wavelet Transform (CWT) analysis to the periodic structures of the image. The obtained direction has been used as input of the CMOD5 model to calculate the intensity. The large swath width of an Envisat WSM image corresponds to incidence angles ranging from 15 to 45, involving specular and Bragg scatter mechanisms; the Normalised Radar Cross Section (NRCS) depends on the position along the range, on the wind intensity as well as on the angle between the radar beam and the wind direction. Here we present an analysis concerning the accuracy of the wind field deduced from Envisat WSM SAR images as a function of the position along the range. We have compared the data obtained in different position along the range, of about 100 images, with external data of wind. The accuracy of the outputs of the wind as a function of the different zones of the range are presented and the backscatter mechanisms and sensitivity concerning NRCS are discussed.

1. INTRODUCTION

The interaction between wind and sea surface is a widely analysed phenomenon. Sea surface roughness is produced by wind when a threshold is reached [1]. Information about sea surface roughness can be derived from backscattered radio signal at short (1–10 cm) wavelengths. The detected signal is due to constructive interference coming from these short wavelengths. The basic mechanism involved is the NRCS which, for incidence angles higher than 20°, is proportional to the spectral energy density of the sea waves having wavelength Λ that obey the Bragg resonance condition:

$$\Lambda = \frac{\lambda}{2 \sin \theta} \quad (1)$$

where λ is the radar wavelength and θ the incidence angle of radar beam. For low incidence angles the backscatter is due to specular reflection.

Figure 1(a) shows an example of Envisat WSM SAR image in descending orbit; the right of the image correspond to the near range and NRCS is higher than in the far range. This is due to the backscatter mechanism between radio waves and the sea surface roughness strongly dependent from incidence angle; in Figure 1(b)) the theoretical NRCS as a function of incidence angle (at fixed wind speed and direction) is plotted, showing the strong difference between near and far range.

According to different satellites features, the range of incidence angles can vary from small intervals (e.g., ERS, from 19° to 26°) to wide extensions as Envisat: WSM images span from 16°, where specular reflection is still important, to 42°; therefore different angles lead to different sea wavelengths inducing Bragg condition.

The aim of this work is to evaluate the performances, at different incidence angles, of wind extraction algorithms from SAR images.

2. WIND FIELD EVALUATION

Starting from NRCS the wind intensity is evaluated by using a geophysical model function (GMF), i.e., a nonlinear function who describes the NRCS as a function of wind speed (normalized to 10 m height), wind direction, incidence angle and beam azimuth angle with respect to wind direction. For satellite-borne SAR operating at C-band and vertical (VV) polarization, several empirical GMFs have been developed and validated.

To correctly measure wind speeds from SAR images using such kind of methods, it is often necessary to get the wind direction from an external source; the methodology of wind extraction

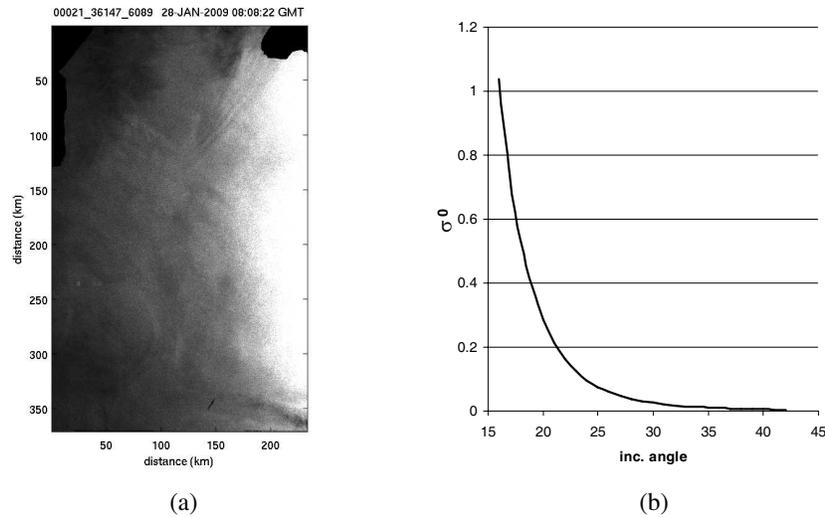


Figure 1: (a) SAR image quicklook; (b) N RCS versus incidence angle.

used here, is based on a two-dimensional implementation of the continuous wavelet transform (2D-CWT), as described in [2].

With the knowledge of the radar incidence angle, the normalized radar cross section and the angle between the radar looking and the wind direction, the latter retrieved with the above methodology, the wind speed is derived using the CMOD5 model [3].

3. DATASET ANALYSIS

The analysed dataset is composed by 100 SAR images, each one containing about 1500 wind vectors. Wind velocities computed by CMOD5 have 0.1° spatial resolution; the direction is evaluated by wavelet analysis. The reference wind (ECMWF data) has been interpolated in the same points starting from a coarser resolution grid.

For every image, wind velocities from CMOD5 and ECMWF have been compared, in order to evaluate the ability of CMOD5 to correctly evaluate wind over the whole incidence angles range. For every wind intensity the difference between the two values has been evaluated and represented as function of the corresponding incidence angle in order to carry out an evaluation of CMOD5 performances along the range.

4. DISCUSSION

The results of the above mentioned analysis are reported in Figure 2.

The graph in Figure 2 represents the difference between CMOD5 and ECMWF evaluated wind velocities. Blue line is the mean difference for all wind vectors included in a 3° interval of incidence angle. Purple and red line represent the difference 2–3 m/s and 5–6 m/s wind speed respectively. These data show that CMOD is correlated to ECMWF with a slightly underestimation in the near range where specular reflection is the main contribution, while overestimates wind values in the far range. Higher winds (red line) have a better agreement up to about 35° . A strong dependence from incidence angle is however noticeable.

An explanation of this features can be found in the variability of radar backscatter. Figure 3 shows theoretical values of σ_0 for different wind values at different incidence angles; it easy to note how the σ_0 drastically decrease in the far range, reaching very low values. In Figure 4 are plotted the differences in σ_0 for a variation of 1 m/s between different wind speeds, showing a decreasing sensitivity in the far range. At 45° , for example, a difference of 1 m/s is almost undetectable because the radiometric σ_0 values at low wind speeds have nearly the same order of magnitude of these differences.

Values shown in Figure 3 can be explained by supposing that, for low winds and large incidence angles, the SAR signal averaged in a single cell shows a large variability and does not fall under a certain threshold value. For low winds this threshold is reached at low angles, while at stronger winds the threshold is reached at larger incidence angles.

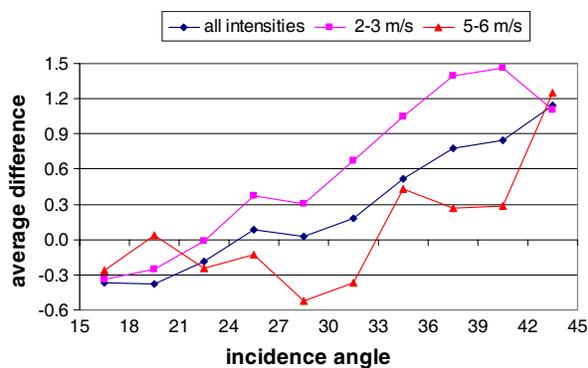


Figure 2: Difference between CMOD5 and ECMWF data, versus incidence angle.

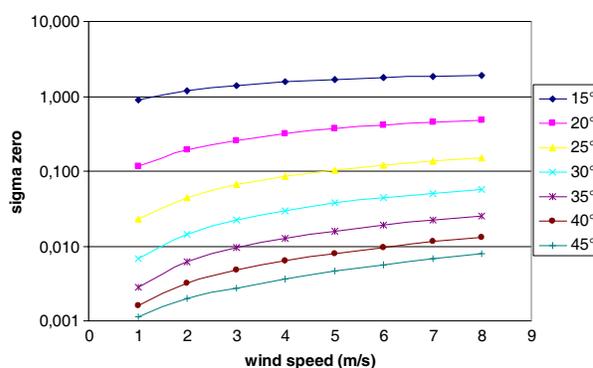


Figure 3: Theoretical NRCS versus wind speed at different incidence angles for a beam azimuth angle (with respect to wind direction) of 90°.

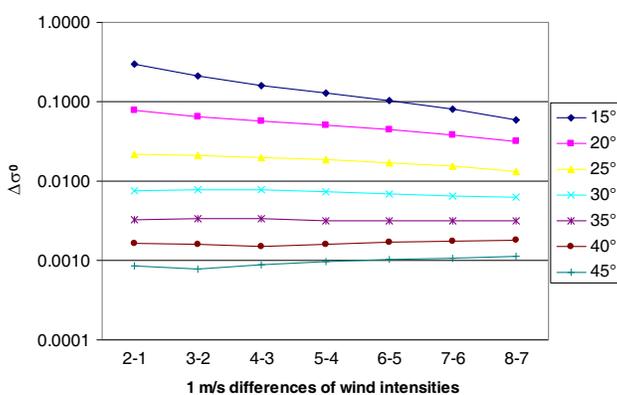


Figure 4: Differences in σ_0 between different winds (from Figure 3).

ACKNOWLEDGMENT

The Envisat ASAR WS images data have been provided by the European Space Agency, on the framework of the ESA Project Start Up C1P.5404 and C1P.7102.

REFERENCES

1. Pierson, W. J. and R. A. Stacy, "The elevation, slope, and curvature spectra of wind roughened sea surface," NASA C.R. 2646, December 1973
2. Zecchetto, S., F. de Biasio, and P. Trivero, "A wavelet technique to extract the backscatter signatures from SAR images of the Sea," *PIERS Online*, Vol. 5, No. 7, 696–700, 2009.
3. Hersbach, H., A. Stoffelen, and S. de Haan, "An improved scatterometer ocean geophysical model function: CMOD5," *J. Geophys. Res.*, Vol. 112, 5767–5780, 2007.

Combined Direct and Remote Sensing Measurements of Wave Parameters at the off-shore Research Platform in the Black Sea

N. Y. Komarova¹, F. De Blasio², A. S. Kuznetsov³, M. N. Pospelov¹, and S. Zecchetto²

¹Space Research Institute, Moscow, Russia

²Institute of Atmospheric Sciences and Climate, Italy

³Marine Hydrophysical Institute, Ukraine

Abstract— The paper presents the results of the experiment CAPMOS'05 performed at an offshore oceanographic platform in the Black Sea. The experiment was aimed at air-sea coupling investigations by means of direct and remote measurements. A specialized research platform managed by the Marine Hydrophysical Institute is located on the shelf slope approximately 600 m to the south of Crimea coast, Ukraine. Spectral parameters of wind and waves were estimated from direct and remote measurements. The peak frequency of the wind retrieved from radar measurements varied from 0.002 to 0.007 Hz, corresponding to wavelengths between 500 and 7000 m. The spectral peak frequency of gravity waves varied in the range from 0.2 to 0.7 Hz that corresponded to dominating wave lengths from 50 down to several meters. Comparison of two different techniques of wave spectrum retrieval from radar data, and also comparison of radar and radiometric data shows satisfactory agreement.

1. INTRODUCTION

Satellite-based measurements are broadly used in modern oceanographic research. Accuracy and consistency of remote sensing data depend on the quality of models, which relate geophysical parameters to the parameters of electromagnetic waves emitted/scattered by natural environment. Detailed and highly accurate measurements are necessary for developing and testing of these models. Offshore platforms provide a unique opportunity to perform long-term measurements of oceanic and atmospheric parameters in a fixed position using various kinds of remote and contact sensors [1, 2].

The paper presents some results of the field experiments CAPMOS'05, performed at an offshore oceanographic platform in the Black Sea. The experiment was carried out in framework of the INTAS (International Association for the promotion of co-operation with scientists from the New Independent States of the former Soviet Union) project “Combined Active/Passive Microwave Measurements of Wind Waves for Global Ocean Salinity Monitoring (CAPMOS)”. The major goal of the experiment was to compare the results of synchronous direct and remote sensing measurements of waved sea surface and atmosphere boundary layer with specific emphasis on the ocean wave spectrum and wind speed retrieval.

Remote measurements of wave parameters are of particular importance for oceanographic studies. Whereas traditional contact sensors (like string wave gauges or pressure sensors) are still being widely used, recent trends are toward increased use of remote sensors. Several sea wave spectrum models have been constructed recently basing upon remote sensing (mainly, radar) data. It is worth noting, however, that these models differ one from another, justifying the necessity of field measurements of wave parameters under variety of weather and sea conditions using various kinds of sensors.

2. EXPERIMENT

The experiment CAPMOS'05 was performed on an offshore oceanographic platform in the Black Sea in June 2005. The research platform managed by the Marine Hydrophysical Institute is located approximately 600 m to the south of Crimea coast near Katsiveli, Ukraine (Figure 1). The sea depth at the site is about 30 m, so the deep water and long fetch conditions were ensured for prevailing winds from the south, south-east and south-west.

The platform has several working levels. The lowest deck at 4 meters above the surface is used for various instruments installation. At the main deck at 14 meters, the living rooms and laboratories are situated. The mast on the roof is used for the meteorological equipment installation.

List of research instruments and equipment used in the experiment included:

- below the surface:
 - five current meters at a depth of 3, 5, 10, 15 and 20 m;



(a)



(b)

Figure 1: (a) Map of the south coast of Crimea, the platform position is marked with an arrow. (b) Research platform.

- bottom to top — CTD sections every 3 hours;
- water temperature and turbulence sensors at 1 m;
- CTD floating at a depth of 0.3 m;
- 6-strings wave gauge;
- above the surface:
 - 3-component sonic anemometer and air temperature sensor — at 1.5 m;
 - a set of microwave and IR radiometers mounted on an automatic rotator, 3-component sonic anemometer and air temperature sensor, water vapor and carbon dioxide sensor — at 4 m;
 - Ku-band scatterometer (polarizations VV , HH or cross), air pressure sensor — at 14.5 m;
 - 3-component sonic anemometer, air temperature and humidity sensor — at 21.5 m.

The measurements were carried out continuously 24 h/day from June 2 to June 20, 2005. The mean wind speed ranged from 0 to 13 m/s; two episodes of high wind speeds were observed, with gusts well above 20 m/s. The boundary layer conditions were mostly stable (60%) or neutral (30%). Wind directions were essentially from three sectors: from 350 degrees (from land, short fetch), from East (90 degrees, long fetch) and from West, North-West (long fetch).

3. SCATTEROMETER MEASUREMENTS

The Ku-band scatterometer is pulsed and coherent radar, emitting 40 ns long pulses of 15.4 GHz e.m. radiation at 156 kHz repetition frequency. Azimuth and incidence angle could be chosen manually. Azimuth direction was usually either upwind or downwind. Almost 100 hours of radar backscatter at VV -polarization were obtained during the whole experiment: the 44% at upwind direction, the 56% at downwind. The typical record time length was of 1800 s. Incidence angle (and corresponding distance from the radar to the surface) varied from measurement to measurement from 25 up to 65 degrees.

Since the radar is a coherent device, it is possible to derive simultaneous values of backscatter intensity (expressed in terms of the Normalized Radar Cross Section — NRCS) and radar Doppler frequency f_d . The radar Doppler frequency may result positive or negative according to the radar-target distance variation.

Further, standard techniques, like spectral analysis, may be used to extract information on the frequency content of both the waves (from the NRCS and f_d time series) and the energy containing sub range in the atmospheric boundary layer (from the NRCS time series only) [3], as the frequency domain of the gravity waves roughly lies between 0.08 Hz and 0.5 Hz, while the energy containing sub range extends “at time scales between tens of seconds and tens of minutes” [4]. Figure 2 shows an example of the spectra of NRCS and f_d of a record performed at incidence angle 46 degrees, normalized for comparison. The power spectrum of the NRCS series exhibits the interesting feature of describing the frequency structure of the marine atmospheric boundary layer fingerprint on the sea surface as well as that of the gravity waves, also described by the spectrum of the Doppler frequency.

To extract the wind spectral characteristics, we must filter out a non-stationary component of the process, due to the presence of periodicities larger than the 1800 s measurements windows. The

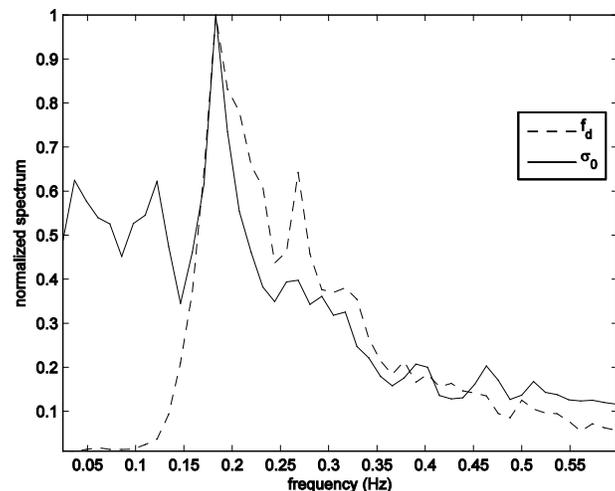


Figure 2: Power spectra of NRCS σ_0 (solid line) and the radar Doppler frequency f_d (dashed line) at incidence $\theta = 46^\circ$. The spectra have been normalized for comparison.

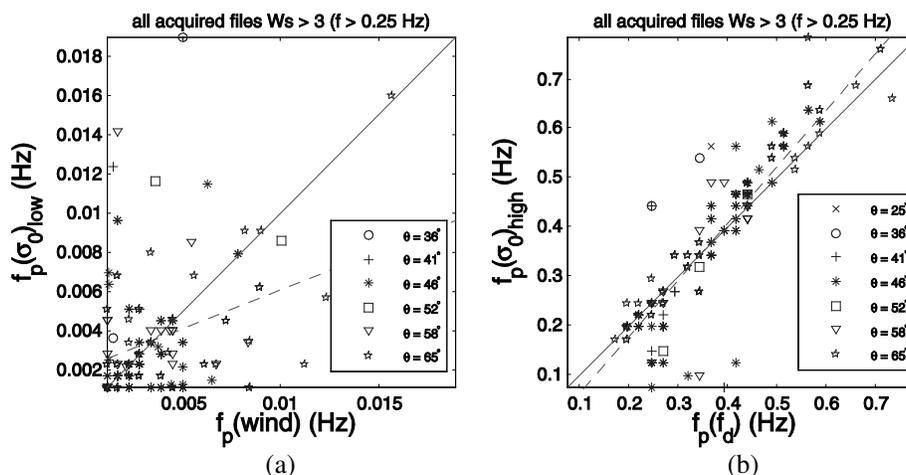


Figure 3: (a) Scatter plot of wind spectral characteristics derived from the up-zero crossing analysis of the wind and the NRCS time series. (b) Scatter plot of wave spectral characteristics derived from the f_d and NRCS time series. Solid line: perfect agreement. Dashed line: best fit.

peak frequency of the wind obtained from the NRCS time series resulted from 0.002 to 0.007 Hz (relevant period of 150 to 500 s), corresponding to wavelengths between 500 to 7000 m. Wind speed and direction was measured by sonic anemometers installed at different heights. In the Figure 3(a) the spectral parameters of wind obtained from the radar NRCS and the wind time series have been compared. The scatter is large and thus the correlation is low (0.4). The Figure 3(b) shows the comparison of the gravity wave peak frequencies obtained from the spectral analysis of NRCS and f_d . The agreement is very good, with correlation equal to 0.9.

4. RADIOMETER MEASUREMENTS

List of radiometers enabled during the experiments included: S-band radiometer (V-pol.); K- and Ka-band polarimeters (3 Stokes parameters); W-band radiometer (V- and H-pol.); IR-radiometer. All radiometers were mounted on an automatic rotator which allowed changing the angle of observation azimuthally and vertically. The rotator was controlled by a computer. Typical algorithm of rotation consisted of vertical scanning from 20 to 153 degrees and backward at 0.2 rpm at six fixed azimuths from 60 to 240 degrees stepped by 36 degrees.

Further, the brightness temperature dependence on incidence angle was used for the curvature spectrum of gravity-capillary waves retrieval. These results are presented in a separate paper [5]. Hereafter we will compare the results of gravity wave spectrum retrieval from radiometer and scatterometer data.

A chance for a comparison appeared on the 9th of June when the radiometers performed measurements in a non-scanning mode. All radiometers together with the scatterometer were pointed in upwind direction at the same observation angle, and registered synchronously microwave brightness and backscatter signals. The total duration of measurements was about 70 minutes at observation angle about 65 degrees, and then about 95 minutes at 40 degrees. The wind speed during the measurements gradually decreased from 11 down to 5 m/s.

Figure 4 shows a 60-s fragment from the record at 40 degrees observation angle of the first three Stokes parameters of partially polarized microwave radiation at Ka-band: $((T_v + T_h)/2, T_v - T_h, T_{+45} - T_{-45})$, where T_v , T_h , T_{+45} , and T_{-45} are brightness temperatures at basic linear polarizations, respectively: vertical, horizontal, +45 and -45 degrees from the vertical. The brightness temperature of water surface is known to be highly dependent on observation angle, due to angular dependence of the Fresnel coefficients. The radiometer antenna footprint on the surface at 40 degrees observation angle was approximately 1.3×1 m that is much less than the length of energy-carrying gravity waves. While every single wave is crossing the antenna beam, the local observation angle changes causing respective change of the brightness temperature. So, the spectral analysis of

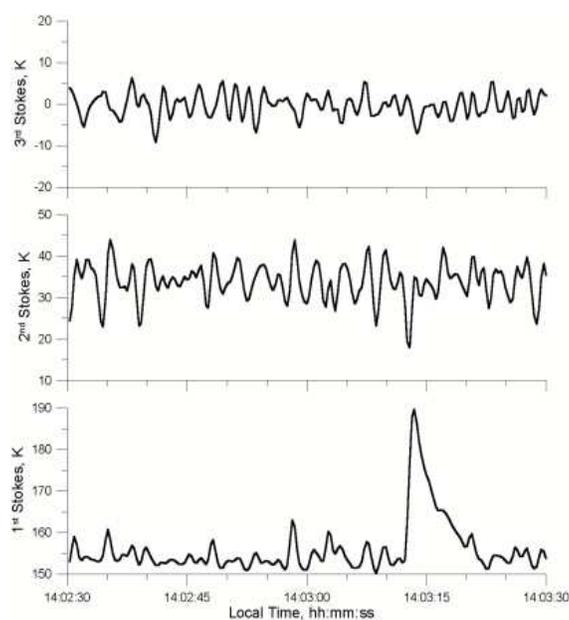


Figure 4: A 60-s fragment of the three Stokes parameters (in Kelvins) record at 40 degrees observation angle. June 9, 2005. Wind speed at the moment was about 8 m/s.

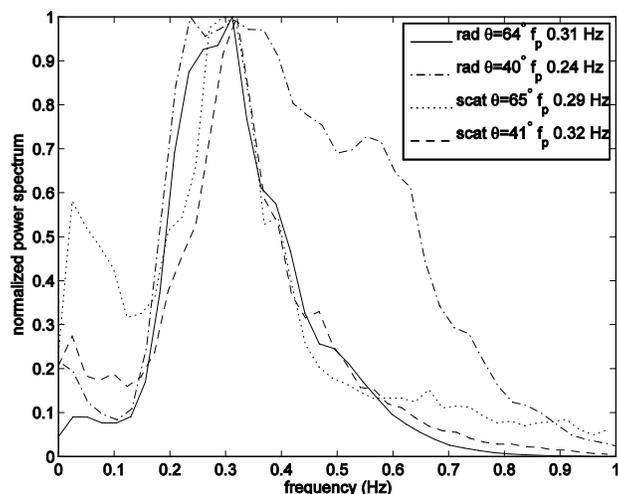


Figure 5: Normalized power spectrum of gravity waves retrieved from NRCS and radiometer (the third Stokes parameter) data at observation angles about 65 and about 40 degrees from nadir. June 9, 2005.

the second or third Stokes parameter may provide information about gravity wave slope spectrum (we mean the waves with lengths exceeding the antenna footprint).

Figure 5 shows normalized power spectrum retrieved from NRCS at Ku-band and from the third Stokes parameter at Ka-band during the mentioned series of measurements. The agreement of scatterometer and radiometer data is evident. The broadest line width of the spectrum retrieved from radiometer data at 40 degrees of incidence is a result of the observation geometry. It corresponds to the smallest antenna footprint on the surface among all four cases and hence the range of the observed waves is extended to smaller lengths.

5. CONCLUSIONS

Wind and wave spectral parameters were measured from an offshore oceanographic platform using microwave scatterometer and radiometers. Spectral parameters of the wind were estimated using NRCS measurements at Ku-band. The peak frequency of the wind resulted from 0.002 to 0.007 Hz, corresponding to wavelengths between 500 to 7000 m. Comparison of spectral maxima in the atmospheric boundary layer motions derived from direct measurements by a sonic anemometer and NRCS time series does not show perfect correlation. The reason for this disagreement is still unknown, but we are aware that to compare quantities obtained from different time series, like wind record at 21.5 m of height and radar backscatter at sea surface is very delicate. Gravity wave spectrum was retrieved from scatterometer data using two different approaches: NRCS and Doppler frequency analysis. The spectral peak frequency of gravity waves varied in the range from 0.2 to 0.7 Hz that corresponded to dominating wave lengths from 50 down to several meters. Comparison of two different techniques, and also comparison with radiometric data shows satisfactory agreement.

ACKNOWLEDGMENT

The study was supported by INTAS project 03-51-4789 and in part by RFBR project 09-02-00780-a.

REFERENCES

1. Keller, W. C. and W. J. Plant, "Cross section and modulation transfer function at L and Ku bands measured during the tower ocean wave and radar dependence experiment," *J. Geophys. Res.*, Vol. 95, 16277–16289, 1990.
2. Camps, A., J. Font, J. Etcheto, V. Caselles, A. Weill, I. Corbella, M. Vall-Ilossera, N. Duffo, F. Torres, R. Villarino, L. Enrique, A. Julia, C. Gabarro, J. Boutin, E. Rubio, and S. Reising, "Sea surface emissivity observations at L-band: First results of the wind and salinity experiment WISE 2000," *IEEE Trans. Geosci. Remote Sensing*, Vol. 40, No. 10, 2117–2130, 2002.
3. Zecchetto, S., "Effects of modulation of σ_0 with implications on physical scatterometer modelling," *The Air-sea Interface, Radio and Acoustic Sensing, Turbulence and Wave Dynamics*, M. A., Donelan, W. H. Hui, and W. J. Plant, Editors, 729–733, University of Miami, 1996.
4. Panofsky, H. A. and J. A. Dutton, *Atmospheric Turbulence*, John Wiley & Sons, 1984.
5. Pospelov, M. N., A. V. Kuzmin, and I. N. Sadovsky, "Using of multi-angular radiometric measurements for short wind wave parameters estimate," this issue.

RCS Simulations on Wet Corner Reflectors with SBR Code SIGRAY

E. Kemptner

DLR, Microwaves and Radar Institute, Germany

Abstract— Corner reflectors play a major role in the calibration of radar systems. The influence of geometry deformations, caused by insufficient production accuracy or outside exposure, and caused by dirt, water film, snow or snow slush on the RCS (Radar Cross Section) is often neglected. In this work simulation results on trihedral corner reflectors with wet base areas are presented. The ‘Shooting & Bouncing of Rays’ (SBR) code SIGRAY developed at DLR has been applied. Depending on the polarization strong reduction of the RCS was detected.

1. INTRODUCTION

The accuracy of remote sensing radar systems has been improved in the past. Therefore the requirements on the system’s calibration are also increasing. Corner reflectors play a major role in the calibration of radar systems. The influence of geometry deformations, caused by insufficient production accuracy or outside exposure, and the influence of dirt, water film, snow or snow slush on the RCS is often neglected. This topic has been investigated by applying the ‘Shooting & Bouncing of Rays’ (SBR) [1] code SIGRAY which has been developed at DLR. SIGRAY has been successfully applied in the past to deformed dry corner reflectors [2] and to several other targets. In this work simulation results on trihedral corner reflectors with wet base areas are presented. Because of adhesion a water layer occurs not only on a horizontal base panel of the corner reflector but also on slightly tilted ones. This results in the need to account for a water film on the base panel of corner reflectors in radar calibration.

2. RCS MODEL SIGRAY

SIGRAY is based on the ‘Shooting & Bouncing of Rays’ method and models multiple reflections by using a combination of Geometrical Optics (GO) and Physical Optics (PO). This high frequency approximation method is well suited for RCS calculations on 50λ corner reflectors. In the framework of SIGRAY the triple reflection of rays, which is the dominant reflection order on electrically large corner reflectors, is simulated by twofold application of GO and single-use of PO which leads to good results on the mentioned radar targets. SIGRAY can separate scattering contributions corresponding to the order of reflection.

Instead of an incident plane wave a great number of rays are shot towards the reflector corresponding to a fine grid. The propagation lengths are measured from and back to a virtual plane, which is a kind of substitute for the antenna in infinity, perpendicular to the viewing direction. All reflected rays interfere at this receiving plane and accumulate to the total scattered field.

For each ray reflection on an imperfectly conducting surface Fresnel’s reflection coefficients, which depend on frequency, angle and polarization, are applied. The table of these coefficients is either calculated at the beginning of a SIGRAY simulation or is imported, if the values have been measured in advance.

3. SIMULATIONS

A trihedral corner reflector with an edge length of 1.5 m was investigated at a frequency of 10 GHz. Typically several millions of rays hit the target and are tracked. Double and triple bounce contributions have been taken into account because of the wide elevation angle range. All calculations have been performed in a θ -plane at $\varphi = 45^\circ$. The base panel of the reflector is located in the x - y -plane at $z = 0$ (Fig. 1). At the well-known elevation angle of $\theta = 54.74^\circ$ the maximum RCS value for a PEC corner reflector occurs.

The base panel of the reflector was assumed to be wet. Water films of different thicknesses have been investigated. Fresnel’s reflection coefficients have been applied.

Figure 2 shows Fresnel’s reflection coefficients of a PEC plate coated with 1 mm water at $\nu = 10$ GHz. The permittivity of water has been taken as $\text{Re } \varepsilon = 55$, $\text{Im } \varepsilon = -30$ [3]. E - and H -polarization relate to the base panel.

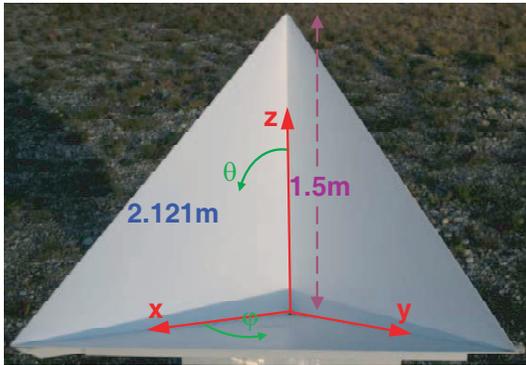


Figure 1: Geometry of the investigated corner reflector.

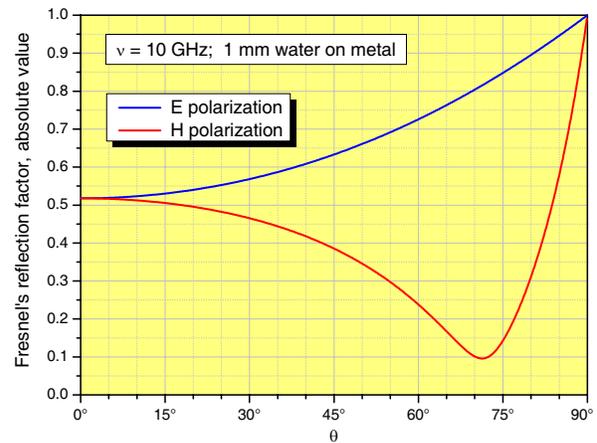


Figure 2: Fresnel's reflection coefficients.

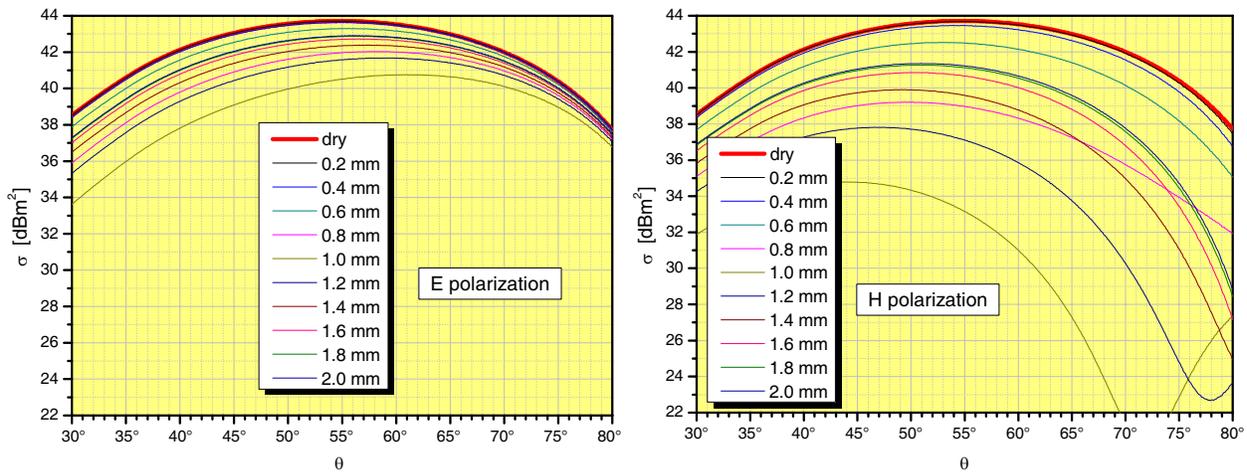


Figure 3: RCS as a function of the elevation angle.

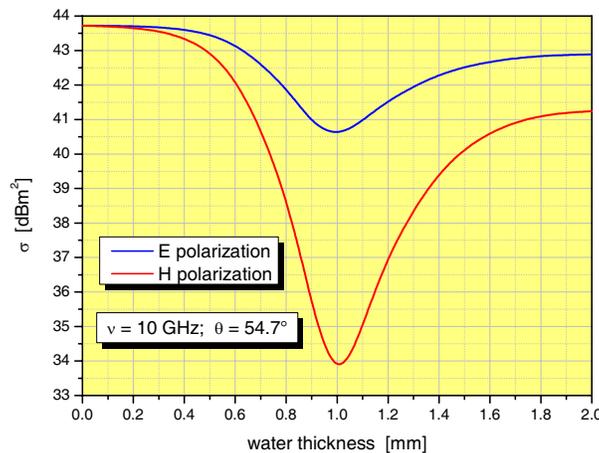


Figure 4: RCS as a function of the water film thickness.

4. RESULTS

As shown in Fig. 2 the dependence of Fresnel's reflection coefficients on the elevation angle at 10 GHz is very strong. According to this the corner reflector's RCS changes dramatically and there is a significant difference between the polarization cases. In Fig. 3, the results of the RCS calculations for both polarizations are shown in dependence on the elevation angle. Several water film thicknesses have been considered.

According to the well-known high-frequency approximation formula for the RCS of PEC tri-

hedral corner reflectors $\sigma = (\pi l^4)/(3\lambda^2)$ the maximum RCS value is 43.72 dBm² ($l = 2.121$ m, $\lambda = 0.03$ m). Applying Fresnel's reflection coefficients for 54.74° and 1 mm water layer ($R_E = 0.69$ and $R_H = 0.3$) yields 40.50 dBm² and 33.26 dBm². These three values correspond perfectly to those calculated by SIGRAY (see Fig. 3 for $\theta = 54.74^\circ$). Fig. 4 shows the RCS values for $\theta = 54.74^\circ$ as a function of the water film thickness. Obviously the attenuating effect is highest if the water layer is 1 mm. The influence of water on the base panel is not only the reduction of RCS, but also a shift of the elevation angle with the highest RCS value, a declination in elevation. Both polarizations show strongly different effects. For example in the case of 1 mm water layer the highest RCS values occur at $\theta = 61.4^\circ$ (E -polarization) and at $\theta = 44.0^\circ$ (H -polarization).

5. CONCLUSIONS

The presented results show that a thin water film on the base area of corner reflector can cause a significant reduction of the RCS depending on the applied frequency, the thickness of the water layer and the polarization. A consequential recommendation consists in avoiding incidence angles in the calibration process which correspond to a roughly horizontal base area of the corner reflector. Coatings of the metal plates which handicap the drain off of water should also be avoided.

REFERENCES

1. Ling, H., R. Chou, and S. Lee, "Shooting and bouncing rays: Calculating RCS of an arbitrary cavity," *IEEE Trans. on Antennas and Propagation*, Vol. 37, 605–607, May 1987.
2. Kemptner, E., "RCS simulations on deformed corner reflectors applying SBR code SIGRAY," *EuCAP 2010*, Barcelona, Spain, April 12–16, 2010.
3. Meinke, H. H. and F. W. Gundlach, *Taschenbuch der Hochfrequenztechnik*, K. Lange, K.-H. Löcherer, Editors, E4, Springer-Verlag, 1986.