Circuit Simulation of Varactor Loaded Line Phase Shifter

M. Ould-Elhassen[1], M. Mabrouk[1], A. Ghazel[1], and P. Benech[2]

[1]CIRTACOM, SUPCOM, ISETCOM de Tunis, Tunisia
[2]IMEP-LAHC (CNRS-INPG-UJF), France

Abstract— This paper describes circuit simulation of analogue phase shifter based on distributed CPW transmission lines loaded by Varactor diodes. The expression of phase shifting is obtained using the global (S_{ij}) matrix of 9 units of proposed phase shifter. The simulations are carried out on ADS simulator. Comparison between our simulated results and published measurements of the studied phase shifter is made and a good agreement is obtained.

1. INTRODUCTION

Over the last decades, several advances have been made in analogue and digital. These devices are used to change the insertion phase of transmitted signal. The main and interesting phase shifters are those providing low insertion and return loss, and equal amplitude in all phase states. These criteria are becoming very important for several wireless communications. Most of phase shifters are reciprocal networks, meaning that they work effectively on signals passing in either direction. Phase shifters can be controlled electrically, magnetically or mechanically [1]. The main important application is within a phased array antenna system in which the phase of a large number of radiating elements can be controlled to force the electromagnetic signal to add up at a particular angle to the array.

In this paper, distributed phase shifter consists of a high impedance line (180 Ω) capacitively loaded by the periodic placement of varactors. By applying a single bias voltage on the line, the distributed capacitance can be changed, which in turn changes the velocity of the line and creates a phase shift [2]. The phase shift can be varied in a large variation range depending on the bias voltage and the length of the distributed line.

2. LOADED LINE THEORY

The first step in understanding loaded phase shifter is the basic "electrically NonLinear Transmission Line" (NLTL). NLTL is consisting of coplanar waveguide (CPW) periodically loaded with reverse biased varactor diodes. Nonlinearity is created by the voltage controlled capacitance. The Figure 1 presents two models of our phase shifter unit: the circuit model (Figure 1(a)) and the equivalent model (Figure 1(b)).

Basically, the transmission line model using series L ($\mathrm{H \cdot m^{-1}}$) and shunt C ($\mathrm{F \cdot m^{-1}}$) lumped elements [4], has a phase velocity defined by (1).

$$v_p = \frac{1}{\sqrt{LC}} \tag{1}$$

With a line of constant physical length, a phase shift can then be introduced by varying the phase velocity. A variable L or C is needed to vary v_p. In a transmission line shunt loaded with diodes shown in Figure 1 the total capacitance, and hence the phase velocity become a function of DC bias voltage defined and shown in Equation (2) where l_{sec} is the physical length of a transmission

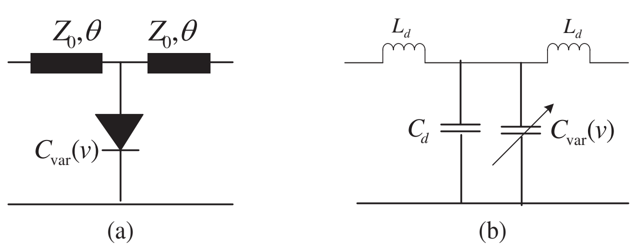


Figure 1: Circuit model of NLTL unit [3].

line section in meters [2]. The parameters C_d and L_d are the inductance and capacitance of each unit cell. L_d, C_d and C_{var} have the same units as L and C respectively.

$$v_p = \frac{1}{\sqrt{L_d\left(C_d + \frac{C_{var}}{l_{\sec}}\right)}} \tag{2}$$

The model of periodic sections transmission line has a Bragg frequency defined and shown in Equation (3), it is similar to optical Bragg diffraction.

$$f_{Bragg} = \frac{1}{\pi\sqrt{L_d\left(C_d + C_{var}\right)}} \tag{3}$$

The first decision is, which transmission line topology have we to use. Coplanar wave guide (CPW) was chosen for our phase shifter structure. CPW has some immediate advantages. First, both ground and signal lines are on the same plane affording easy access for shunt mounting of elements without drilling. Second, CPW has a canonical closed form model from [7] that can be used to obtain design equations.

To maintain a balanced CPW line, a balanced shunt loading topology was chosen with two shunt diodes per transmission line cell unit, one to each of the ground planes.

The variable capacitance parameters are set by the choice of diodes. From (3), a larger C_{var} will cause a reduction in the maximum operating frequency. Also, the range from $C_{var\_\max}$ to $C_{var\_\min}$ will affect the variability of the phase velocity of Equation (2) and hence the phase shift. This affects some of variations in the characteristic impedance Z_0 of the line versus C_{var} [5] as shown in Equation (4), which is desired to be $50\,\Omega$ or impedance matching.

$$Z_0 = \sqrt{\frac{L_d[H]}{C_d[F] + C_{var}[F]}} \tag{4}$$

3. DIODE VARACTOR MODEL

After studying several diodes models, and matching schemes in simulation, the given model by Equation (5) was chosen. With specified $C_{var\_\max}$ to $C_{var\_\min}$ ratio this diode affords reasonable phase shift while allowing the transmission line to be well matched to 50 ohms without additional circuitry.

Diodes have two origins of nonlinearity: conductive and reactive [3]. The conductive nonlinearity is shown in the $I(v)$ curves and the reactive nonlinearity is shown in the $C(v)$ curves.

This model of diode varactor has a series resistance R_s, parasitic series inductance L_s, and parasitic parallel capacitance C_p. Equation (5) gives the mathematical model of simulated varactor diode [8].

$$C_j(V_j) = \frac{C_{j0}}{\left(1 - \frac{V_j}{\phi}\right)^M} \tag{5}$$

where C_j is the fitted junction capacitance, C_{j0} is the zero-bias junction capacitance, V is the junction potential, ϕ is the fitted potential barrier and M is the grading coefficient.

4. LOADED LINE PHASE SHIFTER

After choosing an appropriate diode model the remaining degrees of freedom are substrate choice and loading factor as defined by [5] and shown in the Equation (6). For our simulation RG4003 substrate was chosen to apply the ideal closed form CPW equations.

$$x = \frac{C_{\max}/l_{\sec}}{C_d} \tag{6}$$

From that choice, the required CPW line parameters can be computed to give a $50\,\Omega$ matched line. From [5], the relation between loading factor and the characteristic impedance of the CPW line is shown in (7). From $C_{var\_\max}$ and x from (8), the desired C_d is obtained. The open CPW closed form expressions given by [7] can be used to compute the C [F/m] of the CPW line shown in Equation (11), given Z_i from CPW Equations (9) and (10). The length of each T-line section cell, $l_{\rm sec}$, is given by (12). Z_i is the characteristic impedance of each section.

$$Z_i[\Omega] = 50\sqrt{1+x} \tag{7}$$

$$C_d = \frac{C_{\max}}{x} \tag{8}$$

$$\varepsilon_e = \frac{\varepsilon_r + 1}{2} \tag{9}$$

$$KK = Z_i \frac{4\sqrt{\varepsilon_e}}{120\pi} \tag{10}$$

$$C = \frac{4\varepsilon_0 \varepsilon_e}{KK} \tag{11}$$

$$l_{\rm sec} = \frac{C_d\,[F]}{C\left[\frac{F}{m}\right]} \tag{12}$$

For $x = 5$, the simulation parameters are: $Z_i = 122\,\Omega$, $C_d = 0.4\,\text{pF}$, and $l_{\rm sec} = 8.9\,\text{mm}$. These parameters are then introduced into ADS simulator, and the diode model given by the Figure 2 with a grading coefficient $M = 0.5$, for simulating this phase shifter. The corresponding dimensions of $Z_i = 122\,\Omega$ (CPW) are $W = 2.5\,\text{mm}$ for a conductor width, and $G = 3\,\text{mm}$ for a gap. The input line CPW ($Z_i = 50\,\Omega$) for the biasing sections was computed to have the dimensions $W = 2.5\,\text{mm}$ and $G = 0.25\,\text{mm}$. The Figure 3 gives the circuit model of our studied phase shifter.

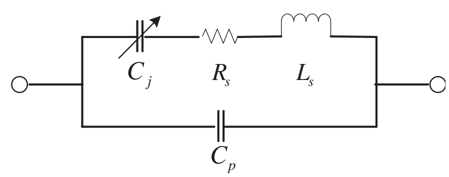


Figure 2: Circuit model of diode Varactor [8].

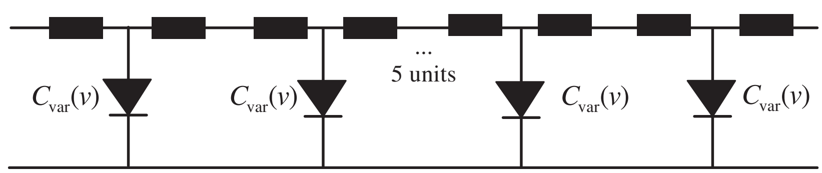
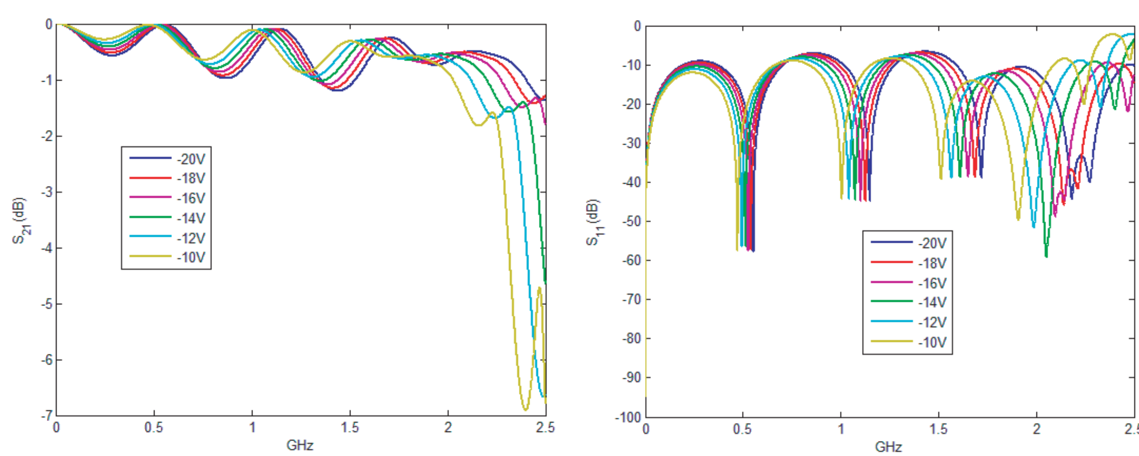


Figure 3: Circuit model of phase shifter.



(a) (b)

Figure 4: Insertion and return losses of our phase shifter.

The number of segments is chosen for the desired phase shift at operating frequency, and we have in our case nine sections.

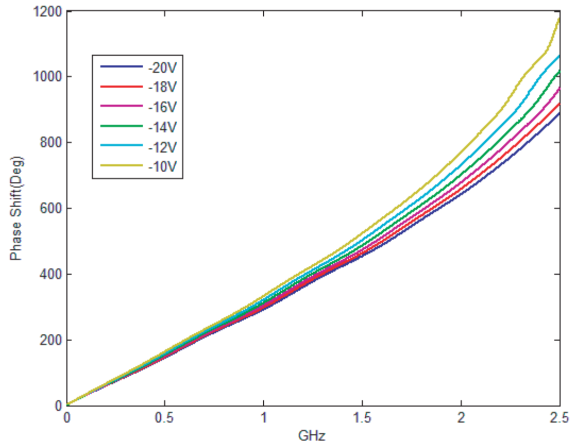


Figure 5: Phase variation over voltage.

5. SIMULATION RESULTS

ADS simulator of Agilent was chosen as circuit simulator. All of the components can be modeled in a circuit simulator. As shown in Figure 4, the return loss S_{11} (Figure 4(b)) of our studied phase shifter is no less than 10 dB up to 2 GHz and the insertion loss S_{21} (Figure 4(a)) is no more than 1 dB.

Figure 5 shows the phase shift versus bias voltage. In our simulation, we can see 1dB of insertion loss at $V_{bias} = -10\,\mathrm{V}$, the phase shift can reach approximately 800° at 2 GHz.

6. CONCLUSIONS

In this work, we have developed a circuit modeling for analogue distributed phase shifter, the measurements of which were published. We used varactors diodes controlled by bias voltages. We have also shown that significant phase shift can be generated using a loaded line phase shifter. The phase shift obtained was linear from 100 MHz up to 1.5 GHz. Then the phase shift has a quadratic variation from 1.5 GHz up to 2.5 GHz. The values obtained by authors of [6] extend up to 800 rad/s according to varactors polarisation voltage. The obtained results are good, but the main drawback is the non linear variation of phase shift as function of frequency due to varactors.

REFERENCES

1. Rebeiz, G. M. and J. B. Muldavin, "RF MEMS switches and switch circuits," *IEEE Microwave Magazine*, 59–71, December 2001.
2. Nagra, A. S. and R. A. York, "Distributed analog phase shifters with low insertion loss," *IEEE Transactions on Microwave Theory and Tequniques*, Vol. 47, No. 9, 1705–1711, September 1999.
3. Case, M. G., "Nonlinear transmission lines for picosecond pulse, impulse and millimeter-wave harmonic generation," Ph.D. Dissertation, University of California Santa Barbara, Santa Barbara, CA, 1993.
4. Popovic, Z. and E. Keuster, "Principles of RF and microwave measurements (lecture notes for ECEN 4634)," 11–13, Electromagnetics Laboratory, University of Colorado, Boulder, CO, 2005.
5. Nagra, A. S., "Varactor loaded transmission lines for linear applications," Ph.D. Comprehensive Exam Presentation, University of California at Santa Barbara, Santa Barbara, CA, 1999.
6. Bourtoutian, R. and P Ferrari, "Tapered distributed analogue tunable phase shifter with low insertion and return loss," *Electronics Letters*, Vol. 41, No. 15, 660–664, July 21, 2005.
7. Collin, R. E., *Foundations for Microwave Engineering*, 2nd Edition, 175–180, McGraw Hill, New York, NY, 1992.
8. Mottonen, V. S., J. Mallat, and A. V. Raisanen, "Characterisation of european millimetre-wave planar diodes," *34th European Microwave Conference (EuMC)*, 921–924, Amsterdam, Holland, October 11–15, 2004.

Development of a Double-clad Fiber Laser Simulator for the Design of Laser Cavities with Specific Applications

D. Mgharaz[1], **A. Boulezhar**[2], **and M. Brunel**[1]

[1]UMR 6614 CORIA CNRS, Université de Rouen, France
[2]Université Hassan II Ain Chock, Casablanca, Morocco

Abstract— We have developed a numerical simulator to design actively Q-switched Yb-doped Double Clad fiber lasers. Based on our simulator, we can design specific fiber laser cavities for various applications: a cavity able to emit a pair of sub-nanosecond pulses separated by more than 500 ns for Particle Imagery Velocimetry applications; a second cavity that can emit long 150 ns pulses exceeding a few millijoules per pulse; a third system composed of coupled-cavities to increase the pulse-energy.

1. INTRODUCTION

In the last years, double-clad (DC) fibers have shown their potentiality for the development of low cost, compact, high power fiber lasers [1]. Many operating regimes have been demonstrated experimentally from Continuous Wave operation to Q-switched, or mode-locked regimes [2]. In parallel, numerical simulators have been developed to describe quantitatively the behaviors observed experimentally [3–5]. It is now possible to simulate precisely DC fiber laser cavities and to design cavities for specific applications. We are particularly interested in the development of Q-switched DC fiber lasers that can emit a pair of nanosecond pulses separated by more than 500 ns. The application of such lasers is the Particle Image Velocimetry (PIV) technique. PIV is commonly involved in flow measurements. A second domain of application concerns the emission of long nanosecond pulses. The peak power of such pulses remains relatively low while their energy is important. Industrial and scientific applications are coherent anti-Stokes Raman spectroscopy, emission spectroscopy of laser ablation, texturing and coloring surface in the domain of metal surface treatments.

We have developed a numerical simulator to design actively Q-switched Yb-doped Double Clad fiber lasers [6, 7]. Based on our simulator, we can design specific fiber laser cavities for various applications: the first example that will be presented is a cavity able to emit a pair of sub-nanosecond pulses separated by more than 500 ns for Particle Imagery Velocimetry applications. The second cavity designed allows to emit long 150 ns pulses exceeding a few millijoules per pulse. Applications concern in this case materials science and combustion. In both cases, the rise time of the Q-switch modulator is an essential parameter.

In a second part, we show that our simulator can be used to describe laser combining. High efficiency coherent combining of CW fiber lasers has been demonstrated in the last years [8]. The combining method can be based on the use of a multi-arm resonator in an interferometric configuration. Michelson and Mach-Zehnder type resonators have been successfully used to reach nearly 100% combining efficiency with two fiber lasers. This concept, which is adapted to the use of double clad doped fibers, has brought some novel perspectives for scaling the output power of the CW monomode fiber lasers. In a similar way, the power rising of an actively Q-switched erbium-doped fiber laser by using two coupled cavities with amplifying fibers could be demonstrated. The pulse peak power obtained could be 1.7 higher than in the case of a unique laser. This concept brings some novel perspectives for scaling the output peak power of single mode Q-switched fiber lasers, where the number of industrial applications is particularly important. In the last section, we present the development of a numerical simulator that can predict precisely the output pulses emitted by a laser system composed of two coupled Q-switched fiber laser cavities.

2. THEORETICAL MODEL

The fiber laser that will be considered is described schematically in Figure 1. The laser medium is an Yb-doped Double-Clad fiber, noted YDCF. The YDCF is pumped with a laser diode. Other elements are an electro-optic modulator (EOM), an optical isolator, a 90/10 output coupler and a wavelength division multiplexer (WDM). The laser diode pump light is injected into the cavity through the WDM. Q-switching is ensured by the EOM. Unidirectional oscillation is obtained with

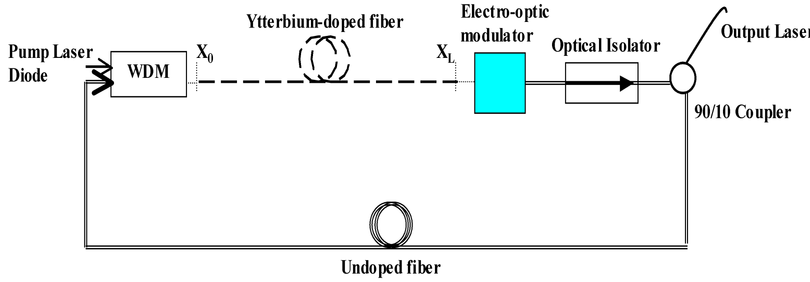


Figure 1: Set-up of the fiber laser cavity.

the optical isolator. This configuration eliminates backward Stimulated Brillouin Scattering that could modify dramatically the pulse emission [2]. An un-doped optical fiber is inserted to adjust the ISL of the cavity. The output pulses are extracted with the 90/10 coupler.

The modeling of Q-switched lasers can be done using the traveling-wave model [4]. It has been applied to different pulsed solid-state lasers and amplifiers, and led to great success in explanations of experimental observations, and optimization of systems. The amplifying medium and the pump and signal powers are described versus time, along the fiber, using the rate equations. The Yb ions are described by two level-atoms. The total ytterbium density is assumed to be uniform in the doped fiber. We resolve numerically the equations governing the evolution of the level populations and of the intensities of the pump and signal lights along the fiber. We neglect the chromatic dispersion effect. The doped and undoped fibers are further characterized by attenuation constants for the pump and signal lights. At time $t < 0$, the pump power is applied to the laser but the EOM is off. The EOM is then turned on and Q-switching can occur. When the EOM is switched on, its transmission passes linearly from 0 to 95%. The EOM opening time is called τ_0. In a next step, the EOM is switched off. This procedure is then repeated at a given repetition rate to obtain the final form of the pulse emitted [4, 6, 7].

3. EMISSION OF A PAIR OF SUB-NANOSECOND PULSES FOR PIV APPLICATIONS

Our first application concerns the emission of a pair of subnanosecond pulses for Particle Image Velocimetry applications. The emission of pulses composed of multiple peaks has been studied in detail in recent works (see for example [5]). Our work consists in the optimization of the different parameters of the cavity, to favour the emission of a pair of peaks. After several optimization steps, we consider the following configuration: the undoped fiber length is 80 m. This fiber allows to increase the temporal interval between the multiple peaks without increasing the pulse duration of each peak. The total cavity length is 110 m. The EOM has a short rise time of 10 ns to generate short pulses, and an opening time of 1.7 µs. Figure 2 shows the output pulses emitted by our fiber ring cavity at a 30 kHz-repetition frequency. Figure 2 shows clearly the emission of a pair of pulses separated by approximately 530 ns (i.e., the round trip time of our 110 m-ring cavity) [6]. The FWHM and the energy of each pulse are 190 ps and 0.15 mJ for the first pulse, 120 ps and 0.14 mJ for the second pulse. The two pulses are not rigorously identical. Note however that the energy difference between them is only 6%. This results show the possibility to design a Q-switched DC fiber laser for PIV applications. The main advantage of this system is that there is only one cavity (and not 2 as in Nd:YAG PIV systems). The two pulses are thus systematically spatially aligned. Combined to the low-cost of our device, it offers an important range of applications to DC fiber lasers.

4. EMISSION OF LONG ENERGETIC PULSES

In this section, we consider another case, i.e., the emission of long energetic pulses for applications in spectroscopy and materials science. We consider the ring cavity of Figure 1. The length of the Yb-doped fiber is 3.5 m, and there is no undoped fiber in this case. The Yb-dopant concentration is $2 \cdot 10^{25}\,\mathrm{m}^{-3}$. The fiber is pumped by a 14 W laser diode at 940 nm. The EOM parameters are: a rise time of 380 ns, and an opening time of 5 µs. 10% of the signal that propagates in the cavity is extracted through a 90/10 coupler. Figure 3 shows the pulse predicted in this case. We obtain a long smoothed envelope. The pulse length is 150 ns at full width half time, and the pulse energy is 4.8 mJ. It is thus possible to develop Q-switched DC fiber lasers that can emit 150 ns pulses

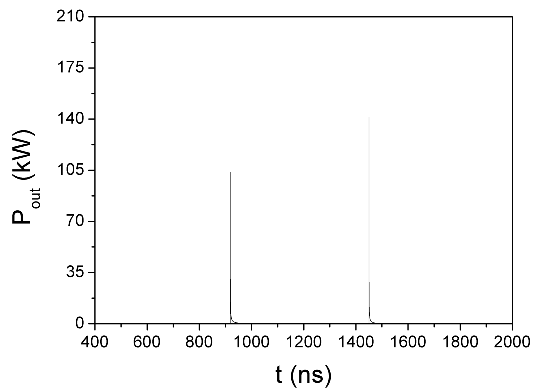


Figure 2: Emission of a pair of sub-nanosecond pulses.

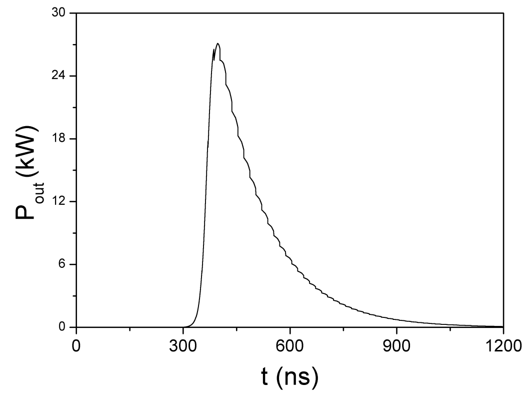


Figure 3: Emission of a long nanosecond pulse.

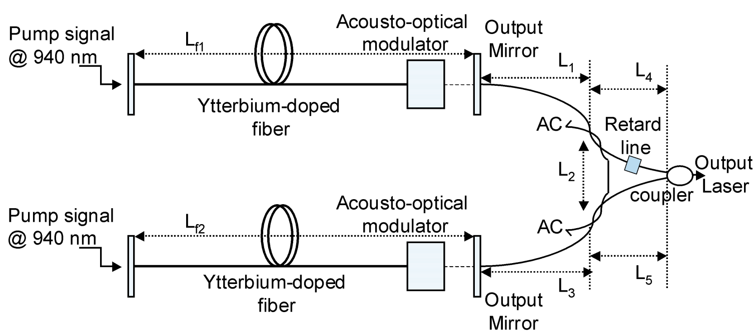


Figure 4: Set-up for coupled-cavities.

exceeding the millijoule per pulse.

5. MODELING OF COUPLED CAVITIES

In this last section we show that it is now possible to use our simulator to develop a modeling of combined cavities. High efficiency coherent combining of CW fiber lasers has been demonstrated in the last years [8]. The laser configuration that will be considered is presented in Figure 4. It is composed of two elementary cavities numbered 1 and 2. Each cavity consists of a laser medium (an YDCF), an acousto-optic modulator (noted AOM), and two mirrors of reflectivities R_1 on the left-hand side and R_2 on the right-hand side. The Yb-concentration into the fiber-core is $4 * 10^{25}\,\mathrm{m}^{-3}$. The core diameter of the YDF is 16 μm, while the inner-cladding diameter is 400 μm. Other parameters describing the YDCF correspond to those of classical double-clad fibers. The pump power is injected on the left-hand side of the cavity through dichroic mirror. Both cavities are mutually injected through two identical unbalanced fiber couplers with a 70:30 splitting ratio (resp. 60/40) The 70% (resp. 60) ports are spliced to the two branches of the elementary lasers, whereas the 30% (resp. 40) ports of both couplers are connected to ensure the radiation transfer between lasers 1 and 2. The unused coupler ports are angle cleaved to avoid any parasitic feedback. Pump powers for both cavities are $P_{p1} = P_{p2} = 10.5\,\mathrm{W}$. The length of the amplifying fibers are $L_{f1} = L_{f2} = 5\,\mathrm{m}$. The lengths of the intermediate fibers between the different couplers are $L_1 = 3\,\mathrm{m}$, $L_2 = 1\,\mathrm{m}$, $L_3 = 3\,\mathrm{m}$, $L_4 = 1\,\mathrm{m}$ and $L_5 = 1\,\mathrm{m}$. We adjust the rise time of the modulators to the value $\tau_r = 1\,\mu\mathrm{s}$. This slow opening time of the AOMs allows to generate long smoothed pulses.

Figure 5 shows the pulses predicted using a sole cavity (solid line), or the bi-cavity system using 70/30 couplers (dotted line) or 60/40 couplers (dashed line). The different pulse shapes are very similar. The FWHM pulse duration is approximately 120 ns is all three cases. However the pulse energies are different. It is only 0.16 mJ using only one cavity wile it equals 0.22 mJ using the 60/40 couplers and reaches 0.24 mJ using 70/30 couplers.

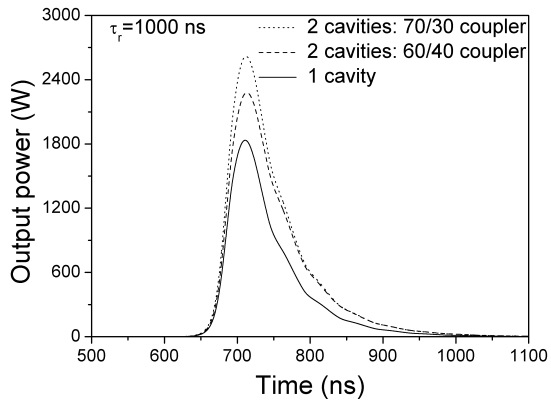


Figure 5: Comparison between the pulses emitted by a sole cavity and a pair of coupled-cavities.

6. CONCLUSION

In conclusion, we have developed a numerical simulator to design actively Q-switched Yb-doped Double Clad fiber lasers. Based on our simulator, we can design specific fiber laser cavities for various applications: a cavity able to emit a pair of sub-nanosecond pulses separated by more than 500 ns for Particle Imagery Velocimetry applications; a second cavity that can emit long 150 ns pulses exceeding a few millijoules per pulse; a third system is based on coupled-cavities to increase the pulse-energy.

REFERENCES

1. Kelson, I. and A. A. Hardy, "Optimization of strongly pumped fiber lasers," *J. Lightwave Technol.*, Vol. 17, 891–897, 1999.
2. Hideur, A., T. Chartier, M. Brunel, M. Salhi, C. Özkul, and F. Sanchez, "Mode-lock, Q-switch and CW operation of an Yb-doped double clad fiber ring laser," *Opt. Commun.*, Vol. 198, 141–146, 2001.
3. Wang, Y. and C. Q. Xu, "Modeling and optimization of Q-switched double-clad fiber lasers," *Appl. Opt.*, Vol. 45, 2058–2071, 2006.
4. Huo, Y., R. T. Brown, G. G. King, and P. K. Cheo, "Kinetic modeling of Q-switched high-power ytterbium-doped fiber lasers," *Appl. Opt.*, Vol. 43, 1404–1411, 2003.
5. Adachi, S. and Y. Koyamada, "Analysis and design of Q-switched erbium-doped fiber lasers and their application to OTDR," *Appl. Opt.*, Vol. 43, 1506–1511, 2002.
6. Mgharaz, D., M. Brunel, and A. Boulezhar, "Design of bi-fiber laser for particle image velocimetry applications," *Open Optics J.*, Vol. 3, 56–62, 2009.
7. Mgharaz, D.,N. Rouchdi, A. Boulezhar, and M. Brunel, "Double-clad fiber laser design for particle image velocimetry and material science applications," *Opt. Las. Eng.*, Vol. 49, 1–7, 2011.
8. Sabourdy, D., V. Kermène, A. Desfarges-Berthelemot, L. Lefort, A. Barthélémy, P. Even, and D. Pureur, "Efficient coherent combining of widely tunable fiber lasers," *Opt. Exp.*, Vol. 11, No. 2, 87–97, 2003.

Maize Crop Yield Map Production and Update Using Remote Sensing

Jesus Soria-Ruiz[1] and Yolanda Fernandez-Ordonez[2]

[1]National Institute of Research for Forestry, Agriculture and Livestock (INIFAP), Mexico
[2]Colegio de Postgraduados, Montecillo, México

Abstract— Crop yield mapping is an important endeavor for agricultural policy making in Mexico where it is a main staple crop. In central Mexico, maize is cultivated under different technological regimes ranging from traditional rain water dependency and using native seeds up to irrigation and improved seeds regimes. Yield variation is in the range of 1.0 ton/ha to over 12.0 ton/ha under the different regimes. It is necessary to explain the increase in average state yield for this crop in the past 10 years in view of a notorious decline tendency in cultivated and harvested area. Precision farming is a system of advanced technologies and procedures which merge spatial mapping variables of the terrain and surrounding conditions with specific management actions for crops. PF requires the integration of several basic component systems such as global positioning system (GPS), data collection and processing devices based on remote sensing and geographical information management systems. Measurements provided by these systems are oriented towards assessing terrain characteristics and spatial variability and can help to locate the better areas to orient management actions to best practices.

1. INTRODUCTION

During the last decade an increase in maize production has been observed in Mexico. Maize is the main staple crop in this country. The production increased from 26.2 millions ton in 2007 to 28.7 millions ton in 2008. So, the maize production does not meet the internal demand and annual imports are required, around 5 million tons are imported from the USA. In central Mexico, specifically in the State of Mexico maize is cultivated under different technological regimes ranging from traditional rain water dependency and native seeds producing yields of under 1.0 ton/ha, up to irrigation and improved seeds regimes with yields above 12.0 ton/ha. The average state yield harvested area for this crop in the past 10 years has been 549,000 ha with an average yield of 3.22 tons/ha and an overall average of 1.8 million tons of grain [1]. On the other hand there is a notorious decline tendency in cultivated area.

Precision farming (PF) is a system of advanced technologies and procedures which merge spatial mapping variables of the terrain and surrounding conditions with specific management actions for crops. PF requires the integration of several basic component systems such as global positioning system (GPS), data collection and processing devices based on remote sensing and geographical information management systems. Measurements provided by these systems are oriented towards assessing the terrain characteristics and spatial variability and orienting the management actions to the best practice in manner, time and place.

Crop monitoring involves estimating crop yield within a geographic area and over a certain time and producing final maps depicting with varied colors and tones the expected yield ranges of different areas for the crop. These results can then be subjected to post-harvest accuracy assessment to produce the maps that are distributed to government agencies, producer associations and research institutions. Crop monitoring and evaluation towards PF objectives include acquisition, analysis and synthesis of crop yield data and their precise location within the areas of interest using satellite images and specialized software. In developing countries the official agricultural statistics such as cultivated areas, yield and production volumes lack credibility due to the shortcomings of traditional on the field sampling and producer survey methods which are employed. The objective of this study was to produce yield maps with geo-statistics and geographic information systems techniques with data from a yield database, which was obtained by continued sampling of real producers' plots over 6 agricultural cycles.

2. MATERIAL AND METHODS

2.1. Crops Monitoring

To this end this six year study aimed at generating a methodology to obtain yield maps of maize crops (*Zea mays*) from SPOT satellite images and geo-referencing of pilot plots with PDA-GPS and other equipments for crop monitoring and yield assessment on the field. The study was carried out over the period 2004–2009.

2.2. Satellite Images

To obtain the cultivated area, SPOT panchromatic and multispectral satellite images were processed over the growth and development stages of the plants. In the cultivated areas sample yield data were collected, geo-referencing the collection sites with geographic information management products [2].

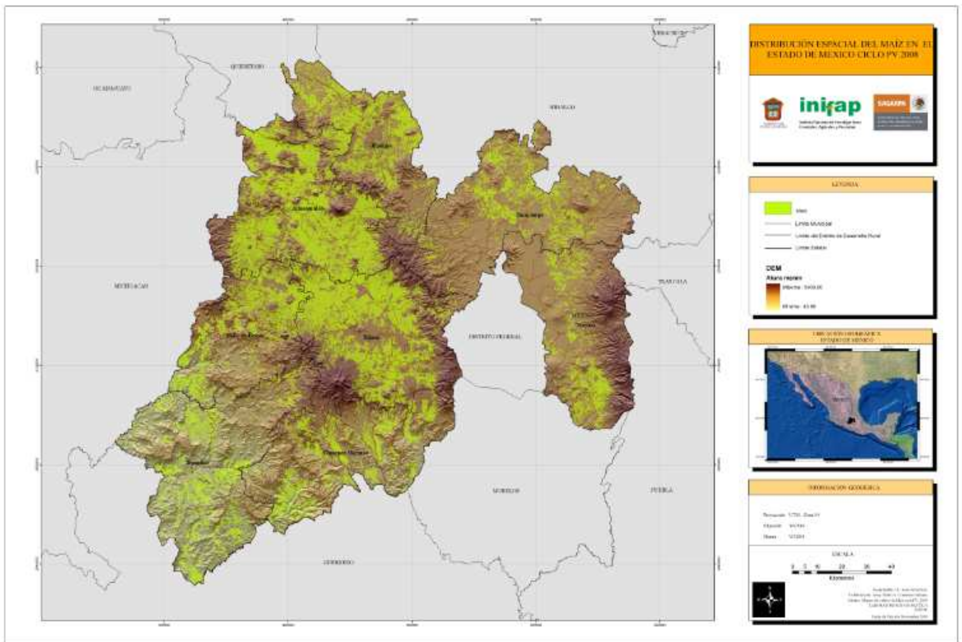


Figure 1: Cultivated surface of maize in the State of Mexico.

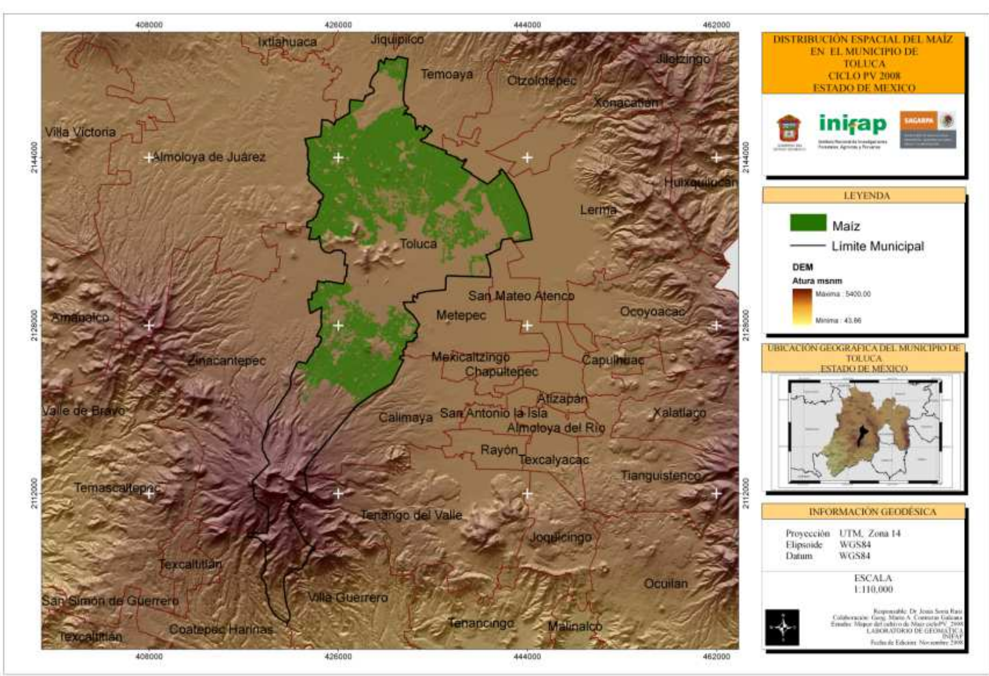


Figure 2: Cultivated surface of maize at municipality level.

2.3. Interpolation

These data were spatially represented via interpolation using the inverse weighted distance method [3]. The mathematical expression of the inverse distance weighting (IDW) interpolation method is expressed as follows [4]:

$$\hat{\beta} * \partial x_o f = \frac{\sum_{i=1}^{n} \frac{1}{d_i^p} \beta \partial x_i f}{\sum_{i=1}^{n} \frac{1}{d_i^p}}$$

where $\hat{\beta} * \partial x_o f$ is the estimated value for a non-sampled site at location X_0, and $\beta(x_i)$ is the observed value at location X_i; d_i is the distances of each of the observed sites to the non-sampled point; p is the value of the exponent of the distance; n is the number of sampled sites.

The final products were yield maps at different cartographic scales. A summary of the methodology is graphically shown in Figure 1 [5].

3. RESULTS AND DISCUSSION

3.1. Cultivated Areas of Maize Crops

A yield map graphically represents the collected data resulting from the direct field monitoring for which we used in first instance information layers of the spatial distribution of the crop derived from the remote sensing data analysis, the Digital Elevation Model (DEM) and terrain slope information, as well as 2400 geo-referenced sampling points providing yield data from 2004 through 2009. The cultivated surfaces of maize at state and municipality level are shown in Figures 1 and 2.

3.2. Yields Maps

125 yield maps were obtained at the municipal level, eight regional maps at rural development district (DDR) level and a state yield map. The 14 yield ranges are separated into three levels: less than 3.0 ton/ha, which are considered low yield; from 4.0 to 6.0 ton/ha, which are considered as medium yield; and over 7.0 ton/ha, which are labeled as high yield.

The yields maps of maize at state and municipality level, shown in Figures 3 and 4.

Yield maps are useful as quick references with a 91% degree of accuracy since the information source is backed by a data base of historic yields collected over six full agricultural cycles. This information source supports decision making for three levels of government in Mexico (federal,

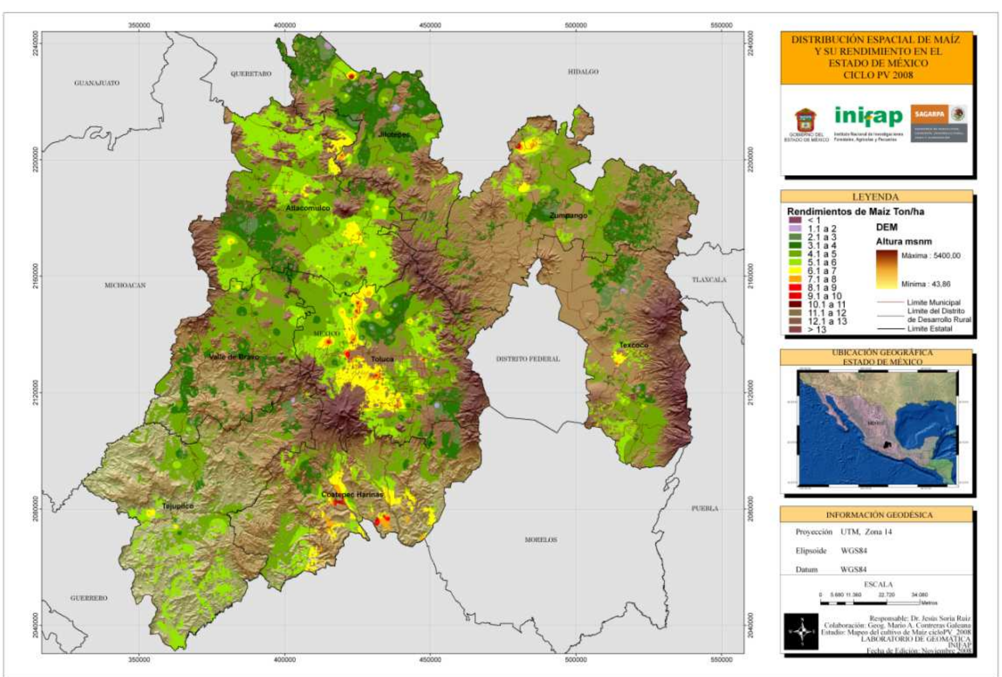


Figure 3: Yield map of maize in the State of Mexico.

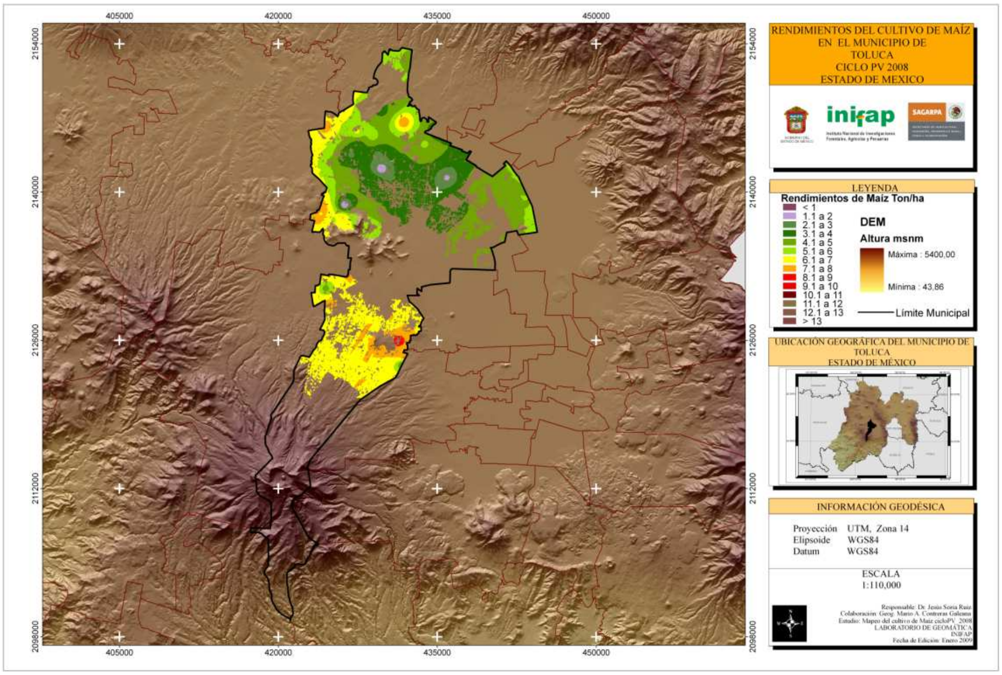


Figure 4: Yield maps of maize atmunicipal level.

state and municipal) mainly for the aid programs to the rural sector and more specifically for maize producers [6].

4. CONCLUSIONS

The inverse distance weighting (IDW) interpolation method permits the analyses of yield from geo-referenced data from on the field sampling. Yield analysis is related to the number and quality of the field samples taken.

With yield maps obtained from spatial modeling and the overlay of other thematic layers related to climate, soil and terrain topography the simulation of dynamic processes is feasible. At the same time homogeneous response areas where maize exhibits its maximum production and productivity can be delimited.

The yield maps obtained by this study and shown here exhibit an accuracy of over 90% due to the information source: a database of geo-referenced yield data which have been collected continuously over six agricultural cycles.

REFERENCES

1. Soria-Ruiz, J., Y. Fernandez-Ordonez, H. McNairn, and J. Bugden-Storie, "Corn monitoring and crop yield using optical and Radarsat-2 images," *Geoscience and Remote Sensing Symposium 2007, IGARSS 2007*, Barcelona, Spain, 2007, ISBN: 978-1-4244-1211-2.
2. ESRI, ArcMap$^{\mathrm{TM}}$ 9.0, License Type: Arc Info Copyright ©1999–2004 ESRI Inc. 2004.
3. Mitas, L. and H. Mitasova, "Spatial interpolation," *Geographical Information Systems*, P. Longley, M. F. Goodchild, D. J. Maguire, D.W. Rhind (Eds.), 481–492, 1999.
4. Flores, G. G., "Modeling the spatial variability of forest fuel arrays," 201, Ph.D. Dissertation, Dept. of Forest Sciences, Colorado State University, USA, 2001.
5. Soria-Ruiz, J., "Mapeo de rendimientos de maíz en el Estado de México, informe final," Laboratorio de Geomatica, INIFAP, Toluca, México, 267, 2008 (in Spanish).
6. Soria-Ruiz, J., Y. Fernandez-Ordonez, and H. McNairn, "Corn monitoring and crop yield using optical and microwave remote sensing," *Geoscience and Remote Sensing*, Ho, P. P. (Ed.), 405–419, In-Tech, Vukovar, Croatia, 2009, ISBN: 978-953-307-003-2.

Adaptive RF Power Amplifier Tuned with Ferroelectric BST Varactor

Yulan Zhang and T. S. Kalkur
Microelectronics Research Laboratories
Department of Electrical and Computer Engineering
University of Colorado at Colorado Springs
CO 80933-7150, USA

Abstract— Recent advancements in wireless communication systems necessitate improvements in system functionality and performance with reduced cost and size. Present communication systems use fixed band width amplifiers, antennas and filters which increases the need for large RF components. The development of tunable capacitors using ferroelectrics gives us the opportunity to develop tunable RF blacks. This paper presents the design and characterization of an adaptive matching network which is made of thin-film barium-strontium-titanate ($Ba_{0.7}Sr_{0.3}TiO_3$ (BST)) for a class A RF amplifier. A 14.479 dB gain has been achieved with both input and output matching networks. Maximum output power load performance is improved compared to a fixed-impedance amplifier. Control bias variation of 4 V results in about 5.8 dB gain difference at 900 MHz. It provides the adaptation for frequency bands of 600, 700, 800 and 900 MHz and different power levels.

1. INTRODUCTION

Currently, RF front-end applications support various standards and frequency bands. The replacement of fixed frequency band circuits with reconfigurable circuits will save space and cost. BST thin films as higher-capacitance density varactors can be used in a variety of adaptive RF building blocks such as tunable matching networks, duplexers, voltage-controlled oscillators and baluns [1–6]. BST thin films have the advantages of high permittivity, tunability, and relatively lower loss than other materials. In order to transfer maximum power to a load, RF amplifiers need matching networks. Ferroelectric capacitors based on materials like Barium Strontium Titanate (BST) offer the required flexibility. By changing the applied voltage, the band width of the BST capacitors can be changed significantly and thus the capacitance changed too.

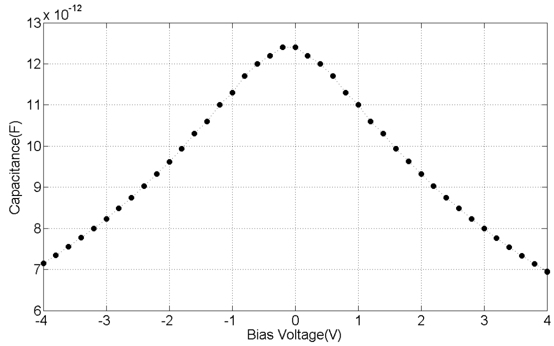
Ferroelectric capacitors in RF applications have been studied and investigated [7–9]. Although power amplifiers with tunable matching network have been designed with ferroelectric capacitors, these analyses have not been investigated in detail. Also, they only consider output matching networks without discussing input matching networks [1–6]. In this paper, we will discuss the impact of input and output matching networks on improving output power and efficiency has been discussed. The simulated power amplifier characteristics have been compared with the fabricated amplifier on RF PCB board.

2. ANALYSIS

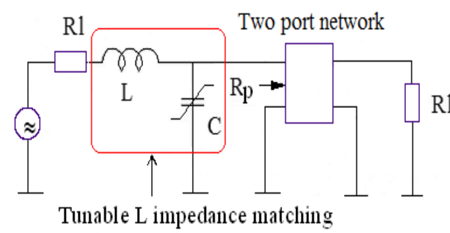
The currently available BST capacitor in Figure 1(a) has tuning range of 6.95 pF to 12.4 pF when applied voltage is from −4 V to 4 V with fairly symmetric characteristics with respect to the bias voltage of 0 V.

These device characteristics place the following design constraints on the adaptive matching network: the effective capacitance tuning range is Cmax/Cmin < 12.4/6.95 and control voltage is within ±4 V. This is used in the input matching network circuits. In the output matching network circuit, the currently available BST device has a tuning range of 4.28 pF to 7.64 pF when applied voltage is from −4 V to 4 V and fairly symmetric through the 0 V crossing. These device characteristics place the following design constraints on the adaptive matching network: the effective capacitance tuning range is Cmax/Cmin < 7.64/4.28. The C-V characteristic shown in Figure 1(a) was measured using HP 4275 LCR meter at frequency 1 MHz at a signal level of 100 mV. An L matching network is used for its simplicity. It has a low cost and a small size compared to other topologies [10]. Figure 1(b) shows the L matching network for input matching connecting the input port and the amplifier input. Rl is 50 ohms and C is the tunable ferroelectric capacitor. When this resistance is viewed across the capacitor C, it is transformed to an equivalent Rp.

$$\text{If} \quad Q^2 \gg 1, \quad Rp \approx \frac{L}{C \cdot Rl} \tag{1}$$



(a)



(b)

Figure 1: (a) C-V characteristics for BST capacitor in the input matching network circuit. (b) Tunable matching network for input matching.

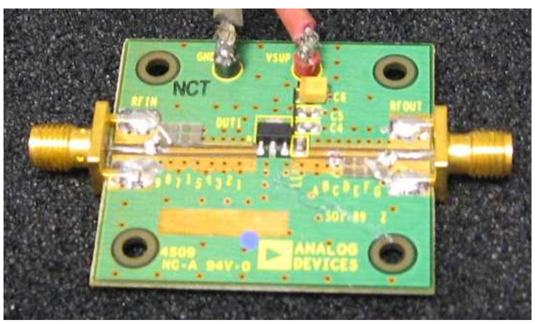


Figure 2: Matching network prototype board.

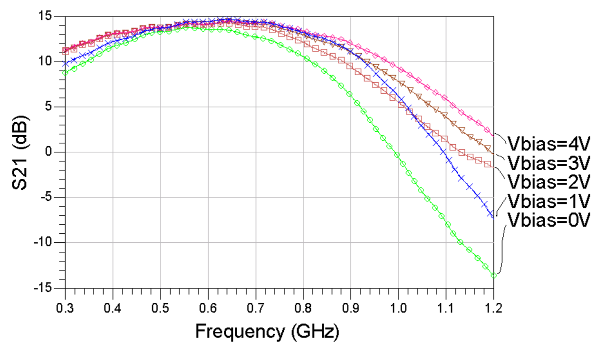


Figure 3: Measured S_{21} versus frequency from 0 V to 4 V bias variations.

The scattering parameter, S_{11} can be given by the equations below:

$$S_{11} = \frac{Rp - Rl}{Rp + Rl} \tag{2}$$

As bias voltage increases, the ferroelectric capacitance decreases. According to Equations (1) and (2), Rp increases and S_{11} (dB) decreases. More power has been transferred into the input of amplifier.

For a class A amplifier, the output power is linear to input power before saturated. C' is the capacitance after tuned. So the change of output power can be expressed as:

$$\Delta P_{out} = \Delta P_{in} = C/C' \tag{3}$$

Similarly, with a tunable output matching network more power can be transferred to the output load.

3. SIMULATIONS, IMPLEMENTATION AND MEASUREMENTS

ADL5320 RF Amplifier has been used in the design. The device can be used in a wide variety of wired and wireless applications, including ISM, WLL, PCS, GSM, CDMA and WCDMA. The ADL5320 operates with 5 V supply voltage with a supply current of 104 mA.

S2P file for this amplifier can be obtained from measurement in network analyzer. The S-parameters of the matching network were measured at different power levels and frequencies at 600, 700, 800 and 900 MHz. Figure 2 shows the matching network and amplifier prototype board. The measured S-parameter with control bias voltage of 0 to 4 V is shown in Figure 3. S_{21} peak shifts from 562.1 MHz of 13.788 dB to 643 MHz of 14.695 dB. Control bias variation of 4 V of input and output matching networks results in about 5.8 dB gain difference at 900 MHz. The unity bandwidth also expands when larger control voltage is applied.

S_{22} have the best performance at $V_{bias} = 3\,\mathrm{V}$ as shown in Figure 4. This reflects that more power has been transferred to the load at the frequency of interest than under other bias conditions.

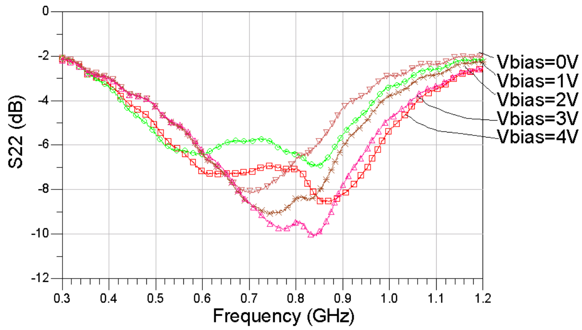


Figure 4: Measured variation of S_{22} with bias voltage.

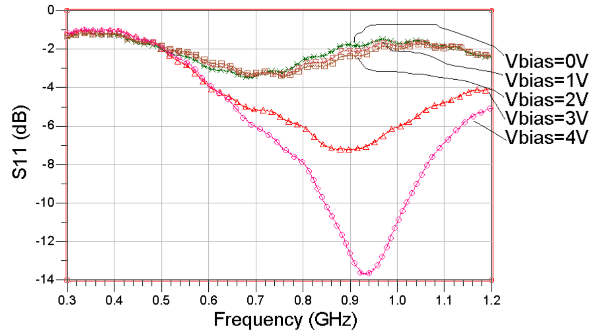
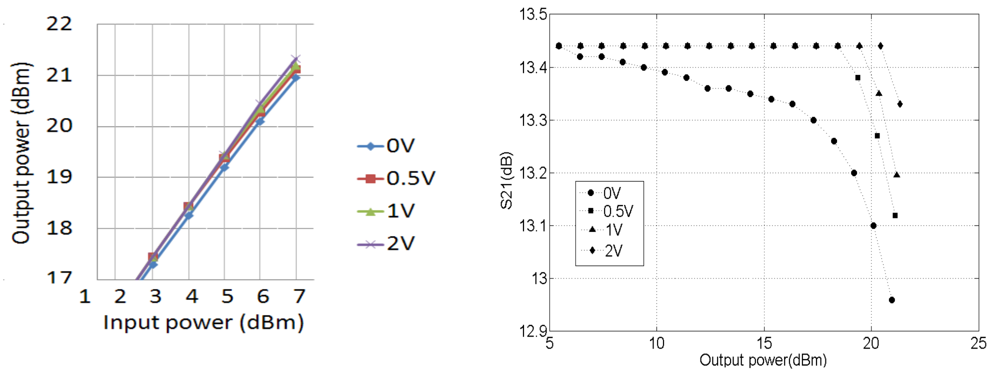


Figure 5: Measured varaition of S_{11} with bias voltage.



(a) (b)

Figure 6: (a) Measured output power versus input power for the different bias voltage at 700 MHz which reflects adaptation for different input power levels. (b) Measured S_{21} versus output power for different bias voltage at 600 MHz which shows adaptation for different power levels.

S_{11} improves when Vbias increases while the ferroelectric capacitance value increases as shown in Figure 5. It provides more power into the input of the amplifier at the frequency of interest.

Figure 6(a) shows the improvement on linearity with adaptive load when bias voltage is varied from 0 to 2V. This results in the shift of maximum output power point as shown in the Figure 6(b). The input capacitance and output capacitance both drop as bias voltage increases as reflected in improved power gain (S_{21}) of the amplifier shown in Figure 6(b). If the adaptive load is applied, the gain of the circuit can be maintained at about 13.44 dB until output power reaches 21 dBm which is significantly higher than without bias voltage. An adaptive load can increase the power gain over a wider range of output power.

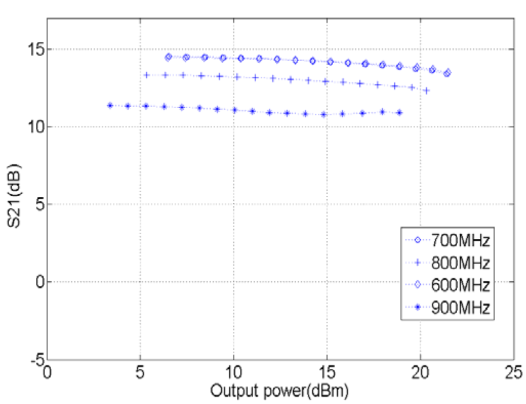


Figure 7: Measured S_{21} versus output power for optimum tuning conditions.

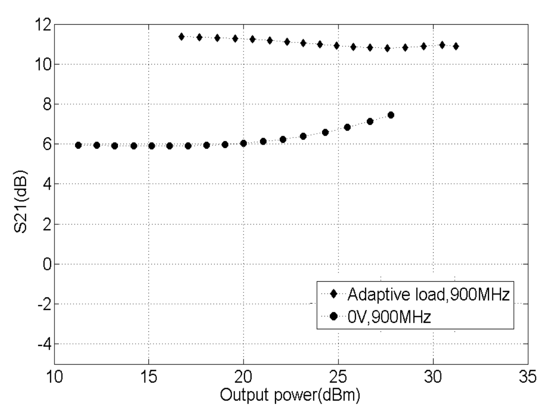


Figure 8: Measured S_{21} versus output power for the optimum tuning conditions for static and adaptive load at 900 MHz.

Figure 7 shows the gain versus output power at 600, 700, 800 and 900 MHz. The results are measured when the amplifier matching network is fixed to its maximum output power settings for each band which shows adaptation for different desired frequencies. This could be used for further applications to provide tunability for different communication systems operating at various frequency bands.

Figure 8 shows the S_{21} improvement of the optimum performance when compared to static tuning conditions. At 900 MHz, 5.7 dB improvement is achieved at output power of 17 dBm; 3.5 dB improvement is achieved at output power of 28 dBm. The maximum effciency reflects improvement of 5% of 600 MHz, 5% of 600 MHz, 7% of 800 MHz and 8% of 800 MHz, which matches the S_{21} improvements.

4. CONCLUSION

It is demonstrated that the RF power amplifier with BST capacitor based matching network can be tuned for bands at 600, 700, 800 and 900 MHz and at different power levels. The measurements show the band-switching and impedance transformation capabilities for input and output matching networks. Ferroelectric capacitance brought in the advantage of control voltage of 4 V limit which requires little DC power. The gain at maximum output power is linear to capacitance change.

REFERENCES

1. Chen, L. Y., R. Forse, D. Chase, and R. A. York, "Analog tunable matching network using integrated thin-film BST capacitor," *IEEE MTT-S Int. Microwave Symp. Dig.*, Vol. 1, 261–264, 2004.
2. Scheele, P., F. Goelden, A. Giere, S. Mueller, and R. Jakoby, "Continuously tunable impedance matching network using ferroelectric varactors," *IEEE MTT-S Int. Microwave Symp. Dig.*, Vol. 1, 4, 2005.
3. Katta, H., H. Kurioka, and Y. Yashima, "Tunable power amplifier using thin-film BST capacitors," *IEEE MTT-S Int. Microwave Symp. Dig.*, Vol. 1, 564–567, 2006.
4. Tombak, A., "A ferroelectric-capacitor-based tunable matching network for quad-band cellular power amplifiers," *IEEE Trans. Microwave Theory Tech.*, Vol. 55, 370–375, 2007.
5. Kabir, A., A. Jamil, Y. Zhang, and T. S. Kalkur, "Voltage controlled oscillators with ferroelectric capacitors," *IEEE Int. Frequency Control Symp.*, Vol. 12, 414–417, 2008.
6. Zhang, Y., T. S. Kalkur, and T. Mulcahy, "Ferroelectric capacitor based adaptive matching network for RF amplifier," *Integrated Ferroelectrics*, Vol. 112, No. 1, 60, 2010.
7. Schmidt, M., E. Lourandakis, A. Leidl, S. Seitz, and R. Weigel, "A comparison of tunable ferroelectric π and T-matching networks," *Proc. 37th Eur Microw. Conf.*, 98–101, Munich, Germany, October 2007.
8. Zhang, Y. and T. S. Kalkur, "Analysis of distortion in tunable ferroelectric capacitors," *IEEE Ultrasonics, Ferroelectrics and Frequency Control Society*, Vol. 57, No. 6, 1263–1266, 2010.
9. Zhu, X., J. Zhu, S. Zhou, Z. Liu, N. Ming, S. Lu, H. L. Chan, and C. Ghoy, "Recent progress of $(Ba,Sr)TiO_3$ thin film of tunable microwave devices," *J. Electron. Materials*, Vol. 32, No. 10, 1125–1134, 2003.
10. Neo, W. C. E., et al., "Adaptive multi-band multi-mode power amplifier using integrated varactor-based tunable matching networks," *IEEE J.Solid-State Circuits*, Vol. 41, 2166–2176, 2006.

Practical Use of the Kramers-Kronig Relation at Microwave Frequencies. Application to Photonic Like Lines and Left Handed Materials

J. Lucas[1], **E. Géron**[1], **T. Ditchi**[2], **and S. Holé**[2]

[1]Laboratoire d'Electricité Général, ESPCI-Paris Tech, France
[2]Laboratoire de Physique et d'Etude des Matériaux, UPMC Univ Paris 06, ESPCI-Paris Tech
CNRS UMR8213, France

Abstract— As early as 1914 Arnold Sommerfeld published two papers in *Annalen der Physik* discussing the question of signal velocity in dispersive media. This work was at this time of pure academic interest. Nowadays photonic crystals, and left handed materials are common objects at microwave frequencies, and the questions raised by Arnold Sommerfeld, and addressed by Léon Brillouin in his famous book of 1960[1] are practically encountered in everyday physics at the laboratory. The famous well known second order differential equation known as the wave equation is indeed not satisfied in dispersive medium such as photonic crystals and left handed materials. Group velocities greater than the speed of light in vacuum and negative phase velocities are encountered respectively in photonic crystals and left handed materials. These phenomenons arise many questions especially insofar the theory of relativity is concerned and promise many applications. A. Sommerfeld stated the problem considering the propagation of a signal terminated on one side, and L. Brillouin carried out the calculus trying to find out a general signal velocity. The results that can be obtained using this methodology which is equivalent to considering the causality of a signal is more widely known particularly in optics through the Kramers-Kronig relation. This relation has many applications and allows to link for instance the imaginary part to the real part of the dielectrical constant of materials for instance. It is also often used to explain the so-called anomalous dispersion phenomenon of light encountered in some materials.

In this work, after a brief presentation of the implication of attenuation on the interpretation of the group velocity, we demonstrate the use of the Kramers-Kronig relation for instance to obtain the dispersion from an attenuation measurement at microwave frequencies in microstrip lines. The measurements obtained using a Vector Network Analyser are compared to those calculated from a single spectrum analyzer measurement on a photonic microwave crystal. We discuss at this point the different nature of the attenuation required by causality, and the one due to the losses in the media and how they interfere with each other. The method is also successfully applied to periodical structures presenting evanescent waves as well as to left handed materials.

1. INTRODUCTION

It is commonly accepted that the group velocity is the speed of energy and information. It is true most of the time. Nevertheless there are numerous reported situations where it appears not to be the case. Though Brillouin and Sommerfield [1] early stated the conditions required for the group velocity to be the velocity of the energy, there are still numerous papers presenting and discussing superluminal group velocities [2–4]. Exotic propagation phenomena called anomalous dispersion at optics frequencies have been reported long ago. Those phenomenon often imply superluminal group velocity such as in photonic crystals. It is seldom stated though probably because it is not easy to perform speed measurements of wave envelopes at optical frequencies and is consequently rarely noticed. Kramers Kronigs relations are commonly used at optics [5, 6] or TeraHertz [7] frequencies to verify the consistency of performed measurements or to extrapolate for instance transmission coefficient measurement at frequencies where they are difficult to perform. This tool which stems from causality considerations should consequently also apply to address superluminal group velocity considerations efficiently. More recently the left handed material lines or structures introduced by Velesago [8] and Pendry [9] came into focus. They are typically media where some anomalous dispersion occurs. The thread binding those phenomenon is the relation which links the phase shift to the amplitude of the signal. When the phase shift is due to propagation one has to consider the dispersion relation along with the amplitude. At lower frequencies when no propagation is involved it is referred to as the frequency response.

[1]Wave Propagation and Group Velocity, Academic Press Inc.

In this work we propose a method to apply the Kramers-Kronig relations even when they cannot be straightforwardly applied. Its applications to S parameter measurements as a causality criterion or for data extrapolation are then theoretically discussed. The method is then practically applied to some interesting typical cases such as photonic crystals, right and left handed media, and media where the waves can be evanescent.

2. THEORETICAL BACKGROUND

2.1. Kramers-Kronig Relations

Whatever the introductory phenomenon presented (photonic crystal, left handed material and so on) they occur in linear and time invariant media. In that case, the media response $r(t)$ to an excitation $e(t)$ is:

$$r(t) = e(t) * h(t) \tag{1}$$

where * denotes the convolution product, and $h(t)$ the response to an impulse. Taking the Fourier transform denoted $\mathcal{F}$ of Equation (1), one obtains:

$$\mathcal{F}(r) = R(\omega) = \mathcal{F}(e) \times \mathcal{F}(h) = E(\omega) \times H(\omega) \tag{2}$$

where ω is the circular frequency. Upper case letters are used to denote Fourier transform of signals. $H(\omega) = \mathcal{F}(h(t))$ is referred to as the frequency response of the system.

In a real thus causal media, when an impulse occurs at $t = 0$, the response of the system must occur at $t \geq 0$. Consequently one has $h(t) = u(t) \times h(t)$ where $u(t)$ denotes the unity step function. With the Fourier transform, one obtains:

$$H(\omega) = -\frac{2j}{\omega} * H(\omega) = j \times \mathcal{H}\big(H(\omega)\big) \tag{3}$$

In this equation, $j = \sqrt{-1}$ and $\mathcal{H}$ is the Hilbert transform.

The relation (3) is known as the Kramers-Kronig (K-K) relation. It is often referred to as the K-K relation pair by presenting separately the imaginary part H_i and the real part H_r of $H(\omega)$:

$$\begin{aligned} H_i(\omega) &= \mathcal{H}(H_r(\omega)) \\ H_r(\omega) &= -\mathcal{H}(H_i(\omega)) \end{aligned} \tag{4}$$

Since Hilbert transform is the convolution product with $-\frac{2}{\omega}$, the numerical Hilbert transform HT can be calculated on a given sampled signal:

$$\mathrm{HT}(S) = \mathrm{IFFT}\Bigg(j \times \mathrm{SGN(S)} \times \mathrm{FFT}(S)\Bigg) \tag{5}$$

In this equation FFT is the Fast Fourier Transform algorithm, IFFT the inverse FFT and SGN(S) the sign function.

One must keep in mind the limitations of both the FFT algorithm and its application to sampled signals. Firstly, the signal which transform is calculated is periodized by the algorithm. Secondly, the resolution of the method depends on number of samples.

2.2. Usages

The K-K relations are of great theoretical importance because it proves that the imaginary part of a frequency response is not independent from its real part. One of the general consequence is that imaginary and real parts of the complex dielectric constant ($\epsilon = \epsilon' + j\epsilon''$) are not independent [10].

Equation (3) can be used as a causality criterion to verify the consistency of a measurement for instance. Considering Equation (3) one can compute:

$$\mathcal{C} = H(\omega) - (j \times \mathcal{H}\big(H(\omega)\big) \tag{6}$$

This criterion must equal to zero over all frequencies to satisfy causality.

Computing a phase shift from an amplitude measurement and vice-versa is the main practical application of the K-K relations. Nevertheless it cannot be applied straightforwardly. Indeed $H(\omega) = G(\omega)e^{j\theta(\omega)}$ is often determined by measuring its amplitude $G(\omega)$ and its argument $\theta(\omega)$. The K-K relation relates $G(\omega)\cos(\theta(\omega))$ to $G(\omega)\sin(\theta(\omega))$. Notice that consequently the causality criterion can be easily computed from those two measurements. But it would be interesting from an experimental point of view to obtain $G(\omega)$ from the measurement of $\theta(\omega)$ or vice-versa especially when one of the measurements is difficult or impossible to perform such as $\theta(\omega)$ when using a spectrum analyzer.

In order to separate $G(\omega)$ and $\theta(\omega)$, it is usual to consider the logarithm of $H(\omega)$. If the K-K relation could be applied, one would obtain the pair of relation:

$$\begin{aligned} &\ln\big(G(\omega)\big) = -\mathcal{H}\big(\theta(\omega)\big) \\ &\theta(\omega) = \mathcal{H}\big(\ln\big(G(\omega)\big) \end{aligned} \tag{7}$$

To apply the K-K relation to $\ln(H(\omega)) = \ln\big(G(\omega)\big) + j\theta(\omega)$, the system whose frequency response is $\ln(H(\omega))$ must be causal.

This is unfortunately not the general case. Let us consider a system consisting in a piece of ideal coaxial cable of length L. For the TEM mode, the group and phase velocities are equal and constant: $V_\varphi = V_g = \text{cste}$. Therefore one has: $H(\omega) = \exp(-j\frac{\omega}{V_\varphi}L)$. $H(\omega)$ corresponds to a causal system since the transit time in such a system is $\frac{L}{V_g} \geq 0$. One can also check that the causality criterion of Equation (6) is verified by $H(\omega)$. Indeed as $\theta = \frac{\omega}{V_\varphi}L$, it comes:

$$\mathcal{C} = e^{-j\theta} - j \times \mathcal{H}(e^{-j\theta}) = 0 \tag{8}$$

Considering $\ln(H(\omega))$, one obtains:

$$\begin{aligned} &\ln\big(G(\omega)\big) = 0 \\ &\theta(\omega) = -\tfrac{\omega}{V_\varphi}L \end{aligned} \tag{9}$$

Consequently $\mathcal{H}(\ln(G(\omega))) = 0$ which is obviously different from $\theta(\omega)$. The K-K relations are not respected. In this case $H(\omega)$ corresponds to a causal system when $\ln(H(\omega))$ does not.

One can find in the literature methods inspired from the direct integration to demonstrate K-K pairs of relations between $\ln(G(\omega))$ and $\theta(\omega)$. For instance Wooten [10] assuming that:

$$|H(z)| < |b \times z^{-s}| \tag{10}$$

where b and s are real strictly positive constants, and z the analytical extension to the complex plane of ω with $\Re(z) = \omega$ and $\Im(z) \leq 0$, proves that:

$$\theta(\omega) = \frac{2}{\pi}\int \frac{\ln\big(G(\omega)\big)}{u^2 - \omega^2}\mathrm{d}u \tag{11}$$

Considering that $G(\omega)$ is a even function, it can be rewritten as $\theta(\omega) = \mathcal{H}(\ln(G(\omega)))$. Finally by taking the Hilbert transform of this equation one has:

$$\begin{aligned} \theta(\omega) &= \mathcal{H}\big(\ln(G(\omega))\big) \\ \ln(G(\omega)) &= -\mathcal{H}\big(\theta(\omega)\big) \end{aligned} \tag{12}$$

In the case of plain propagation, such as the propagation of TEM waves in the piece of cable previously presented, the condition (10) does not apply. Consequently, the pair of relations (12) does not apply nether. Nevertheless, as propagation cannot be neglected when the frequency gets sufficiently high, it is necessary to consider the case.

It is interesting at this point to shift from an ideal coaxial wire model to a more realistic model where the wire presents dielectric losses. These losses result from the time response of the dielectric used in the wire to the electric field. When the losses are sufficiently small the frequency response is modified as follows:

$$H(\omega) = G(\omega) \times \exp(-j\frac{\omega}{V_\varphi}L) = \exp(-\frac{1}{2}\frac{\omega}{C}\sqrt{\epsilon_r}\tan(\delta)L) \times \exp(-j\frac{\omega}{V_\varphi}L). \tag{13}$$

In this equation c is the speed of light in vacuum, ϵ_r the relative dielectric constant and δ the loss angle. It yields:

$$\ln\big(H(\omega)\big) = j\times\omega - \frac{\omega}{\Omega} \tag{14}$$

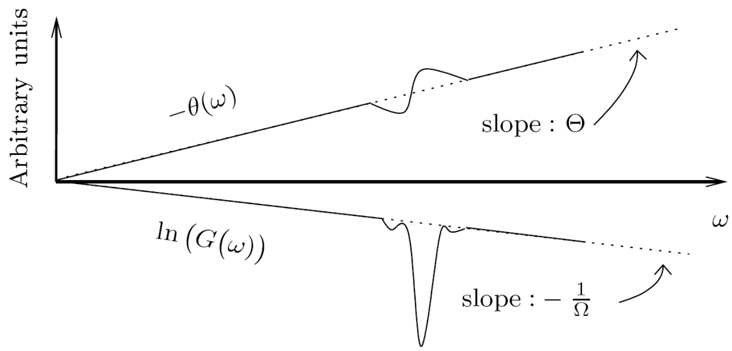


Figure 1: Dispersion and attenuation. Standard propagation in dotted line and anomalous dispersion in solid line.

where $\frac{1}{\Omega} = \frac{1}{2c}\sqrt{\epsilon_r}\tan(\delta)L$ and $\Theta = \frac{L}{V_\varphi}$. These considerations allow to obtain the aspect of the plots presented in Figure 1 for a plain propagation.

In this figure dispersion curves with and without anomalous dispersion are presented. That kind of phenomenon is often reported at optical frequencies. It is the kind of curves that are obtained in photonic crystals for instance.

Finally it is physically interesting and realistic to consider media where the two following limits exist:

$$\begin{cases} \lim\limits_{\omega\to\pm\infty} -\frac{\theta(\omega)}{\omega} = \Theta \\ \lim\limits_{\omega\to\pm\infty} \frac{\ln\big(G(\omega)\big)}{|\omega|} = -\frac{1}{\Omega} \end{cases} \tag{15}$$

In plain words, it means that both $G(\omega)$ and $\theta(\omega)$ are asymptotic to straight lines whose slopes are respectively $-\frac{1}{\Omega}$ and Θ . Using those two limits, One can obtain by integrating in the complex plane using Cauchy's theorem the following new relations:

$$\ln\big(G(\omega)\big) + \frac{|\omega|}{\Omega}\big) = -\mathcal{H}\big(-\theta(\omega) - \omega\Theta\big) \tag{16}$$

$$-\theta(\omega) - \omega\Theta = \mathcal{H}\Big(\ln\big(G(\omega)\big) + \frac{|\omega|}{\Omega}\Big) \tag{17}$$

In order to used these two equations one must have some knowledge of the system studied. Θ or $\frac{1}{\Omega}$ must be known. Note that $\Theta = \lim_{\omega\to+\infty}(-\frac{\theta(\omega)}{\omega})$, is sometime called the front velocity.

Notice the use of $|\omega|$ in Equation (15) in order to turn $\ln(G(\omega))/\omega$ into an even function. It is required to correctly compute the Hilbert transform using the FFT algorithm.

Practically, to compute a phase shift from an amplitude measurement, the measured amplitude must be symmetrized over the negative frequencies as an even function before calculating the first FFT. In the same way, a phase measurement has to be symmetrized as an odd function.

3. APPLICATIONS

The results presented in the precedent section are applied to various cases chosen to cover some of the most typical situations that can be encountered when working at microwave frequencies.

3.1. Photonic Crystal at Microwave Frequencies

A photonic crystal as the one presented in the Figure 2 has been realized. This structure is made of microstrip lines whose characteristic impedance is alternately $Z_0 = 50\,\Omega$ and $Z_1 = 100\,\Omega$. The substrate used is standard FR4 substrate 0.8 mm thick. The length of each segment has been chosen to be equal to a quarter of the wave length for each characteritic impedance at 2 GHz. This way at 2 GHz the wave reflected at the end of each Z_0 line segment interferes destructively with the wave entering the segment thus creating the photonic effect. This is often referred to as the Bragg condition. It can be written for this line $\omega = \omega_0 + 2n\omega_0$. Where ω_0 is the frequency for which the length of each segment is $\frac{\lambda}{4}$ (2 GHz) and n is a positive integer. At these frequencies, the amplitude of the waves is greatly diminished which leads to frequency bands where the transmission coefficient is very low.

3.1.1. Practical Application for Media Where Log(G) and θ Are Asymptotic

The relation linking θ to $\ln(G)$ can be efficiently applied to obtain the argument of S_{21} from its modulus. This is particularly interesting because $|S_{21}|$ can be measured using a spectrum analyzer (SPA) and a generator when a vectorial analyzer (VNA) is required to measure both θ and $\ln(G)$. In order to demonstrate the efficiency of the method we have performed the measurement of $\ln(G) = \ln(|S_{21}|)$ with both instruments. The results are presented Figure 3(a). The results match up to 12 GHz. Beyond this frequency a difference is visible. This difference can probably be accounted for the calibration procedure (simple two ports SOLT) used for the VNA which is certainly not the most adapted for such a large bandwidth. The measurement from the SPA were compensated using a THRU measurement from the generator.

One can verify in Figure 3(a) that $\log(|S_{21}|)$ behaves asymptotically. By measuring this asymptote a "rectified" gain as defined by the left member of Equation (16) can be obtained. This gain has been filled in with zeros for frequencies below the first effectively measured frequency by the analyzer in order to extrapolate the signal down to zero. It has also been symmetrized over the negative frequencies to take into account the considerations presented in Section 2.2. Finally the argument of S_{21} has been calculated using Equation (17). Figure 3(b) presents both the calculated and the measured argument of S_{21}. Of course to compare both results one needs to know the asymptotic behavior of θ as well. To obtain a figure with an interesting zoom factor the arguments are presented without the asymptotic behavior.

The results are identical as long as the amplitudes measured by the VNA and the SPA are indentical which validates the method.

3.2. Left/right Handed Hybrid Media Presenting Evanescent Waves

A left handed media is a media where the phase velocity $V_\varphi = \frac{\omega}{k}$ is negative over a frequency band [8]. A simple model of left handed line is presented in Figure 4. In this model the capacitors and the inductors yield the left handed properties of the line, and the transmission lines take into account the propagation that cannot be neglected at high frequencies, that is to say when the physical length of the line is not small compared to the wavelength. The results of the simulation of such a line are presented in Figure 5.

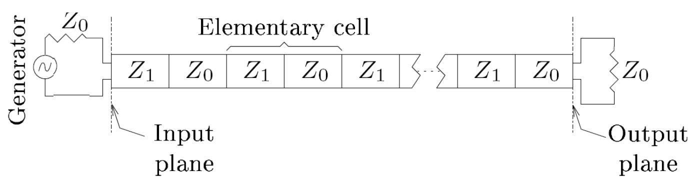


Figure 2: Photonic crystal built up from microstrip lines.

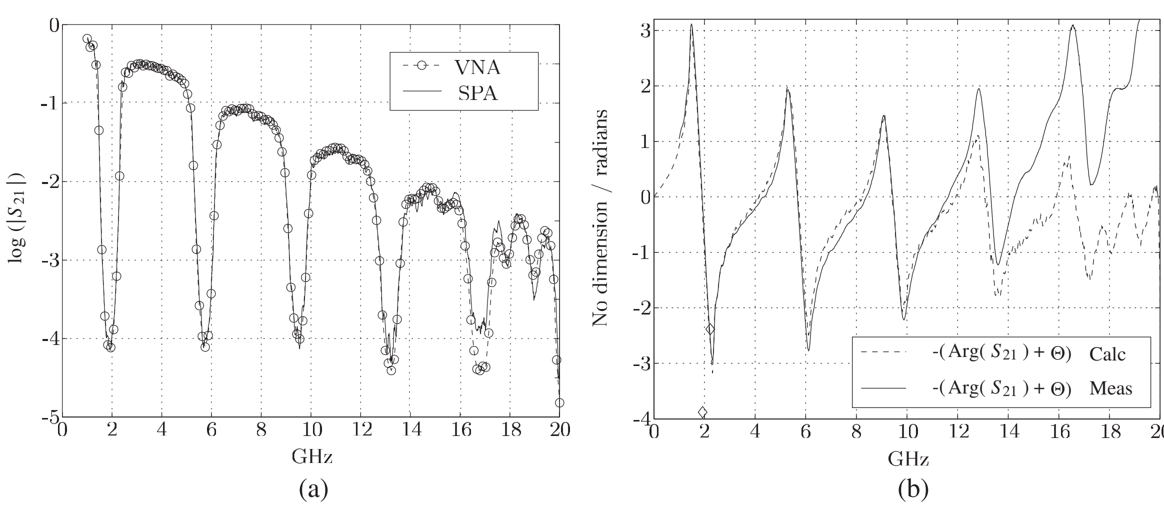


Figure 3: (a) $|S_{21}|$ measured using a Vectorial Network Analyzer(VNA) and a Spectrum Analyzer (SPA). (b) Phase measured using a VNA compared to the phase calculated from a spectrum analyzer measurement using K-K.

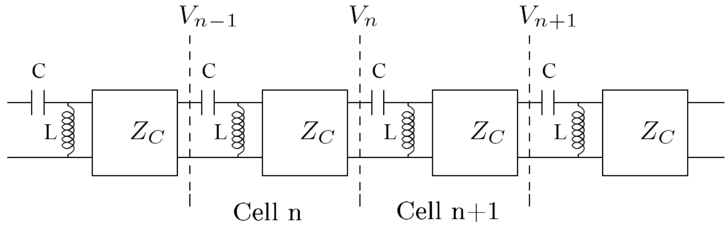


Figure 4: Left handed line model with lumped components and transmission line segments.

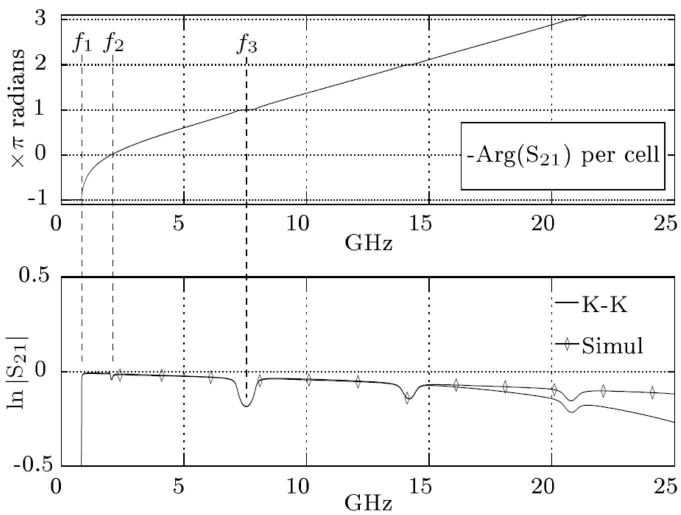


Figure 5: Application of the K-K relations over a simulated hybrid left/right handed line.

At low frequencies, when the effect of line segments are negligible, the capacitor and the inductor induce a phase shift that can be calculated using the Kirchhoff's Laws: $\cos(\theta) = 1 - \frac{1}{LC\omega^2}$. At low frequencies when $\frac{1}{LC\omega^2} \gg 2$, i.e., $f < f_1$ in Figure 5, the waves in the structure are evanescent and no phase shift occurs as long as $\frac{1}{LC\omega^2} > 2$. Consequently the phase shift remains fixed at $-\pi$.

When the frequency gets higher, the capacitance and inductance per unit length of the line can, if chosen correctly, introduce a second frequency band (centered on f_2 in Figure 5) where the waves are evanescent. After this band which occurs for $\arg(S_{21}) = 0$, the line turns back to a standard right handed behavior. At even higher frequencies the propagation occurring in the line segments can no longer be neglected, and a classical propagation phenomenon occurs yielding an asymptotic behavior of the dispersion curve. In addition the line presents regularly spaced forbidden frequency band gaps where the waves are again evanescent. The usage of such lines to build up couplers is presented in [11]. This line also is interesting regarding the application of K-K relations, indeed, the waves are successively evanescent, left-handed, evanescent and right handed.

As $\arg(S_{21})$ is asymptotical, one must apply the K-K relations of Equation (16).

At high frequencies one can see that some discrepancy occurs between the results obtained for the simulated value of $\ln(|S_{21}|)$ and its value calculated using K-K. This can be improved if required by computing the Hilbert transform over a larger frequency span filling in the missing data using the asymptotical behavior.

4. CONCLUSION

In this work, we have firstly reminded the underlying theory that yields the Kramers-Kronig relation. We have proposed a simple way to compute the Hilbert transform that subtend the Kramers-Kronig relation. This method has been successfully applied to all presented examples. We have proposed a criterion that can be used to verify the consistency of measurements or to help analyzing causality considerations.

Though the causality criterion is an interesting tool, its interest is limited to consistency check. It is often required to be able to extrapolate the phase shift from the attenuation or vice versa. In a general way, it is not possible to apply the Kramers-Kronig relation to perform such calculation.

Nevertheless, by means of some hypothesis on the modulus of the frequency response, it is possible to apply the Kramers-Kronig relation on many physical systems. Furthermore we have shown that the case of plain propagation has to be dealt with in a particular way.

We have applied successfully the method to various special cases chosen because of their representatives behaviors. we have verified that method can be applied whether there is propagation or not. In the case of propagation the waves can be evanescent, right or left handed, or be the steady state resulting from interferences as in the case of photonic crystals.

The precision reached for data extrapolation is very good. It greatly depends nevertheless on the knowledge of the asymptotical behavior of the media at high frequencies.

REFERENCES

1. Brillouin. *Wave propagation and group velocity.* Pure and applied physic. ACADEMIC PRESS INC., 111 Fifth avenue New york, third edition, 1960.
2. Haché, A. and L. Poirier, "Anomalous dispersion and superluminal group velocity in a coaxial photonic crystal: Theory and experiment," *Physical Review E*, Vol. 65, 036608, 2002.
3. Chen, K., L.-A. Wu, and Y.-H. Shih, "Is 'superluminal' light propagation possible in dispersive media?," *Chin. Phys. Lett.*, Vol. 21, No. 5, 770, 2004.
4. Ojha, S. P., K. B. Thapa, and S. K. Singh, "Superluminal propagation in plasma photonic band gap materials," *Optics*, Vol. 119, 81–85, 2008.
5. Young, R. H., "Validity of the Kramers-Kronig transformation used in reflection spectroscopy," *J. Optical Society of America*, Vol. 67, No. 4, 520–523, 1977.
6. Gornov, E. and K.-E. Peiponen, "Dispersion relations for testing the validity of linear and nonlinear optical spectra," *Applied Optics*, Vol. 46, No. 24, 6081–6083, 2007.
7. Tuonomen, H., E. Gornov, J. A. Zeitler, J. Aaltonen, and K.-E. Peiponen, "Using modified Kramers Kronig relations to test transmission spectra of porous media in THz TDS," *Optics Letters*, Vol. 35, No. 5, 631–633, 2010.
8. Veselago, V. G., "The Electrodynamics of substances with simultaneous negative values of ϵ and μ," *Soviet Physics Uspekhi*, Vol. 10, No. 4, 509–514, 1968.
9. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 47, No. 11, 2075–2084, 1999.
10. Wooten, F., *Optical Properties of Solids*, Academic Press Inc, 111 fifth Avenue, New York, 1972.
11. Géron, E., T. Ditchi, J. Lucas, and S. Holé, "Asymetric microstrip right/left-handed line coupler with variable coupling ratio," *Progress In Electromagnetics Research Symposium*, Marrakesh, Morocco, March 20–23, 2011.

Coaxial Quasi-elliptic Filter Using a Suspended Resonator and Vertically Stacked Coaxial Lines

Aline Jaimes-Vera[1], **Ignacio Llamas-Garro**[2], **and Alonso Corona-Chavez**[3]

[1]Signal Theory and Communications Department, Technical University of Catalonia
Barcelona 08034, Spain
[2]Centre Tecnologic de Telecomunicacions de Catalunya, CTTC
Barcelona 08860, Spain
[3]National Institute for Astrophysics, Optics and Electronics
Puebla 72840, México

Abstract— In this paper a four pole quasi-elliptic filter at X-band is presented. The filter is composed of two vertically stacked rectangular coaxial lines, where one pair of resonators is placed on the lower coaxial line and another pair is located on the upper line. Coupling between coaxial lines is achieved through an iris in the common coaxial ground plane. The filter has been designed at 9.1 GHz with a 4% fractional bandwidth. Two transmission zeros located at the sides of the passband have been successfully achieved with the proposed filter topology. A measured insertion loss of 1.7 dB has been obtained.

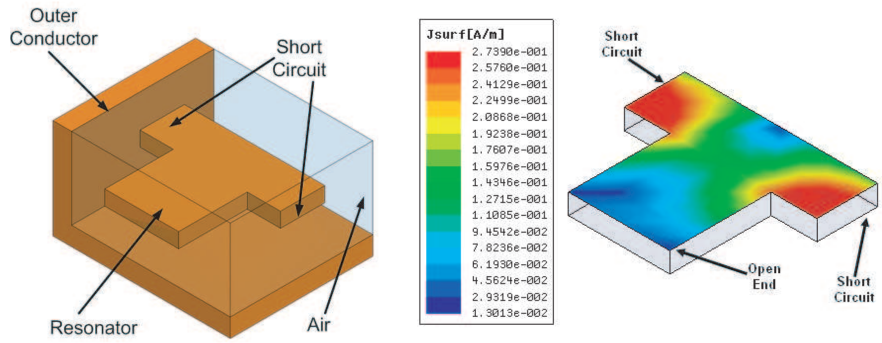
1. INTRODUCTION

The rapid advance in wireless communications has led to the development of high performance filters, with high selectivity and good out of band rejection, including device miniaturization. Introducing transmission zeros on a bandpass filter response can achieve high selectivity with good out of band rejections. Commonly coaxial-combline cavity filters with transmission zeros have been designed by means of coupling probes embedded in the filter fixture [1], also the use of extra cavities or small metal plates among the resonators [2, 3] have been used. The structure presented in this paper uses a suspended resonator in different coupling arrangements to produce electric, magnetic and mixed couplings, enabling quasi-elliptic function approximation or flat group delay responses, instead of using extra coupling probes, extra cavities or folded structures.

The use of narrowband filters between amplifier stages with insertion loss (< 2 dB) can be tolerated, because compact size and low fabrication costs are more important to keep the filter competitive [4]. The proposed structure was designed at 9.1 GHz where only few works can be found, since most of combline and interdigital filters are designed between 1 and 2 GHz. The introduction of cross-couplings has been used to obtain a pair of transmission zeros on the sides of the passband. The unloaded quality factor of a single resonator was measured and has been found to fall in-between the high Q obtained through optimized conventional coaxial-combline designs and interdigital or microstrip-combline filters. The planar implementation of the filter (made by 9 stacked planar metal layers) allows scaling the designs to the millimeter wave frequencies range, using micromachined structures [5].

2. SUSPENDED COAXIAL RESONATOR

This section presents a quarter wavelength resonator suspended by short circuits inside an air filled rectangular coaxial cable, shown in Fig. 1(a). The resonator has the maximum magnetic field density next to the short circuited stubs, and the maximum electric field at the opposite side, the surface current distribution at the resonant frequency of 9.1 GHz is plotted in Fig. 1(b). Electric, magnetic and mixed couplings can be obtained by choosing adequate coupled resonator configurations. By placing the two short circuited sides of the resonator facing each other, a magnetic coupling can be obtained. Electric coupling can be attained by placing the open end of the resonators facing each other. The responses of the electric and magnetic coupling arrangements are out of phase and have been used to produce the quasi-elliptic filter presented in this paper. Mixed couplings have been obtained by placing resonators on different coaxial lines, coupled by an iris on the common coaxial ground plane between them.



(a) (b)

Figure 1: Suspended quarter wavelength resonator. (a) Schematic of the quarter wavelength resonator (side and top walls removed for clarity). (b) Surface current distribution at the resonant frequency of 9.1 GHz.

3. NARROWBAND QUASI-ELLIPTIC COAXIAL FILTER USING VERTICALLY STACKED COAXIAL LINES

This filter was designed using two vertically stacked coaxial cables, as illustrated in Fig. 2, where two resonators are placed on the upper coaxial line, and two others on the lower line, along with the feed lines that provide the input/output to the device. The entire topology is formed by nine conductive layers stacked and compressed together to obtain the two coaxial transmission lines. An iris in the common coaxial ground plane allows the cross coupling arrangement between resonators.

The design procedure for this filter follows the methodology provided in [6], which consists in calculating the coupling coefficients between resonators (K_{ij}) and the external quality factor (Q_e), achieved by full wave simulations using HFSS. The equations to obtain the theoretical couplings between resonators and the external quality factor can be found in [6]. The design data and parameters used for this filter are summarized in Table 1, which contains the lowpass quasi-elliptic element g values, the required K_{ij} and Q_e for the design. The filter was designed at 9.1 GHz with a 0.01 dB passband ripple and a 4% fractional bandwidth, having a quasi-elliptic response.

After obtaining the optimum spacing between resonators and feed lines, the filter can be realized. Overall dimensions of the filter are $29.8 \times 48.7 \times 20\,\mathrm{mm}^3$. Simulation and measurements of the filter are shown in Fig. 3. This simulation assumes the effect of layer misalignment, where 12 simulations

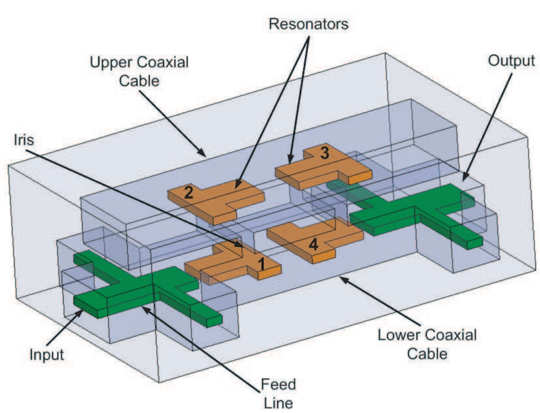


Figure 2: Quasi-elliptic filter topology using vertically adjacent coaxial lines.

Table 1: Quasi-elliptic filter design parameters.

Lowpass filter element g values for $\Omega_d = 2.00$			
$g_1 = 0.95449$	$g_2 = 1.38235$	$J_1 = -0.16271$	$J_2 = 1.06062$
Q_e and K_{ij}			
$Q_e = 23.862$	$K_{12} = K_{34} = 0.0348$	$K_{23} = 0.0307$	$K_{14} = -6.8 \times 10^{-3}$

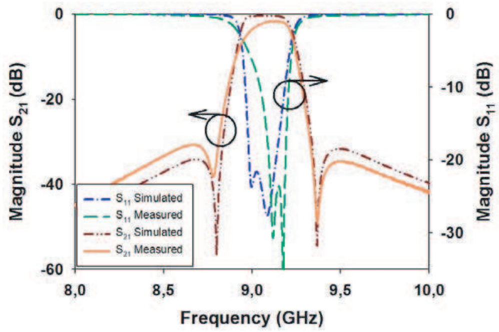


Figure 3: Simulated and measured results for the quasi-elliptic filter implemented on stacked coaxial lines.

have been done displacing the layers that compose the filter arbitrarily using values ranging from 100 to 500 μm, and one of them was arbitrarily selected to be included in Fig. 3. The 12 simulations showed a reduction in filter bandwidth due to the change in coupling coefficients caused by the overlapping of air spacing layers with the main central coaxial conductor. This causes a reduction in coupling coefficients and results in narrower bandwidths. The variations between simulation and measurement results can be attributed to the misalignment of the stacked coaxial layers, fabrication tolerances and the effect of using eight brass tuning screws to obtain the measured filter response. The measured bandwidth decreased from 4% considered for the design to 2.88% experimentally, the simulation with misaligned layers showed a bandwidth of 3.13%. The transmission zero on each side of the passband was successfully achieved using the proposed vertically integrated coaxial filter.

4. CONCLUSIONS

A new type of narrowband quasi-elliptic filter using vertically stacked coaxial lines has been presented. The filter uses cross-couplings to produce a quasi-elliptic response with a pair of transmission zeros, using an inline suspended resonator. The filter implementation using planar machined metal layers allows scaling the design to millimeter wave frequencies using micromachined techniques.

ACKNOWLEDGMENT

This work was supported by research project PIB2010BZ-00585 from the Spanish ministry of science and innovation. Jaimes V. Aline wishes to thank CONACYT for scholarship No. 198264°. All authors wish to thank Nahu Pérez Pérez at the Aspheric Surfaces Laboratory, Large Millimeter Telescope at the National Institute for Astrophysics, Optics and Electronics, Mexico for machining the prototype filter. All authors wish to thank Joaquim Giner at the Signal Theory and Communications Department, Technical University of Catalonia, Barcelona, Spain, for polishing and assembling the device.

REFERENCES

1. Wang, C. and K. A. Zaki, "Full-wave modeling of electric coupling probes in comb-line resonators and filters," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 48, No. 12, 2459–2464, Dec. 2000.
2. Cogollos, S., R. J. Cameron, R. R. Mansour, M. Yu, and V. E. Boria, "Synthesis and design procedure for high performance waveguide filters based on nonresonanting nodes," *IEEE MTT-S International Microwave Symposium Digest*, 1297–1300, Honololu, Jun. 2007.
3. Wang, Y. and M. Yu, "True inline cross-coupled coaxial cavity filters," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 57, No. 12, 2958–2965, Dec. 2009.
4. Wang, Z., Q. Wang, and Y. Zhai, "Design of an economical compact combline filter," *International Conference on Microwave and Millimeter Wave Technology*, Vol. 1, 308–310, 2008.
5. Lancaster, M. J., J. Zhou, M. Ke, Y. Wang, and K. Jiang, "Design and high performance of a micromachined K-band rectangular coaxial cable," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 55, No. 7, 1548–1553, Jul. 2007.
6. Hong J.-S. and M. J. Lancaster, *Microstrip Filters for RF/Microwave Applications*, John Wiley and Sons Inc, 2001.

Asymmetric Microstrip Right/Left-handed Line Coupler with Variable Coupling Ratio

E. Géron[1], **T. Ditchi**[2], **J. Lucas**[1], **and S. Holé**[2]

[1]Laboratoire d'Electricité Général, ESPCI-ParisTech, France
[2]Laboratoire de Physique et d'Etude des Matériaux, UPMC Univ Paris 06, ESPCI-ParisTech
CNRS UMR 8213, 10, rue Vauquelin, Paris 75005, France

Abstract— This work presents a microstrip coupler with a variable couling ratio taking advantage of structures presenting both LH and RH behaviour. After some considerations on the LH line theory, the asymmetric RH/LH coupler is presented focusing on the parameters yielding the coupling ratio and the frequency range. This type of couplers exhibits a high coupling ratio though the gap between the two coupled lines is relatively large compared to the one of the classical RH couplers. Furthermore, the frequency range of these couplers does not depend on their length. Both simulations and measurements of the coupling ratio versus the number of cells constituting the coupler are presented. From these results, we explain why such a structure is a good way to realize a coupler presenting a variable ratio. Finally, the feasibility of electronically controlled ratio couplers presenting a large bandwidth is discussed.

1. INTRODUCTION

Since the nineties, the feasibility of materials presenting simultaneously a negative permeability and permittivity are demonstrated in built up macro-structures named metamaterials. A negative phase velocity for the electromagnetic waves is observed in such structures. Consequently, the wave vector $\vec{k}$, the electrical field vector $\vec{E}$ and the magnetic field vector $\vec{H}$ constitute an indirect trihedron. Thus, such metamaterials are called Left Handed materials (LH) in opposition to the standard Right Handed propagation (RH) [8]. The singularity of this negative phase velocity has triggered a great interest and a lot of works have been carried out lately to develop these metamaterials in 1D, 2D and 3D structures.

As 2D and 3D structures are interesting optical and radiofrequency radiated applications [4, 5], 1D structures such as hybrid transmission lines are potentiality relevant for microwave circuit applications [6, 12]. Symmetrical coupling structures using two LH coupled lines [3] or asymmetrical structures using one LH line and one RH line coupled together [2] are excellent candidates to develop new circuits for telecommunications.

In a great number of applications both a large coupling bandwidth and a high coupling ratio are required. At high frequencies these electromagnetic couplers are often designed by closing up two RH lines. The frequency range of such couplers is limited since it is directly dependent on the length of the two lines in regard. In this case, the ratio between the coupled output and the transmit output can only be adjusted by the gap between the two lines [7]. In some applications, a variable coupling ratio is necessary.

This work presents a microstrip coupler with a variable coupling ratio. After some considerations on the LH line theory, the asymmetric RH/LH coupler is presented focusing on the parameters yielding the coupling ratio and the frequency range. We will present results on the coupling factor function of the number of cells constituting the coupler. From these results, we explain why such a structure is a good way to realize a coupler presenting a variable coupling ratio. Finally, the feasibility of electronically controlled ratio couplers presenting a large bandwidth is discussed.

2. THEORETICAL APPROACH OF THE LH/RH COUPLER

A LH line is a 1D structure in which the roles of the capacitor and of the inductor have been swapped compared to the elementary model of a conventional transmission line. As these RH lines do not exist naturally, a LH line must be artificially realized using capacitors C_L connected in series and inductors L_L connected to the ground. Such a line is in fact unrealisable because short RH lines are necessary to connect the lumped components. The transmission line used in our prototype is therefore an hybrid RH/LH line [1]. This line exhibits a typical transmission response (Figure 2) with the three following characteristic frequencies.

$$f_{c_1} = \frac{1}{4\pi\sqrt{L_L C_L}}$$
$$f_{c_2} = \frac{1}{2\pi\sqrt{C_R L_L d_R}}$$
$$f_{c_3} = \frac{1}{2\pi\sqrt{L_R C_L d_R}}$$

In these equations, C_L and L_L are lumped components inducing the LH behaviour. C_R and L_R are the inductance and capacitance per unit length of the Right Handed lines model. The distance d_R is the length of RH line used to connect C_L and L_L.

The targeted applications are telecommunications systems in the 2.4 GHz ISM band. Thus, the central frequency of the coupling range is chosen equal to $f_{coupling} = 2.45$ GHz. The circuit board is realized with an ordinary FR4 substrate ($\epsilon_r = 4.3$, $h = 0.8$ mm, copper35 µm, $\tan(\delta) = 0.02$) where the wavelength of a RH microstrip line at the central frequency is $\lambda_{line} = \frac{1}{\sqrt{L_R C_R}}/f = 70.34$ mm.

The hybrid line consists in a sequence of the elementary cells presented Figure 1 connected in series. Each cell is composed of two capacitors $2C_L = 2.4$ pF serially connected, separated by a self $L_L = 5.1$ nH connected in parallel to the ground. The two RH line sections used to connect these components are $\frac{d_R}{2} = 2.35$ mm length and their characteristic impedance is $Z_0 = \sqrt{L_R/C_R} = 71\,\Omega$. The parameters of the RH lines are $C_R = 82\,\text{pF}\cdot\text{m}^{-1}$ and $L_R = 410\,\text{nH}\cdot\text{m}^{-1}$. Using these parameters the 3 characteristic frequencies are $f_{c_1} = 1$ GHz, $f_{c_3} = 3.34$ GHz, $f_{c_2} = 3.63$ GHz. This structure is equivalent to a line assuming that the cell length d_r is small compared to the wave length in the microstrip wave guide. The RH/LH material can be considered as homogeneous for a dimension d_r lower than $\frac{\lambda}{10}$. The full elementary length of a cell is $d_r = 5$ mm. As the wave length of microstrip line at this frequency on such substrate is $\lambda_{line} = 70.34$ mm, we have $d_r \approx \frac{\lambda}{14}$ and the condition $d \leq \frac{\lambda}{10}$ is respected.

The hybrid line presents two typical bandpass filter behaviours. Between f_{c_1} and f_{c_3}, it has been demonstrated that the phase velocity is negative while it is positive beyond f_{c_2}.

The coupler is composed of a succession of cells containing lumped components to induce left behaviour as explained above and two small classical RH couplers as illustrated in Figure 3.

In such a coupler, measurements and simulations of the Poynting vector show a backward propagation of the power in the hybrid line (port 2 to 3) when it is classically forward in the standard RH line (port 1 to 4). This forward/backward propagation yields a greater power transfer between the two lines [10] even if the gap between the two lines of the coupler is large. 3D electromagnetic simulations show that the coupling effect is the result of a vortex-like interface mode between the hybrid RH/LH and the classical RH line.

As explained in [9, 10], the coupling effect between the RH line and the hybrid line is noticeable when the hybrid line exhibits its LH characteristic. According to simulations [11], the coupling

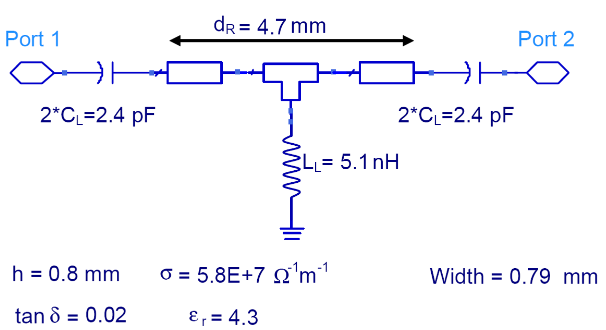


Figure 1: The hybrid RH/LH transmission line cell.

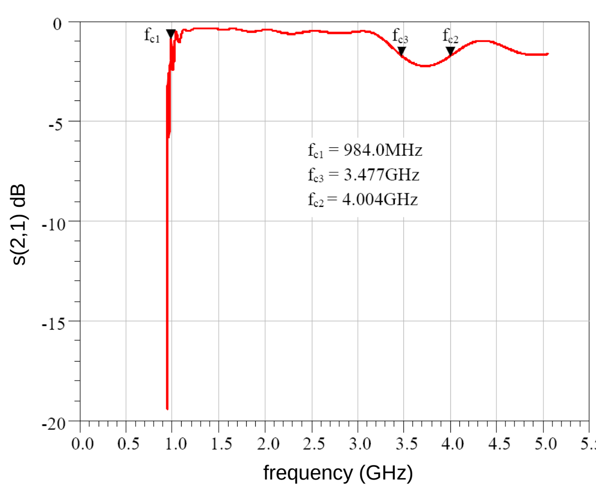


Figure 2: The combined RH/LH transmission line.

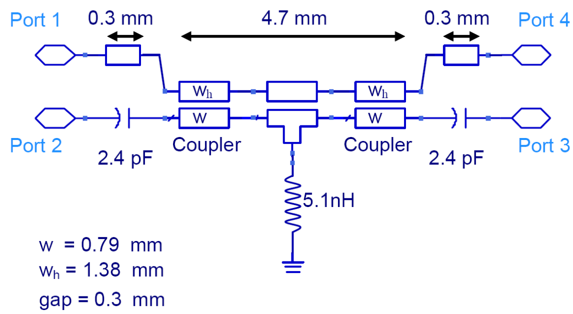


Figure 3: The elementary hybrid coupler cell.

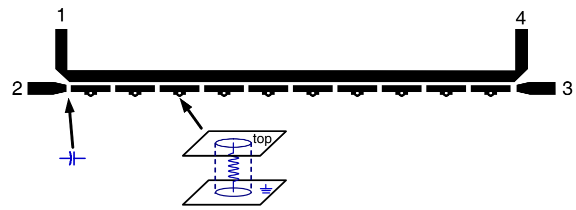


Figure 4: Layout of the hybrid coupler composed of 10 cells.

effect occurs as expected between the RH line sections of the hybrid line that are close to the RH line. The coupling ratio is enhanced by the combination of the right and left handed behaviours of the hybrid line allowing to achieve a high ratio for the power coupling with a small number of cells. In [3], the authors have studied the coupling ratio as a function of the gap between the two lines showing a strong coupling effect compared to that of a classical RH/RH coupler even for a large gap. This characteristic is very interesting in this coupler design because it reduces one of the constraint. As the coupling effect is clearly due to the LH behaviour of the hybrid RH/LH line, its frequency bandwidth is defined by the characteristic frequencies controlled by the lumped capacitors and the inductor of the elementary cell of the line. Finally the bandwidth does not depend on the number of elementary cells of the structure but only on their length. Consequently, this work is focused on the relation between the number of cells and the coupling ratio. Indeed, we intend to introduce some electronic switches in the structure to modify the number of coupled cells thus driving the coupling ratio.

3. THE COUPLING RATIO AS A FUNCTION OF THE NUMBER OF CELLS

For an application in the frequency range of telecommunications systems in the 2.4 GHz ISM band, the central frequency of the coupling range defined at $f_{coupling} = 2.45\,\text{GHz}$ is pertinent. As explained before, the LH line can be assimilated to an homogeneous LH material.

In order to study coupling ratio as a function of the number of cells, a simple hybrid RH/LH microstrip line coupler composed of 10 cells (Figure 4) was realized. In [11], the input port is located on the LH line. In the coupler studied in this work, the input and the transmit ports (port 1 and 4) have been chosen on the RH microstrip line while the coupled port (#2) and the isolated port (#3) are on the LH/RH line. With such a choice it is possible to leave some of the last cells disconnected while maintaining the transmit power way. It is important to note that each cell is symmetric (Figure 3). As specified for the hybrid RH/LH line, the 2 capacitors of each cell are equal to 2.4 pF which leads to an equivalent capacitor $C_L = 1.2\,\text{pF}$, the inductor is equal to $L = 5.1\,\text{nF}$ and the two RH section lines used for the coupling effect of the hybrid line are 2.3 mm long. The coupling gap is set to $s = 300\,\mu\text{m}$.

In order to validate the control of the coupling ratio by varying the number N of cells connected, simulations and measurements were performed for various values of N.

Simulations were performed with the structure of the coupler presented Figure 4. Assuming that the coupling effect which depends on LH behaviour is localized in the small RH couplers, it is relevant to use microwave circuits simulator like ADS from Agilent Technologies instead of a 3D electromagnetic simulator. The results are as accurate but obtained with a shorter computational time. A good agreement between measurements and simulations validates this choice. Simulation results are presented in Figure 5.

In a complementary way, measurements were performed with 1 up to 10 cells by populating them with the lumped components one by one on the prototype presented in Figure 4. The 5 cells measurement, for instance, was performed with all the capacitors and inductors of the 5 first cells soldered leaving the remaining component footprints empty. Figure 6 presents the coupling ratio and the transmit ratio of the coupler versus the frequency. Measurements were performed using an HP8722ES network analyser. The bandwidth of the coupler is nearly as large as 500 MHz centered at 2.45 GHz. This bandwidth is adequate for Ultra Wide Band Telecommunication applications.

Figure 7 presents on the same graph, at the central frequency, the growth of the coupling ratio with the number of populated cells and its counterpart decrease of the transmit ratio. The graph

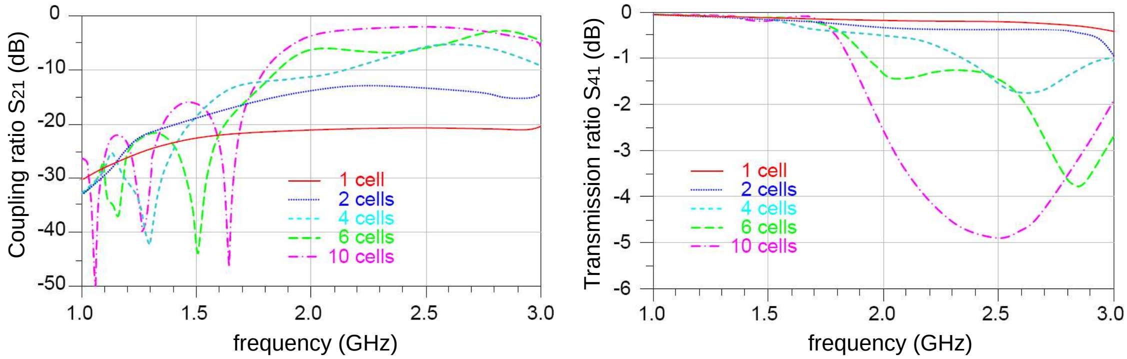


Figure 5: Coupling and transmission ratio simulated for 1 to 10 cells connected.

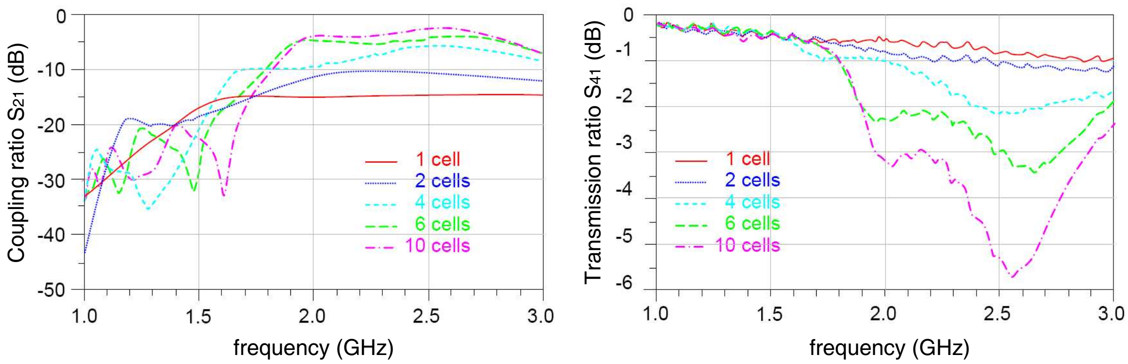


Figure 6: Coupling and transmission ratio for 1 to 10 cells populated.

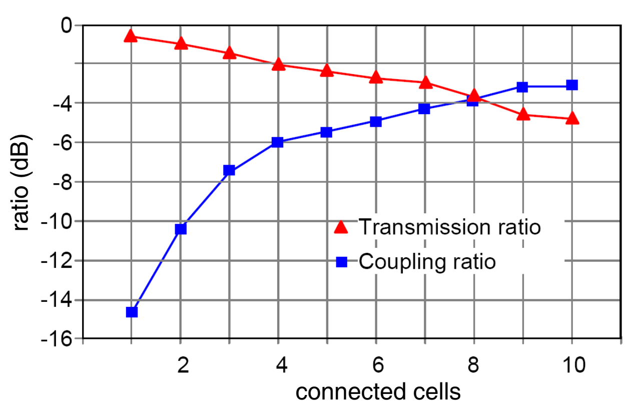


Figure 7: Coupling and transmit ratio at 2.45 GHz as a function of the number of component populated cells.

presents an asymptotic behaviour for more than 9 cells populated. It shows that the maximal coupling that can be obtained is limited. The obtained coupling ratio goes from −15 dB (1 cell) up to −3 dB (10 cells).

On can deduce from these measurements that the hybrid coupler presents some power losses resulting of propagation along the 10 cells. This power losses are due to the FR4 substrate which presents dielectric losses at the frequency used for the measurements.

4. FEASIBILITY OF AN ELECTRONIC CONTROLLED COUPLER

As shown by simulations and verified with measurements, it is clearly possible to modify the coupling ratio by modifying the number of cells involved in the coupling process. The problem consists of implementing one or more electronic switches on the hybrid RH/LH line at high frequency while changing as little as possible its characteristics. The LH behaviour of the line is very dependent on the RH section line necessary to integrate the switch, ie a PIN diode, its polarization and control circuit for instance. Inserting some RH additional section lines will modify the f_{c_2} and f_{c_3}

characteristic frequencies. New simulations must then be performed to validate such structure.

Various switches can be used, preferentially MEMS switches or PIN diode switches. The isolation provided by the electronic switch must be high enough to well control the coupling ratio. The necessary minimum isolation has to be evaluated on the prototype with different active elements. Standard RF PIN diodes exhibit a 10 dB isolation at 2 GHz which is enough to realize RF front-end TX/RX switches. In the case of the coupler, such isolation seems to be just enough. A prototype is being realized to test all these essential points for the validation of the coupler with an electronically controlled coupling ratio.

The great overall length of such couplers can be invoked as a limiting factor for their industrial use. Nevertheless, for typical industrial application where the maximal coupling ratio can be less than −3 dB, a smaller number of cells is necessary which reduces drastically the length of the coupler. The length of such couplers is therefore not the blocking factor to their industrial use.

5. CONCLUSION

The hybrid LH/RH line is a very efficient component to realize specific telecommunications circuits using its Left Handed behaviour. When well dimensioned, the coupling ratio of such structure can be modified by driving the number of cells of the coupler serially connected which has been demonstrated by the performed simulations and measurements. We have discussed the feasibility of electronically controling the number of cells: It essentially depends on the ability to maintain the hybrid behaviour of the LH/RH line once the electronic switches have been introduced. To definitively validate the feasibility of a driven power ratio coupler, a prototype with an electronic switch is being investigated.

REFERENCES

1. Caloz, C., "Dual composite right/left-handed (D-CRLH) transmission line metamaterial," *IEEE Microwave and Wireless Components Letters*, Vol. 16, No. 11, 585–587, November 2006.
2. Caloz, C. and T. Itoh, "A novel mixed conventional microstrip and composite right/left-handed backward-wave directional coupler with broadband and tight coupling characteristics," *IEEE Microwave and Wireless Components Letters*, Vol. 14, No. 1, 31–33, January 2004.
3. Caloz, C., A. Sanada, and L. Liu, "A broadband left-handed (LH) coupled-line backward coupler with arbitrary coupling level," *IEEE MTT-S Digest*, 317–320, 2003.
4. Coulombe, M., H. V. Nguyen, and C. Caloz, "Substrate integrated artificial dielectric (SIAD) structure for miniaturized microstrip circuits," *IEEE Antennas and Wireless Propagation Letters*, Vol. 6, 575–579, 2007.
5. Pendry, J. B., A. J. Holden, and D. J. Robbins, "Low frequency plasmons in thin wire structures," *Journal of Physics Condensed Matter*, Vol. 10, 4785, 1998.
6. Scher, A. D. and C. T. Rodenbeck, "Compact ring resonators using negative-refractive-index microstrip line," *Microwave and Optical Technology Letters*, Vol. 45, No. 4, 294–295, May 20, 2005.
7. Tsui, J. B.-Y., *Microwave Receivers and Related Components*, Peninsula Publishing, 1985.
8. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of ϵ and μ," *Soviet Physics Uspekhi*, Vol. 10, No. 4, 509–514, 1968.
9. Wang, Y. and Y. Zhang, "Direct observation of negative phase velocity and positive group velocity in time domain for composite right/left-handed transmission lines," *Journal of Applied Physics*, Vol. 100, No. 113503, 1–4, 2006.
10. Wang, Y., Y. Zhang, and L. He, "Discussion on coupling mechanism of asymmetric CRLH/RH TL coupler," *Journal of Zhejiang University Science A*, Vol. 7, No. 1, 95–98, 2006.
11. Wang, Y., Y. Zhang, and L. He, "Time-domain study of vortexlike interface mode in metamaterials," *Applied Physics Letters*, Vol. 91, No. 221907, 1–3, 2007.
12. Zhang, D., F. Liu, and Y. Zhang, "Experimental investigation of a power divider based on microstrip and metamaterials with L-C lumped-elements," *Microwave Journal*, November 2005.

A Dual-band Wilkinson Power Divider Utilizing EBG Structure

Hsin-Hao Chen and Yi-Hsin Pang
Department of Electrical Engineering, National University of Kaohsiung, Taiwan, R.O.C.

Abstract— This paper presented a novel dual-band Wilkinson power divider. This dual-band Wilkinson power divider loaded with a mushroom-like electromagnetic band-gap (EBG) structure. By carefully designed EBG structure, the required phase constants β at the operated dual frequencies could be acquired, and a dual-band Wilkinson power divider was then achieved. A dual-band power divider operated at 1 GHz and 2 GHz was implemented and measured for verification.

1. INTRODUCTION

With the progress of technology, wireless communication systems developed at radio and microwave frequency are increasingly significant. A Wilkinson power divider is one of the key components in microwave circuits, and is widely used in balanced amplifiers, the feeding networks of an antenna array, etc [1]. A dual-band or multi-band power divider is required for seamless communications. A dual-band power divider was realized by two-section transmission line with an additional lumped inductor and capacitor [2]. The additional surface-mounted devices, however, increase the cost and decrease the reliability of the power divider. By attaching two transmission line stubs to the conventional Wilkinson power divider, a dual-band operation was obtained [3]. In addition, the quarter wavelength transformer in the conventional Wilkinson power divider has been replaced by two-section cascaded couple lines for the dual-band operation [4]. A tri-band Wilkinson power divider has also been presented [5].

EBG structures have been widely used in antenna engineering, such as the suppression of surface wave in substrate-loaded antennas, efficiency enhancement of low profile antennas [6]. Its applications in dual-band operation have attracted people's eyes in recent years [7, 8]. Diodes were suggested to switch on and off the EBG structure to have different dispersion diagram and make a microstrip antenna dual-band [7]. This paper employed the nonlinear dispersion relation to the design of a dual-band Wilkinson power divider loaded with a mushroom-like EBG structure. For verification, a power divider operated at 1.0 GHz and 2.0 GHz was designed and implemented on an FR4 substrate. The measured data was presented and showed a good consistency with the simulation results.

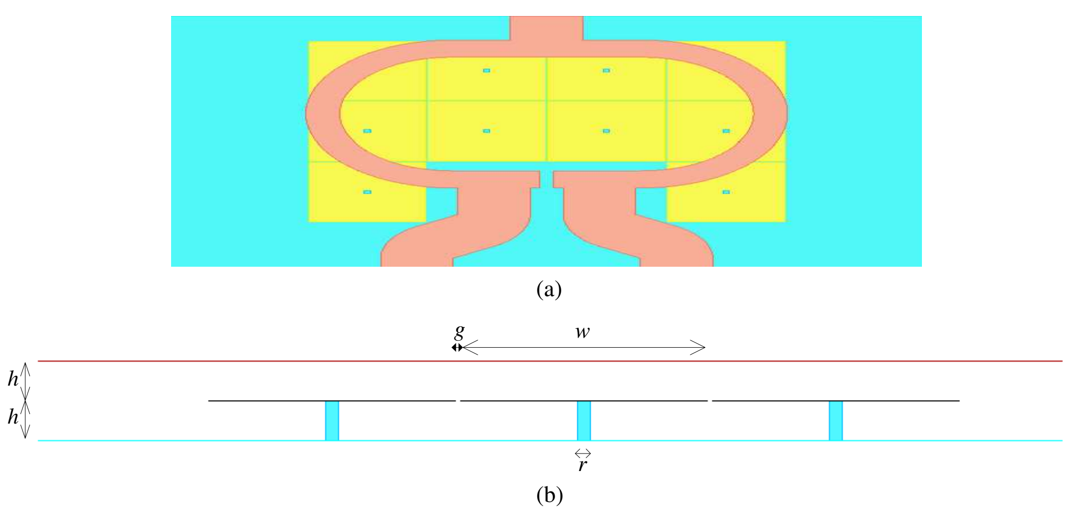


Figure 1: Geometry of the proposed dual-band power divider over the EBG structure. (a) Top view. (b) Side view.

2. ANALYSIS AND DESIGN

Figure 1 shows the proposed dual-band Wilkinson power divider. This power divider consists of a conventional Wilkinson power divider and an EBG structure. The conventional Wilkinson power divider is composed of two transmission lines, each having a characteristic impedance $70.7\,\Omega$. A lump resistor of resistance $100\ \Omega$ is connected between the output ports for good isolation. The conventional Wilkinson power divider can be operated at frequency f_n where the electrical length of the transmission line is $\theta = n\pi/2$, $n = 1, 3, 5, \ldots$. The Wilkinson power divider therefore inherently possesses the multiband characteristics [1].

A mushroom-like periodic metallic pattern was implemented for EBG structure in this paper, as shown in Figure 1(a). Each unit cell comprises two elements: (1) a patch and (2) a via. The EBG structure generates a band-gap in the specific frequency range. Thus, no signal will be received at output ports. The via is connected between the patch and ground plane. The size of patch, gap between patches and the radius of via affect the band-gap frequency [6]. In this paper, the width of metal patch is $w = 10\,\mathrm{mm}$, gap between patches is $g = 0.2\,\mathrm{mm}$ and the diameter of via $r = 0.5\,\mathrm{mm}$, as shown in Figure 1(b). Figure 2 shows the dispersion relations for a microstrip line with and without the EBG structure. For the microstrip line with the EBG structure, a band-gap is presented and the phase constant β increases nonlinearly with frequency. Besides, the phase constant β is almost not influenced by the band-gap at low frequency. Based on the above reasons, at a lower frequency f_1, the electrical length of the transmission line of Wilkinson power divider with the EBG structure is designed to be $\theta = \pi/2$. By adjusting the size of the EBG structure, the electrical length of the transmission line of Wilkinson power divider would be $\theta = 3\pi/2$ at the desired second frequency f_2 which is less than $3f_1$ due to the rapid change of β near the band-gap. The required dual-band operation is achieved.

3. RESULTS

A dual-band Wilkinson power divider was designed and fabricated on an FR4 printed circuit board (PCB). The parameters of FR4 substrate are thickness $h = 1.6\,\mathrm{mm}$, relative permittivity $\varepsilon_r = 4.33$ and loss tangent $\tan\delta = 0.022$. Two FR4 substrates were stacked for the implementation of the EBG structure. Figure 3 shows the photograph of the proposed power divider. The power divider is of size $60\,\mathrm{mm} \times 65\,\mathrm{mm}$ and measured by Agilent E5071C network analyzer.

Figure 4 shows the simulation and measured results. The simulation results were obtained by using the full-wave simulator Agilent Momentum, and showed good consistency with the measured data. A dual band operation is observed by return loss S_{11} shown in Figure 4(a), and the $|S_{11}|$ is less than $-15\,\mathrm{dB}$ at the two designed frequencies, 1 GHz and 2 GHz. Figure 4(b) shows the measured insertion loss, S_{21} are $-3.56\,\mathrm{dB}$ and $-4.57\,\mathrm{dB}$, S_{31} are $-3.46\,\mathrm{dB}$ and $-4.77\,\mathrm{dB}$ at 1 GHz and 2 GHz respectively. The band-gap occurs at 2.3 GHz where $|S_{21}|$ is less than $-15\,\mathrm{dB}$. From Figures. 4(c) and (d), the output ports are well matched and have good isolation between each other at the designed frequencies.

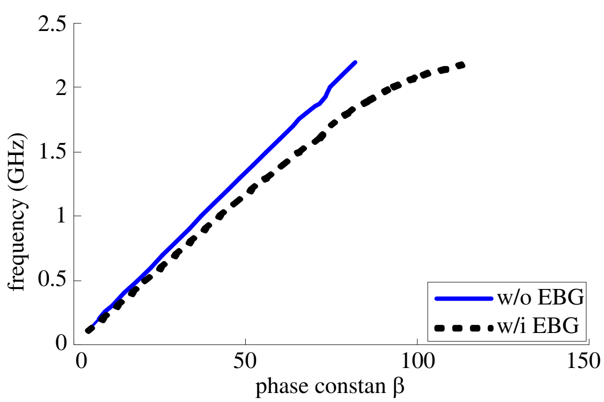


Figure 2: Dispersion diagrams of a microstrip line with and without the EBG structure. Unit of phase constant = rad/m.

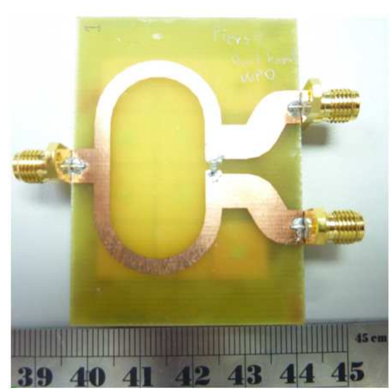


Figure 3: The photograph of the fabricated dual-band Wilkinson divider.

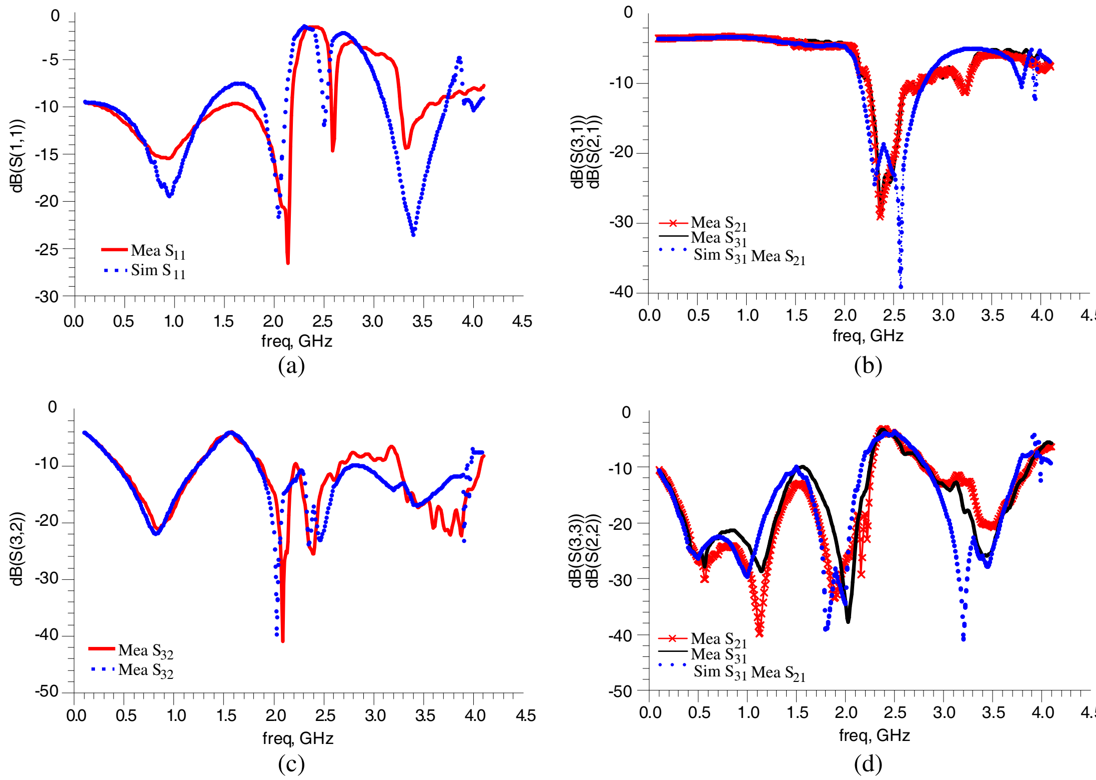


Figure 4: Comparison of simulated and measured magnitudes of the proposed dual-band power divider. (a) S_{11}. (b) S_{21} and S_{31}. (c) S_{32}. (d) S_{22} and S_{33}.

4. CONCLUSION

A dual-band Wilkinson power divider loaded with a mushroom-like EBG structure has been proposed. By carefully designing the size of the EBG structure, the required dispersion relation can be obtained to make the power divider operated at two arbitrary frequencies. Both the simulated and measured data show that the proposed circuit has good matching and isolation properties at the designed 1 GHz and 2 GHz.

ACKNOWLEDGMENT

This work was supported by the National Science Council, Republic of China, under Grant NSC 99-2221-E-390-018.

REFERENCES

1. Pozar, D. M., *Microwave Engineering*, 3rd Edition, Wiley, New York, 2005.
2. Wu, L., Z. Sun, H. Yilmaz, and M. Berroth, "A dual-frequency Wilkinson power divider," *IEEE Trans. Microw. Theory Tech.*, Vol. 54, No. 1, 278–284, Jan. 2006.
3. Park, M.-J. and B. Lee, "A dual-band Wilkinson power divider," *IEEE Microw. Wireless Compon. Lett.*, Vol. 18, No. 2, 85–87, Feb. 2008.
4. Park, M.-J., "Two-section cascaded coupled line Wilkinson power divider for dual-band applications," *IEEE Microw. Wireless Compon. Lett.*, Vol. 19, No. 4, 216–222, Apr. 2009.
5. Chongcheawchamnan, M., S. Patisang, M. Krairiksh, and I. D. Robertson, "Tri-band Wilkinson power divider using a three-section transmission-line transformer," *IEEE Microw. Wireless Compon. Lett.*, Vol. 16, No. 8, 452–454, Aug. 2006.
6. Yang, F. and Y. Rahmat-Samii, *Electromagnetic Band Gap Structures in Antenna Engineering*, Cambridge University Press, 2009.
7. Liang, J. and H. Y. D. Yang, "Microstrip patch antennas on tunable electromagnetic band-gap substrates," *IEEE Trans. Antennas Propag.*, Vol. 57, No. 6, 1612–1617, Jun. 2009.
8. Ayop, O., M. K. A. Rahim, and T. Masri, "A dual band gap slotted patch electromagnetic band gap for dual band microstripantenna," *2008 IEEE International RF and Microwave Conference*, 322–324, 2008.

Large Scale Measurement of Microwave Electric Field Using Infrared Thermography and Electromagnetic Simulation

D. Prost, F. Issac, and P. Reulet
ONERA — The French Aerospace Lab, 2 avenue E. Belin, Toulouse, France

Abstract— We present an original method for measuring the microwave electric field and achieving large scale field cartography. This is based on the use of a conductive and thermoemissive thin film of large dimensions, where induced currents due to incident field are responsible for ohmic losses and therefore create thermal heating. This heating is recorded by an infrared camera. The obtained thermal frame is obviously directly linked to the electromagnetic energy dissipated inside the film. This frame can be used qualitatively, for example to detect field leakage near electromagnetic junctions or to get antenna near field patterns, as it has been used for many years at French Aerospace Lab ONERA. To get a quantitative field evaluation, a thermal computation is necessary.

Moreover, with the help of simulation tools, more quantitative information, like field amplitude or even component values of the field, can be obtained, making the film become an electric field sensor. The simulation work addresses electromagnetism and thermal physics. We present such approach in the particular case of the near field of a ZOR (zero order) antenna. The measured field frame is compared to the computed one and good matching of the two cartographies is obtained.

1. INTRODUCTION

The electromagnetic infrared (EMIR$^{\mathrm{TM}}$) method was developed and has been used at ONERA [1] in order to visualize and/or measure microwave electric fields. It consists in putting conductive films of eventually large dimension (up to more than one square meter) in the electromagnetic field. Induced currents in the films create ohmic losses and consequently an heating that can be filmed using an infrared camera. This method provides thermal frames that can be linked to the square of the electric field amplitude (absorbed electric power); it gives a qualitative information (for example the electric field profile near an antenna) but to get an estimation of the electric field amplitude, a calibration, using a field sensor, is necessary. However, in the cases where the field source is known, a coupled electromagnetic and thermal computation is another way to get that quantitative information (without calibration). We first present the measurement set-up and results in Paragraph 2. The thermal computation and the electromagnetic simulation are presented in Paragraph 3.

2. MEASUREMENTS SET-UP

The EMIR$^{\mathrm{TM}}$ film is a very thin conductor sheet (almost 50-μm thick), with resistivity of the order of 50 mΩ·m. Therefore the skin depth is much larger than the thickness d at our (microwave) frequencies. This suggests that the characteristic value for the film is its surface impedance, Z_s (equivalent to impedance per square), given by:

$$Z_s = \frac{\rho}{d} \tag{1}$$

Typical values for EMIR$^{\mathrm{TM}}$ films, made of carbon loaded polymers or metallic deposit, are in the range 500 to 3000 Ω. Regarding the absorbed power, this corresponds (in the plane wave approximation, with the wave vector perpendicular to the film) to a proportion between 40% (highly intrusive film) to 10% (low interaction with the field), as illustrated by the Figure 1 [2].

The film used in the present case has a (measured) impedance $Z_s = 1600\,\Omega$, leading to a 20% absorption rate.

This film is located in the near field region of the antenna. As the microwaves induce currents heating the film in a continuous waves (CW) mode, a low frequency modulation is applied to avoid convection process that would hide the field image. The antenna is thus supplied with a 0.5 Hz modulation, the camera being demodulated at the same frequency. The measurement set-up is represented below on Figure 2.

Using that test bench we can visualize the induced currents due to the antenna near electric field, through induced heating. We noticed an influence (capacitive coupling) of the film on the resonance frequencies of the antenna, which are slightly shifted. The Figure 3 shows the thermal frame obtained for a frequency corresponding to the "zero order mode" of a metamaterial-based micro strip antenna.

3. THERMAL TRANSFER AND ELECTROMAGNETIC SIMULATION

The electromagnetic power is thermally dissipated by convection process. In first approximation, we would have $P_{abs} = 2h \cdot \Delta T$, where h is the heat transfer coefficient (a factor 2 is occurring

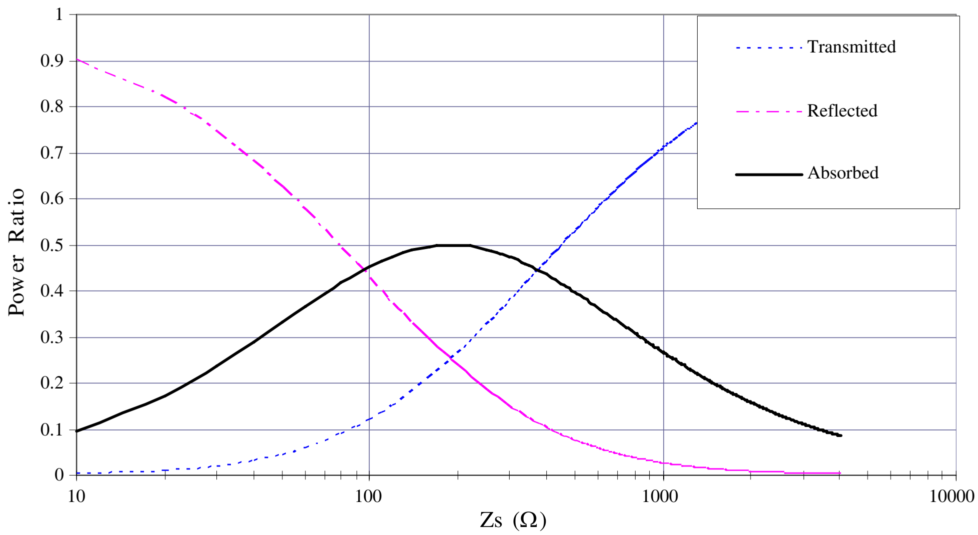


Figure 1: Absorbed, transmitted and reflected power as a function of thin film surface impedance.

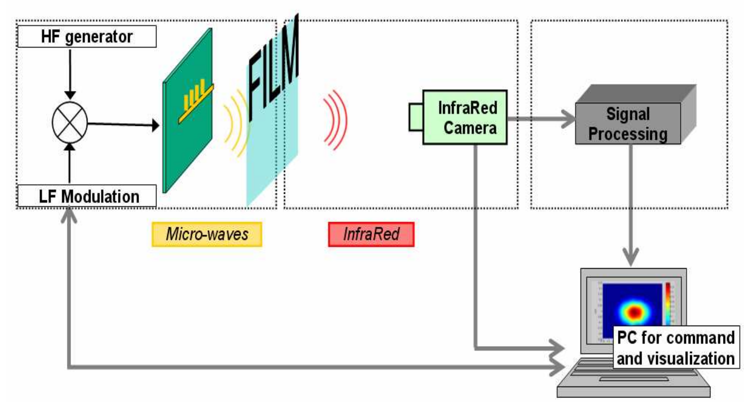
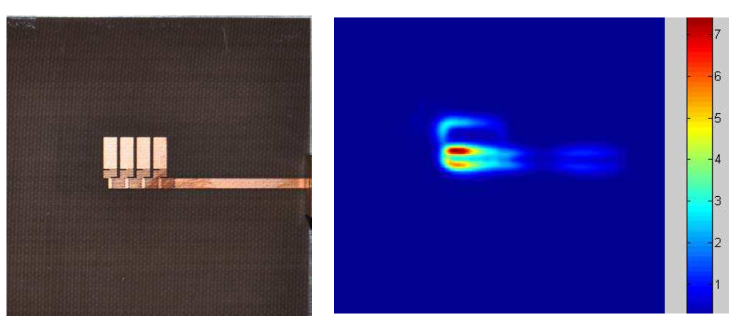


Figure 2: Thermography measurement setup.



(a) (b)

Figure 3: (a) View of the antenna, (b) thermal heating on the EMIR film, in °C.

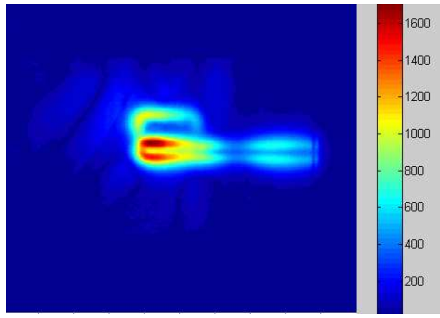


Figure 4: Electric field cartography obtained by EMIR measurement (V/m).

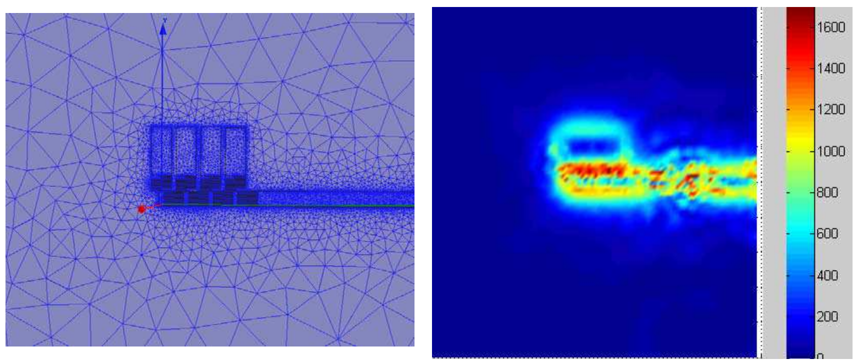


Figure 5: Electric field cartography obtained by finite elements computation (V/m).

to take onto account the two sides of the film) and ΔT the measured heating. However, because of the modulation at low frequency, the heating is modulated itself and further computations are necessary. This has been done using complete thermal film characteristic (using the thermal conductivity C, the thickness d and the density ρ). A thermal steady state is reached after a few second, corresponding to the time constant of the film $\tau = \rho Ce/2h$. The heating amplitude is then more precisely linked to the absorbed power by:

$$P_{abs} = \Delta T \sqrt{4h^2 + (\rho C d\omega)^2} \tag{2}$$

This equation allows us to obtain the absorbed power and therefore the electric field since we have:

$$P_{abs} = \frac{1}{2}\frac{E^2}{Z_s} \tag{3}$$

The "measured" electric field map is therefore obtained (see Figure 4); this corresponds however to the near field antenna frame in presence of the conductive film.

This frame is proportional to the square root of the values of the Figure 3, and is therefore less contrasted.

The electromagnetic simulation of the antenna has been achieved using the HFSS Finite Element Analysis tool [4]. The meshing is adapted to the resonant frequency (slightly below 5 GHz), the antenna is supplied with a nominal power on his input port. The film is represented as an ohmic surface with the surface impedance $Z_s = 1600\,\Omega$. This allows us to get the electric field frame on the film, as given on the Figure 5; a coefficient factor is applied on the input power to make the measured and computed maximum field match.

The computed and measured field frames are in good accordance. The finite element model is validated and can now be used for further analysis. Moreover, the field map when the film is not present can be computed.

4. CONCLUSION

We exhibit an original way to measure electric field on plane surface using a special thermoemissive conductive film. This allows direct measurement of the field cartography, without moving any probe.

Moreover, sub-wavelength details are visible, and the measurement is relatively low intrusive (the perturbation due to the film itself can be evaluated, and corrected). Finite elements simulation has been also achieved, that give good concordance with experiments results. In this particular case of the ZOR antenna near field, computation is usually the only available way to get electric field maps; thus the present measurement experimentally confirms the validity of the simulation approach. Moreover, this measurement method can be used on larger system whose finite elements modeling would require huge computational resource and would therefore be problematic.

REFERENCES

1. Levesque, P., J. Grando, F. Issac, M. Nacitas, and D. Balageas, "Mesures des champs électromagnétiques par thermographie infrarouge," *Journée D'étude SEE sur les Capteurs électromagnétiques et Leurs Applications*, Toulouse, 1997.
2. Nacitas, M., "Contribution au développement de méthodes de contrôle non destructif et de caractérisation de matériaux par interférométrie microondes associée à la thermographie infrarouge," *Note Technique ONERA*, 1997.
3. Parmantier, J.-P., F. Issac, F. Lemaître, J.-L. Lasserre, A. Paupert, M. Wilbanks, M. Harrison, and H. Pohle, *ANTEM/AMEREM 2010*, Ottawa, Canada, 2010.
4. Ansoft, www.ansoft.com.

Numerical Study of a Coplanar Zeroth-order Resonator on YIG Thin Film

Aziza Zermane[1,2,3], **Bruno Sauviac**[1,2], **Bernard Bayard**[1,2], **and Abdelmadjid Benghalia**[3]

[1]Université de Saint-Etienne, F42023, France
[2]TELECOM Saint-Etienne, école associée de l'Institut TELECOM, DIOM
EA 3523, F-42023, 25 rue Dr. Rémy Annino, Saint-Etienne 42000, France
[3]Laboratoire LHS, Université Mentouri de Constantine
Route d'Ain EL Bey 25000 Constantine, Algérie

Abstract— The objective of this work is to study numerically the behaviour of a coplanar composite right-left handed zeroth order resonator (CRLH ZOR) realized on a YIG (Yttrium Iron Garnet) thin film. We are focusing on the effects on the resonant frequency when changing the magnetic bias and the thickness of the YIG. This study presents for the first time a tunable CRLH ZOR with a coplanar structure on YIG thin film. The length of the proposed design (device) is 5.2 mm, which is very small compared to a traditional half wave resonator. The proposed resonator is designed to be tuned over the 5–6 GHz frequency band; insertion loss is lower than 1 dB and return loss is better than 10 dB.

1. INTRODUCTION

The concept of the Composite Right-Handed Transmission Line (line with traditional propagation) and Left-Handed Transmission Line (line having a negative phase velocity) was proposed by UCLA group [1]. An equivalent circuit approach was also developed by Eleftheriades and al [2]. This approach takes into account the parasitic Right-Handed (RH) effects naturally occurring in a practical Left-Handed (LH) structure. So the CRLH notation is more general than the LH notation. The CRLH approach describes in a simple manner the fundamentally right-handed and left-handed nature of metamaterials [1, 3].

The first zeroth-order resonator (ZOR) is proposed by Sanada and al [4]; it is realized in microstrip configuration. Based on this unit cell, a practical application is realized for antennas [5]. A coplanar waveguide (CPW) presents the advantage that line and ground are located in the same plane, and it can be easily fabricated at low cost by using a lithography process. Many researchers studied and realized the zeroth-order resonators in a coplanar waveguide technology [6–8].

Recently, ferrite materials have been used in metamaterials and in meta-line too. Tsutsumi and Abdalla proposed a nonreciprocal left-handed transmission line in microstrip and in coplanar waveguide configurations with a ferrite substrate [9, 10]. All of them employed a bulk ferrite. Our objective is to prove that it is possible to make a tunable CRLH ZOR with a 15 μm ferrite film only. As the manufacturing process developed in DIOM laboratory allows making YIG thin films, then it will be possible for us to prospect for collective fabrication in order to reduce the production costs.

2. THEORY

Zeroth order resonator (ZOR) is one of the novel applications of LHMs (Left-handed materials). The interest lies in the fact that the resonant frequency is independent from the physical length of the resonator. Due to metamaterials properties, negative and zero resonances are possible. The zeroth-order CRLH resonance ($\ell/\lambda_g = 0$, $m = 0$) is particularly unique and interesting [4], since it has infinite guided wave-length at a specific frequency. The proposed resonator is realized in coplanar waveguide configuration constructed using an interdigital capacitor (IDC) and a short-circuited stub inductor (SSI). Figure 1(a) illustrates the complete layout of the CRLH CPW ZOR which is coupled by air gaps. It can be divided into three parts: $50\,\Omega$ access lines (one for each side), the CRLH CPW cell and two CPW series gaps as coupling structures. In this work, specific physical parameters have been used: The YIG film has been modeled with a relative dielectric permittivity close to 15, a saturation magnetization equal to $Ms = 1780$ Gauss and a ferromagnetic resonance (FMR) line width $\Delta H = 20$ Oe. For conductor lines made of copper, the conductivity is 0.610^8 S. The YIG layer is magnetized in a direction perpendicular to the propagation direction (Figure 1(b)) and we suppose that the ferrite is saturated and that the internal bias field is uniform.

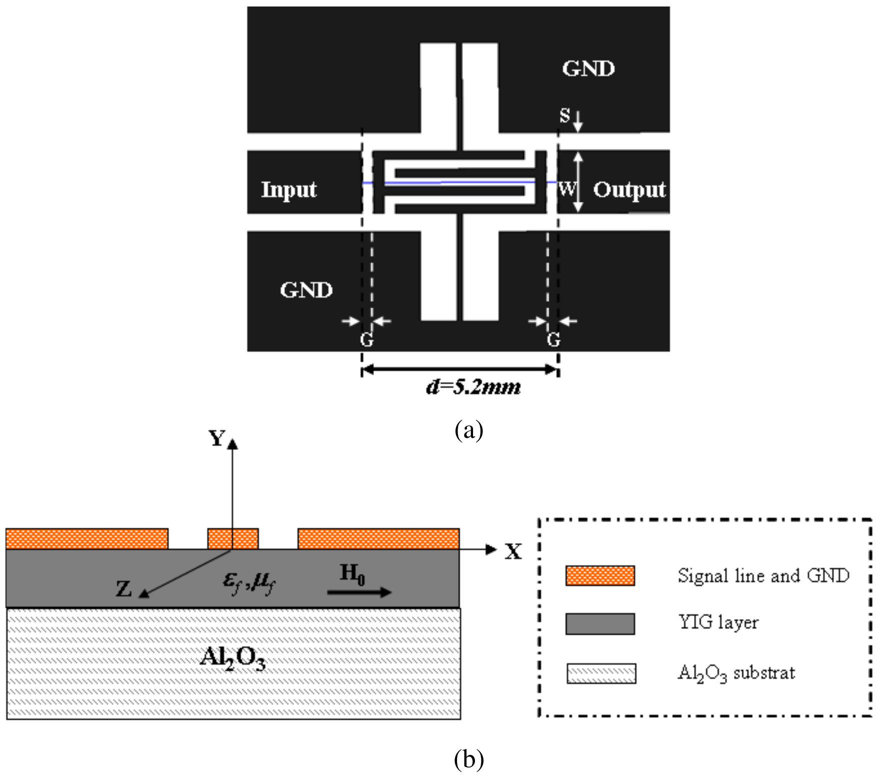


Figure 1: (a) Physical configuration of the CRLH CPW ZOR, (b) its cross section.

The physical parameters of our resonator and the theoretical developments used for modeling were presented in reference [6]. According to the direction of the DC magnetic bias presented in the previous figure, the permeability tensor has the form of:

$$[\bar{\bar{\mu}}_f] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \mu_f & +jk \\ 0 & -jk & \mu_f \end{bmatrix} \tag{1}$$

The elements of the permeability tensor depend on the frequency:

$$\mu_f = 1 + \frac{\omega_0 \omega_M}{\omega_0^2 - \omega^2} \tag{2}$$

$$k = \frac{\omega \omega_M}{\omega_0^2 - \omega^2} \tag{3}$$

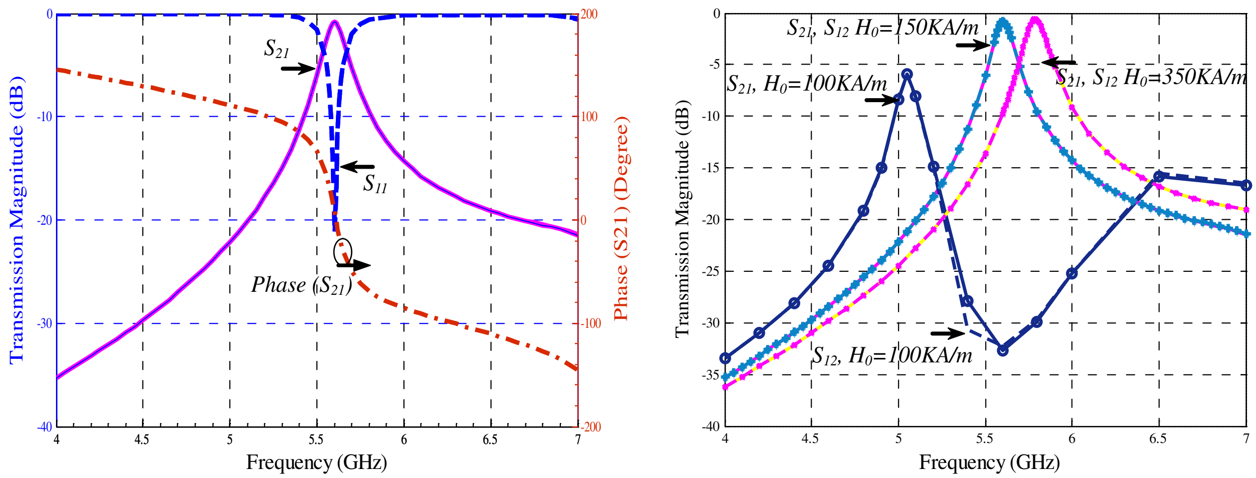
where $\omega_M = \gamma\mu_0 M_S$, $\omega_0 = \gamma\mu_0 \mathrm{H}_0$ and $\gamma = 176 \times 10^9\,\mathrm{rad\,s^{-1}\,T^{-1}}$ ($\Leftrightarrow 28\,\mathrm{GHz/T}$).

As the material is supposed to be saturated, the ferrite is modeled using Polder's tensor as in Equation (1).

3. NUMERICAL RESULTS

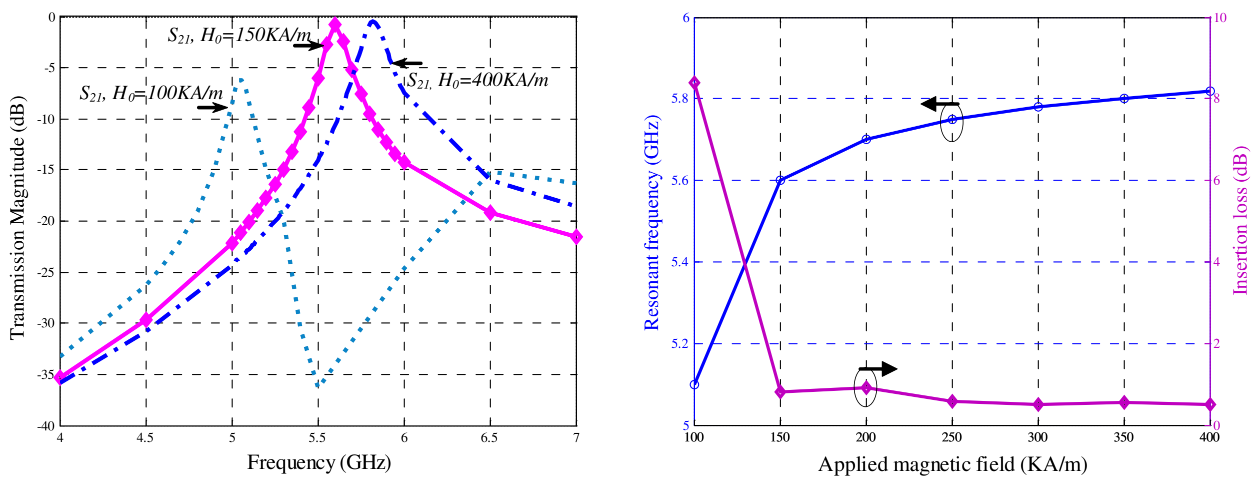
The results which enabled us to launch manufacture thereafter were obtained starting from 3D electromagnetic simulations. First, to show the tunable property of the proposed CRLH CPW ZOR, we present the results obtained from a structure using 20 µm YIG layer thickness. Our resonator has been analyzed numerically for different values of DC magnetic bias. For example in the case of the applied field of 150 KA/m, the resonant frequency is equal to 5.6 GHz, at this frequency the insertion loss is about 0.81 dB and the return loss is better than 21 dB. We observe that the phase of transmission coefficient is close to zero near the resonant frequency (Figure 2(a)). The transmission characteristics are shown in Figure 2(b) for three values of the DC magnetic field.

Second, to prove that it is possible to make a tunable CRLH CPW ZOR with a 15 µm ferrite film only, a numerical study was then performed using HFSS for different values of the magnetic field. Figure 3(a) shows the modelled transmission characteristics of this device.



(a) (b)

Figure 2: Numerical results for CRLH CPW ZOR on 20 µm ferrite film. (a) Transmission characteristics and the phase of S_{21} of the resonator under $H_0 = 150\,\mathrm{KA/m}$. (b) Transmission characteristics of resonator at $H_0 = 100\,\mathrm{KA/m}$, $H_0 = 150\,\mathrm{KA/m}$ and $H_0 = 350\,\mathrm{KA/m}$.



(a) (b)

Figure 3: Numerical results for CRLH CPW ZOR on 15m ferrite film only. (a) Transmission characteristics of the resonator for different values of the applied field. (b) Variation of the resonant frequency and insertion loss of the resonator with the applied magnetic field.

The variation of the resonant frequency and the insertion losses according to the applied magnetic field in the case of the CRLH CPW ZOR with a 15 µm ferrite film is shown in Figure 3(b). We observe that the resonant frequency varies with applied field. Insertion loss seems also to depend on the external applied field; nevertheless for values above 250 KA/m losses remain constant.

4. CONCLUSIONS

A CRLH CPW ZOR on the YIG thin film has been studied numerically using a three dimensional electromagnetic simulation. The tunable propriety has been validated numerically for the air gap coupled line CRLH CPW ZOR with a 15 µm ferrite film only. This result shows the possibility of modifying the agreement of the device by an external intervention (magnetic field). The proposed device can be applied in many filters, for example when multiband devices are necessary.

REFERENCES

1. Caloz, C., A. Sanada, and T. Itoh, "Microwave circuits based on negative refraction index material structures," *33rd European Microwave Conference*, 105–110, Munich, 2003.
2. Eleftheriades, G. V., A. K. Iyer, and P. C. Kremer, "Planar negative refractive index media

using periodically L-C loaded transmission lines," *IEEE. Trans. Microw. Theory Tech.*, Vol. 50, 2702–2712, 2002.

3. Caloz, C., A. Sanada, and T. Itoh, "Microwave applications of transmission-line based negative refractive index structures," *Microwave Conference*, Asia Pacific, 2003.
4. Sanada, A., C. Caloz, and T. Itoh, "Novel zeroth-order resonance in composite right/left-handed transmission line resonators," *Microwave Conference*, Asia Pacific, 2003.
5. Caloz, C., T. Itoh, and A. Renning, "CRLH metamaterial leaky-wave and resonant antennas," *IEEE Antennas and Propagation Magasing*, Vol. 50, No. 5, 25–39, October 2008.
6. Zermane, A., B. Sauviac, B. Bayard, B. Payet-Gervy, J. J. Rouseau, and A. Benghalia, "Modeling and measurement of a zeroth-order-resonator and a composite right-left-handed transmission line in coplanar technology," *Progress In Electromagnetic Research C*, Vol. 14, 33–43, 2010.
7. Li, C., K. Y. Liu, and F. Li, "Analysis of composite right/left-handed coplanar waveguide zeroth-order resonators with application to a band-pass filter," *PIERS Online*, Vol. 3, No. 5, 599–602, 2007.
8. Li, L. and D. Uttamchandani, "Demonstration of tunable RF MEMS bandpass filter using silicon foundry process," *Journal of Electromagnetic Waves and Application*, Vol. 23, Nos. 2–3, 405–413, 2009.
9. Tsutsumi, M. and T. Ueda, "Nonreciprocal left-handed microstrip lines using ferrite substrate," *IEEE MTT-S Int. Microwave Sym. Dig.*, Vol. 1, 249–252, 2004.
10. Abdalla, M. and Z. Hu, "A nonreciprocal left handed coplanar waveguide on ferrite substrate with only shunt inductive load," *Proceeding of Metamaterials*, October 2007.

Metallic Absorptivity at Normal Incidence above Far-infrared

F. E. M. Silveira
Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André, Brazil

Abstract— The absorptivity of monochromatic radiation at normal incidence on metals is described by the Hagen-Rubens (H-R) relation, provided the wave energy is not higher than $\sim 100\,\text{meV}$ (wavelength not shorter than $\sim 10\,\mu\text{m}$), that is, it is restricted to the far-infrared and downwards. The H-R relation is a consequence of Maxwell's equations together with Ohm's law. This work proposes an extension of the H-R relation above the far-infrared by taking into account inertial effects due to charge carriers. The influence of inertia is described by introducing a relaxation time for the current density in the framework of a generalized Ohm's law. In the extended approach, it is found that the unique characteristic length scale of the problem, the penetration depth of the wave into the conductor, splits in two such that two distinct regimes for the skin effect can be identified: at low frequencies, the fields exhibit the usual attenuated oscillations; at high frequencies, they become purely attenuated in space. A transition frequency between the regimes is also identified. It is shown as well that the relaxation time can be fully determined from the metallic permittivity and angular frequency. The extended formulation is applied to aluminum and it is found that its relaxation time $\sim 6.91\,\text{fs}$ at room temperature. It is also shown that the absorptivity by aluminum is in excellent agreement with the extended theory up to the near-infrared ($\sim 200\,\text{meV}$).

1. INTRODUCTION

Investigations pursued by Hagen and Rubens more than a century ago [1] established that the absorptivity A of monochromatic radiation (angular frequency ω) by many metals (static electric conductivity σ), as calculated from classical electrodynamics,

$$A = 2\sqrt{\frac{2\epsilon_0\omega}{\sigma}}, \tag{1}$$

where ϵ_0 denotes the vacuum electric permittivity, is in good agreement with experimental results, provided the wave energy is not higher than about $100\,\text{meV}$. This means that the domain of validity of Eq. (1), commonly referred to as the Hagen-Rubens (H-R) relation, is restricted to the far-infrared region and downwards, that is, to a wavelength not shorter than about $10\,\mu\text{m}$. Since Eq. (1) describes the observed high reflectivity by most metals at small frequencies, the limit $\sigma \gg \epsilon_0\omega$ holds, that is, the metallic refractive index n and extinction coefficient k satisfy the so-called Hagen-Rubens (H-R) condition [2], $k = n \gg 1$, where the unit denotes the vacuum refractive index. Therefore, precise measurements of metallic optical constants are required to determine the limits for consistency of the H-R condition with Eq. (1), and even to attempt extensions of the classical approach for the optics of metals.

This work proposes an extension of the classical H-R relation beyond the far-infrared region in the framework of a generalized Ohm's law. Within the extended approach, it is found that the penetration depth of the wave into the conductor splits in two such that two distinct regimes for the skin effect are identified as well as a transition frequency between them. The theory is then applied to aluminum and it is shown that the behavior of its observed absorptivity at normal incidence is in excellent agreement with the extended formulation up to the near-infrared region.

2. DOUBLE SCALE FOR SKIN EFFECT

Let us start by considering the generalized Ohm's law

$$\left(1 + \tau\frac{\partial}{\partial t}\right)\vec{J} = \sigma\vec{E}, \tag{2}$$

where τ is the characteristic time scale describing the relaxation of the induced current density $\vec{J}$ due to the inertia of charge carriers as a response to the applied electric field $\vec{E}$. By assuming that the fields vary in time as $\sim \mathrm{e}^{-\imath\omega t}$, the curl Maxwell's equations can be written as

$$\nabla \times \vec{E} = \imath\mu_0\omega\vec{H}, \quad \nabla \times \vec{H} = \left[\sigma\cos\varphi\left(\cos\varphi + \imath\sin\varphi\right) - \imath\epsilon\omega\right]\vec{E}, \tag{3}$$

where ϵ is the material electric permittivity, μ_0 is the vacuum magnetic permeability, $\vec{H}$ is the magnetic field and $\vec{J}$ lags $\vec{E}$ in time by the dephasing angle [3] $\varphi = \tan^{-1}(\omega\tau)$. From Eq. (3), we see that, when the inertia of charge carriers is taken into account, the standard good conductor limit $\sigma \gg \epsilon\omega$ must be extended to $\sigma\cos\varphi \gg \epsilon\omega$.

A widely known experimental result states that the variations of both magnetic and electric fields perpendicular to the surface of the conductor are much more rapid than those parallel to it. Therefore, if $\hat{n}$ is the unit normal exterior to the metallic surface and x is a coordinate in the opposite direction of $\hat{n}$, with origin at the surface, the gradient operator can be identified with $\nabla \equiv -\hat{n}\partial_x$. Then, when the extended good conductor limit holds true, Eq. (3) approach

$$\hat{n} \times \frac{\partial \vec{E}}{\partial x} = -\imath\mu_0\omega\vec{H}, \quad \hat{n} \times \frac{\partial \vec{H}}{\partial x} = -\sigma\cos\varphi\,(\cos\varphi + \imath\sin\varphi)\,\vec{E}. \tag{4}$$

The electric field may be eliminated from Eq. (4), leading to

$$\frac{\partial^2 \vec{H}}{\partial x^2} = -\imath\mu_0\omega\sigma\cos\varphi\,(\cos\varphi + \imath\sin\varphi)\,\vec{H}, \tag{5}$$

whose solution can be written as $\vec{H} = \vec{H}_0 \mathrm{e}^{-x/\delta}\mathrm{e}^{\imath x/\xi}\mathrm{e}^{-\imath\omega t}$, where the ascribed length scales are

$$\left.\begin{matrix}\delta \\ \xi\end{matrix}\right\} = \sqrt{\frac{\tau}{\mu_0\sigma}}\sqrt{\frac{2}{\sin\varphi\,(1 \pm \sin\varphi)}}. \tag{6}$$

From Eq. (6), we see that, due to the introduction of the inertial relaxation time for the current density into the problem, the most basic correction to the classical approach for the skin effect is the split of its formerly unique characteristic length scale in two.

Figure 1 shows the behavior of the length scales $\hat{\delta}$ and $\hat{\xi}$ (normalized to $\sqrt{\tau/\mu_0\sigma}$) as functions of the time dephasing angle φ.

3. EXTENDED HAGEN-RUBENS RELATION

Equation (5) can be interpreted as a diffusion type equation for the magnetic field, $\partial^2\vec{H}/\partial x^2 = -\kappa^2\vec{H}$, where the complex wave number is

$$\kappa = \sqrt{\frac{\mu_0\sigma}{\tau}}\left[\sqrt{\frac{\sin\varphi\,(1-\sin\varphi)}{2}} + \imath\sqrt{\frac{\sin\varphi\,(1+\sin\varphi)}{2}}\right]. \tag{7}$$

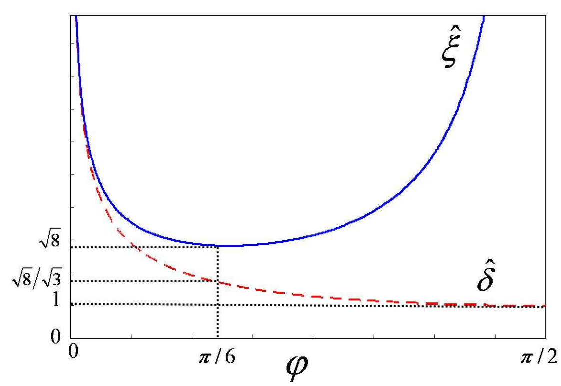


Figure 1: Normalized length scales $\hat{\delta}$ and $\hat{\xi}$ as functions of the time dephasing angle φ. In the limit $\varphi \to \pi/2$, the penetration depth $\delta \to \sqrt{\tau/\mu_0\sigma}$. The oscillation scale ξ exhibits its minimum value at $\varphi = \pi/6$, for which the angular frequency of the monochromatic wave attains $\omega_\mathrm{t} = 1/\tau\sqrt{3}$. Such a quantity is the transition frequency between two distinct regimes for the skin effect: at low frequencies, the fields exihibit the usual attenuated oscillations; at high frequencies, they become purely attenuation in space (see [4]).

Such an interpretation enables one to ascribe to the system the usual transverse dispersion relation [2] $\kappa^2 = \omega^2\mu_0\epsilon_0(\epsilon_r + \imath\epsilon_i)$, where $\epsilon_r + \imath\epsilon_i$ denotes the metallic relative permittivity. Therefore, Eq. (7) leads to

$$\epsilon_r + \imath\epsilon_i = \imath\frac{\sigma\tau}{\epsilon_0}\left(\frac{\cos^2\varphi}{\tan\varphi} + \imath\cos^2\varphi\right). \tag{8}$$

A much interesting result of the present theory is that it provides a simple closed formula for the inertial relaxation time. As one may easily check, Eq. (8) leads to

$$\tau = -\frac{\epsilon_r\hbar}{\epsilon_i\varepsilon}, \tag{9}$$

where, to comply with experimental practice, use has been made of Einstein's relation $\varepsilon = \hbar\omega$ for the photon energy ε, with $\hbar$ denoting the normalized Planck's constant. This means that the inertial relaxation time is fully determined from the metallic permittivity and photon energy. We emphasize that our macroscopic formulation is totally independent from other microscopic approaches, such as Drude's theory of metallic conduction. Within the latter scheme, the relaxation time of the current density is the averaged interval between two successive collisions of free electrons with (essentially) stationary ions (classical collisional regime). On the other hand, our relaxation time is due to inertial effects of charge carriers, and that can be rigorously justified in the realm of extended irreversible thermodynamics (see [3, 4], and references therein).

Now, we calculate the complex refractive index. It can be determined from the complex relative permittivity [2], $n + \imath k = \sqrt{\epsilon_r + \imath\epsilon_i}$. Therefore, the former quantity may be promptly computed from the complex wave number, $\kappa = \omega\sqrt{\mu_0\epsilon_0}(n + \imath k$. Eq. (7) leads to

$$n + \imath k = \sqrt{\frac{\sigma\tau}{\epsilon_0}}\left[\cos\varphi\sqrt{\frac{1-\sin\varphi}{2\sin\varphi}} + \imath\cos\varphi\sqrt{\frac{1+\sin\varphi}{2\sin\varphi}}\right]. \tag{10}$$

From Eq. (10), we see that, when the inertia of charge carriers is taken into account, the standard H-R condition must be extended to [5] $k \geq n \gg 1$.

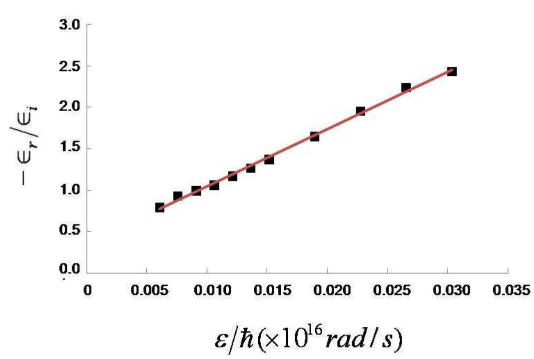


Figure 2: Absolute value for the ratio of the real (negative) to imaginary (positive) parts of the complex relative permittivity for aluminum as a function of the angular frequency. The squared points are plotted according to tabulated values for that metal extracted from [6]. The straight line is drew from linear regression (least squares) of the plotted points. The slope of the straight line gives the inertial relaxation time for aluminum at room temperature, $\tau \sim 6.91$ fs (see Eq. (9)).

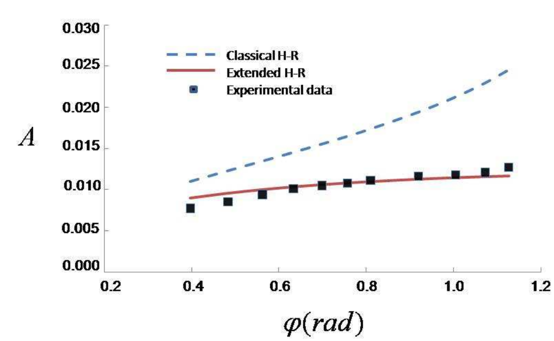


Figure 3: Absorptivity, A, by aluminum as a function of the time dephasing angle, φ. The squared points are plotted according to tabulated values for that metal extracted from [6]. The dashed line gives the prediction for A due to the classical H-R relation, Eq. (1), and the continuous line, the same prediction due to the extended H-R relation, Eq. (12). As it appears, the behavior of the observed absorptivity at normal incidence by aluminum is in excellent agreement with our theory up to the near-infrared region.

Finally, we derive an extended formula for the H-R relation. The absorptivity of monochromatic radiation at normal incidence can be determined from the metallic optical constants [2],

$$A = \frac{4n}{(n+1)^2 + k^2}, \tag{11}$$

where the unit denotes the vacuum refractive index. Then, when the extended H-R condition holds true, Eq. (10) leads to

$$A = 2\sqrt{\frac{\epsilon_0}{\sigma\tau}}\sqrt{\frac{2\sin\varphi}{1+\sin\varphi}}. \tag{12}$$

That is the sought extended H-R relation.

Figure 2 shows the absolute value for the ratio of the real (negative) to imaginary (positive) parts of the complex relative permittivity for aluminum [6] as a function of the angular frequency.

Figure 3 shows the absorptivity A by aluminum [6] as a function of the time dephasing angle φ.

4. CONCLUSION

Applications of our theory to other systems can lead to advances in extending the classical theory for the optics of metals. An interesting possibility is the analysis of electronic correlations in metallic transport. Indeed, within Landau's theory of Fermi liquid, those processes are described by assuming an effective mass, as well as an effective scattering time, for charge carriers.

REFERENCES

1. Hagen, E. and H. Rubens, "Über die Beziehung des Reflexions- und Emissionsvermögens der Metalle zu ihrem elektrishen Leitvermögen," *Ann. d. Physik*, Vol. 4, No. 11, 873, 1903.
2. Born, M. and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th Edition, Cambridge University Press, Cambridge, 1999.
3. Silveira, F. E. M. and J. A. S. Lima, "Attenuation and damping of electromagnetic fields: Influence of inertia and displacement current," *J. Phys. A: Math. Theor.*, Vol. 42, 095402, 2009.
4. Silveira, F. E. M. and J. A. S. Lima, "Skin effect from extended irreversible thermodynamics perspective," *Journal of Electromagnetic Waves and Applications*, Vol. 24, Nos. 2–3, 151–160, 2010.
5. Silveira, F. E. M. and S. M. Kurcbart, "Hagen-Rubens relation beyond far-infrared region," *EPL*, Vol. 90, 44004, 2010.
6. Weaver, J. H. and H. P. R. Frederikse, "Optical properties of selected elements," *CRC Handbook of Chemistry and Physics*, 90th Edition, edited by Lide, D. R., CRC Press, Boca Raton, 2010.

Enhanced SBS Instability Growth Rate of Extraordinary Electromagnetic Waves in Strongly Coupled, Magnetized Plasma

M. S. Bawa'aneh[1, 2], **Ibrahim Y. Abualhaol**[1], **and Feras Al-Dweri**[2]

[1]Khalifa University of Science, Technology and Research, Sharjah, P. O. Box 573, United Arab Emirates
[2]Department of Physics, The Hashemite University, Zarqa, Jordan

Abstract— Stimulated Brillouin Backscattering (SBBS) of a large amplitude, extraordinary electromagnetic wave travelling across an ambient dc-magnetic field in a strongly coupled plasma is investigated. Using a magnetohydrodynamic model, a system of coupled equations that describes the problem is derived and then solved for a strongly coupled regime. The SBBS maximum growth rate in a magnetized plasma is obtained numerically and compared to that for non-magnetized plasma known in literature. Result shows an enhancement in the instability growth rate in the presence of the magnitic field.

1. INTRODUCTION

Parametric instabilities occur when the incident electromagnetic wave resonantly decays into a scattered electromagnetic wave and an electrostatic plasma mode, which, in laser-produced plasma, could lead to possible absorption, as well as scattering of the wave [1–7]. Stimulated Brillouin scattering (SBS) of electromagnetic waves in plasma is the parametric instability, where an incident light wave with frequency ω_0 and wavenumber $\vec{k}_0$ decays resonantly into a scattered light wave with frequency ω_1 and wavenumber $\vec{k}_1$ and an ion acoustic wave with $(\omega, \vec{k})$. The matching conditions for the frequencies and wave numbers are given by the following equations;

$$\omega_0 = \omega_1 + \omega \quad \text{and} \quad \vec{k}_0 = \vec{k}_1 + \vec{k}, \tag{1}$$

where the equation of the wave number matching condition for stimulated Brillouin bachscattering (SBBS) reads $k_1 = k_0 + k$.

In inertial confinement fusion (ICF), since the driving energy for the implosion is provided by the incident light wave, the occurance of backscattered light could greatly reduce the laser energy absorption efficiency. The full understanding and controll of SBBS instability growth is still of a fundamental concern in more understanding of ICF. This makes SBBS recieve a great deal of attention both theoretically [8–11] and experimentally [12–14]. Megagauss magnetic fields, that have been known to exist for a long time in laser-produced plasmas [15–19], have significant effects on the dynamics of the plasma; The dispersion relations of electromagnetic waves inside the plasma are greatly modified by the presence of such fields, hence magnetized plasma. In previous studies [7, 9], Bawa'aneh et al. have dealt analytically with the SBBS and filamentation problem considering weak field coupling in magnetized plasma. In the present work, SBBS instability in strongly coupled, magnetized plasma is considered. The problem is formulated in Section 2 and a system of coupled equations that describes SBBS in magnetized palsma is derived. The dispersion relation of SBBS in strongly coupled, magnetized plasma is obtained, compared with previous result in the literature for the non-magnetized case and solved for SBBS groth rate in Section 3, where the effect of the dc-magnetic field is obtained numerically. Finally, the results are discussed in Section 4.

2. THE MODEL

To study the coupling of a large amplitude laser beam, propagating in a plasma across a magnetic field, into a scattered electromagnetic wave and an ion acoustic wave we begin with Ampere's and Faraday's laws, given respectively by $\vec{\nabla} \times \vec{B} = \mu_0 \vec{J} + \frac{1}{c^2}\frac{\partial \vec{E}}{\partial t}$ and $\frac{\partial \vec{B}}{\partial t} = -\vec{\nabla} \times \vec{E}$, and the definition of the electric current density in plasma, given by $\vec{J} = q_j n_j \vec{v}_j$, where j stands for electrons and ions, q_j, n_j and $\vec{v}_j$ are the charge, density and velocity for the jth species of the plasma, respectively. This yields the following equation governing the electric field in the plasma;

$$\frac{\partial^2 \vec{E}}{\partial t^2} + c^2 \vec{\nabla} \times \left(\vec{\nabla} \times \vec{E}\right) = -\frac{q_j}{\epsilon_0} \frac{\partial \left(n_j \vec{v}_j\right)}{\partial t} \tag{2}$$

In this equation, the electric field is coupled with the charge, density and velocity as seen on the right hand side, which follow from the force equation and the continuity equation, given for the jth specieis, respectively, by

$$\frac{\partial n_{\mathrm{j}}}{\partial t} + \vec{\nabla}\cdot(n_{\mathrm{j}}\vec{v}_{\mathrm{j}}) = 0 \tag{3}$$

and

$$m_j\left(\frac{\partial \vec{v}_j}{\partial t} + \left(\vec{v}_j\cdot\vec{\nabla}\right)\vec{v}_j\right) = q_j\vec{E} + q_jn_j\vec{v}_j\times\vec{B} - \frac{\vec{\nabla}p_j}{n_j} + m_j\nu\vec{v}_j, \tag{4}$$

where ν is the effective electron-ion collisional frequency. The closed system of equations, namely Equations (2)–(4), can be linearized following the standard procedure known in literature [see for example Refs. [6, 7]] considering immobile, neutralizing background ions in the fast time scale of the electron motion and the coupling of the incident laser beam with the plasma to come through the electrons [6]. Considering magnetized plasma (with a static dc-magnetic field be given along the z-axis), both longitudinal and transverse components of the electron perturbation velocity take high and low frequency components $[\vec{v}_{1e} = (v^h_{ex} + v^l_{ex})\hat{x} + (v^h_{ey} + v^l_{ey})\hat{y}]$, where the number 1 in the subscript denotes a perturbed quantity and the superscripts l and h denote low and high frequency, respectively. The ion perturbation velocity is $v_{ix}\hat{x} + v_{iy}\hat{y}$, where v_{ix}, v_{iy} are the low frequency ion velocity components in the x and y-directions, respectively. Also, the electric field is given by $\vec{E}_1 = E_x\hat{i} + E_y\hat{j}$, where both E_x and E_y have low and high frequency components. The presence of such high and low components for the electron and ion perturbation velocities and the electric field are well explained in Ref. [7].

We consider now the geometry where electrostatic fluctuations are along the x-axis, the incident laser beam is polarized along the y-direction, where both the pump field and the direction of propagation are perpendicular to the static dc-magnetic field. At modest pump power, the nonlinear behaviour would be negligible and linear theory would be good enough to describe the system. The linearized form of Equations (2) and (4) give the x and y-components for the low and high frequency electric field, and for the elctron low and highy frequency velocity components, respectively, and the linearized form of the force equation, namely Equation (4), gives the x and y-components for the ion low frequency velocity [see Ref. [7]]. A very lengthy process of vector algebra done by simplifying the electron and ion velocities to the first order that leads to first order iterational solution of the equations describing the low and high frequency perturbation in the electric field and the low frequency density perturbation that is used in literature [4–7], ignoring collisions, considering $Zn_{0i} = n_{0e} \to n_0$, where the number 0 in the subscript denotes an equilibrium quantity, the plasma approximation $Zn_{1i} \approx n_{1e} \to n_1$, a linearly polarized pump of the form $\vec{E}_0(x,t) = \hat{j}\ E_0\cos(k_0x - \omega_0t) = \frac{1}{2}\left(E_0\ e^{i(k_0x-\omega_0t)} + c.c.\right)\ \hat{j}$, where $cc.$ denotes complex conjugate, also considered the dispersion relations of the X-waves and ion cyclotron waves that occure in magnetized plasma, and finally considering resonant terms only, the equation that governs the high frequency electric field perturbations of the backscattered wave yields

$$\left[\frac{\partial^2}{\partial t^2} - c^2\frac{\partial^2}{\partial x^2} + \omega^2_{\mathrm{pe}}\frac{\omega_1^2 - \omega^2_{\mathrm{pe}}}{\omega_1^2 - \omega^2_{UH}}\right]E_1 = \frac{-iev_{0e}\omega_1}{\epsilon_0}n_1^*\,, \tag{5}$$

and that governing the low frequency ion density perturbation yields

$$\left[\frac{\partial^2}{\partial t^2} - c_s^2\frac{\partial^2}{\partial x^2} - \Omega_i^2\right]n_1 = \frac{iZen_0k^2v_{0e}}{m_i}\frac{E_1^*}{\omega_1}\frac{\omega_1^2 - \omega^2_{\mathrm{pe}}}{\omega_1^2 - \omega^2_{UH}}\left(1 + \frac{\Omega_i^2}{c_s^2k^2}\right)\,, \tag{6}$$

where $c_s = \sqrt{(\gamma_ek_BT_e + \gamma_ik_BT_i)/m_i}$ is the ion acoustic speed, $\omega^2_{UH} = \omega^2_{\mathrm{pe}} + \Omega^2_e$ the upper hybrid frequency and γ_j the specific heat ratio for the jth species. Equations (5) and (6) form a closed system of equations. As seen in the system of equations, eleminating the magnetic field, the equations will reduce to the well known system of coupled equations for SBS in nonmagnetized plasmas known in literature [6].

3. SBS GROWTH RATE IN STRONGLY COUPLED PLASMA

In Section 2, we have derived the system of equations that describes SBBS instability growth rate in weakly magnetized plasma, namely Equations (5) and (6). Upon using the fourier transformation in space and time, making use of the frequency and wave number matching conditions of Equation (1), and solving the resulting two linear equations for the electric field and density simultaneously, our system of equations yields the following dispersion relation

$$\left(\omega_1^2 - c^2k_1^2 - \omega_{pe}^2\frac{\omega_1^2-\omega_{\rm pe}^2}{\omega_1^2-\omega_{UH}^2}\right)\left(\omega^2 - c_s^2k^2 - \Omega_i^2\right) = \frac{1}{4}\omega_{pi}^2k^2\upsilon_{os}^2\frac{\omega_1^2-\omega_{\rm pe}^2}{\omega_1^2-\omega_{UH}^2}\left(1+\frac{\Omega_i^2}{c_s^2k^2}\right) \qquad (7)$$

This dispersion relation reduces, when the magnetic field is ignored, to that formula well known in literature for nonmagnetized plasma [see Refs. [5, 6] for example], which has been analyzed analytically for weak and strong plasma coupling. Equation (7) has been analyzed for weakly coupled plasma [9], where the instabiliy growth rate is much smaller than the ion acoustic frequency. However, in this work we analyze Equation (7) analytically for strongly coupled plasma.

Following [6], we consider the strong coupling condition $|\omega| \gg kc_s$. The first term on the left hand side of Equation (7) reduces to $-2\omega_0\omega - (4\omega_0^2c_s^2/c^2)(1+(\Omega_e^2\omega_{pe}^2)/(\omega_0^2-\Omega_i^2))$, and the second term on the left hand side reduces to $\omega^2+\Omega_i^2$. Also, the maximum growth rate occures when the electromagnetic wave is resonant, with k taking the following values; $k \approx 2k_0 - (2\omega_0c_s/c^2)(1+\alpha)$, where $\alpha = \left(\Omega_e\omega_{\rm pe}/(\omega_0^2-\Omega_e^2)\right)^2$. Back substitution in Equation (7), considering $\beta = (\omega_1^2 - \omega_{\rm pe}^2)/(\omega_1^2-\omega_{UH}^2)$, yields the following cubic equation;

$$\omega^3 - D_2\omega^2 - D_1\omega + D_0 = 0\ , \qquad (8)$$

where

$$D_2 = k_0c_s\frac{\Omega_e}{\omega_0}(1+\alpha)\left(1+\frac{\omega_0}{k_0c}(1+\alpha)\right) \qquad (9)$$

$$D_1 = \Omega_i^2 \qquad (10)$$

$$D_0 = D_2D_1 + \frac{k_0^2\upsilon_{os}^2}{2}\frac{\omega_{pi}^2}{\omega_0}\beta\left(1+\frac{\Omega_i^2}{c_s^2k^2}\right) + \frac{1}{2}\omega_{pi}^2\beta\left(\frac{c_s\upsilon_{os}}{c^2}\right)^2(1+\alpha)(\omega_0-4ck_0)\left(1+\frac{\Omega_i^2}{c_s^2k^2}\right), \qquad (11)$$

and α and β are given by the expressions in the paragraph just above Equation (8). In Equation (8), ignoring the magnetic field will yield the frequency for the strongly coupled, nonmagnetized plasma that is well known in the literature [6] with one difference that we use $k \approx 2k_0 - (2\omega_0c_s/c^2)(1+\alpha)$, while Ref. [6] uses $k \approx 2k_0$.

The numerical solution of the cubic Equation (8) shows one real root and two complex conjugates, as expected. Figure 1 shows the imaginary part of the unstable root of Equation (8). It represents the growth rate (normalized to γ_0, the imaginary part of ω for non magnetized plasma [see Ref. [6]]) versus the magnetic field given in terms of the electron cyclotron frequency Ω_e normalized to incident laser frequency ω_0. for the different plasma densities $n/n_{cr} = 0.1, 0.4, 0.8$, respectively,

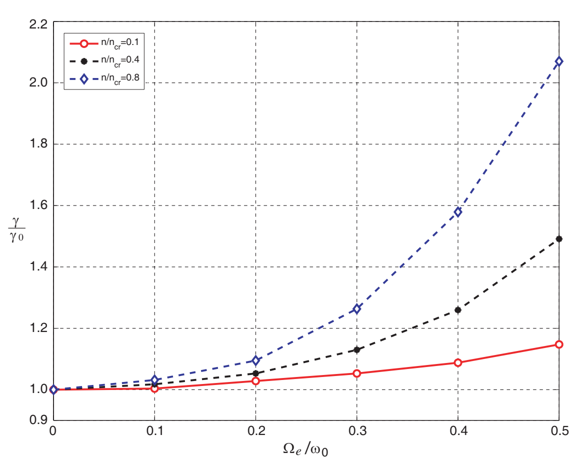


Figure 1: Normalized maximum growth rate versus normalized electron cyclotron frequency for different densities. $I_0 = 10^{15}\,\mathrm{W/cm^2}$, $\lambda_0 = 1.06\,\mu\mathrm{m}$, $T_e = 3\,\mathrm{keV}$ and $T_i = 0.3T_e$.

where the solid line corresponds to $n = 0.1n_{cr}$. In obtaining the figure, the following parameters, typical for ICF, are used for the incident laser intensity and wavelength; $I_0 = 10^{15}\,\mathrm{W/cm^2}$, $\lambda_0 = 1.06$ microns, and the following for the electron and ion temperatures; $T_e = 3\,\mathrm{keV}$ and $T_i/T_e = 0.3$, respectively. The figure shows enhancement of the instability by the magnetic field. This effect is stronger for plasma with higher density.

4. CONCLUSION

Stimulated Brillouin Backscattering (SBBS) of a large amplitude, extraordinary electromagnetic wave travelling across an ambient dc-magnetic field in a strongly coupled plasma is investigated. In this study, we consider a static magnetic field perpendicular to the direction of the electrostatic density fluctuation and to the direction of polarization of the incident laser field. Using a magneto-hydrodynamic model, a system of coupled equations that describes the problem is derived [7]. This system of equations is then solved for a strongly coupled regime to obtain a modified expression for the maximum growth rate in magnetized plasmas, where it is found to be the solution of a cubic equation. In the absence of the magnetic feld, both the system of coupled equations and the cubic equation for the maximum growth rate, reduce to the corresponding expressions known in the literature for SBBS in non-magnetized plasmas [6].

The numerical solution of the cubic equation shows a real solution and two complex conjugates. For the unstable mode, the static magnetic field is found to enhance the instability growth rate. The growth rate doubles at a value of Ω_e/ω_0 just below 0.5 for $n = 0.8n_{cr}$.

REFERENCES

1. Perkins, F. W. and J. Flick, "Parametric instabilities in inhomogeneous plasmas," *Phys. Fluids*, Vol. 14, 2012–2019, 1971.
2. Forslund, D. W., J. M. Kindel, and E. L. Lindman, "Parametric excitation of electromagnetic waves," *Phys. Rev. Lett.*, Vol. 29, 249–252, 1972.
3. Rosenbluth, M. N., "Parametric instabilities in inhomogeneous media," *Phys. Rev. Lett.*, Vol. 29, 565–567, 1972.
4. Chen, F. F., *Introduction to Plasma Physics*, Plenum, New York, 1985.
5. Banmgärtel, K. and K. Sauer, *Topics on Nonlinear Wave — Plasma Interaction*, Birkhäuser, Berlin, Germany, 1987.
6. Kruer, W. L., *The Physics of Laser Plasma Interactions*, Westview, Boulder, CO, 2003.
7. Bawa'aneh, M. S., G. Assayed, and S. Al-Awfi, "Filamentation instability of electromanetic radiation in magnetized plasma," *IEEE Transactions on Plasma Science*, Vol. 38, No. 5, 1066–1072, 2010.
8. Tikhonchuk, V. T., J. Fuchs, C. Labaune, S. Depierreux, S. Hller, J. Myatt, and H. A. Baldis, "Stimulated brillouin and raman scattering from a randomized laser beam in large inhomogeneous collisional plasmas. II. Model description and comparison with experiment," *Phys. Plasmas*, Vol. 8, 1636–1350, 2001.
9. Bawa'aneh, M. S., "Stimulated brillouin backscattering in magnetized plasma," *Contrib. Plasma Phys.*, Vol. 43, 447–455, 2003.
10. Maximov, A. V., J. Myatt, W. Seka, R. W. Short, and R. S. Craxton, "Modeling of stimulated Brillouin scattering near the critical-density surface in the plasmas of direct-drive inertial confinement fusion targets," *Phys. Plasmas*, Vol. 11, No. 6, 2994–3000, Jun. 2004.
11. Bawa'aneh, M. S. and T. J. M. Boyd, "Enhanced levels of stimulated brillouin reflectivity from non-maxwellian plasmas," *Journal of Plasma Physics*, Vol. 73, 159–167, 2007.
12. Fuchs, J., C. Labaune, S. Depierreux, H. A. Baldis, and A. Michard, "Modification of spatial and temporal gains of stimulated brillouin and raman scattering by polarization smoothing," *Phys. Rev. Lett.*, Vol. 84, 3089–3092, 2000.
13. Malka, V., J. Faure, S. Hüller, V. T. Tikhonchuk, S. Weber, and F. Amiranoff, "Enhanced spatiotemporal laser-beam smoothing in gas-jet plasmas," *Phys. Rev. Lett.*, Vol. 90, 075002–075006, 2003.
14. Maximov, A. V., I. G. Ourdev, D. Pesme, W. Rozmus, V. T. Tikhonchuk, and C. E. Capjack, "Plasma induced smoothing of a spatially incoherent laser beam and reduction of backward stimulated Brillouin scattering," *Phys. Plasmas*, Vol. 8, 1319–1329, 2001.
15. Stamper, J. A. and B. H. Ripin, "Faraday-rotation measurements of megagauss magnetic fields in laser-produced plasmas," *Phys. Rev. Lett.*, Vol. 34, 138–141, 1975.

16. Boyd, T. J. M., G. J. Humphreys-Jones, and D. Cooke, "Structure of magnetic fields generated in laser produced plasmas," *Phys. Lett. A*, Vol. 88, 140–143, 1982.
17. Pukov, A. and J. Meyer-ter-Vehn, "Laser hole boring into overdense plasma and relativistic electron currents for fast ignition of ICF targets," *Phys. Rev. Lett.*, Vol. 79, No. 14, 2686–2689, Oct. 1997.
18. Boyd, T. J. M., A. Tatarinov, M. S. Bawa'aneh, and A. Dyson, "Magnetic field generation and penetration into dense, laser-produced plasma," *Proceedings of the 1996 International Conference on Plasma Physics*, edited by H. Sugai and T. Hayashi, Vol. 2, 1714–1718, Japan Society of Plasma Science and Nuclear Fusion Research, Nagoya, 1997.
19. Bawa'aneh, M. S., "Magnetohydrodynamics of megagauss magnetic fields generated in plasma," *IEEE Transactions on Plasma Science*, Vol. 38, No. 8, 1808–1814, 2010.

On the Electrodinamics of Counter Propagating Transverse-electric and Transverse-magnetic Waves in the Absorbing Plate in a Waveguide

E. A. Gevorkyan
Moscow State University of Economics, Statistics and Informatics
7 Nezhinskaya Str., Moscow 119501, Russia

Abstract— Propagation of two coherent counter propagating transverse-electric (TE) and transverse-magnetic (TM) electromagnetic waves in the absorbing plate placed in a waveguide of arbitrary cross section is considered. It is supposed that the electromagnetic waves of various initial phases incident on the plate from two sides. Fresnel formulas and analytical expressions for the power reflection, transmission and transmission interference coefficients are obtained. Some physical features of tunnel interference effect in the absorbing plate are discussed.

1. INTRODUCTION

In the articles [1, 2] various aspects of interaction of two coherent counter propagating electromagnetic waves with different initial phases in a plane-parallel plate in unlimited space are studied. In this article the similar problem for two transverse-magnetic (TM) and transverse-electric (TE) waves, incident on absorbing plate in a waveguide from two sides, is solved.

Suppose the generatrix of regular waveguide of arbitrary cross section coincides with OZ axis of some rectangular system of coordinates and the absorbing plate with thickness d occupies the region $-d/2 \leq Z \leq d/2$. We shall assume that the nonmagnetic ($\mu = 1$) plate has the constant permittivity ε, while outside of the plate in a waveguide $\varepsilon = 1,\ \mu = 1$. Let's consider propagation of two counter synchronous transverse-magnetic (TM) and transverse-electric (TE) waves **a** and **b** with frequency ω and initial phases φ and ψ in an absorbing plate in a waveguide, when the wave a incident on the plate from the side $z \leq -d/2$ and the wave **b** — from the side $z \geq d/2$ (Fig. 1).

Note, that if, as in our earlier articles (see, e.g., [8, 9]) the transverse-electric (TE) and transverse-magnetic (TM) fields in the absorbing plate in the waveguide are described through the longitudinal components of the magnetic and electric vectors ($H_z(x, y, z)$ and $E_z(x, y, z)$), then from Maxwell's equations it is possible to receive the following wave equations for H_z and E_z [3, 4]

$$\Delta_\perp H_z + \frac{\partial^2 H_z}{\partial z^2} - \frac{\varepsilon}{c^2}\frac{\partial^2 H_z}{\partial t^2} = \frac{4\pi\sigma}{c^2}\frac{\partial H_z}{\partial t}, \quad \Delta_\perp E_z + \frac{\partial^2 E_z}{\partial z^2} - \frac{\varepsilon}{c^2}\frac{\partial^2 E_z}{\partial t^2} = \frac{4\pi\sigma}{c^2}\frac{\partial E_z}{\partial t}, \tag{1}$$

where $\Delta_\perp = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the two-dimensional Laplace operator, c is the velocity of light in vacuum, σ is the electrical conductivity of the absorbing plate. The solutions of the Equation (1) we shall look for in the form

$$H_z(x, y, z, t) = \sum_{n=0}^{\infty} H_n(z, t)\,\widehat{\psi}_n(x, y) = \sum_{n-0}^{\infty} H_n(z)\,e^{-i\omega t}\widehat{\psi}_n(x, y), \tag{2}$$

$$E_z(x, y, z, t) = \sum_{n=0}^{\infty} E_n(z, t)\,\psi_n(x, y) = \sum_{n-0}^{\infty} E_n(z)\,e^{-i\omega t}\psi_n(x, y), \tag{3}$$

where the orthonormal eigenfunctions $\widehat{\psi}_n(x, y)$ and $\psi_n(x, y)$ of the second and first boundary-value problems for the cross section of a waveguide satisfy the Helmholtz equations (see formula (10) and (11) in [3]). Substitution (2) and (3) in (1) leads us to the following equations

$$\frac{d^2 H_n(z)}{dz^2} + \widehat{P}_n^2 H_n(z) = 0, \quad \frac{d^2 E_n(z)}{dz^2} + P_n^2 E_n(z) = 0, \tag{4}$$

where

$$\widehat{P}_n^2 = \frac{\omega^2}{c^2}\varepsilon - \widehat{\lambda}_n^2 + i\frac{4\pi\sigma\omega}{c^2}, \quad P_n^2 = \frac{\omega^2}{c^2}\varepsilon - \lambda_n^2 + i\frac{4\pi\sigma\omega}{c^2}, \tag{5}$$

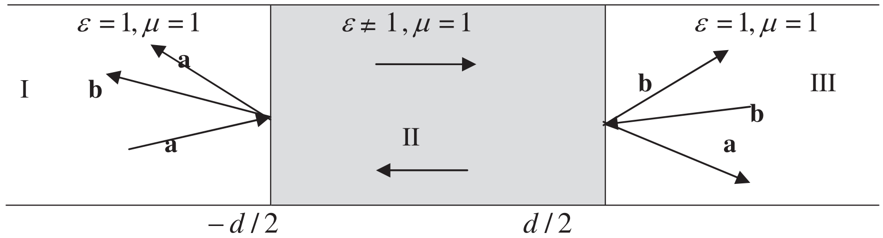


Figure 1: Two opposing coherent waves **a** and **b** propagate in the absorbing plate in a waveguide.

$\widehat{\lambda}_n$ and λ_n are the eigenvalues of the second and first boundary-value problems for the cross section of the waveguide, corresponding eigenfunctions $\widehat{\psi}_n(x, y)$ and $\psi_n(x, y)$. Note that the transverse components of the TE and TM fields in this case are expressed by formulas (6), (7), (8) and (9) of [3], if in them to substitute $\mu = 1$. The Equation (4) have the solutions in the form

$$H_n(z) = \widehat{A}_n e^{i\widehat{p}_n z}, \quad E_n(z) = A_n e^{ip_n z}, \tag{6}$$

where $\widehat{A}_n$ and A_n are the known complex amplitudes. Let the complex wave numbers $\widehat{P}_n$ and P_n present as $\widehat{P}_n = \widehat{\alpha}_n + i\widehat{\beta}_n$ and $P_n = \alpha_n + i\beta_n$. Then, taking into account (5) we have

$$\widehat{\alpha}_n = \frac{1}{\sqrt{2}}\left[\sqrt{\left(\frac{\omega^2}{c^2}\varepsilon - \widehat{\lambda}_n^2\right)^2 + \frac{16\pi^2\sigma^2\omega^2}{c^4}} + \left(\frac{\omega^2}{c^2}\varepsilon - \widehat{\lambda}_n^2\right)\right]^{\frac{1}{2}},$$

$$\widehat{\beta}_n = \frac{1}{\sqrt{2}}\left[\sqrt{\left(\frac{\omega^2}{c^2}\varepsilon - \widehat{\lambda}_n^2\right)^2 + \frac{16\pi^2\sigma^2\omega^2}{c^4}} - \left(\frac{\omega^2}{c^2}\varepsilon - \widehat{\lambda}_n^2\right)\right]^{\frac{1}{2}}, \tag{7}$$

$$\alpha_n = \frac{1}{\sqrt{2}}\left[\sqrt{\left(\frac{\omega^2}{c^2}\varepsilon - \lambda_n^2\right)^2 + \frac{16\pi^2\sigma^2\omega^2}{c^4}} + \left(\frac{\omega^2}{c^2}\varepsilon - \lambda_n^2\right)\right]^{\frac{1}{2}},$$

$$\beta_n = \frac{1}{\sqrt{2}}\left[\sqrt{\left(\frac{\omega^2}{c^2}\varepsilon - \lambda_n^2\right)^2 + \frac{16\pi^2\sigma^2\omega^2}{c^4}} - \left(\frac{\omega^2}{c^2}\varepsilon - \lambda_n^2\right)\right]^{\frac{1}{2}}. \tag{8}$$

The electromagnetic field of TE and TM waves in different regions of a waveguide can be represented as (see Fig. 1).

I region, $z \le -d/2,\ \varepsilon = 1,\ \mu = 1$

$$H_n^I(z) = \widehat{A}_n^a e^{i\widehat{\Gamma}_n z} + \widehat{D}_n^{ab} e^{-i\widehat{\Gamma}_n z}, \quad E_n^I(z) = A_n^a e^{i\Gamma_n z} + D_n^{ab} e^{-i\Gamma_n z}, \tag{9}$$

where

$$\widehat{D}_n^{ab} = \widehat{B}_n^a + \widehat{C}_n^b, \widehat{\Gamma}_n = \sqrt{(\omega/c)^2 - \widehat{\lambda}^2}, \quad D_n^{ab} = B_n^a + C_n^b, \quad \Gamma_n = \sqrt{(\omega/c)^2 - \lambda^2}, \tag{10}$$

$\widehat{A}_n^a = \left|\widehat{A}_n^a\right| e^{i\varphi}$ and $A_n^a = |A_n^a|\, e^{i\varphi}$ are the known complex amplitudes of the wave **a**, incident on the plate, $\widehat{B}_n^a$ and B_n^a are yet unknown complex amplitudes of the wave **a**, reflected from the plate in region I, $\widehat{C}_n^b$ and C_n^b are yet unknown complex amplitudes of the wave **b**, transmitted in region I.

II region, $-d/2 \le z \le d/2,\ \varepsilon = \text{const},\ \mu = 1$

$$H_n^{II}(z) = \widehat{F}_n^{ab} e^{i\widehat{p}_n z} + \widehat{F}_n^{ba} e^{-i\widehat{p}_n z}, \quad E_n^{II}(z) = F_n^{ab} e^{ip_n z} + F_n^{ba} e^{-ip_n z}, \tag{11}$$

where

$$\widehat{F}_n^{ab} = \widehat{M}_n^a + \widehat{N}_n^b, \quad \widehat{F}_n^{ba} = \widehat{M}_n^b + \widehat{N}_n^a, \quad F_n^{ab} = M_n^a + N_n^b, \quad F_n^{ba} = M_n^b + N_n^a, \tag{12}$$

$\widehat{M}^a_n$, $\widehat{M}^b_n$, $\widehat{N}^a_n$, $\widehat{N}^b_n$, M^a_n, M^b_n, N^a_n and N^b_n are yet unknown complex amplitudes of the waves in the absorbing plate.

III region, $z \geq d/2$, $\varepsilon = 1$, $\mu = 1$

$$H^{II}_n(z) = \widehat{F}^{ab}_n e^{i\widehat{p}_n z} + \widehat{F}^{ba}_n e^{-i\widehat{p}_n z}, \quad E^{III}_n(z) = A^b_n e^{-i\Gamma_n z} + D^{ba}_n e^{i\Gamma_n z}, \tag{13}$$

where $\widehat{D}^{ba}_n = \widehat{B}^b_n + \widehat{C}^a_n$, $D^{ba}_n = B^b_n + C^a_n$, $\widehat{A}^b_n = |\widehat{A}^b_n| e^{i\psi}$ and $A^b_n = |A^b_n| e^{i\psi}$ are the known complex amplitudes of the wave **b**, $\widehat{B}^b_n$ and B^b_n are yet unknown complex amplitudes of the wave **b**, reflected from the plate in a region III, $\widehat{C}^a_n$, C^a_n are yet unknown complex amplitudes of the wave **a**, transmitted in region III.

Now requiring that (9), (11) and (13) satisfied to following boundary conditions at $z = \pm d/2$

$$H^I_n(z) = H^{II}_n(z), \quad \frac{dH^I_n(z)}{dz} = \frac{dH^{II}_n(z)}{dz}, \quad H^{II}_n(z) = H^{III}_n(z), \quad \frac{dH^{II}_n(z)}{dz} = \frac{dH^{III}_n(z)}{dz}, \tag{14}$$

$$E^I_n(z) = \varepsilon E^{II}_n(z), \quad \frac{dE^I_n(z)}{dz} = \frac{dE^{II}_n(z)}{dz}, \quad E^{III}_n(z) = \varepsilon E^{II}_n(z), \quad \frac{dE^{III}_n(z)}{dz} = \frac{dE^{II}_n(z)}{dz}, \tag{15}$$

we shall receive the system of algebraic equations for determining the unknown amplitudes. Solving this system and having received analytical expressions for unknown amplitudes, we shall find the power reflection and transmission coefficients for the wave **a** in the form of

$$\widehat{R}^a_n = \frac{\widehat{S}^{ref\cdot a}_{zn}}{\widehat{S}^{inc\cdot a}_{zn}} = \frac{\left|\widehat{B}^a_n\right|^2}{\left|\widehat{A}^a_n\right|^2} = \frac{ch 2\widehat{\beta}_n d - \cos 2\widehat{\alpha}_n d}{ch\left(\widehat{\psi}_n + 2\widehat{\beta}_n d\right) - \cos\left(\widehat{\varphi}_n - 2\widehat{\alpha}_n d\right)}, \tag{16}$$

$$R^a_n = \frac{S^{ref\cdot a}_{zn}}{S^{inc\cdot a}_{zn}} = \frac{|B^a_n|^2}{|A^a_n|^2} = \frac{ch 2\beta_n d - \cos 2\alpha_n d}{ch\left(\psi_n + 2\beta_n d\right) - \cos\left(\varphi_n - 2\alpha_n d\right)}, \tag{17}$$

$$\widehat{T}^a_n = \frac{\widehat{S}^{trans\cdot a}_{zn}}{\widehat{S}^{inc\cdot a}_{zn}} = \frac{\left|\widehat{C}^a_n\right|^2}{\left|\widehat{A}^a_n\right|^2} = \frac{8\widehat{\Gamma}^2_n\left(\widehat{\alpha}^2_n + \widehat{\beta}^2_n\right)}{\left[\left(\widehat{\alpha}^2_n + \widehat{\beta}^2_n + \widehat{\Gamma}^2_n\right)^2 - 4\widehat{\alpha}^2_n\widehat{\Gamma}^2_n\right]\left[ch\left(\widehat{\psi}_n + 2\widehat{\beta}_n d\right) - \cos\left(\widehat{\varphi}_n - 2\widehat{\alpha}_n d\right)\right]}, \tag{18}$$

$$T^a_n = \frac{S^{trans\cdot a}_{zn}}{S^{inc\cdot a}_{zn}} = \frac{|C^a_n|^2}{|A^a_n|^2} = \frac{8\varepsilon^2\Gamma^2_n\left(\alpha^2_n + \beta^2_n\right)}{\left[\left(\alpha^2_n + \beta^2_n + \varepsilon^2\Gamma^2_n\right)^2 - 4\alpha^2_n\varepsilon^3\Gamma^2_n\right]\left[ch\left(\psi_n + 2\beta_n d\right) - \cos\left(\varphi_n - 2\alpha_n d\right)\right]}, \tag{19}$$

where

$$tg\widehat{\varphi}_n = \frac{4\widehat{\beta}_n\widehat{\Gamma}_n\left(\widehat{\Gamma}^2_n - \widehat{\alpha}^2_n - \widehat{\beta}^2_n\right)}{\left(\widehat{\Gamma}^2_n - \widehat{\alpha}^2_n - \widehat{\beta}^2_n\right)^2 - 4\widehat{\beta}^2_n\widehat{\Gamma}^2_n}, \quad tg\widehat{\psi}_n = \frac{4\widehat{\alpha}_n\widehat{\Gamma}_n\left(\widehat{\Gamma}^2_n + \widehat{\alpha}^2_n + \widehat{\beta}^2_n\right)}{\left(\widehat{\Gamma}^2_n + \widehat{\alpha}^2_n\widehat{\beta}^2_n\right)^2 + 4\widehat{\alpha}^2_n\widehat{\Gamma}^2_n}, \tag{20}$$

$$tg\varphi_n = \frac{4\beta_n\varepsilon\Gamma_n\left(\varepsilon^2\Gamma^2_n - \alpha^2_n - \beta^2_n\right)}{\left(\varepsilon^2\Gamma^2_n - \alpha^2_n - \beta^2_n\right)^2 - 4\beta^2_n\varepsilon^2\Gamma^2_n}, \quad tg\psi_n = \frac{4\alpha_n\varepsilon\Gamma_n\left(\varepsilon^2\Gamma^2_n + \alpha^2_n + \beta^2_n\right)}{\left(\varepsilon^2\Gamma^2_n + \alpha^2_n\beta^2_n\right)^2 + 4\alpha^2_n\varepsilon^2\Gamma^2_n}. \tag{21}$$

Now by means of Umov-Poynting vector and in view of (13) we can find streams of electromagnetic energy in the region III. Calculations result in the following

$$\widehat{S}^{III}_{zn} = \frac{1}{8\pi}\cdot\frac{\omega\widehat{\Gamma}_n}{\widehat{\lambda}^2_n}\left|\widehat{C}^a_n\right|^2 + \frac{1}{8\pi}\cdot\frac{\omega\widehat{\Gamma}_n}{\widehat{\lambda}^2_n}\left|\widehat{B}^b_n\right|^2 + \frac{1}{4\pi}\cdot\frac{\omega\widehat{\Gamma}_n}{\widehat{\lambda}^2_n}\mathrm{Re}\left(\widehat{C}^a_n\bar{\widehat{B}}^b_n\right), \tag{22}$$

$$S^{III}_{zn} = \frac{1}{8\pi}\cdot\frac{\omega\Gamma_n}{\widehat{\lambda}^2_n}|C^a_n|^2 + \frac{1}{8\pi}\cdot\frac{\omega\Gamma_n}{\widehat{\lambda}^2_n}\left|B^b_n\right|^2 + \frac{1}{4\pi}\cdot\frac{\omega\Gamma_n}{\widehat{\lambda}^2_n}\mathrm{Re}\left(C^a_n\bar{B}^b_n\right), \tag{23}$$

where the last terms are the interference streams, resulting from the interaction of waves **a** and **b** in the absorbing plate in the waveguide and the quantities $\bar{\widehat{B}}_n^b$ and $\bar{B}_n^b$ are the complex-conjugate expressions of $\widehat{B}_n^b$ and B_n^b. So we have

$$\widehat{S}_{zn}^{inter.} = \frac{1}{4\pi}\cdot\frac{\omega\widehat{\Gamma}_n}{\widehat{\lambda}_n^2}\mathrm{Re}\left(\widehat{C}_n^a\bar{\widehat{B}}_n^b\right), \quad \widehat{S}_{zn}^{inc\cdot a} = \frac{1}{8\pi}\cdot\frac{\omega\widehat{\Gamma}_n}{\widehat{\lambda}_n^2}\left|\widehat{A}_n^a\right|^2, \quad \widehat{S}_{zn}^{in\cdot b} = \frac{1}{8\pi}\cdot\frac{\omega\widehat{\Gamma}_n}{\widehat{\lambda}_n^2}\left|\widehat{A}_n^b\right|^2, \tag{24}$$

$$S_{zn}^{inter.} = \frac{1}{4\pi}\cdot\frac{\omega\Gamma_n}{\lambda_n^2}\mathrm{Re}\left(C_n^a\bar{B}_n^b\right), \quad S_{zn}^{inc\cdot a} = \frac{1}{8\pi}\cdot\frac{\omega\Gamma_n}{\lambda_n^2}\left|A_n^a\right|^2, \quad S_{zn}^{inc\cdot b} = \frac{1}{8\pi}\cdot\frac{\omega\Gamma_n}{\lambda_n^2}\left|A_n^b\right|^2. \tag{25}$$

Let's define the power interference transmission coefficients $\widehat{T}_n^{inter.}$ and $T_n^{inter.}$ under the formulas

$$\frac{\widehat{S}_{zn}^{inter.}}{\sqrt{\widehat{S}_{zn}^{inc\cdot a}\cdot\widehat{S}_{zn}^{inc\cdot b}}} = \widehat{T}_n^{inter.}\sin(\varphi-\psi+\delta), \quad \widehat{T}_n^{inter.} = \sqrt{\widehat{T}_{1n}^2+\widehat{T}_{2n}^2}, \quad tg\delta = \frac{T_{2n}}{T_{1n}}, \tag{26}$$

$$\frac{S_{zn}^{inter.}}{\sqrt{S_{zn}^{inc\cdot a}\cdot S_{zn}^{inc\cdot b}}} = T_n^{inter.}\sin(\varphi-\psi+\delta), \quad T_n^{inter.} = \sqrt{T_{1n}^2+T_{2n}^2}, \quad tg\delta = \frac{T_{2n}}{T_{1n}}, \tag{27}$$

where

$$\widehat{T}_{1n} = -\left(16\widehat{\Gamma}_n/\left|\widehat{\Delta}_n\right|^2\right)\cdot\left[\widehat{\beta}_n\left(\widehat{\Gamma}_n^2+\left|\widehat{P}_n\right|^2\right)sh\widehat{\beta}_nd\cdot\cos\widehat{\alpha}_nd+\widehat{\alpha}_n\left(\widehat{\Gamma}_n^2-\left|\widehat{P}_n\right|^2\right)ch\widehat{\beta}_nd\cdot\sin\widehat{\alpha}_nd\right], \tag{28}$$

$$\widehat{T}_{2n} = \left(16\widehat{\Gamma}_n/\left|\widehat{\Delta}_n\right|^2\right)\cdot\left[\widehat{\alpha}_n\left(\widehat{\Gamma}_n^2-\left|\widehat{P}_n\right|^2\right)sh\widehat{\beta}_nd\cdot\cos\widehat{\alpha}_nd-\widehat{\beta}_n\left(\widehat{\Gamma}_n^2+\left|\widehat{P}_n\right|^2\right)ch\widehat{\beta}_nd\cdot\sin\widehat{\alpha}_nd\right], \tag{29}$$

$$T_{1n} = -\left(16\varepsilon\Gamma_n/|\Delta_n|^2\right)\cdot\left[\beta_n\left(\varepsilon^2\Gamma_n^2+|P_n|^2\right)sh\beta_nd\cdot\cos\alpha_nd+\alpha_n\left(\varepsilon^2\Gamma_n^2-|P_n|^2\right)ch\beta_nd\cdot\sin\alpha_nd\right] \tag{30}$$

$$T_{2n} = \left(16\varepsilon\Gamma_n/|\Delta_n|^2\right)\cdot\left[\alpha_n\left(\varepsilon^2\Gamma_n^2-|P_n|^2\right)sh\beta_nd\cdot\cos\alpha_nd-\beta_n\left(\varepsilon^2\Gamma_n^2+|P_n|^2\right)ch\beta_nd\cdot\sin\alpha_nd\right], \tag{31}$$

$$\widehat{\Delta}_n = \left(\widehat{\Gamma}_n-\widehat{P}_n\right)^2e^{i\widehat{P}_nd}-\left(\widehat{\Gamma}_n+\widehat{P}_n\right)^2e^{-i\widehat{P}_nd}, \quad \Delta_n = (\varepsilon\Gamma_n-P_n)^2e^{iP_nd}-(\varepsilon\Gamma_n+P_n)^2e^{-iP_nd}. \tag{32}$$

Calculating $\widehat{T}_n^{inter.}$ and $T_n^{inter.}$ under the formulas (26) and (27) in view of (28)–(32), we shall receive

$$\widehat{T}_n^{inter.} = \frac{8\widehat{\Gamma}_n\sqrt{\left(\widehat{\alpha}_n^2+\widehat{\beta}_n^2\right)\left(ch2\widehat{\beta}_nd-\cos2\widehat{\alpha}_nd\right)}}{\sqrt{2}\sqrt{\left(\widehat{\alpha}_n^2+\widehat{\beta}_n^2+\widehat{\Gamma}_n^2\right)-4\widehat{\alpha}_n^2\widehat{\Gamma}_n^2\left[ch\left(\widehat{\psi}_n+2\widehat{\beta}_nd\right)-\cos\left(\widehat{\varphi}_n-2\widehat{\alpha}_nd\right)\right]}}, \tag{33}$$

$$T_n^{inter.} = \frac{8\varepsilon\Gamma_n\sqrt{(\alpha_n^2+\beta_n^2)(ch2\beta_nd-\cos2\alpha_nd)}}{\sqrt{2}\sqrt{\alpha_n^2+\beta_n^2+\varepsilon^2\Gamma_n^2-4\alpha_n^2\Gamma_n^2}\left[ch(\psi_n+2\beta_nd)-\cos(\varphi_n-2\alpha_nd)\right]}. \tag{34}$$

Apparently, according to (33), (34) and (16)–(19) it are true the ratios

$$\widehat{T}_n^{inter.} = 2\sqrt{\widehat{T}_n^a\cdot\widehat{R}_n^a}, \quad T_n^{inter.} = 2\sqrt{T_n^a\cdot R_n^a}. \tag{35}$$

Now, considering that $\widehat{R}_n^a = \widehat{R}_n^b = \widehat{R}_n$, $\widehat{T}_n^a = \widehat{T}_n^b = \widehat{T}_n$, $R_n^a = R_n^b = R_n$ and $T_n^a = T_n^b = T_n$, the streams of electromagnetic energy in the regions I and III in the waveguide can be presented in the form of

$$\widehat{S}_{zn}^{I} = (1/8\pi)\,\mathrm{Re}\left(\omega\widehat{\Gamma}_n/\widehat{\lambda}_n^2\right)\left[\widehat{R}_n\left|\widehat{A}_n^a\right|^2 + \widehat{T}_n\left|\widehat{A}_n^b\right|^2 + 2\left|\widehat{A}_n^a\right|\left|\widehat{A}_n^b\right|\sqrt{\widehat{T}_n\widehat{R}_n}\sin\left(\psi-\varphi+\delta\right)\right], \quad (36)$$

$$\widehat{S}_{zn}^{III} = (1/8\pi)\,\mathrm{Re}\left(\omega\widehat{\Gamma}_n/\widehat{\lambda}_n^2\right)\left[\widehat{T}_n\left|\widehat{A}_n^a\right|^2 + \widehat{R}_n\left|\widehat{A}_n^b\right|^2 + 2\left|\widehat{A}_n^a\right|\left|\widehat{A}_n^b\right|\sqrt{\widehat{T}_n\widehat{R}_n}\sin\left(\varphi-\psi+\delta\right)\right], \quad (37)$$

$$S_{zn}^{I} = (1/8\pi)\,\mathrm{Re}\left(\omega\Gamma_n/\lambda_n^2\right)\left[R_n\left|A_n^a\right|^2 + T_n\left|A_n^b\right|^2 + 2\left|A_n^a\right|\left|A_n^b\right|\sqrt{T_nR_n}\sin\left(\psi-\varphi+\delta\right)\right], \quad (38)$$

$$S_{zn}^{III} = (1/8\pi)\,\mathrm{Re}\left(\omega\Gamma_n/\lambda_n^2\right)\left[T_n\left|A_n^a\right|^2 + R_n\left|A_n^b\right|^2 + 2\left|A_n^a\right|\left|A_n^b\right|\sqrt{T_nR_n}\sin\left(\varphi-\psi+\delta\right)\right]. \quad (39)$$

Analyzing expressions 1-2 we come to conclusions:

1. When $\varphi-\psi+\delta=-\pi/2$ and $\left|\widehat{A}_n^b\right|^2/\left|\widehat{A}_n^a\right|^2=\widehat{T}_n/\widehat{R}_n$, then $\widehat{S}_{zn}^{III}=0$ and $\widehat{S}_{zn}^{I}\neq 0$, (40)
2. When $\psi-\varphi+\delta=-\pi/2$ and $\left|\widehat{A}_n^b\right|^2/\left|\widehat{A}_n^a\right|^2=\widehat{R}_n/\widehat{T}_n$, then $\widehat{S}_{zn}^{I}=0$ and $\widehat{S}_{zn}^{III}\neq 0$, (41)
3. When $\varphi-\psi+\delta=-\pi/2$ and $\left|A_n^b\right|^2/\left|A_n^a\right|^2=T_n/R_n$, then $S_{zn}^{III}=0$ and $S_{zn}^{I}\neq 0$, (42)
4. When $\psi-\varphi+\delta=-\pi/2$ and $\left|A_n^b\right|^2/\left|A_n^a\right|^2=R_n/T_n$, then $S_{zn}^{I}=0$ and $S_{zn}^{III}\neq 0$. (43)

The results received above, show, that choosing in appropriate way initial phases φ and ψ, phase shift δ and amplitudes of counter propagating waves, passing through absorbing plate in the waveguide, we can control the streams of electromagnetic energy in the regions I and III in the waveguide.

REFERENCES

1. Sidorenkov, V. V. and V. V. Tolmachyov, "Effects of electromagnetic interference in metallic plates," *Pisma V Zhurnal Tekhnicheskoy Fiziki,* Vol. 16, No. 3, 20–25, 1990.
2. Sidorenkov, V. V. and V. V. Tolmachyov, "Effect of tunneling electromagnetic interference of fields of arbitrary physical nature and its technical applications," http://www. Sciteclibrary.ru/texsts/rus/stat/st. 2251.pdf.
3. Gevorkyan, E. A., "On the electrodinamics of space — time periodic mediums in the waveguides of arbitrary cross section," *Uspekhi Sovremennoy Radioelektroniki*, No. 1, 3–29, 2006.
4. Gevorkyan, E. A., "Electromagnetic waves in a waveguide with a periodically modulating filling," *Proceedings of International Symposium on Electromagnetic Theory*, Vol. 1, 69–70, Thessaloniki, Greece, May 1998.

Harmonic Imaging through Nonlinear Metamaterial Surfaces

Zhiyu Wang[1], **Yu Luo**[1,3], **Tao Jiang**[1], **Zheng Wang**[2],
Jiangtao Huangfu[1], **and Lixin Ran**[1]

[1]Department of Information and Electronic Engineering, Zhejiang University, Hangzhou 310027, China
[2]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]Department of Physics, Imperial College, London SW7 2AZ, UK

Abstract— We experimentally demonstrate a microwave far-field imaging modality with the transverse resolution exceeding the diffraction limit through the use of a single layer of highly nonlinear metamaterial. The harmonic fields of the nonlinear metamaterial surface allow the far-field propagation of wavefronts with spatial frequencies several times higher than that of the fundamental field. Near-field images can therefore be mathematically recovered from the far-field patterns of the harmonic fields.

1. INTRODUCTION

In linear media, the spatial resolution of an EM imaging system has long been considered to be limited by the Rayleigh limit, approximately at one half of the wavelength [1]. Recent studies in left-handed metamaterials and nonlinear microscopy have resulted in several super-resolution schemes that promise sub-wavelength imaging. Using only linear media, artificial left-handed metamaterials have been proposed [2] and experimentally verified [3, 4] to restore the evanescent waves that carry sub-wavelength spatial information in near-field and far-field configurations. Nevertheless linear metamaterials are intrinsically lossy and higher spatial resolution often comes at the price of much reduced signal-to-noise ratio.

To apply such subwavelength imaging schemes by nonlinear media, one can conceivably extend the concept of metamaterials to synthesize artificial nonlinearities from active circuit elements. Indeed, enhanced electromagnetic nonlinearities from microwave metamaterials have been proposed by Pendry et al. [5] and have been realized in [6, 7]. In this paper, we propose and demonstrate experimentally that a diode-based nonlinear metamaterial surface, when placed in the near-field of the object, generates harmonic fields that allow for far-field imaging of details at a resolution double the diffraction limit of the fundamental frequency.

2. PRINCIPLES

To demonstrate the principle of harmonic imaging with a nonlinear surface, we start by considering a highly nonlinear surface is placed adjacent to the object plane, like the one shown in Figure 1. In the absence of the nonlinear surface, the system is diffraction limited: the maximum diffraction limit can be simply estimated by $d_{\max} = \lambda/2$, where λ denotes the wavelength corresponding to the frequency f_0 of the incident EM wave. In Figure 1(a), the contour map of the emitted electric intensity by two dipole antennas in free space with an interval $D = \lambda/2$ along the x direction is shown. We see that in the region immediately surrounding the dipoles, known as "reactive zone" generally estimated by $\lambda/2\pi$ (or about 0.2λ), the two dipoles can still be distinguished if the total intensity is measured by a detector scanning along the x direction. However, in a distance beyond the reactive zone, the evanescent waves (or reactive waves) that carry sub-wavelength information about the sources have been rapidly vanished, and the detector can only find one peak and lose the distinguishability, which clearly obeys the diffraction limit.

When a nonlinear surface is put inside the reactive zone, for instance at $y = 0.05\lambda$, the schematic contour map of electric intensity is shown in Figure 1(b). Since the nonlinear surface is in the reactive zone and senses the reactive fields carrying the sub-wavelength information, the second and higher order harmonics generated from the surface carry the sub-wavelength information either. If a detector working at these harmonic frequencies are used to scanning the corresponding harmonics in the region to the right of the surface, the dipoles can be distinguished even in a place beyond the former reactive zone. Figure 1(b) illustrates the schematic intensity distribution of the field of the second harmonic ($2f_0$), showing that the diffraction limit no longer holds.

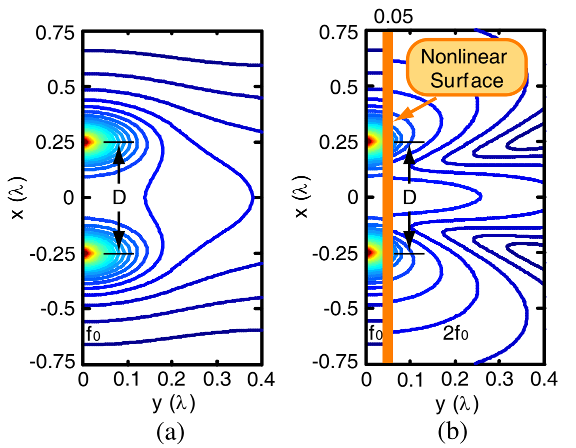


Figure 1: (a) The calculated near-field intensity distribution of two dipole sources (in-phase) located with distance D closer than the diffraction limit. (b) The schematic intensity distribution when a nonlinear surface is inserted in the near-field region, 0.05λ from the sources.

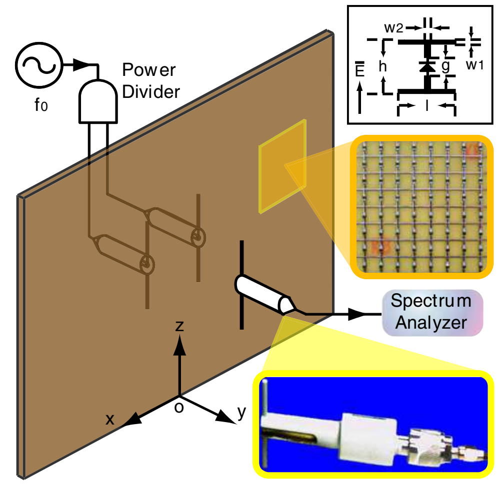


Figure 2: Setup of the experiment while two dipole sources are adjacent to the nonlinear metamaterial surface. The near-field mapping is performed with a dipole reception antenna (bottom-right inset), while the far-field pattern is detected with a horn antenna (not shown).

3. NONLINEAR MATERIALS AND EXPERIMENTAL SETUP

In this paper, we will use the metamaterial sample reported in [8]. The right inset of Figure 2 shows the photograph of the nonlinear sample. The sample is made of a pair of two-dimensional arrays of I-shaped metallic traces printed and aligned on both sides of a 1-mm-thin FR4 substrate with a relative permittivity of 4.6. A small gap exists in the center of the I-pattern in each unit cell in order for each half to establish an electrical contact with a microwave diode (Infineon's BAT15-03W), as shown in the top-right inset of Figure 2. A direct current (DC) source is used to bias the diodes, choosing a strongly nonlinear region of the volt-ampere characteristic curve of the diodes to obtain a strong nonlinearity.

The detailed dimensions for each unit cell are $l = h = 6\,\text{mm}$, $g = 1.6\,\text{mm}$, $w_1 = 0.3\,\text{mm}$, $w_2 = 1\,\text{mm}$, and there are 40 and 48 unit cells along the x and z directions, respectively, yielding a 288-mm-long, 240-mm-wide, 1-mm-thick thin flat sample. For an incident electric field polarized along the z direction, electric resonance can be induced by the metallic resonant patterns and, due the existence of the diodes, enhanced nonlinear electric response can be obtained [8].

The experimental setup for the observation of resolution improvement is shown in Figure 2. Two standard dipole antennas polarized along the z direction driven by equal-amplitude and in-phase monochromatic waves serve as sources, and a third identical dipole antenna serves as a detector. The interval D between the sources is set to be $\lambda/2$ within the diffraction limit. The sample is put between the sources and the detector, with a distance very close to the sources. The input monochromatic wave is generated by a Vector Signal Generator (Agilent E8267C) and the output spectrum is detected by a Spectrum Analyzer (Advantest R3271A).

4. EXPERIMENTAL RESULTS

The experiments are conducted in a microwave anechoic chamber in both near- and far-field ranges. The frequency and power of the input monochromatic wave is selected to be 3 GHz and 20 dBm, respectively. In the near-field measurement, we move the third dipole antenna along x axis at four different distances, i.e., $y = 10\,\text{mm}$, 20 mm, 30 mm and 40 mm, respectively. In the far-field measurement, the sources as well as the sample are placed in the quiet zone of the chamber and rotated, meanwhile a broadband horn antenna 6λ (600 mm) away from the sources is served as a receiver to measure the far-field radiation pattern.

For comparison, we firstly perform a control experiment without the insertion of the sample. When the incidence is at 3 GHz and the interval $D = \lambda/2$, i.e., 50 mm, the measured electric fields are shown in Figures 3(a)–(c).As expected, we find from the near-field distribution in Figure 3(a) and 3(b) that when the detector is leaving from the sources, the distinguishability to the two sources drastically degrades, and after $y > 20\,\text{mm}$, or about 0.2λ, which is just around the boundary of

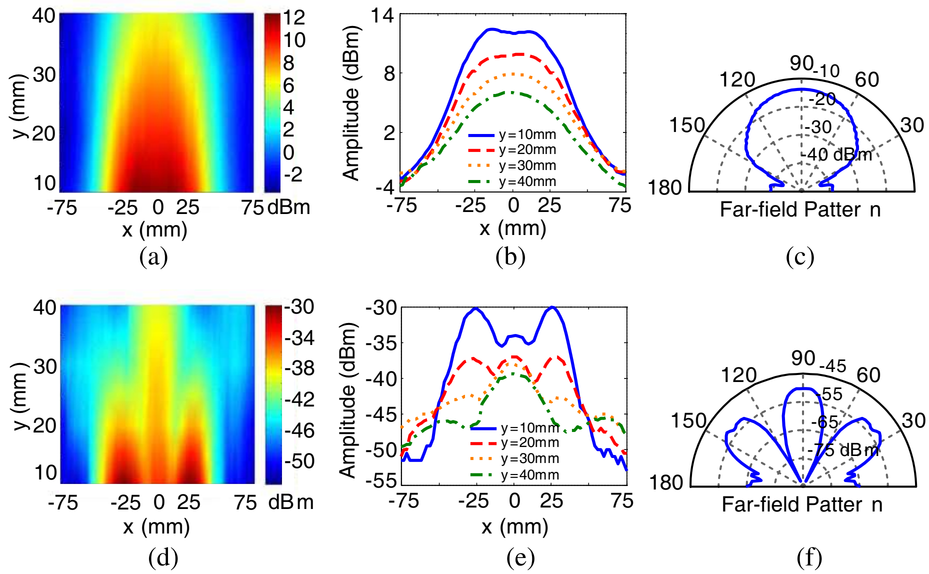


Figure 3: Measured near-field and far-field intensity patterns before and after the insertion of the nonlinear surface.

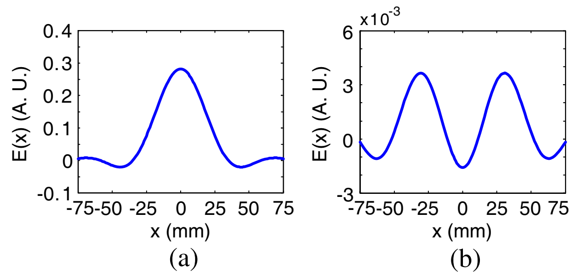


Figure 4: IFT of near- and far-field data with 50 mm interval between two sources. (a) IFT of far-field data at 3 GHz without sample. (b) IFT of far-field data at 6 GHz with sample.

the reactive zone, the distinguishability completely loses. Similarly, from the far-field pattern in Figure 3(c), we can only find one lobe, implying that we cannot recover the position information of the two sources from the far-field data either.

Then we insert the sample at the place where $y = 0.05\lambda$, or 5 mm away from the sources, and perform the same measurement but at the second-harmonic frequency, i.e., 6 GHz. The data are shown in Figures 3(d)–(f). Compared with the control experiment, we see clearly from Figures 3(d) and 3(e) that the near-field distribution changes obviously, and even when $y = 40$ mm, we may still observe the variation of the near-field. For far-field, the pattern shown in Figure 3(f) now has three lobes, which is also completely different from that in Figure 3(c). These clearly indicate that in this case the diffraction limit no longer holds. We will show later that the position information, or image, of the sources can be retrieved from the far-field data by appropriate algorithm.

5. IMAGE RETRIEVAL

According to antenna theory, the source field distribution in an aperture can be calculated by performing an inverse Fourier transformation (IFT) to its far-field pattern [9]. In our case, we can calculate the electric field distribution $E(x)$ at $y = 0$ using a normalized approximate equation

$$E(x) = \int_0^{\pi} E(\varphi) e^{-j2\pi x \cos\varphi/\lambda_m} d\varphi, \tag{1}$$

where $E(\varphi)$ is the far-field pattern, φ is the azimuth angle and λ_m is the wavelength corresponding to the frequency at which the pattern is measured. Applying Equation (1) to the patterns in Figures 3(c) and 3(f), we obtain the field distribution, or virtual image, of the source antennas shown in Figures 4(a) and 4(b), respectively.

The phase information is not measured, so in the calculation, the side lobe level (SLL) is treated to be negative, knowing that the pattern is caused by two linear antennas [9]. We see that without the sample (Figure 4(a)), the image only implies a single source, whereas with the sample (Figure 4(b)), the image clearly illustrates two discrete sources with $D = 61.5\,\mathrm{mm}$. Here, the IFT actually acts a role as a virtual "digital" lens.

6. CONCLUSION

In conclusion, we show by experiments that metamaterial included with active elements can easily behave strong nonlinearity under weak incident EM powers, and by covering a thin flat nonlinear metamaterial on the sources, the enhanced resolution can be achieved by measuring the far-field radiation of the sources at the harmonic frequencies and calculating the corresponding IFT. The application of the nonlinear sample proposed in this paper would have important potential in improving the sub-wavelength resolution in the near future.

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, the ZJNSF (No. Y1080715).

REFERENCES

1. Goodman, J. W., *Introduction to Fourier Optics*, 2nd Edition, Chapter 6.5, McGraw-Hill Companies, Inc., 1996.
2. Pendry, J. B., "Negative refraction makes a perfect lens," *Phys. Rev. Lett.*, Vol. 85, 3966, 2000.
3. Fang, N., H. Lee, C. Sun, and X. Zhang, "Sub-diffraction-limited optical imaging with a silver superlens," *Science*, Vol. 308, 534, 2005.
4. Taubner, T., D. Korobkin, Y. Urzhumov, G. Shvets, and R. Hillenbrand, "Near-field microscopy through a SiC superlens," *Science*, Vol. 313, 1595, 2006.
5. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE Trans. Microw. Theory and Tech.*, Vol. 47, 2075, 1999.
6. Klein, M. W., C. Enkrich, M. Wegener, and S. Linden, "Second-harmonic generation from magnetic metamaterials," *Science*, Vol. 313, 502, 2006.
7. Shadrivov, I. V., A. B. Kozyrev, D. W. van der Weide, and Y. S. Kivshar, "Tunable transmission and harmonic generation in nonlinear metamaterials," *Appl. Phys. Lett.*, Vol. 93, 161903, 2008.
8. Wang, Z., Y. Luo, L. Peng, T. Jiang, D. Wang, H. Chen, J. Huangfu, and L. Ran, "Second-harmonic generation and spectrum modulation by an active nonlinear metamaterials," *Appl. Phys. Lett.*, Vol. 94, 134102, 2009.
9. Kraus, J. D. and R. J. Marhefka, *Antennas: For All Applications*, 3rd Edition, Chapter 15.2, McGraw-Hill Companies, Inc., 2003.

Generation of Waves by a Neutron Beam in a Quantum Plasma of Nonzero Spin. An Influence of the Spin-orbit Interaction

P. A. Andreev[1] **and L. S. Kuz'menkov**[2]

[1]Department of General Physics, Physics Faculty, Moscow State University, Russian Federation
[2]Department of Theoretical Physics, Physics Faculty, Moscow State University, Russian Federation

Abstract— In the present paper we model the electromagnetic process in a plasma placed in an external magnetic field. This plasma is an anisotropic medium with strongly expressed electromagnetic properties. For the modeling of the electromagnetic quantum properties, we use the Schrodinger equation. Based on a Hamiltonian of a charged particle system with an intrinsic magnetic moment in an external electromagnetic field quantum, hydrodynamic equations are derived. The Coulomb, spin-spin, spin-current and spin-orbit in the Hamiltonian are included. The equations for number of particles, momentum and magnetic moment are obtained. The self-consistent field approximations of these equations are considered. Based on the quantum hydrodynamic equations the process of generation of waves by a neutron beam in a dense quantum plasma of nonzero spin is considered. Existence of new wave solutions of the quantum hydrodynamic equations is shown (with respect to our previous works as described in [1, 2]). One type of such waves propagates perpendicularly to the direction of the external magnetic field. Another type is the spin waves which propagate at arbitrary angle with respect to the direction of the external magnetic field. Dispersion relations for waves in a medium traversed by a beam of neutrons whose velocity has a nonzero constant component are derived. Extreme cases of waves propagation parallel or transverse to the direction of the external magnetic field are considered. Generation of plasma, electromagnetic and self-consistent spin waves is investigated. The analytic formulas for the increments of instability are obtained. The contributions of spin-orbit interaction in the process of generation waves are investigated. The spin-spin, spin-current and spin-orbit interactions are the physical mechanism of instabilities in two-component quantum plasma traversed by a beam of neutrons. The collective dynamic of the magnetic moment of particles and methods of generation of this process can influence the properties of spintronic devices.

1. INTRODUCTION

In this work we theoretically consider the generation of waves in plasma contained in external magnetic field. For generation of waves we use the neutron beam. We take attention to the own magnetic moment of particles. For modeling of the electromagnetic process in this anisotropic media we use the method of quantum hydrodynamic for the many particle systems. The physical mechanism of arising of the instabilities is the spin-spin, spin-current and spin-orbit interaction.

We consider the eigenwaves problem in electron-ion plasma and influence of dynamic of own magnetic moments of particles to the dispersion properties of well-known branches of dispersion and the problems of arising new branches of dispersion. Moreover we interest the problem of existence of a spin waves in that systems. A spin waves in the plasma we mention the waves which propagate without of excitation of collective electric field in the wave. That waves propagate by means of magnetic field. The dynamic of magnetic moments of particles where is influence on dispersion relations for well-known waves in plasma [1, 3] (see also references on [3]). Dispersion properties of the system electron beam–plasma in whole k-space is considered in [4].

In 1999 a new method of first-principle derivation of equations of many-particles dynamic, directly from Schrodinger equation was presented [5, 6]. This equations arise in the form of hydrodynamic-like equations.

2. MODEL; SYSTEM OF QUANTUM HYDRODYNAMIC EQUATIONS

In this paper we consider generation of waves by neutron beam for the case when velocity of beam is parallel to the direction of external magnetic field. For solving of this problem we use the system of quantum hydrodynamic equation in the self-consistent field approximation. That system of equation consist of the continuity equation:

$$\partial_t n(\mathbf{r},t) + div(n(\mathbf{r},t)\mathbf{v}(\mathbf{r},t)) = 0;$$

the momentum balance equations (an analog of the Euler equation):

$$mn(\mathbf{r},t)\left(\partial_t+v^\beta(\mathbf{r},t)\nabla^\beta\right)v^\alpha(\mathbf{r},t)-\frac{\hbar^2}{2m}n(\mathbf{r},t)\partial_\alpha\frac{\triangle\sqrt{n(\mathbf{r},t)}}{\sqrt{n(\mathbf{r},t)}}$$
$$= en(\mathbf{r},t)E^\alpha(\mathbf{r},t)+\frac{e}{c}\varepsilon^{\alpha\beta\gamma}n(\mathbf{r},t)v^\beta(\mathbf{r},t)B^\gamma(\mathbf{r},t)+M^\beta(\mathbf{r},t)\nabla^\alpha B^\beta(\mathbf{r},t)+F^\alpha_{s-o}(\mathbf{r},t);$$

where the force field for spin-orbit interaction has form:

$$F^\alpha_{s-o}(\mathbf{r},t) = \frac{1}{c}\frac{2\gamma}{\hbar}\varepsilon^{\alpha\beta\mu}\varepsilon^{\beta\gamma\delta}B^\gamma(\mathbf{r},t)M^\delta(\mathbf{r},t)E^\mu(\mathbf{r},t)-\frac{1}{c}\varepsilon^{\alpha\beta\gamma}M^\beta(\mathbf{r},t)\partial_t E^\gamma(\mathbf{r},t)$$
$$-\frac{1}{c}\varepsilon^{\alpha\beta\gamma}\partial^\delta E^\gamma(\mathbf{r},t)J_M^{\beta\delta}(\mathbf{r},t)-\frac{1}{c}\varepsilon^{\beta\gamma\mu}J_M^{\beta\gamma}(\mathbf{r},t)\partial^\alpha E^\mu(\mathbf{r},t);$$

the equation of evolution of the magnetic moment is:

$$\partial_t M^\alpha(\mathbf{r},t)+\nabla^\beta J_M^{\alpha\beta}(\mathbf{r},t)=\frac{2\gamma}{\hbar}\varepsilon^{\alpha\beta\gamma}\left(M^\beta(\mathbf{r},t)B^\beta(\mathbf{r},t)+\frac{1}{c}\varepsilon^{\beta\mu\nu}J_M^{\gamma\nu}(\mathbf{r},t)E^\mu(\mathbf{r},t)\right).$$

That equations take place for each sorts of particles. They connected by means the Maxwell equations:

$$div\mathbf{B}(\mathbf{r},t)=0,\quad div\mathbf{E}(\mathbf{r},t)=4\pi\sum_a e_a n_a(\mathbf{r},t),$$

$$rot\mathbf{E}(\mathbf{r},t)=-\partial_t\mathbf{B}(\mathbf{r},t),\quad rot\mathbf{B}(\mathbf{r},t)=\partial_t\mathbf{E}(\mathbf{r},t)+\frac{4\pi}{c}\sum_a\mathbf{j}_a(\mathbf{r},t)+4\pi\sum_a rot\mathbf{M}_a(\mathbf{r},t).$$

The tensor of the magnetic moments flow $J_M^{\alpha\beta}$ may be approximately present in the form $J_M^{\alpha\beta} = M^\alpha v^\beta$.

Starting from this system of equation we consider an eigenwaves and a problem of they generation in linear approximation. The non-perturbed state is described by equilibrium concentration n_{0a}, magnetic field $\mathbf{B}_0 = B_0\mathbf{e}_z$, magnetic moment of the media $\mathbf{M}_{0a} = \chi_a\mathbf{B}_0$. The electric field $\mathbf{E}$ and velocity field of the electrons and the ions $\mathbf{v}$ in equilibrium state is equal to zero. The velocity field of neutron beam is $v^\alpha_{0b} = U^\alpha = U_z\delta^{\alpha\beta}$ Where using parameter $\chi_a = \kappa_a/\mu_a$ is the ratio between equilibrium magnetic susceptibility κ_a and magnetic permeability μ_a. We suppose that for the perturbation of the pressure $p^{\alpha\beta} = p\delta^{\alpha\beta}$ to take place formula $\delta p_a = mv^2_{sa}$.

3. PARALLEL PROPAGATION

The dynamic of magnetic moment of particles is the cause of arising a new wave solution in both cases the propagation of waves parallel and perpendicular to the direction of the external magnetic field. In [1], the existence of two new solutions, in the case of propagation transverse to external magnetic field was obtained. Frequencies of that solutions located around ion and electron cyclotron frequencies. Moreover in [1] the existence of the spin-waves in plasma was shown for the case of parallel propagation.

For the case parallel propagation the dispersion equation is the sampling of three independent equation. Two equation has new solutions with respect to absence of magnetic moments of particles. This equations has form:

$$\frac{\omega^2}{c^2}-k_z^2-\sum_c\left(\frac{\omega^2_{Lc}}{c^2}\frac{\omega}{\omega\pm\Omega_c}\mp 4\pi\lambda\frac{\omega^2}{c^2}\frac{W_{\gamma c}}{\omega\pm\Omega_c}\mp 4\pi k_z^2\frac{W_{\gamma c}}{\omega\mp\Omega_c}\right)$$
$$\pm 4\pi\lambda\frac{\omega W_{\gamma b}}{c^2}\frac{(\omega-k_zU_z)^2}{(\omega-k_zU_z)^2-\Omega^2_{\gamma b}}\mp 4\pi\frac{\omega}{c}k_z\frac{W_{\gamma b}}{\omega-k_zU_z\mp\Omega_{\gamma b}}\left(\frac{k_zc}{\omega}-\lambda\frac{U_z}{c}\right)=0 \quad (1)$$

In this equation a follows notation is used $\omega^2_{La} = 4\pi e^2_a n_{0a}/m_a$ — is the Langmuir frequency, $\Omega_c = e_cB_0/(m_cc)$, $\Omega_{\gamma a} = 2\gamma_a B_0/\hbar$ that two parameters it is the cyclotron frequency which arise from the motion of the charge e_c and the magnetic moment in external magnetic field B_0 correspondingly, γ_a is the gyromagnetic ratio, for example, for neutrons $\gamma_b = -1.91\mu_{nuc}$, where μ_{nuc} is the nuclear magneton, $W_{\gamma a} = \chi_a \Omega_{\gamma a}$. $\lambda = 1$ — this parameter indicate the terms which arise from spin-orbit interaction.

In the absence of neutron beam Equation (1) has two new solutions:

$$\omega = |\Omega_e| \left(1 + \frac{8\pi k_z^2 c^2 \chi_e}{\omega_{Le}^2 + 2\omega_{Li}^2 \frac{\Omega_i}{|\Omega_e|} + 2k_z^2 c^2 - 2\Omega_e^2 + 4\pi\lambda|\Omega_e| W_{\gamma e} + 4\pi k_z^2 c^2 \frac{W_{\gamma i}}{|\Omega_e|}} \right), \tag{2}$$

from the second equation in (1), and

$$\omega = \Omega_i \left(1 - \frac{4\pi k_z^2 c^2 \chi_i}{0.5\omega_{Li}^2 - k_z^2 c^2 + \Omega_i^2 + 2\pi\lambda\Omega_i^2(\chi_i + 2\chi_e) - 4\pi k_z^2 c^2 \chi_e} \right), \tag{3}$$

from the first equation.

In the absence of media we can obtain dispersion relations for the eigenwaves in neutron beam. Where is three solutions of each equations in (1).

$$\omega^2 = k^2 c^2 \mp 4\pi\lambda k_z c W_{\gamma b} \frac{k_z^2 (c - U_z)^2}{k_z^2 (c - U_z)^2 - \Omega_{\gamma b}^2} \pm 4\pi k_z^2 c^2 \frac{W_{\gamma b}}{k_z(c - U_z) - \Omega_{\gamma b}} \left(1 - \lambda \frac{U_z}{c} \right), \tag{4}$$

$$\omega = k_z U_z \pm \Omega_{\gamma b} - \frac{2\pi\lambda W_{\gamma b} \Omega_{\gamma b}(k_z U_z \pm \Omega_{\gamma b}) - 4\pi(k_z c)^2 W_{\gamma b} + 4\pi\lambda k_z U_z W_{\gamma b}(k_z U_z \pm \Omega_{\gamma b})}{(k_z U_z \pm \Omega_{\gamma b})^2 - (k_z c)^2}, \tag{5}$$

$$\omega = k_z U_z \mp \Omega_{\gamma b} + \frac{2\pi\lambda(k_z U_z \mp \Omega_{\gamma b}) W_{\gamma b} \Omega_{\gamma b}}{(k_z U_z \mp \Omega_{\gamma b})^2 - k_z^2 c^2 (1 - 2\pi\chi_b) - 2\pi k_z U_z \chi_b (k_z U_z \mp \Omega_{\gamma b})}. \tag{6}$$

Formulas (4) it is dispersion relation for the light which propagate through neutron beam.

Further we consider the process of generation of the waves by means the neutron beam.

Here we take up the first equation in (1). One of the solution of the first equation in (1) is the fast magneto-sound wave. The condition of the resonance interaction of the fast magneto-sound wave and beam mode with dispersion dependence approximately equal $k_z U_z + \Omega_{\gamma b}$ (5) is $\omega_0 = k_z U_z + \Omega_{\gamma b}$. Under that condition arise instabilities, increment of instabilities is obtained with formula

$$\delta\omega^2 = -2\pi\omega_0 |W_{\gamma b}| \frac{\lambda|\Omega_{\gamma b}| - 2\lambda k_z U_z + 2(k_z c)^2/\omega_0}{2\omega_0 - \sum_c \omega_{Lc}^2 \Omega_c/(\omega_0 + \Omega_c)^2} < 0, \tag{7}$$

where $\omega_0 = \omega_0(k_z)$ — dispersion dependence of fast magneto-sound wave in the absence of beam. For the fast magneto-sound wave where is $\omega_0 \subset (0, |\ \Omega_e\ |)$ and $2\omega_0 - \sum_c \omega_{Lc}^2 \Omega_c/(\omega_0 + \Omega_c)^2 > 0$.

Resonance interaction of the wave (3) with beam mode (5) on condition $\Omega_i(1 + \delta\Omega/\Omega_i) = k_z U_z + \Omega_{\gamma b}$ lead to the instability with

$$\delta\omega^2 = -\pi \frac{|W_{\gamma b}| \left(\lambda|\Omega_{\gamma b}| + 2k_z c(k_z c/\Omega_i - \lambda U_z/c)\right) + 2\chi_i (k_z c)^2}{1 + \frac{3}{8} \frac{\omega_{Le}^2}{|\Omega_e|\Omega_i}} < 0. \tag{8}$$

Further, we investigate the second equation in (1). On condition $\omega_A = k_z U_z - \Omega_{\gamma b}$ (ω_A dispersion dependence of Alfven wave in absence of beam) where is the generation of Alfven waves

$$\delta\omega^2 = -2\pi \frac{\lambda\omega_A W_{\gamma b}\Omega_{\gamma b} + 2(k_z c)^2 W_{\gamma b} - 2\omega_A \lambda k_z U_z W_{\gamma b}}{2\omega_A - \sum_c \omega_{Lc}^2 \Omega_c/(\omega_A + \Omega_c)^2} < 0. \tag{9}$$

by means resonance with beam mode (5).

In the case $|\Omega_e|(1 + \delta\Omega/|\Omega_e|) = k_z U_z + \Omega_{\gamma b}$ where is resonance interaction of wave (2) with beam mode (6) arise the frequency shift:

$$\delta\omega = \pm\sqrt{\frac{32\pi|\Omega_e|(\lambda|\Omega_e \Omega_{\gamma b} W_{\gamma b}| - 2k_z^2 c^2 |W_{\gamma e}|)}{\omega_{Le}^2 - 8\Omega_e^2}}. \tag{10}$$

For the dense plasma and conditions $k_z^2 c^2 > |\chi_b/\chi_e| \Omega_{\gamma b}^2$, $|\Omega_{\gamma e}| + |\Omega_{\gamma b}| = k_z U_z$ the solution (10) become imaginary, since arise condition for instabilities.

4. TRANSVERSE PROPAGATION

Under the condition of transverse propagation of waves to the direction of the external magnetic field from dynamics of magnetic moments arise four wave solutions. Two of them was obtained in our previous article [1], and dispersion dependence of that wave has form:

$$\omega = |\Omega_a| \left(1 - \frac{2\pi k^2 c^2 \chi_a}{\omega_e^2 + k^2 c^2 - \Omega_a^2}\right). \tag{11}$$

In this paper we report about another two solutions. In approximation motionless ions one of that waves has follows dispersion relation

$$\omega = \sqrt{\Omega_e^2 + v_{qe}^2 k_\perp^2} + \frac{8\pi^2 \chi_e^2 \Omega_e^2 k_\perp^4 c^4}{\sqrt{\Omega_e^2 + v_{qe}^2 k_\perp^2}} \times \frac{1}{(\omega_e^2(\omega_e^2 + k_\perp^2 c^2 - 2\Omega_e^2 - v_{qe}^2 k_\perp^2) - 8\pi\chi_e \Omega_e^2 k_\perp^2 c^2)}. \tag{12}$$

where $v_{qsa}^2 = v_{sa}^2 + \frac{\hbar^2 k^2}{4m_a^2}$. If take attention to the motion of ions we obtain two solutions (for one solution to each sorts of particles):

$$\omega = \sqrt{\Omega_a^2 + (v_{sa}^2 + \hbar^2 k_\perp^2/(4m_a^2))k_\perp^2} + \frac{8\pi^2 \chi_a^2 \Omega_a^2 k_\perp^4 c^4}{\sqrt{\Omega_a^2 + (v_{sa}^2 + \hbar^2 k_\perp^2/(4m^2))k_\perp^2}}$$
$$\times \frac{1}{\omega_a^2\left(\omega_a^2 + k_\perp^2 c^2 - 2\Omega_a^2 - k_\perp^2 v_{sa}^2 - \frac{\hbar^2}{4m_a^2} k_\perp^4\right) - 8\pi k_\perp^2 c^2 \chi_a \Omega_a^2 + \frac{\omega_a^2(\Omega_a - \Omega_b)(\omega_b^2 \Omega_a - 8\pi\chi_b \Omega_b k_\perp^2 c^2)}{\Omega_b^2 - \Omega_a^2 + k_\perp^2 (v_{sb}^2 - v_{sa}^2) + k_\perp^4 \hbar^2 (\frac{1}{4m_b^2} - \frac{1}{4m_a^2})}}. \tag{13}$$

Solutions (11), (12), (13) exist only on conditions $\chi_a \neq 0$. They arise from equation the form of which is analogous to Equation (1), and new solutions due to new terms which proportional to the χ_a.

5. SPIN WAVES

In the case of parallel propagation of spin waves to the direction of external magnetic field three solution is exist. Two solution is the oscillations with constant frequencies:

$$\omega = |\Omega_a|(1 - 4\pi\chi_a). \tag{14}$$

The dispersion relation of the third waves arise like solution of follows equation:

$$\frac{\chi_e |\Omega_e|}{\omega^2 - (v_{se}^2 + \frac{\hbar^2}{4m_e^2} k^2) k^2} - \frac{\chi_i |\Omega_i|}{\omega^2 - (v_{si}^2 + \frac{\hbar^2}{4m_i^2} k^2) k^2} = 0. \tag{15}$$

This equation has no solutions at $\chi_i = \chi_e$. At $\chi_i \neq \chi_e$ the solution is

$$\omega^2 = k^2 \frac{(v_{si}^2 + \frac{\hbar^2}{4m_i^2} k^2)\chi_e |\Omega_e| - (v_{se}^2 + \frac{\hbar^2}{4m_e^2} k^2)\chi_i |\Omega_i|}{\chi_e |\Omega_e| - \chi_i |\Omega_i|}. \tag{16}$$

This formula represents the dispersion of self-consistent spin waves in the system of electrons and ions with nonzero intrinsic magnetic moments.

In the case of spin wave propagation on arbitrary angle to the external magnetic field where is five solution, for two of them we analytically obtain dispersion dependence. For both waves the dispersion relation may be present in the form:

$$\omega = |\Omega_a| + \delta\omega,$$

with $a = e, i$. For wave with frequency around electron cyclotron frequency we have

$$\delta\omega = \frac{\pi\chi_e \Omega_e \left(2\chi_e \Omega_e \left(\Omega_e^2 - v_i^2 k^2\right) - \chi_i \Omega_i (v_{Se}^2 + \frac{\hbar^2}{4m^2 k^2})(2k_z^2 + k_\perp^2)\right)}{\chi_e \Omega_e (\Omega_e^2 - v_i^2 k^2) - \chi_i \Omega_i (v_{Se}^2 + \frac{\hbar^2}{4m^2} k^2) k^2}, \tag{17}$$

and around ion cyclotron frequency

$$\delta\omega = 2\pi\chi_i\Omega_i \frac{-2\chi_i\Omega_i\left((\Omega_e^2k_z^2-\Omega_i^2k_\perp^2)(v_{Se}^2+\frac{\hbar^2}{4m^2}k^2)-\Omega_i^2\Omega_e^2\right)+\chi_e\Omega_e^2(\Omega_ik_\perp^2-2\Omega_ek_z^2)v_i^2}{\chi_i\Omega_i\left((\Omega_e^2k_z^2-\Omega_i^2k_\perp^2)(v_{Se}^2+\frac{\hbar^2}{4m^2}k^2)-\Omega_i^2\Omega_e^2\right)+\chi_e\Omega_e(\Omega_e^2k_z^2-\Omega_i^2k_\perp^2)v_i^2}. \quad (18)$$

Under condition $k_\perp = 0$ solutions (17) and (18) pass to (14).

Spin-orbit interaction has no influence on dynamic of spin waves in plasma. We investigate the spin waves like waves in propagation of which the electric field take no part. But force of spin-orbit interaction is proportional to electric field. Problem of the generation of the spin-waves which propagate parallel to the direction of external magnetic field was presented in [2].

6. CONCLUSION

In this paper we obtained contribution of spin-orbit interaction in many-particle quantum-hydrodynamic equation. Using the development theoretical apparat we investigated the opportunity of existence of new waves in plasma. We shown the possibility of existence of two waves in the case propagation of wave parallel to direction of the external magnetic field. Moreover, we revealed existence of four waves solution in the case propagation perpendicular to the external magnetic field. Apart from we consider spin-waves in whole k-space, and obtain analytical solution for two waves, which degenerate to oscillations on cyclotron frequencies under condition of propagation along external magnetic field.

Also we investigated the problem of generation of waves with neutron beams, by means spin-spin, spin-current and spin-orbit interactions. We considered neutron beam with the velocities directional parallel to external magnetic field. In this case where is resonance interaction of beam with waves which propagate along direction of external magnetic field. Where is no resonance interaction with waves which propagate transverse external magnetic field. The neutron beam interact with waves which propagate in transverse direction if where is component of beam velocity which perpendicular to the direction of magnetic field [2].

REFERENCES

1. Andreev, P. A. and L. S. Kuz'menkov, "Eigenwaves in a two-component system of particles with nonzero magnetic moments," *Moscow University Physics Bulletin*, Vol. 62, No. 5, 271–276, 2007.
2. Andreev, P. A. and L. S. Kuz'menkov, "Generation of waves by a neutron beam in a two-component system formed by charged particles of nonzero spin," *Physics of Atomic Nuclei*, Vol. 71, No. 10, 1724–1729, 2008.
3. Marklund, M. and G. Brodin, "Dynamics of spin-1/2 quantum plasmas," *Phys. Rev. Lett.*, Vol. 98, 025001(4), 2007.
4. Bret, A., M.-C. Firpo, and C. Deutsch, "Collective electromagnetic modes for beam-plasma interaction in the whole k space," *Phys. Rev. E*, Vol. 70, 046401(15), 2004.
5. Kuz'menkov, L. S. and S. G. Maksimov, "Quantum hydrodynamics of particle systems with coulomb interaction and quantum bohm potential," *Theoretical and Mathematical Physics*, Vol. 118, 227–240, 1999.
6. Kuz'menkov, L. S., S. G. Maksimov, and V. V. Fedoseev, "Microscopic quantum hydrodynamics of systems of fermions: Part 1," *Theoretical and Mathematical Physics*, Vol. 126, 110–120, 2001.

Horn Antennas Loaded with Metamaterial for UWB Applications

M. Lashab[1], **H. I. Hraga**[2], **Read Abd-Alhameed**[2], **C. Zebiri**[1], **F. Benabdelaziz**[1], **and S. M. R. Jones**[2]

[1]Electronics Department, Skikda University, Algeria
[2]Mobile and Satellite Communication Research Centre, University of Bradford, Bradford, UK

Abstract— In this paper, a conical horn antenna has been designed for Ultra-Wideband applications by loading its section with a metamaterial. The work aims first to compare results obtained by the wavelet-moment method to a simulation performed using HFSS. Secondly the conical horn is loaded with a very thin layer of metamaterial to enhance the radiation pattern and the bandwidth performance of the conical horn antenna and reduce the size of the antenna. The operating bandwidth of the proposed antenna is in the range of 10–13 GHz. The results obtained from HFSS and moment method are in good agreement.

1. INTRODUCTION

Artificial materials such as metamaterials and chiral media have recently been of great interest, both theoretically [1, 2], and experimentally [3, 4]. Metamaterials, for instance, exhibit either negative permittivity or negative permeability. If both of them are negative at a given frequency, the material is characterised by an effective negative index of refraction, so it is often referred to as a left handed metamaterial (LHMs). This type has interested many researchers, e.g., [5, 6]. The main objective of research on LHMs is improvement of the radiation pattern, directivity and bandwidth, and antenna size reduction. However in this paper a low index of permittivity is used to characterize the metamaterial as introduced by [7].

Horn antennas loaded with dielectrics or ferrite materials [8], have desirable properties such as increased directivity, reduced side lobe level, wide bandwidth, low loss, and ease of fabrication [9, 12]. These properties are particularly attractive for applications such as ultra-wideband (UWB) ground penetrating radars (GPR) [13, 14]. However, the characterization of such antennas with increasingly complex designs using analytical techniques is often not possible. On the other hand, a numerical model can provide a virtual test bench to explore different design possibilities before any costly prototyping. Although many numerical techniques can be used to model and study the characteristics of such antennas, the moment method is well known to provide good accuracy [15, 16]. In this paper, an improvement has been made by the introduction of wavelets.

This paper deals firstly with a comparison between an improved moment method and Ansoft's HFSS, then an observation is made of the effect of loading the horn antenna.

2. FORMULATION

2.1. Moment Method Formulation

2.1.1. Integral Equation

The Conical Horn is studied in 3D as shown in Figure 1, the construction of this horn is considered to be from any type of material. Using the boundary conditions, the scattered field may be written as an integral magnetic equation in two dimensions for a PEC structure as:

$$K(J(r)) = \frac{1}{2}J(r) - \hat{n} \times \int_S J(r') \times \nabla' G(r, r') \cdot ds' = \hat{n} \times H^i(r) \tag{1}$$

Here $G(r, r')$ is Green's function and $J(r)$ is the current density, this can be expressed in terms of the tangential components. Because the antenna is a body of revolution, the current may be expanded as follow:

$$\vec{J}(t, \varphi) = \sum_{\nu=-\infty}^{+\infty} \left[J_t(t, \varphi) \cdot \hat{t} + J_\varphi(t, \varphi) \cdot \hat{\varphi}\right] \cdot e^{j\nu\cdot\varphi} \tag{2}$$

where $(J_t,\, J_\varphi)$ are the tangential components of the current on the surface of the antenna.

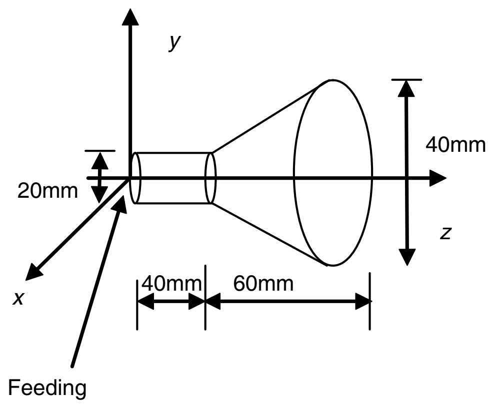


Figure 1: Conical horn in 3D.

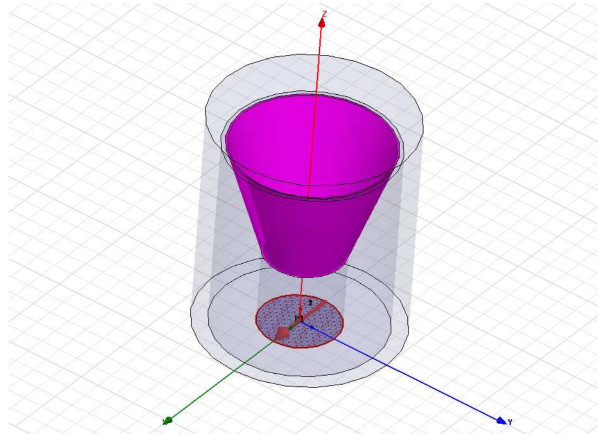


Figure 2: Horn antenna designed by HFSS.

2.1.2. Moment Method

The Moment method is applied on the integral Equation (1), this is discritised by using sets of basis and testing functions [13].

Let W and J denote testing and basis functions, respectively. The integral equation is projected over the two tangential components using the expansion (2). This is done by applying the inner product, denoted by the bracket in (3), to yield:

$$\left\langle \vec{W}, K(J(r)) \right\rangle = \left\langle \vec{W}, \hat{n} \times H^i(r) \right\rangle \tag{3}$$

2.2. Wavelets Expansion

2.2.1. Basis Functions

The basis and testing functions are presented as a superposition of wavelets at several scales and include a scaling function. A Galerkin's method is then applied to transform the integral equation into algebraic equations in the expansion coefficients.

2.2.2. Wavelets Application

The wavelets are applied directly to the integral equation. The current density is expanded as follows

$$J_t(t,\varphi) = \sum_{n=0}^{2^0-1} a_n^t \cdot \phi_{j,n}^t(t,\varphi) + \sum_{m=0}^{j} \sum_{n=0}^{2^{m-1}} c_{m,n}^t \psi_{m,n}^t(t,\varphi) \tag{4}$$

$$J_\varphi(t,\varphi) = \sum_{n=0}^{2^0-1} a_n^\varphi \cdot \phi_{j,n}^\varphi(t,\varphi) + \sum_{m=0}^{j} \sum_{n=0}^{2^{m-1}} c_{m,n}^\varphi \psi_{m,n}^\varphi(t,\varphi) \tag{5}$$

Here ($\psi_{m,n}^t$, $\psi_{m,n}^\varphi$) and ($\phi_{j,n}^t$, $\phi_{j,n}^\varphi$) are the mother and the scaling wavelets, respectively. The corresponding expansion coefficients are a_m^t, $c_{m,n}^t$ and a_n^φ, $c_{m,n}^\varphi$. Using equations (4) and (5) in (3), the following matrix equation is obtained:

$$\begin{bmatrix} Z_{m,n}^{tt} & Z_{m,n}^{t\varphi} \\ Z_{m,n}^{\varphi \cdot t} & Z_{m,n}^{\varphi\varphi} \end{bmatrix} \cdot \begin{bmatrix} c_{m,n}^t \\ c_{m,n}^\varphi \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} \tag{6}$$

The terms a_m^t, a_n^φ are considered very small, thereby they are neglected. The matrix elements are expressed as follow:

$$Z_{pq}^{tt} = \int_t \frac{1}{2} \cdot W_q^t J_p^t \rho \cdot dt - \int_t \int_{t'} W_q^t J_p^t \cdot \hat{\varphi} \times \hat{t}' \cdot I_G \cdot \rho\rho' dt' dt \tag{6a}$$

Here, $I_G = \int_0^{2\pi} \nabla G(r,r') \cdot e^{jv\cdot\varphi'} d\varphi'$.

In more detail, this is given as an integral over the interval $[0, 1]$:

$$Z_{pq}^{tt} = \left\langle \psi_p, \left\langle \psi_q, \frac{1}{2} - T(t,t) \cdot \Omega(t,\xi) \right\rangle \right\rangle \tag{7}$$

where $T(t,t)$ is the term under the double integral of the second part of Equation (6a). In the same manner the other components are given.

$$Z_{pq}^{\varphi\varphi} = \left\langle \psi_p, \left\langle \psi_q, \frac{1}{2} + T(\varphi,\varphi) \cdot \Omega(t,\xi) \right\rangle \right\rangle \tag{8}$$

$$Z_{pq}^{\varphi \cdot t} = \langle \psi_q, \langle \psi_p, T(\varphi,t) \cdot \Omega(t,\xi) \rangle \rangle \tag{9}$$

$$Z_{pq}^{t\varphi} = -\langle \psi_q, \langle \psi_p, T(t,\varphi) \cdot \Omega(t,\xi) \rangle \rangle \tag{10}$$

where $\Omega(t,\varphi,\xi)$ is the calibration of the changing variables, and $D(\xi) = |dt/d\xi|$. The other elements can be written in the same manner. Similarly for the excitation the matrix elements are also expressed as an inner product by:

$$H_1 = \left\langle \psi_q, H^t I_{G2} \cdot \Omega(t,\xi) \right\rangle \tag{11}$$

$$H_2 = -\langle \psi_q, H^\varphi I_{G2} \cdot \Omega(t,\xi) \rangle \tag{12}$$

where $I_{G2} = \frac{1}{2\pi} \int_0^{2\pi} e^{-jv \cdot \varphi} d\varphi$.

The unknowns $[c_{m,n}^t, c_{m,n}^\varphi]$ should be calculated from Equation (6). The current density and the radiation pattern may then be obtained.

3. NUMERICAL RESULTS

In the moment method, the wavelet employed is constructed from the Haar orthogonal wavelet with vanishing moment $N = 7$, the lowest resolution level is chosen Since 128 wavelets are involved, a system of matrices (of 128×128 elements) is generated.

The surface of the taper of the horn is a metamaterial, considered to be an isotropic low index type, in a very thin layer of 1 mm thickness. The permittivity and permeability are respectively $\varepsilon_r = 0.5$, $\mu = 1$. The horn loaded with metamaterial as designed by HFSS is shown in Figure 2. The results obtained by the wavelet-based moment method are in good agreement with the results obtained by HFSS in all the figures of the radiation pattern except the reflection coefficient figure.

The radiation pattern in H-Plane given in Figure 3 and E-Plane in Figure 4 show a slight reduction of the side lobe in the H-plane, and almost no change in the E-plane. A very remarkable reduction in the cross polarization in Figure 5, this is more then 20% of reduction in the side lobes. The directivity and the gain are presented in Figure 6 and Figure 7, the antenna is more directive and better gain when loaded with metamaterial than without.

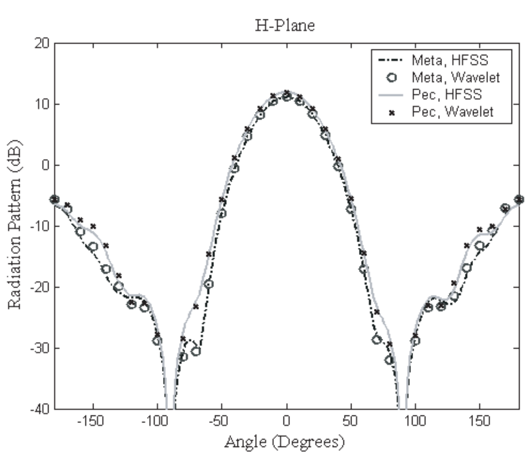


Figure 3: Radiation pattern H-Plane with and without metamaterial at frequency $F = 10$ GHz.

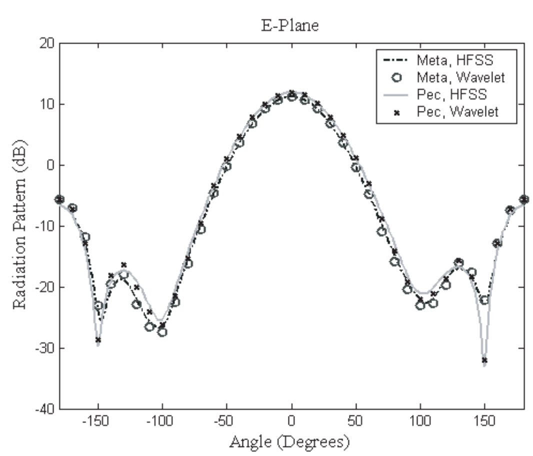


Figure 4: Radiation pattern E-Plane with and without metamaterial at frequency $F = 10$ GHz.

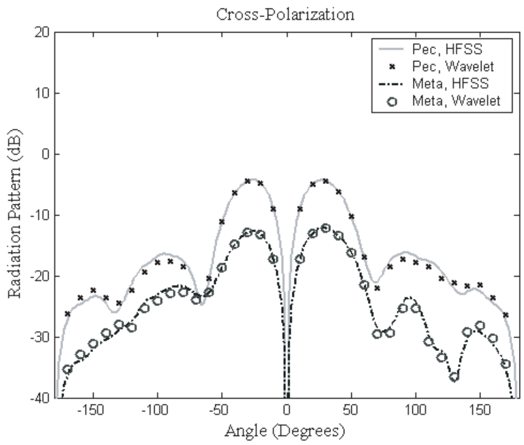


Figure 5: Cross polarisation radiation pattern, effect of the metamaterial, $\varepsilon_r = 0.5$, thickness $d = 1$ mm, $F = 10$ GHz.

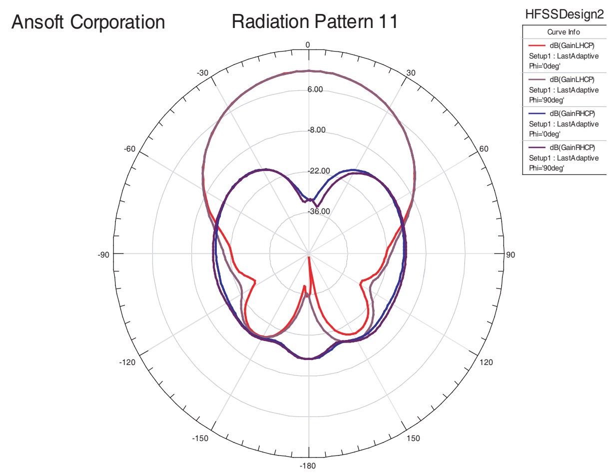


Figure 6: Radiation pattern without metamaterial at frequency $F = 10$ GHz.

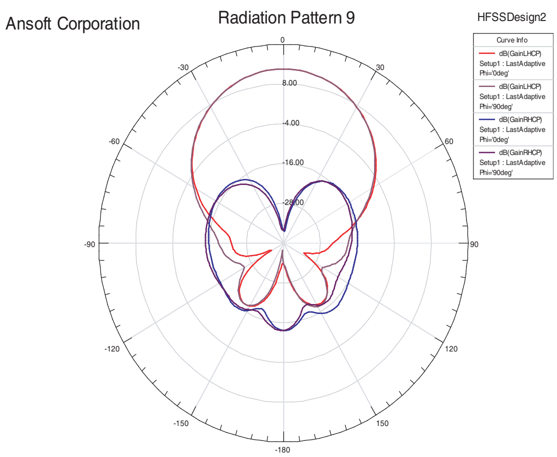


Figure 7: Radiation pattern with metamaterial at frequency $F = 10$ GHz.

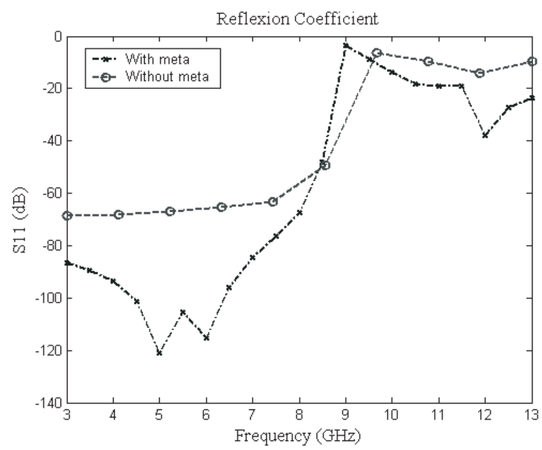


Figure 8: Reflection coefficient with and without metamaterial.

The reflection coefficient in Figure 8 shows a slight displacement of the bandwidth to the lower frequencies, from 9.6 GHz to 10 GHz, i.e., about 10%. This means that one can produce small antenna designs with a reduction in size of about 10%, or simply the bandwidth is enhanced of 10%.

4. CONCLUSIONS

A horn antenna for ultra-wide band (10–13 GHz) has been designed and tested using HFSS and compared to the moment method. The results obtained are in good agreement. The horn loaded with the metamaterial has shown a slight change in the radiation pattern and bandwidth of about 10%, but there is a remarkable effect on the directivity of the antenna. Some antenna miniaturisation is observed but the choice of metamaterial parameters could be further optimized in this respect.

REFERENCES

1. Akyurtlu, A. and D. H. Werner, “A novel dispersive FDTD formulation for modeling transient propagation in chiral metamaterials,” *IEEE Transactions on Antennas and Propagation*, Vol. 52, No. 9, 2267–2276, September 2004.
2. Engheta, N., “An idea for thin subwavelength cavity resonators using metamaterials with negative permittivity and permeability,” *IEEE Antennas and Wireless Propagation Letters*, Vol. 1, 10–13, 2002.

3. Erentok, A. and R. W. Ziolkowski, "Two-dimensional efficient metamaterial-inspired electrically-small antenna," *Microwave and Optical Technology Letters*, Vol. 49, No. 7, July 2007.
4. Zhu, F.-M. and J. Hu, "Improved patch antenna performance by using a metamaterial cover," *Journal of Zhejiang University Science A*, Vol. 8, No. 2, 192–196, 2007.
5. Wang, Y.-S., M.-F. Hsu, and S.-J. Chung, "A compact slot antenna utilizing a right/left-handed transmission line feed," *IEEE Transactions on Antennas and Propagation*, Vol. 56, No. 03, 675–683, March 2008.
6. Wu, B.-I., W. Wang, J. Pacheco, X. Chen, T. Grzegorczyk, and J. A. Kong, "A study of using metamaterials as antenna substrate to enhance gain," *Progress In Electromagnetics Research*, Vol. 51, 295–328, 2005.
7. Lier, E., "Review of soft and hard horn antennas, including metamaterial-based hybrid-mode horns," *IEEE Antennas and Propagation Magazine*, Vol. 52, No. 2, April 2010.
8. Hui, W. W. G., J. M. Bell, M. F. Iskander, and J. J. Lee, "Low-cost microstrip-line-based ferrite phase shifter design for phased array antenna applications," *IEEE Antennas and Wireless Propagation Letters*, Vol. 6, 86–89, 2007.
9. Ikram, S. and G. ahmad, "Design & implementation of a standard gain horn antenna for optimized gain and radiation pattern using mathCAD & HFSS," *Second International Conference on Electrical Engineering*, University of Engineering and Technology, Lahore, Pakistan, March 25–26, 2008.
10. Luo, G. Q., W. Hong, H. J. Tang, J. X. Chen, X. X. Yin, Z. Q. Kuai, and K. Wu, "Filtenna consisting of horn antenna and substrate integrated waveguide cavity FSS," *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 1, 92–98, January 2007.
11. Lee, K.-H., C.-C. Chen, and R. Lee, "UWB dual-linear polarization dielectric horn antennas as reflector feeds," *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 3, 798–804, March 2007.
12. Shafai, L., S. K. Sharma, B. Balaji, A. Damini, and G. Haslam, "Multiple phase center performance of reflector antennas using a dual mode horn," *IEEE Transactions on Antennas and Propagation*, Vol. 54, No. 11, 3407–3417. November 2006.
13. Yin, X.-C., C.-L. Ruan, S.-G. Mo, C.-Y. Ding, and J.-H. Chu, "A compact ultra-wideband microstrip antenna with multipole notches," *Progress In Electromagnetics Research*, Vol. 84, 321–332, 2008.
14. Venkatarayalu, N. V., C.-C. Chen, F. L. Teixeira, and R. Lee, "Numerical modeling of ultrawide-band dielectric horn antennas using FDTD," *IEEE Transactions on Antennas and Propagation*, Vol. 52, No. 5, 1318–1323, May 2004.
15. Lier, E. and A. Kishk, "A new class of dielectric-loaded hybrid-mode horn antennas with selective gain: Design and analysis by single mode model and method of moments," *IEEE Transactions on Antennas and Propagation*, Vol. 53, No. 1, January 2005.
16. Diamantis, S. G., A. P. Orfanidis, and G. A. Kyriacou, "Conical horn antennas employing an offset moment method and mode matching technique," *IEEE Transactions on Magnetics*, Vol. 45, No. 3, 1092–1095, March 2009.

Use of the Neural Net for Road Extraction from Satellite Images, Application in the City of Laghouat (Algria)

F. Benkouider[1], **L. Hamami**[2], **and A. Abdellaoui**[3]

[1]Electrical Engineering Department, Amar Telidji University of Laghouat, Algeria
[2]Electrical and Computer Engineering Department, National Polytechnic School, Algeria
[3]Department of Geography, URF Letters and Human Sciences, Paris XII, France

Abstract— Road extraction in urban areas has been an important task for generating Geographic Information Systems (GIS). In recent years, mainly, the rapid development of urban areas makes it urgent to provide up-to-date road maps. The timely road information is very useful for decision-makers in urban planning, traffic management and car navigation fields, etc. The satellite image is characterized by its big quantity of rich and varied information and constitutes a source of data for generating roads maps by automatic overseeing techniques according to their importance. The interpretation of this one requires mostly, a treatment based on a set of techniques: shaping information, filtering, segmentation and classification, etc. We present in this paper a method based on neuronal net strategy for extracting road networks based on the spectral characteristics of the pixel in satellite image. Normalized spectral information in a window (3×3) around each pixel are used, as 9 red, 9 green and 9 blue (corresponding respectively to red, green and near infra red channels) constitute the input vector of 27 neurons. The origin of the motivation is the homogeneity of roads in high-resolution satellite images, since homogeneity is a characteristic that can be recognized with respect to neighbor pixels, and their spectral information. The system output is represented by one normalized neuron representing the road or not-road characteristics. As the neuronal network requires a large coded data bases in their training stage, we have used a set of road net manually drown using special software. The image used in this application concerns an HVR SPOT image acquired on March 26, 2007 (10 m of resolution) over Laghouat (Algeria), an oasis city located 400 km south of Algiers (Algeria). We have obtained very accurate results with less than 0.022 for the MSE. A set of various applications are presented.

1. INTRODUCTION

The satellite image is characterized by its big quantity of rich and varied information; it constitutes a source of data for generating the road maps by automatic techniques according to their importance. Road extraction in urban areas has been an important task for generating geographic information systems GIS). Nowadays, we are experiencing an explosion in the amount of satellite image data, which provides us with abundant data and also brings challenges to the road extraction task at the same time. The conventional road extraction methods by manual processes are time consuming and tedious, and cannot meet the increasing requirement for such tremendous data. However, automatic extraction of urban roads from high resolution remote sensing imagery is still a challenging problem in digital Photogrammetry and computer vision, the main reason is that the diverse road surfaces and the complex surrounding environments such as trees, vehicles and shadows induced by high buildings make the urban roads take on different textures and gray levels in images. Many researchers have been conducted for this purpose. Using specific operators such as Duda operator for finding linear structures are based on a score function which takes into account the homogeneity and contrast [1]; several road detectors have been applied, the blind operators such as Top to Hat Form (THF), derived from mathematical morphology and designed to extract peak intensity in the image spot [2]; the THF is not very selective and gives noisy results. Followed by several other works in [3, 4]. However, most of them focus on extracting roads in rural or open areas. By contrast, the efforts made for urban road extraction are relatively few [5, 6]. Some works focuse on automatic road extraction in urban area from high resolution satellite images using based machine learning approach [7, 8]. The other semi-automatic methods such as active contour (or snake) [9, 10] and dynamic programming [11] have been the subject of several studies. One can find an excellent survey paper in [12] on road network extraction using the organizing maps applied to classified images. Multilayer neural networks applied in particular for IKONOS images based on the RGB spectral characteristics is presented in [13], those involving texture analysis, fuzzy clustering and genetic algorithms has been treated in [14, 15]. In the present research, road extraction is performed on Spot XS images, using artificial neuronal network algorithms; with a new input

Table 1: Colour composite.

Channel	Wave length	Storage level in the standard composite colour
XS3 (Spot 1 to 3) – B3 (Spot 4 to 5)	0.78 to 0.89 micrometers	Red
XS2 (Spot 1 à 3) – B2 (Spot4 to 5)	0.61 to 0.68 micrometers	Green
XS1 (Spot 1 à 3) – B1 (Spot 4 to 5)	0.50 to 0.59 micrometers	Blue

and architecture (RGB). The format of a window as input is related to that road can be shown as elongated homogenous region with different contrast from the background. We present a different application of the proposed model, and the correspondent MSE error and Kappa coefficient. We use at the end mathematic morphology operations to extract road sides.

2. METHODOLOGY

Road detection can be considered as the first step in road extraction, it is the process of assigning a value to each pixel that can be used as criteria of road and not-road pixels. The problem of road detection from high-resolution satellite images is performed using: 1) artificial neuronal network and 2) pixel spectral characteristics specially the red, green and near infra red channels. The origin of the motivation is the homogeneity of roads in high-resolution satellite images, since homogeneity is a characteristic that can be recognized with respect to neighbor pixels, and their spectral information. In our work, the following hypothesis and input requirements have been considered:

A. Spectral characteristics: The RGB bands of satellite image are calqued on our visual perception, it uses three basic colors: the Red ($\lambda = 700\,\text{nm}$), the green ($\lambda = 546\,\text{nm}$) and the blue ($\lambda = 435.8\,\text{nm}$). In our work we have used the so called Standard Colour Composite (SCC), Table 1.

B. Road appearance: In the RGB standard color composite roads appears in blue. In some cases building appear in blue colour too, this problem appears in classification methods when roads are classified as a building.

C. Road homogeneity: Road networks in high resolution satellite images are presented as elongated homogenous areas having a distinctive brightness pattern compared to their surroundings.

3. PRINCIPLE OF THE ARTIFICIAL NEURAL NETWORK

Artificial Neuronal networks are made up of simple processing units called neurons which are usually organized into layers with full or partial connections. The principal task associated with a neuron is to receive the activation values from its neighbors, compute an output based on its weighted input parameters and send that output to its neighbors. ANNs have already been used in few instances in photogrametry and satellite image processing. In our method they are applied as sophisticated pattern classifiers. We have chosen a feed-forward backpropagation neural network, which is one of the most frequently implemented network types.

Learning (training) neural networks is a time-consuming task. For its efficiency, the Back Propagation learning algorithm which is an iterative gradient decent algorithm, was used. It is designed to minimize error function expressed in Equation (1):

$$E = \frac{1}{2}\sum\nolimits_{j=1}^{N}(D_j - O_j^M)^2 \tag{1}$$

where D_j and O_j are the desired input and the current response of the neurons j in the input layer, respectively and N the number of neurons in the output layer. The iterative method, corrections to weight parameters are computed and added to the previous values as illustrated below:

$$\begin{cases} \Delta w_{i,j} = -\eta \dfrac{\partial E}{\partial w_{i,j}} \\ \Delta w_{i,j}(t+1) = \Delta w_{i,j} + \alpha \Delta w_{i,j}(t) \end{cases} \tag{2}$$

In this equation, $w_{i,j}$ is weight parameter between neurons i and j, η a positive constant that controls the amount of adjustment and is called learning rate, a α momentum factor that can take on values between 0 and 1 and "t" denotes the iteration number. The parameter α can be called smoothing or stabilizing factor as it smoothes the rapid changes between the weights.

4. ARTIFICIAL NEURAL NETWORKS FOR ROAD EXTRACTION

Road detection from satellite images can be considered as a classification process in which pixels are divided into road and background classes. A backpropagation neural network (BNN) with one hidden layer is used. Normalized spectral information in a window of (3×3) around each pixel of RGB images are used, as 9 red neighbours pixels, 9 green neighbours pixels and 9 blue neighbours pixels to constitute the input vector of 27 neurons. The output layer consists of one neuron that represents the networks output by a number between 0 and 1 as not road and road pixel, respectively, Figure 1.

As the neuronal network requires a large coded data bases in the training stage, a set of 100 road net manually drown using special software are chosen as training set in learning stage. It is recommended to have representative pixels of all present objects in the training set.

5. RESULT AND INTERPRETATION

The combination of both 27 input parameters made the network powerful in the detection of road and background, reducing also the request hidden layer and size iteration time. We used as a first step a set of, 10, 15 and 20 neurons in the hidden layer to test the performance of the neural network and we found that a layer of 20 neurons in the hidden layer is sufficient to improve the network's ability in both road and background detection. Figure 2 (a_1), (a_2) shows sections of a satellite image SPOT XS (2007) from the city of Laghouat (Algeria). They were considered as original image and are applied to the input of the neural network. (b_1), (b_2) are their respective produced road map used in accuracy assessment. The results of the neural network for road detection are shown by image (c_1), (c_3).

For accuracy assessment, we consider two parameters: the mean square error (MSE) and the Kappa coefficient, Table 2.

The MSE, proved to be the most reliable parameter to be used as termination condition, improve the accuracy of the results.

The Kappa coefficient, the overall accuracy parameter, is calculated by the same way as classification methods.

The proposed ANN has presented no over-trained problem.

To improve extracted roads, we propose the application of two morphological operations to the ANN extracted roads:

A. Morphological erosion is applied to the gray scale image extracted by the road ANN model

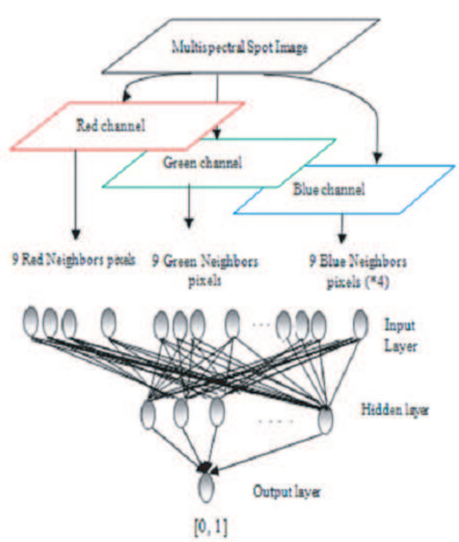


Figure 1: Proposed neural net for road extraction.

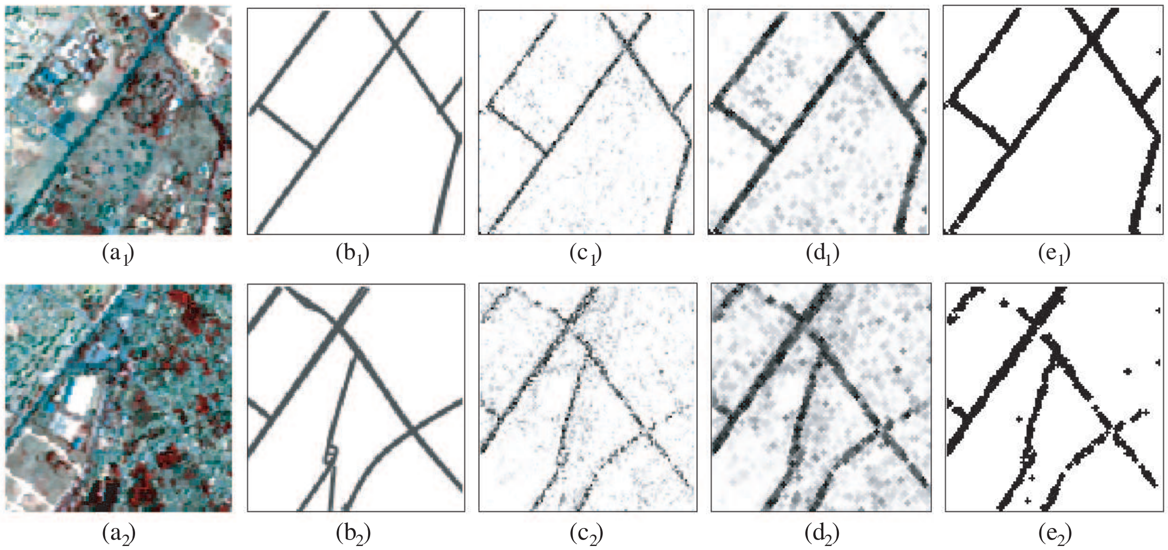


Figure 2: Result of Proposed ANN for Road detection: (a_1), (a_2) Sections of Spot Image (XS, 2007) of the city of Laghouat (Algeria). (b_1), (b_2) Manually produced maps. (c_1), (c_2) Extracted road by ANN model. (c_1) and (c_2) Road extracted by ANN model. (d_1) and (d_2) Road after gray scale erosion. (e_1) and (e_2) binarization by threshod.

Table 2: Accuracy assignment.

Image	**MSE**	**Kappa Coefficient**
(a_1)	0.0206	86.2%
(b_1)	0.0211	84.5%

with a chosen structuring element in order to smooth roads without modifying its wideness. Erosion generally decreases the size of objects and removes small anomalies by subtracting objects with a radius smaller than the structuring element. The result of this algorithm applied to images (c_1) and (c_2) are respectively shown in (d_1) and (d_2) of Figure 2.

B. The obtained images are then transformed to binary image by threshold; the result is shown in images (e_1) and (e_2).

6. CONCLUSIONS

In this paper, a method based on neuronal net strategy for detecting road net in the city of Laghouat (Algeria) is presented. This method is based on spectral characteristics of the pixel from satellite image. As the neuronal net require a large coded data bases in their training stage, we have used a set of road net manually drown using special software. We have obtained very accurate results with less than 0.022 for the MSE. A set of applications are presented. This approach is distinguished from previous works by the choice and the structure of the multilayer neural network input, which is mainly based on the spectral characteristics of the pixel intensity level in the three channels (red , green and blue) and the neighbors of the considered pixel, that influence greatly the quality of output (extracted road network image). These results are very accurate since the method can extract the road despite the resolution of the image (10 m). It is important to notice that, automatic extraction of urban roads from remote sensing imagery is still a challenging problem in digital Photogrammetry and computer vision. The main reason is that the diverse road surfaces and the complex surrounding environments such as trees, vehicles and shadows induced by high buildings make the urban roads take on different textures and gray levels in images. We propose some morphological operations in order to obtain the road cartography: grayscale erosion is applied to the extracted road by the proposed ANN system, followed by a binarization process. For further work we propose using recalled images and geometric characteristics of road for network's training in order to improve network's ability in road detection.

ACKNOWLEDGMENT

Our thanks to Professor André Ozer, Director of the Geomorphology and Geography Laboratory, University of Liege (Belgium) and Doctor Marc Salmon, to provide us the image data needed to develop this work.

REFERENCES

1. Fisher, M. A., J. M. Tenenbaum, and H. C. Wolf, "Detection of road and linear structures in low-resolution areal imagery using a multisource knowledge integration technique," *CGIP*, No. 15, 201–223, 1982.
2. Roux, M., "Recalage images SPOT/carte routière," *International Archives of Photogrammetry and Remote Sensing*, Vol. 29, Tome B3, Commission III, 384–391, 1992.
3. Zhang, C., S. Murai, and E. Baltsairas, "Road network detection by mathematical morphology," *ISPRS Workshop on 3D Geospatial Data Production: Meeting Application Requirements*, 185–200, Paris, 1999.
4. Mohammadzadeh, A., A. Tavakoli, and M. J. Valadan Zoej, "Automaticlinear feature extraction of Iranian roads from high resolution multi-spectral satellite imagery," *20th ISPRS Congress*, 764–768, Istanbul, Turkey, 2004.
5. Hinz, S., A. Baumgartner, C. Steger, H. Mayer, W. Eckstein, H. Ebner, and B. Radig, "Road extraction in rural and urban areas, semantic modeling for the acquisition of topographic information from images and maps," *SM-ATI'99*, 7–27, München, 1999.
6. Geman, D. and B. Jedynak, "An active testing model for tracking roads in satellite images," *IEEE Trans. PAMI*, Vol. 18, No. 1, 1–14, January 1996.
7. Narayanan, P. J., et al., "Extraction of main road high satellite images by machine learning," *ACCV 2006*, LNCS 3851, 236–245, 2006.
8. Sowmya, A. and S. Singh, "Rail: Extraction road segments from areal image using machine learning," *Proc. ICML 99 Workshop on Learning in Vision*, 8–18, 1999.
9. Marikhu, R., M. N. Dailey, S. Makhanov, and K. Honda, "A family of quadratic snakes for road extraction," *ACCV 2007*, Part 1, LNCS 4843, 85–94, 2007.
10. Hinz, S. and A. Baumgartner, "Automatic extraction of urban road networks from multi-view aerial imagery," *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 58, 83–98, 2003.
11. Dal Poz, A. P. and G. M. do Vale, "Dynamic programming approach for semi-automated road extraction from medium- and high-resolution images," *ISPRS Archives*, Vol. 34, Part 3/W8, 17–19, Munich, September 2003.
12. Mokhtarzade, M., H. Ebadi, and M. J. V. Zoej, "Optimisation of road detection from high resolution satellite images using texture parameters in neuronal network classifiers," *Canadian Journal of Remote Sensing*, Vol. 33, No. 6, 484–491.
13. Mokhtarzade, M. and M. J. V. Zoej, "Road detection from high-resolution satellite images using artificial neural networks," *International Journal of Applied Earth Observation and Geoinformation*, Vol. 9, No. 1, 32–40, 2007.
14. Mokhtarzade, M., M. J. V. Zoej, and H. Ebadi, "Automatic road extraction from high resolution satellite images using neuronal net work, texture analysis, fuzzy clustering, Genetic algorithm," *The International archive of Photogrammetry, Remote Sensing and Spacial Information Sciences*, Vol. 37, Part B3b, Beijing, 2008.
15. Mohammadzadeh, A., A. Tavakoli, and M. J. V. Zoej, "Road extraction based on fuzzy logic and mathematical morphology from pan sharpened IKONOS images," *The Photogrammetric Record*, Vol. 21, No. 113, 44–60, 2006.
16. Civco, D. L. and Y. Waug, "Classification of multi-spectral, multitemporal multi-source spatial data using artificial neural networks," *Proceeding of the ASPRS 1994 Annual Convection*, 123–133, Reno, NV, USA, 1994.

Novel Compact RFID Chipless Tag

A. Vena, E. Perret, and S. Tedjni
Laboratoire de Conception et d'Intégration des Systèmes (LCIS), Grenoble Institute of Technology, France

Abstract— In this communication a new RFID chipless tag is presented. The density of coding per surface of this novel tag is important and reaches 3.3 bits/cm^2. For a surface of $1.5 \times 2\,\text{cm}^2$, it is possible to encode 10 bits. The design presented in this paper, based on multiband coplanar strip, brings enhancement in term of miniaturization, configurability and cost of fabrication. Unlike previous solution only one layer of metal is necessary. So, such a tag could be easily printed with conductive inks. Both frequency domain measurement using a VNA and Time domain measurement based on radar approach are compared.

1. INTRODUCTION

RFID is a technology that allows identification and/or authentication of a remote tag by use of radio frequency waves Fig. 1(a). Today RFID system is largely used in fare collection, road toll system and stock managing. But a bigger market for item tracking representing every year 10 trillion items sold [1] is still addressed today by optical barcode. To enter this market, RFID brings some good arguments like increasing reading range and making possible the reading outer of its line of sight. But very low cost per unit has to be achieved before this technology can be largely used. One way for cost reduction is to develop chipless tag solution directly printable on products during fabrication process, as shown in Fig. 1(b).

Since few years several papers can be found on this topic. The temporal approach is based on controlling the presence and position of multiple reflection of a tag in response to an incident field. This technique is used by the SAW chipless tag [2], which operates in ISM band of 2.45 GHz. But tag cost still high due to its piezoelectric substrate. Others solutions have been proposed using a structure with delay line [3] to have a lower coding capacity for a much higher size. The frequency approach, is based on controlling amplitude and phase of multiple resonances in function of a specific code [4, 5], and today the structure that codes the most of data is that used by S. Preradovic et al. [4] with 35 resonators corresponding to 35 bits of data having a size more important than credit card format. The design presented in this paper brings enhancement in term of miniaturization, configurability and fabrication cost. Indeed unit cost can decrease dramatically since this tag is potentially fully printable on only one layer of conductive ink because there is no ground plane.

Measurement was done in frequency domain and in time domain. Frequency domain is used as a first step to get response of the tag with accuracy, in the entire bandwidth. Preradovic et al. is going in this way and made a first prototype of reader working in the frequency domain [6]. Time domain measurement is used as a second step because it brings several advantages. Indeed with a short pulse generator, a very large frequency band can be read in one shot, and very high level of energy can be concentrated on this short time keeping respect of average power of ISM band as UWB reader does. Moreover short pulse can also serves to localize target.

In Section 2, it will be presented the design of the tag based on multiband coplanar strip antenna. Then, in the Section 3, measurement set-up and results in frequency domain and time domain are presented and compared. Discussion on coding capacity and further evolutions are presented in the last section.

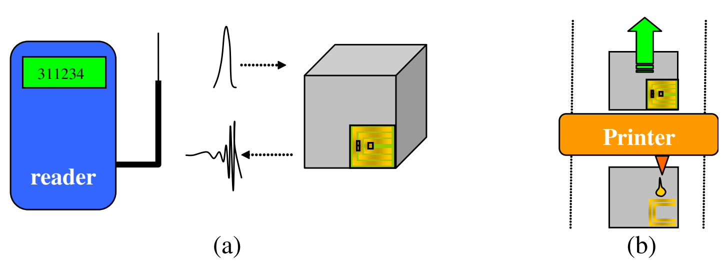


Figure 1: (a) Automatic reading system using RF waves. (b) The chipless tag is directly printed on product in production phase.

2. TAG DESIGN

2.1. Description

The tag designed in this study, Fig. 2, is based on a multiple band coplanar strips resonator with open ended extremity. It uses 4 resonators, denoted 1, 2, 3 and 4 on Fig. 2, having a resonant frequency that can be controlled independently only for resonator 1, 2 and 4. Indeed it was observed in simulation, that resonator 3 is largely influenced by changes in surface current of other modes. As a result it can not be used. It is to be noted that opening the coplanar strip line on one extremity and setting a short circuit on the other side produces a quarter wavelength resonance. Combination of quarter wavelength resonator and path bending, denoted by "C path", makes an improvement in term of size reduction.

2.2. The Encoding

Here, the choice that is used to code information is varying the resonant frequency of each resonator around a reference. This coding technique allows a significant decrease of tag size compared to previous technique [5].

There are two ways to make a frequency shift for a resonant mode. The first one is simply, modifying the length of coplanar strip line. The second way, used in this paper, is to add some short-circuit inside slot as shown in Fig. 3(a). In Fig. 3(b), it can be seen that varying the length of short circuit makes a frequency shift relative to one reference. In this example when there is no short circuit, the resonance frequency of the first mode is equal to 2 GHz while its value is equal to 2.05 GHz and 2.2 GHz for respectively a short circuit length of 1 mm and 3 mm.

3. MEASUREMENT RESULTS

3.1. Frequency Domain Measurement

Measurement was done in frequency domain using a VNA Agilent 8720D in anechoic chamber using the bi-static configuration shown in Fig. 4(a). The two horn antennas used are identical and their gain is between 10 dBi and 12 dBi in the frequency band 1.5 GHz to 6 GHz. The power delivered by VNA approaches 0 dBm in the entire frequency band. To eliminate direct path effect between two antennas, a reference measure has been taken without any tag. Then, systematically, further measurements with presence of tags are subtracted from the reference measurement directly on the VNA.

As it can be seen in Figs. 5(a), (b), (c), level of signal backscattered is very low and approaches -70 dBm, but large dynamic range of VNA produces a reliable result. In Table 1, it is presented the resonant frequency of each mode in function of tag configuration.

Table 1: Resonance frequency measured in function of tag coding.

	Mode 1 (GHz)	Mode 2 (GHz)	Mode 4 (GHz)
Tag 1: $L_1 = 0$ mm, $L_2 = 1$ mm, $L_3 = 0$ mm	1.98	2.49	4.09
Tag 2: $L_1 = 3.5$ mm, $L_2 = 0$ mm, $L_3 = 0$ mm	2.22	2.44	4.09
Tag 3: $L_1 = 0$ mm, $L_2 = 0$ mm, $L_3 = 3.5$ mm	2	2.42	5.53

From Tag 1 to Tag 2, change on short circuit length denoted L_1 and L_2 produces a frequency shift of 280 MHz on mode 1 and 50 MHz on mode 2. From Tag 1 to Tag 3 changes on L_2 and L_3 make a shift of 80 MHz on mode 2 and 1440 MHz on mode 4. To set limit to the frequency

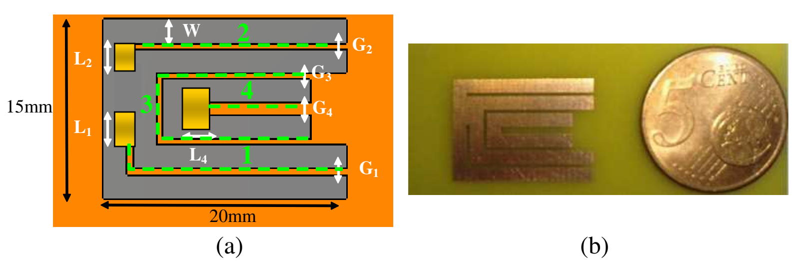
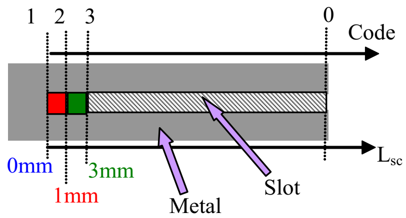
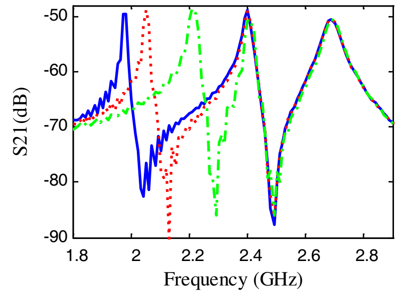


Figure 2: (a) Geometry of the chipless RFID tag. (b) Photography of designed tag 1 (15 mm × 20 mm). Substrate is FR4, and thickness is 0.8 mm.

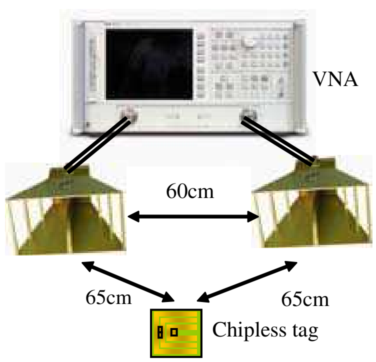


(a)

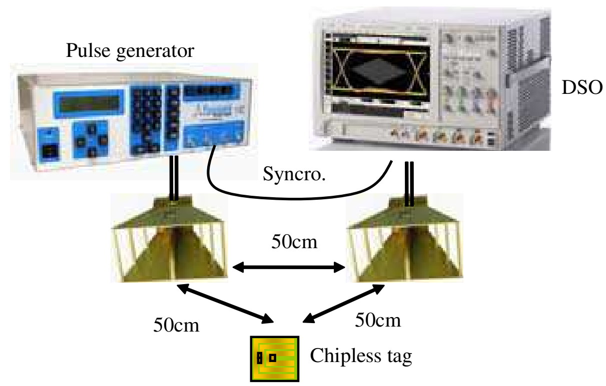


(b)

Figure 3: (a) Short circuit setting in function of code. (b) Simulated resonance shift in function of short circuit setting for mode 1. (_) $L_{sc} = 0$ mm, (- -) $L_{sc} = 1$ mm, (-.-.) $L_{sc} = 3$ mm.



(a)



(b)

Figure 4: Measurement set-up (a) in frequency domain (b) in time domain.

resolution it can be observed that, from Tag 1 to Tag 3, even if L_1 and L_2 are left unchanged it can be noticed a frequency shift of 20 MHz on the mode 1. The same effect can be observed on the mode 2 between Tag 2 and Tag 3. This phenomenon appears because the resonators are not fully isolated, and to make this coding technique reliable, it is necessary to limit the frequency resolution at 5 MHz.

3.2. Time Domain Measurement

In a second time, to compare, a time domain measurement was done. In the Fig. 4(b), it is presented the measurement set-up that was used. DSO used is an Agilent DSO91304A having a bandwidth of 13 GHz and a sampling rate of 40 Gs/s in real time. The pulse generator Picosecond 10060 A that was used is able to send a Gaussian pulse of 110 ps of width, with maximum amplitude of 2 V into a 50 Ω load as shown in Fig. 6(a), that gives a maximum instantaneous power of 19 dBm. The corresponding Power spectral density (PSD) relative to its maximum is given in Fig. 6(b). The curve shows a quasi monotonic decrease from −1 dB at 1.5 GHz to −5 dB at 6 GHz that could be compensated in the receiver module. In the Fig. 6(c) it can be seen the backscattered transient response of the chipless Tag 1, due to the incident Gaussian pulse of Fig. 6(a). This measurement was obtained with repetitive pulse and applying an averaging factor of 64 in order to increase the signal to noise ratio (SNR). Transient response is hard to analyze and to get the frequency response, a FFT has been applied with the window shown in Fig. 6(c).

In Figs. 7(a), (b), (c), it is shown, the frequency response of the three tags used after applying a FFT on each transient response. It can be seen that both results are very close together for the 3 first modes in the band starting from 1.5 GHz to 4 GHz. For higher frequency it is more difficult in case of time domain to see presence of resonance like for the 5.5 GHz resonance of Tag 3 shown in Fig. 7(c). This is mainly due to the decreasing value of PSD of Gaussian pulse as explained previously.

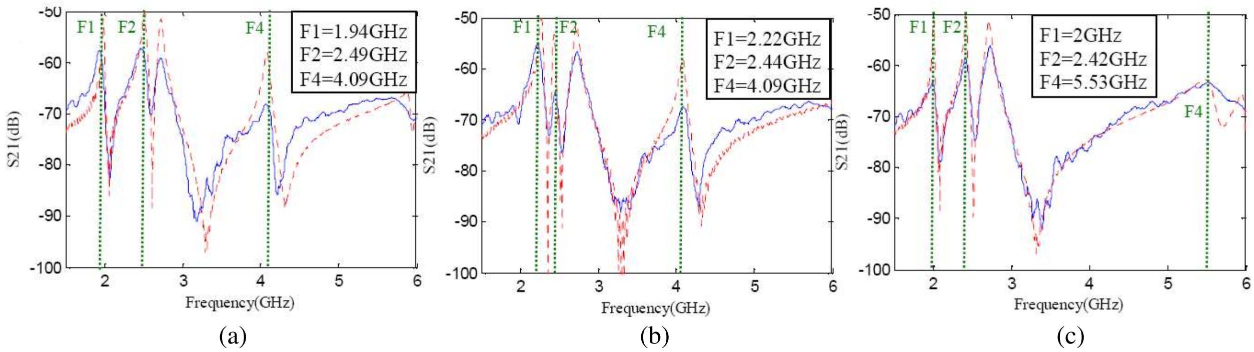


Figure 5: Measurement results in frequency domain (solid curve) and simulation results (dotted curve) for the tags. (a) Tag 1. (b) Tag 2. (c) Tag 3. For each, configuration used is described in Table 1.

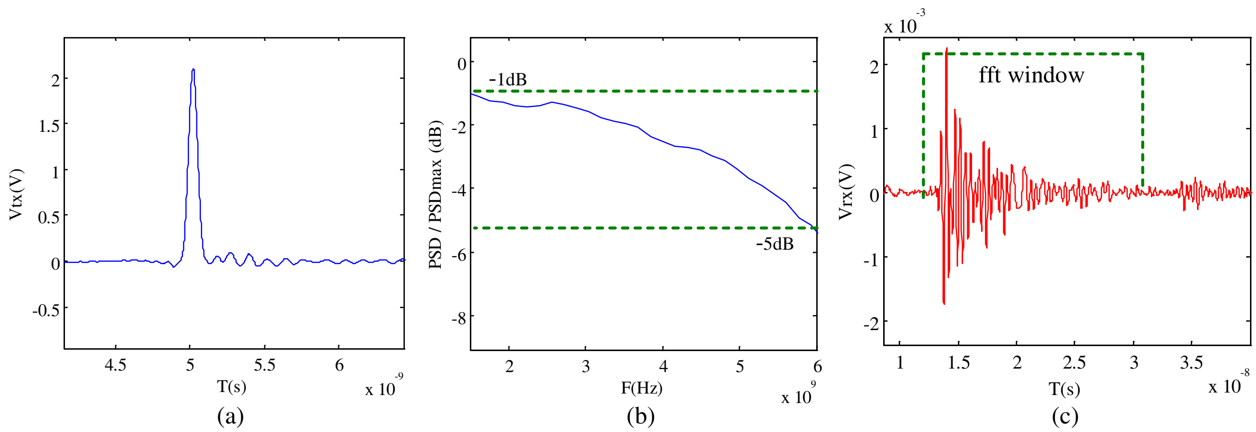


Figure 6: (a) Measured transmitting pulse shape. (b) Measured frequency response of transmitting pulse. (c) Measured transient response of Tag 1.

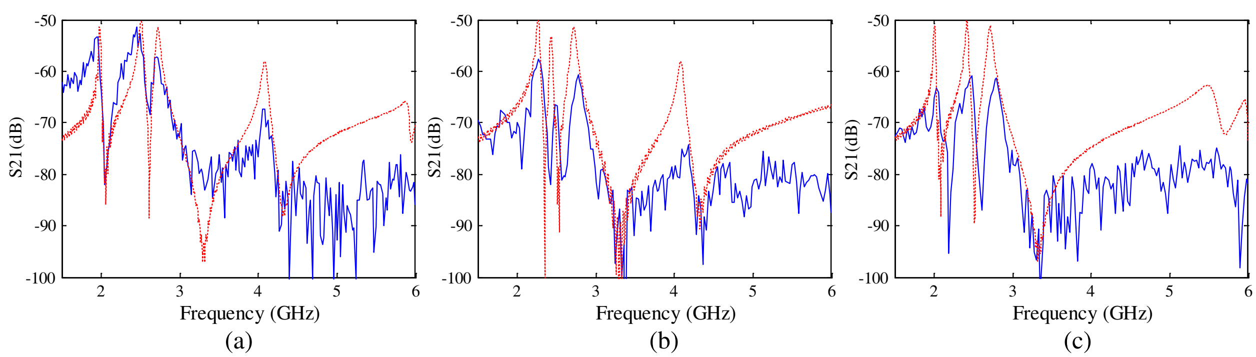


Figure 7: Measurement results in time domain (solid curve) and simulation results (dotted curve) for (a) Tag 1. (b) Tag 2. (c) Tag 3. For each, configuration used is described in Table 1.

4. CAPACITY OF CODING

In the present design, it has been used three short-circuit to configure frequency shift of 3 modes of resonances. For mode 1, the frequency value can vary between 2 GHz and 2.4 GHz, making a bandwidth of 400 MHz. For mode 2, a frequency span between 2.4 GHz and 2.7 GHz, makes a bandwidth of 300 MHz. And Mode 3 can vary between 4 GHz and 5.5 GHz, making a bandwidth of 1500 MHz. Frequency resolution is set to 50 MHz.

$$N = \prod_{i=1}^{k} \frac{BW_i}{\Delta f_i} \tag{1}$$

Using Equation (1) with BW_i the bandwidth of mode k and Δf_i its frequency resolution, it can be found that 8 different frequency values can be encoded for the first mode, 6 frequency values for the second mode and 30 for mode 4, giving a global capacity of $8 * 6 * 30 = 1440$ levels, i.e., 10 bits within a reduced surface of $15\,\mathrm{mm} \times 20\,\mathrm{mm}$. A density of coding per surface of $3.3\,\mathrm{bits/cm^2}$ is reached, which is a good result compared to previous designs which are $0.61\,\mathrm{bits/cm^2}$ for Preradovic et al. [4] and $0.81\,\mathrm{bits/cm^2}$ for Jalaly et al. [5].

5. CONCLUSIONS

It has been shown in this paper that a much reduced size chipless tag can embed a large data capacity that is currently 10 bits. With improvements, a much higher capacity can be obtained. In top of its reduced size, cost of fabrication for this tag is very low mainly because there is no ground plane. Indeed, the chipless tag can be potentially printed with conductive ink, with only one layer. Time domain measurement using a radar approach gives good results even if sensitivity is much lower than for VNA measurement, indeed results obtained are very similar for both techniques. However it has been observed that highest mode frequency is more difficult to read because pulse used has a decreasing power spectral density when frequency rises. In a further work, time domain measurement technique will be improved and transmitting pulse shape will be adjusted in order to limit this effect.

ACKNOWLEDGMENT

The authors would like to thank Prof. L. Duvillaret and Dr. F. Garet, for guidance and fruitful discussions on part of this work. The authors are grateful to Grenoble-inp for supporting this project via the BQR programme.

REFERENCES

1. IDTechEx, *Printed and Chipless RFID Forecasts, Technologies* & *Players 2009–2019*, www.IdtechEx.com.
2. Hartmann, C. S., "A global SAW ID tag with large data capacity," *IEEE Ultrasonics Symposium*, 2002.
3. Zheng, L., S. Rodriguez, L. Zhang, B. Shao, and L.-R. Zheng, "Design and implementation of a fully reconfigurable chipless RFID tag using Inkjet printing technology," *IEEE International Symposium on Circuits and Systems, ISCAS*, 2008.
4. Preradovic, S. and N. C. Karmakar, "Design of fully printable planar chipless RFID transponder with 35-bit data capacity," *Proceedings of the 39th European Microwave Conference*, Rome, Italy, 2009.
5. Jalaly, I. and D. Robertson, "RF barcodes using multiple frequency bands," *IEEE MTT-S Digest*, June 2005.
6. Preradovic, S., N. Karmakar, and M. Zenere, "UWB chipless tag RFID reader design," *Proceedings of the IEEE RFID-TA*, Guangzhou, China, June 2010.

RFID Tag Antenna Design on Metallic Surface by Using Rectangular Micro-strip Feed

Ding-Bing Lin[1], **Chao-Chieh Wang**[1], **I-Tseng Tang**[2], **and Mau-Phon Houng**[3]

[1]Institute of Computer and Communication Engineering
National Taipei University of Technology, Taipei, Taiwan
[2]Department of Environment and Energy, National University of Tainan, Tainan, Taiwan
[3]Institute of Microelectronics, National Cheng Kung University, Tainan, Taiwan

Abstract— A novel UHF RFID antenna structure that can stick on metallic surface is presented. A printed rectangular antenna by using micro-strip feed structure on system ground which has the same size as antenna. By using the inductive property of the micro-strip line with short end to match the capacitive reactance of tag chip, and the performance in free space on metallic surface was measured. The half-power matching bandwidth (VSWR < 3) of the antenna in free space was measured to be in the frequency range 914–934 MHz. The measured reading range in free-space is 5–6 m.

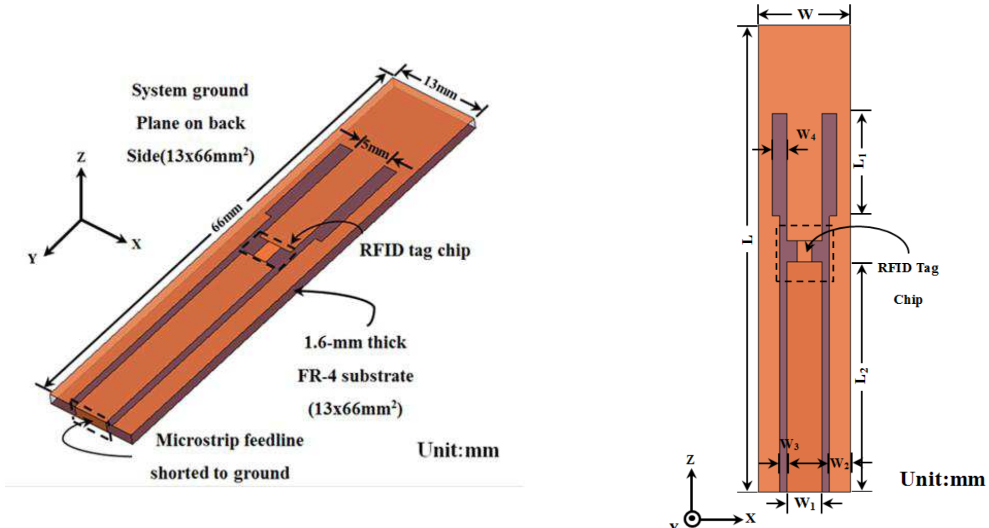
1. INTRODUCTION

In recent years, radio-frequency identification (RFID) technology has become a popular application since it provides a convenient identification information, small size, long reading range and fast reading speed. Numerous papers have been written describing how to construct an antenna for RFID. A good survey can be found in [1], and a more detailed account is given in [2]. The operating band of RFID in UHF is mainly from 860 to 960 MHz. When the operating frequency of RFID rises to microwave region, the antenna design including the impedance matching becomes more important to enhance the system performance. To provide a better impedance matching network between an antenna and a tag chip is the way to improve the chip power and maximize the reading range. As the cost and fabrication requirements, the antenna must directly match to large capacitive reactance and small resistance of tag chip that different from 50 ohm. Because of chip impedance with high Q, the design of a matched antenna is difficult. T-matching networks or inductively coupled [3, 4] feed are commonly used for the efficient matching of UHF tags. Most RFID antennas are currently used to stick on nonmetal material or easy card, but not on metallic material [5–7]. Numerous papers have been written describing how to design a RFID tag antenna on metallic material [8–10].

In this paper, we propose a novel broadband UHF RFID tag antenna. With the structure of patch, the input impedance of the proposed antenna is easy to be tuned. On the other hand, with the structure micro-strip shorted to the ground plane through corresponding shorting strips, the proposed antenna shows an acceptable performance when the tag antenna is attached to a metallic surface. The measured half-power bandwidth of the proposed antenna is 121 MHz (13.5%, 830–951 MHz).

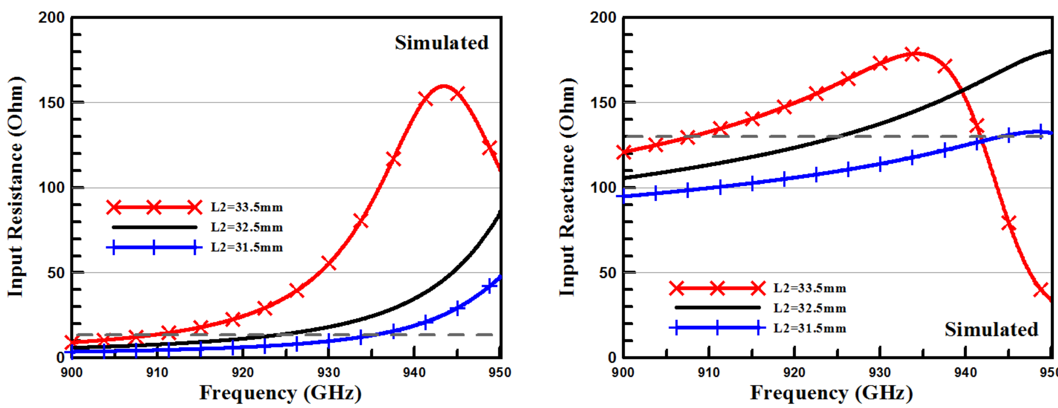
2. RFID TAG DESIGN

The proposed antenna structure is shown in Fig. 1. The antenna discussed in this paper uses both rectangular patch and micro-strip to achieve conjugation match. In this study, a 1.6-mm thick FR4 substrate with relative permittivity (ε_r) 4.4 and area $66 \times 13\,\text{mm}^2$ is considered. On its back side, a system ground plane with the same size is also printed. The micro-strip line is electrically shorted to the ground plane through the shorting strips at the end terminal. Optimized design parameters were found with the following dimensions: $L = 66\,\text{mm}$, $W = 13\,\text{mm}$, $L_1 = 14.5\,\text{mm}$, $L_2 = 32.5\,\text{mm}$, $W_1 = 0.9\,\text{mm}$, $W_2 = 3\,\text{mm}$, $W_3 = 1\,\text{mm}$, $W_4 = 2\,\text{mm}$. The chip conforms to the EPC global Class-1 Gen-2 specification and provides state-of-the-art performance for a broad range of UHF RFID tagging applications, whose input impedance was $11 - j131$ Ohms equal to that of an RC parallel circuit with $R = 1.5\,\text{k}$ and $C = 1.58\,\text{pF}$ at the frequency of 925 MHz. The conjugate matching between the antenna and the micro-strip can be tuned by adjusting the perimeter and the position of the RFID Tag chip. In such a way, a good broadband characteristic can be attained for the antenna.



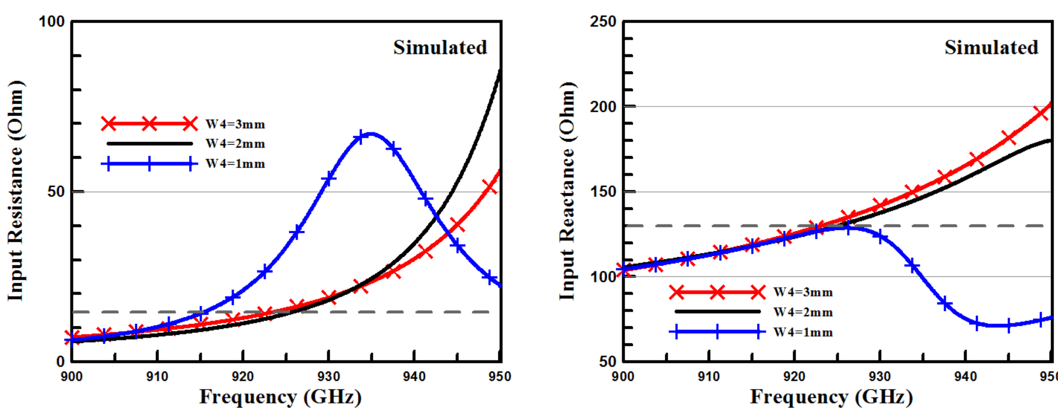
(a) (b)

Figure 1: (a) Geometry of the printed UHF RFID antenna stick on metallic surface operation in the 925 MHz. (b) Dimensions of the metal pattern of the UHF RFID antenna.



(a) (b)

Figure 2: (a) Resistance of the proposed antenna for variation in L_2. (b) Reactance of the proposed antenna for variation in L_2.



(a) (b)

Figure 3: (a) Resistance of the proposed antenna for variation in W_4. (b) Reactance of the proposed antenna for variation in W_4.

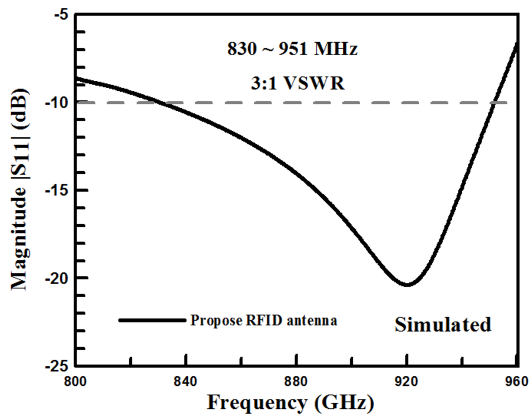


Figure 4: Calculated S_{11} characteristics of the proposed antenna.

3. SIMULATION AND COMPARISON

Ansoft HFSS software [11], FEKO EM Software [12] and Advantest E8362B Network Analyzer were utilized for simulations and measurements. Figs. 2(a), (b) shows the input impedance characteristics of the proposed antenna and the position of the micro-strip L_2 is decided to be 32.5 mm for achieving the resistance and reactance of about $13 + j131$ ohms as shown. Furthermore, Figs. 3(a), (b) show that the required resistance of the tag antenna can be obtained by adjusting W_4. Fig. 4 shows the return loss of the antenna. The half-power bandwidth (return loss, -20.3 dB) of the proposed antenna in measurement is 121 MHz (13.5%), from 830 to 951 MHz.

4. CONCLUSIONS

A novel RFID tag antenna is proposed for the UHF RFID band. The structures of the rectangular patch and micro-strip shorted to ground plane are employed to help getting broadband characteristics and acceptable performance on metallic objects. The return loss performance and the maximum readable distances of the proposed antenna both show that the proposed antenna can work not only in air but also on metallic objects within the UHF RFID band and the maximum readable distances are over 6.2 m.

ACKNOWLEDGMENT

This work was supported by the National Science Council of Taiwan under grant NSC 99-2220-E-027-006 and NSC 99-2515-S-027-001-.

REFERENCES

1. Marrocco, G., "The art of UHF RFID antenna design: impedance-matching and size-reduction techniques," *IEEE Antennas and Propagation Magazine*, Vol. 50, No. 1, 66–79, Feb. 2008.
2. Rao, K. V. S., P. V. Nikitin, and S. F. Lam, "Antenna design for UHFRFID tags: A review and a practical application," *IEEE Transactions and Propagation*, Vol. 53, No. 12, 3870–3876, Dec. 2005.
3. Choo, J., J. Ryoo, J. Hong, H. Jeon, C. Choi, and M. M. Tentzeris, "T-matching networks for the efficient matching of practical RFID tags," *Microwave Conference*, 5–8, Oct. 2009.
4. Son, H. W. and C. S. Pyo, "Design of RFID tag antennas using an inductively coupled feed," *Electronics Letters*, Vol. 41, 994–996, Sep. 2005.
5. Hong, S. J., B. Yu, F. J. Harackiewicz, M. J. Park, Y. S. Chung, and B. Lee, "Crab shaped antenna for RFID tag at UHF band," *Antennas and Propagation Society International Symposium*, 1–4, Sep. 2008.
6. Choi, Y., U. Kim, and J. Choi, "Design of a dipole tag antenna enclosed by a short-stub for UHF RFID application," *Antennas and Propagation Society International Symposium*, 1–4, Sep. 2008.
7. Deavours, D. D., "Analysis and design of wideband passive UHF RFID tags using a circuit model," *IEEE International Conference on RFID*, 283–290, May 2009.
8. Cho, C., H. Choo, and I. Park, "Broadband RFID tag antenna with quasi-isotropic radiation pattern," *Electronics Letters*, Vol. 41, 1091–1092, Oct. 2005.

9. Xu, L., B. J. Hu, and J. Wang, "UHF RFID tag antenna with broadband characteristic," *Electronics Letters*, Vol. 44, 79–80, Jan. 2008.
10. Cho, C., H. Choo, and I. Park, "Design of planar RFID tag antenna for metallic objects," *Electronics Letters*, Vol. 44, 175–177, Feb. 2008.
11. Ansoft HFSS user's Guide-High Frequency Structure Simulator, Ansoft Co., 2003.
12. FEKO User's Manual Suite 5.5, EM Software & Systems-S.A. (Pty) Ltd., South Africa, Jul. 2009, http://www.feko.info.

Simplified Design Approach of Rectangular Spiral Antenna for UHF RFID Tag

Khalil El Khamlichi Drissi[1], **El Mostafa Makroum**[2,3], **Mounir Rifi**[2], **Mohamed Latrach**[4], **and Ali Benbassou**[5]

[1]Laboratoire LASMEA, Université Blaise Pascal, Clermont Ferrand, France
[2]Networks Laboratory, Computer, Telecom & Multimedia, EST, Morocco
[3]CED sciences de l'ingénieurs, ENSEM
Km 7, Route El Jadida, B.P. 8012 Oasis Casablanca Oasis Casablanca, Morocco
[4]Radio & Microwave Group, ESEO, Graduate School of Engineering
4 Rue Merlet de la Boulaye, P. O. 30926, Angers 49009, France
[5]Laboratory of Transmission and Processing of Information EMC/Telecom
School of Technology (TSE), P. O. 2427 FES, Morocco

Abstract— In this paper, we present a method to simplify the calculation of spiral antennas for RFID tag settings without resorting to numerical analysis methods. It saves up to 99% of the time required by the simulation based on the method of moments. Thus we present a theoretical and experimental study for the design of the spiral antennas for RFID label in the UHF band. We present in this study the S_{11} parameter that enables us to evaluate the evolution of current distribution and therefore the resonance frequency of the spiral antennas. This parameter is calculated theoretically by applying the method of moments to wired antenna formed by the rectangular copper spiral printed on a dielectric substrate. The experimental validation of our theoretical models is performed using a network analyzer. The confrontation theory-experience allows us to draw some interesting conclusions concerning the number of loops of the spiral and the choice of dielectric substrate.

1. INTRODUCTION

During these last decades, the technologies of information and communications (ICT) have known unprecedented development. The identification technologies are part of these information technologies. Due to the recent development of microelectronics and wireless systems, new contactless identification technologies have emerged: The radio identification technology (or RFID for Radio-Frequency IDentification). These new technologies, by their greater flexibility, make the exchange of information much faster and efficient.

RFID is a technology to recognize or identify with greater or lesser distance (contact tens of meters) in a minimum of time, an object, an animal or a person with an electronic tag. We can cite, for example, contactless smartcard systems, highway tolls without stopping, parking or building access control, etc..

It falls into the category of automatic identification technologies (AIDC, Automatic Identification and Data Capture), as the bar code character recognition, pattern recognition, or magnetic cards.

RFID systems use mainly four frequency bands [1–3]: 125 kHz (LF band, Low Frequency), 13.56 MHz (HF, High Frequency), 860–960 MHz (UHF, Ultra High frequencies), 2.45 GHz (microwave). In recent years a growing interest in the field of industry and research focused on passive UHF RFID technology. It presents a low cost solution. It also helps to have a data rate higher (around 20 kbit/s) and achieve a reading range greater than other technologies called passive RFID. This interest has helped put in place in most parts of the world regulations and industry standards for market development of this technology.

In the first part of this document, a brief presentation of passive RFID technology is exposed. Then the modeling of a rectangular spiral RFID antenna using the moment method and the results of simulations and measurements are discussed and presented in a second part. Then we develop a method for estimating the peak current and thus the resonant frequency of spiral antennas.

2. THE PRINCIPLE OF PASSIVE RFID TECHNOLOGY

An automatic identification application RFID, as shown in Fig. 1, consists of a base station that transmits a signal at a frequency determined to one or more RFID tags within its field of inquiry. When the tags are "awakened" by the base station, a dialogue is established according to a predefined communication protocol, and data are exchanged.

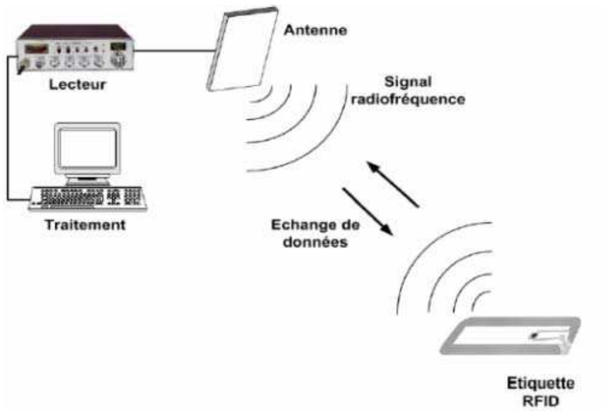


Figure 1: Schematic illustration of an RFID system.

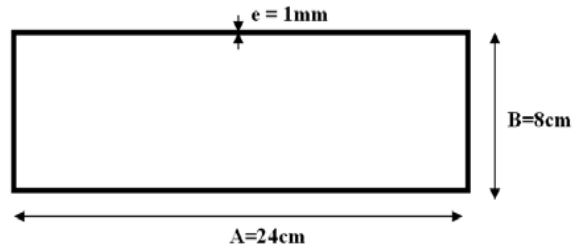


Figure 2: Dimensions of the track used in the simulation.

The tags are also called a transponder or tag, and consist of a microchip associated with an antenna. It is an equipment for receiving an interrogator radio signal and immediately return via radio and the information stored in the chip, such as the unique identification of a product.

Depending on the operating frequency of the coupling between the antenna of the base station and the tag may be an inductive coupling (transformer principle) or radiative (far-field operation). In both cases of coupling, the chip will be powered by a portion of the energy radiated by the base station.

To transmit the information it contains, it will create an amplitude modulation or phase modulation on the carrier frequency. The player receives this information and converts them into binary (0 or 1). In the sense reader to tag, the operation is symmetric, the reader transmits information by modulating the carrier. The modulations are analyzed by the chip and digitized.

3. MODELING OF A SPIRAL ANTENNA IN UHF BAND

To predict the resonant frequencies of the currents induced in the antenna structures forming tags, such as spiral antennas rectangular ICs, we have used initially the model based on the theory of diffraction by thin wires [4]. We arrive at an integro-differential equation, and whose resolution is based on the method of moments [5]. Although a very good result is obtained, the computation time required is a major drawback. It is quite high. In a second step, the study is to find a method to simplify the estimation of the resonance frequency of rectangular spiral antennas in various frequency ranges.

3.1. Method of Moments

It is an integral analysis method used to reduce a functional relationship in a matrix relationship which can be solved by conventional techniques. It allows a systematic study and can adapt to very complex geometric shapes.

This method is more rigorous and involves a more complicated formalism leading to heavy digital development. It applies in cases where the antenna can be decomposed into one or several environments: The electromagnetic field can then be expressed as an integral surface. It implicitly takes into account all modes of radiation.

Moreover, the decomposition of surface current to basis functions, greatly simplifies the solution of integral equations which makes the method simple to implement.

This procedure is based on the following four steps:

- Derivation of integral equation.
- Conversion of the integral equation into a matrix equation.
- Evaluation of the matrix system.
- And solving the matrix equation.

3.2. Formulation of the Method of Moments [5, 6]

We have chosen the configuration shown in Fig. 2, a rectangular metal track, printed on an isolating substrate. It consists of length A, width B and thickness e.

The theory of antennas for connecting the induced current in the metal track to the incident electromagnetic field (Ei, Hi) using integro-differential equation as follows:

$$\vec{t}(l)\vec{E^i}(l) = j\omega\mu \int_0^L I(l')\vec{t}(l')\vec{t}(l)G(R)dl + \frac{j}{\varepsilon\omega}\vec{t}(l)\,\overrightarrow{\mathrm{grad}} \int_0^L \vec{t}(l')\,\overrightarrow{\mathrm{grad}}\, I(l')G(R)dl' \tag{1}$$

The method used to solve such equations is the method of moments [4].

The problem thus is reduced to solving a linear system of the form:

$$[V] = [Zmn][I] \tag{2}$$

with:

$[I]$ representing the currents on each element of the structure.

$[V]$ representing the basic tension across each element m in length Δ given by:

$$Vm = Ei(m)\Delta \tag{3}$$

$[Zmn]$ represents the generalized impedance matrix, reflecting the EM coupling between the different elements of the antenna.

The rectangular loop will be discretized into N identical segments of length Δ:

$$\Delta = 2\frac{(A+B)}{N} \tag{4}$$

The discretization step is chosen so as to ensure the convergence of the method of moments.

$$\Delta = \frac{\lambda}{20} \tag{5}$$

We conclude from this relationship that the number of segments required for convergence of numerical results is:

$$N = 40\frac{(A+B)}{\lambda} \tag{6}$$

λ: Represents the smallest wavelength of EM field incident.

The simulations will focus on frequencies between:

$$100\,\mathrm{MHz} < f < 1.8\,\mathrm{GHz} \quad \text{whether} \quad 17\,\mathrm{cm} < \lambda < 3\,\mathrm{m}$$

The number of segments will be: $N = 76$. The illumination of the loop is done by an plane EM wave, arriving in tangential impact such that the incident electric field Ezi is parallel to the longest track.

The amplitude of the incident field is normalized 1 V/m. we are interested in a loop short-circuited to highlight the resonance phenomena relating to the geometric characteristics of the loop.

We note on the Fig. 3 the presence of two very distinct zones. A first area in which the induced current remains virtually constant. It corresponds to frequencies whose wavelength is greater than about twice the perimeter of the loop.

Beyond these frequencies we observe the appearance of current peaks, showing that the rectangular loop resonates.

To locate the resonance frequencies of the loop, we use the theory of transmission lines (TTL), moderately few adjustments. The induced current $I(z)$ on the track IC has illuminated by a plane wave can be expressed by [4]:

$$I(z) = A_c \sin\left[k\left(\frac{A+B}{2} - |z|\right)\right] + A_d \cos\left[k\left(\frac{A+B}{2} - |z|\right)\right] \tag{7}$$

with:

A_c: The amplitude of common mode current.
A_d: The amplitude of the differential mode current.
$k = \omega\sqrt{\varepsilon\mu}$: wave Vector.

Indeed, A_c and A_d can be written as:

$$A_c = j\frac{e_c(0)}{Z_0\cos\left(k\frac{A+B}{2}\right)} \tag{8}$$

$$A_d = -j\frac{e_d(0)}{Z_0\sin\left(k\frac{A+B}{2}\right)} \tag{9}$$

with Z_0 is the characteristic impedance of a transmission line.

$e_c(0)$ and $e_d(0)$: respectively represent the equivalent electromotive forces at the center of the tracks of common and differential modes.

The coefficients A_c and A_d take an infinite value only if:

$$Z_0\cos\left(k\frac{A+B}{2}\right) = 0 \tag{10}$$

and

$$Z_0\sin\left(k\frac{A+B}{2}\right) = 0 \tag{11}$$

$$\rightarrow \text{ for } Z_0\cos\left(k\frac{A+B}{2}\right) = 0$$

$$\cos\left(k\frac{A+B}{2}\right) = 0 \tag{12}$$

$$\Rightarrow k\frac{A+B}{2} = (2N+1)\frac{\pi}{2} \tag{13}$$

with $N = 0, 1, 2, \ldots$, $k = \omega\sqrt{\varepsilon\mu}$. We have considered the isolated loop, surrounded by air, so $\varepsilon_r = 1$ and $k = \omega\sqrt{\varepsilon_0\mu_0}$

$$k = \frac{\omega}{c} = \frac{2\pi f}{c} \tag{14}$$

(13) becomes: $\frac{2\pi f}{c}\frac{A+B}{2} = (2N+1)\frac{\pi}{2}$.

Hence the resonant frequency common mode:

$$F_{Rc} = (2N+1)\frac{c}{2(A+B)} \tag{15}$$

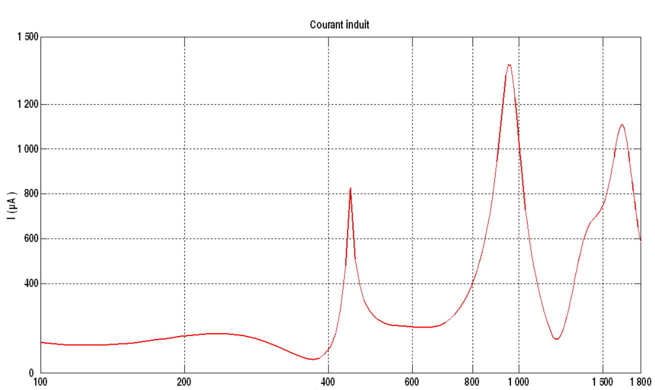


Figure 3: Current induced on the loop 24 × 8 cm simulation by MOM.

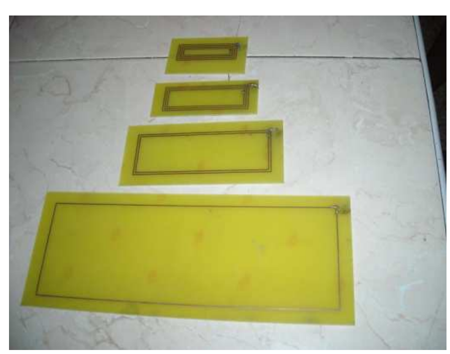


Figure 4: Antenna loops made.

$$\rightarrow \text{ for } Z_0 \sin\left(k\frac{A+B}{2}\right) = 0$$

$$\sin\left(k\frac{A+B}{2}\right) = 0 \tag{16}$$

$$\Rightarrow k\frac{A+B}{2} = N\cdot\pi \tag{17}$$

with $N = 1, 2, 3, \ldots$.

(17) becomes:

$$\frac{2\pi f}{c}\frac{A+B}{2} = N\cdot\pi \tag{18}$$

Therefore, the resonant frequency of the differential mode is:

$$F_{Rd} = N\frac{c}{(A+B)} = 2N\frac{c}{2(A+B)} \tag{19}$$

We therefore find the result observed in Fig. 3, which was obtained by the method of moments.

The resonance frequency of the loop can be easily linked to the length A and width B of the loop by the following approximate relation:

$$F_R = N\frac{c}{2(A+B)} \tag{20}$$

with $c = 3108\,\mathrm{m/s}$, speed of EM waves in vacuum and $N = 1, 2, 3, \ldots$.

4. SIMULATIONS AND MEASUREMENTS

The simulations were done under the MATLAB environment. The experimental validation was performed at the Laboratory LTPI/RUCI[1] in Fez.

4.1. Achievement

We have made various prototypes of antennas as shown in Fig. 4, using as the substrate, glass epoxy, type FR4 with relative permittivity $\varepsilon_r = 4.32$ and 1.53 mm thick.

4.2. Measures and Results

The coefficient of reflection of antennas made, were measured with a vector network analyzer HP-type operating in the 100 Hz–6000 MHz band (Fig. 5).

5. EVALUATION OF PEAKS AND FREQUENCY OF RESONANCE OF INDUCED CURRENTS IN FUNCTION OF GEOMETRIC CHARACTERISTICS OF PRINTED LOOP

The assessment of the size and position of resonance peaks of currents distributed on the printed tracks is crucial for designers of antennas tags. Indeed the action of these peaks can completely change the normal operation of the transponder.

In this part the evolution of the amplitude of these peaks and their resonance frequency will be studied in function of geometric characteristics of loops.

5.1. Evaluation of the Peaks as a Function of the Perimeter of the Printed Loop and the Report (B/A)

In Fig. 6, we found the current to peak resonances for loops with different perimeters (from 40 cm to 120 cm) and reports Width/Length ($B/A = 1/4$, $1/3$ and $2/5$). We find in this figure a linear variation of these peaks as a function of the perimeter of these loops to the same ratio (B/A) thereof. Note that we have made loops whose resonant frequencies are between 400 MHz and 3 GHz.

The peak current amplitude at resonance can be easily connected to the perimeter of the loop by the following equation:

$$I_{pic} = \alpha\cdot P \tag{21}$$

with:

I_{pic}: The amplitude of peak current at resonance.

[1]Laboratory of Transmission and Processing of Information/Regional University Center interface.

Figure 5: Measurement of the reflection coefficient of the antenna using a vector network analyzer.

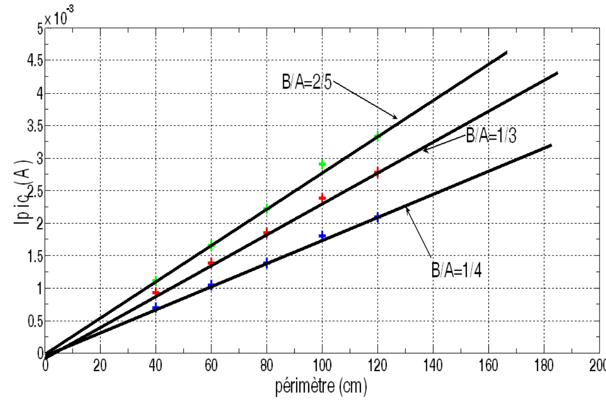
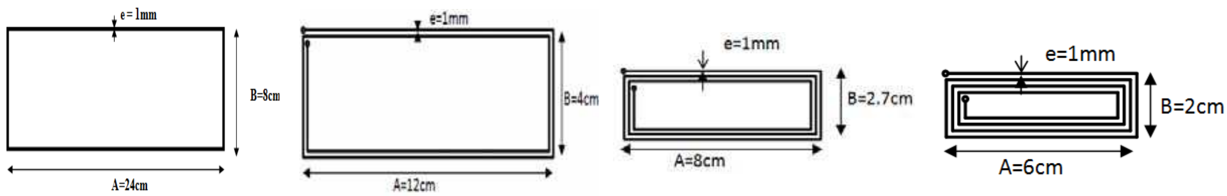


Figure 6: Amplitude of peak current at resonance for loops with different perimeters with a (B/A) constant.



(a) (b) (c) (d)

Figure 7: Dimensions of the Ics tracks used in the simulation and measurements. (a) 1 loop, (b) 2 loops, (c) 3 loops, (d) 4 loops.

α: The slope of the line.
P: Scope of the loop.

- For the $\frac{B}{A}=\frac{1}{4}$ $\quad \alpha_{\frac{1}{4}}=1.735\cdot10^{-3}=\frac{1}{4}\times6.94\cdot10^{-3}$ (22)
- For the $\frac{B}{A}=\frac{1}{3}$ $\quad \alpha_{\frac{1}{3}}=2.353\cdot10^{-3}=\frac{1}{3}\times6.94\cdot10^{-3}$ (23)
- For the $\frac{B}{A}=\frac{2}{5}$ $\quad \alpha_{\frac{2}{5}}=2.776\cdot10^{-3}=\frac{2}{5}\times6.94\cdot10^{-3}$ (24)

From the above, we can write α as follows:

$$\alpha_{\frac{B}{A}}=\frac{B}{A}\times K_t \tag{25}$$

with: $K_t=6.94\cdot10^{-3}$.

The equation can be written as follows:

$$I_{pic\frac{B}{A}}=\frac{B}{A}\times K_t\times P \tag{26}$$

5.2. Evolution of the Resonance Frequencies of the Induced Currents in Function of Numbers of Loops of Rectangular Spiral Antennas (Fixed Perimeter).

The evaluation of the resonance peaks of currents on printed circuit tracks is crucial for designers of printed antennas. Indeed the action of these peaks can completely change the normal operation of the circuit.

We calculated the resonance frequency of the induced currents for four rectangular spiral antennas, respectively having a loop, two loops, three loops and four loops and the same perimeter 64 cm, shown in Fig. 7.

In Fig. 8, we observe that for the same scope the resonance frequencies of induced currents remain almost unchanged in function of the number of loops.

However, we notice a slight difference between the frequencies of resonance peaks of these antennas evaluated theoretically and those measured experimentally. This difference is surely due to the fact that in our theoretical model we do not take account of the dielectric permittivity of the substrate. Indeed, for the simulation we considered a spiral surrounded by air.

To account for the influence of the dielectric permittivity of the substrate on the positions of resonance frequencies, consider the case of a single loop. On Fig. 9, we have plotted the S_{11} parameter for different values of ε_r substrate. We effectively note that this setting actually influence the position of the resonance frequency.

We can take advantage of this finding to try to design antennas that resonate at particular frequencies by playing on the nature of the dielectric substrates. This will allow the miniaturization of spiral antennas for RFID applications.

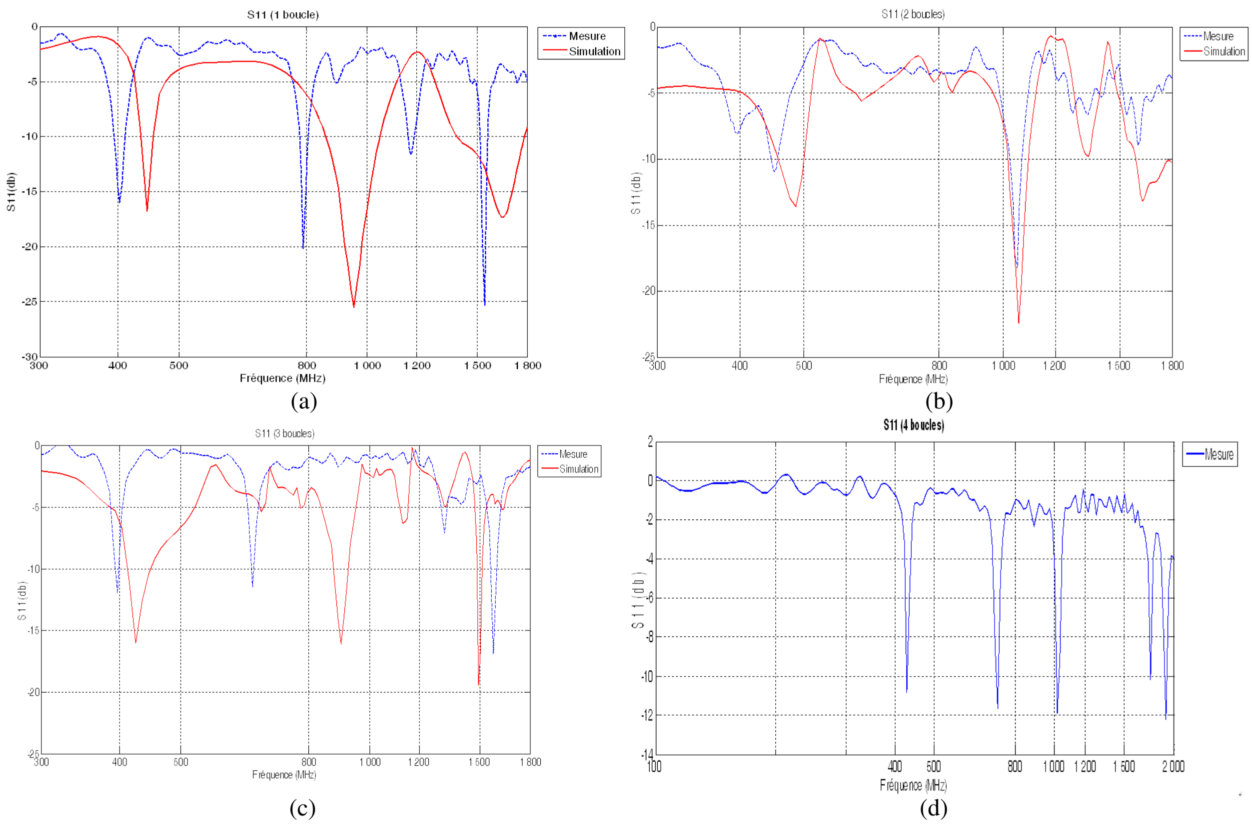


Figure 8: Reflection coefficient S_{11} on the tracks of the ICs in Fig. 7 — simulation by MOM and measurement. (a) 1 loop, (b) 2 loops, (c) 3 loops, (d) 4 loops.

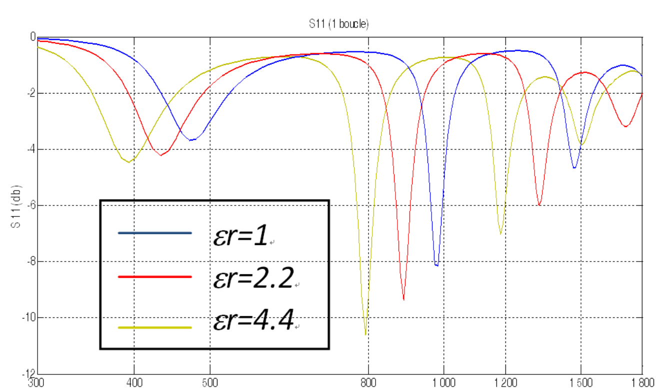
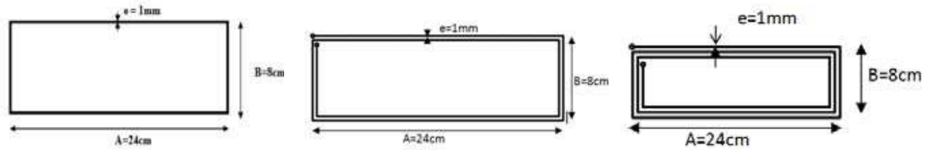


Figure 9: Influence of ε_r on peak resonances of spiral antennas.



(a) (b) (c)

Figure 10: Dimensions of tracks of printed antennas used in the simulation. (a) 1 loop, (b) 2 loops, (c) 3 loops.

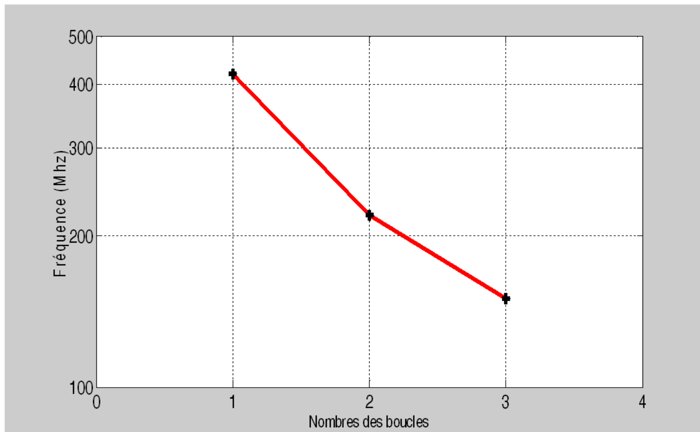


Figure 11: Frequencies of resonance in function of numbers of loops on the tracks of the ICs of the Fig. 10.

5.3. Evolution of the Resonance Frequencies of the Induced Currents in Function of Numbers of Loops of Rectangular Spiral Antennas (A and B Fixed).

We have calculated the resonance frequencies of the induced currents for three rectangular spiral antennas, respectively having a loop, 2 loops and 3 loops and the same lengths ($A = 24\,\text{cm}$) and widths ($B = 8\,\text{cm}$): Fig. 10.

On Fig. 11, we see that for the same lengths ($A = 24\,\text{cm}$) and width ($B = 8\,\text{cm}$) resonance frequencies of the induced currents can be easily represented by the following approximate relation:

$$F_{R(N\,boucles)} = \frac{F_{R(1\,boucles)}}{N} \tag{27}$$

with: $N = 2, 3, 4, \ldots$.

6. CONCLUSION

This paper presents the design of antennas for passive RFID tags. The first part concerned the quick introduction of this technology. It was followed by modeling of a spiral RFID UHF antenna using the theory of the antennas.

Finally we have presented a method of estimating the peak current and resonance frequency of rectangular spiral RFID antennas.

This study showed that the amplitudes of the resonance peaks of rectangular spiral antennas printed vary linearly as a function of geometrical characteristics thereof. This linearity can be used by designers of printed antennas to assess the amplitude of current peaks at resonance with simple graphs that can be drawn as a function of geometrical characteristics of loops. As well as frequencies of resonance of the induced currents are virtually unchanged in function of numbers of loops for the same perimeter.

Currently we are trying to establish a relationship between the resonance frequency of such antennas and the nature of the dielectric substrate on which it is printed. This will surely improve the performance of printed antennas and also to contribute to their miniaturization.

REFERENCES

1. Finkenzeller, K., *RFID Handbook*, 2nd Edition, John Wiley & Sons, Ltd., 2003.

2. Bechevet, D., "Contribution au développement de tags RFID, en UHF et micro ondes sur matériaux plastiques," Ph.D. thesis from INPG, December 2005.
3. Tedjini, S., T. P. Vuong, V. Beroulle, and P. Marcel, "Radiofrequency identification system from antenna characterization to system validation," Invited paper, *Asia-Pacific Microwave Conference*, New Delhi, India, December 15–18, 2004.
4. Rifi, M., "Modélisation du couplage électromagnétique produit par des champs transitoires sur des structures filaires et des pistes de circuits imprimés connectées à des composants non-linaires," Ph.D. thesis, University Mohamed V, 1996.
5. Harrington, R. F., *Field Computation by Moment Methods*, Mac Millman, 1968.
6. Omid, M., "Development of models for predicting the effects of electromagnetic interference on transmission line systems," Ph.D., University of Electro-Communications, Tokyo, Japan, 1997.

Experimental Verification of Snell's Law at Sub-optical Frequencies

Jason Ramage[1], **Paul G. Huray**[1], **Kevin Slattery**[2], **Xiaopeng Dong**[2], **and Mike Resso**[3]

[1]Electrical Engineering, University of South Carolina, Columbia, SC 29208, USA
[2]Intel Corporation, 2111 NE 25th Ave., Hillsboro, OR 97124, USA
[3]Signal Integrity Applications, Agilent Technologies, Inc.
1400 Fountaingrove, Parkway, Santa Rosa, CA 95403, USA

Abstract— In 1621, Willebrord Snellius observed that light in the visible spectrum, incident at angle, θ_i, on a boundary between an isotropic material and Air was refracted by an angle, θ_r, according to the relationship, $n \sin\theta_i = \sin\theta_r$ (called Snell's Law) where n was a dimensionless constant ($n > 1$) he called the index of refraction. With the publication of Maxwell's equations in 1865, it was found that light was a special case of electromagnetic propagation at optical frequencies and that the more general relationship for the angle of refraction could be more generally expressed as $n \sin\theta_i = n_{Air} \sin\theta_r$; where n and n_{Air} were the indices of refraction of the homogeneous material and Air respectively and were related to their relative electric permittivity, ε_r, and permeability, μ_r, by $n = c/u_p = \sqrt{\varepsilon_r \mu_r}$ where $c = 1\big/\sqrt{\varepsilon_0 \mu_0}$ is the speed of light in a vacuum (or in Air) and u_p is the speed of light in the isotropic material. For non-magnetic materials ($\mu_r = 1$), Snell's Law can be written $\sqrt{\varepsilon_r(f)} \sin\theta_i = \sin\theta_r$ and is assumed to be valid for electromagnetic applications at all frequencies, f. But has it been experimentally shown that Snell's Law holds at non-optical frequencies? We have used a near field scanner with propagating electromagnetic fields operating at $f = 300\,\mathrm{MHz}$ and $f = 3\,\mathrm{GHz}$ incident on a non-magnetic liquid FC-40 ($n = \sqrt{\varepsilon_r(f)} \approx$ constant 1.89)/Air interface to verify that the critical angle, $\theta_i = \theta_c = 31.9^\circ$, predicted by Snell's Law for $\theta_r = 90^\circ$ applies at sub-optical frequencies.

1. INTRODUCTION

This paper discusses direct laboratory experiments that attempt to show that Snell's Law applies to sub-optical frequencies as Maxwell's Equations predict by recording the behavior of GHz electromagnetic waves traveling from one medium to another. Since Snell's Law is not frequency dependant (except insofar as the permittivity of materials is frequency dependent), the general law should apply to lower frequencies of electro-magnetic waves in the same manner as has been shown over the past two centuries for the visible light frequencies in the range of 10^{15} Hz. When an electro-magnetic wave travels between air and a different medium, the wave reflects and refracts at the interface boundary. Refraction occurs when the incident ray (direction of propagation) penetrates the boundary and is bent (refracted) at the boundary. Refraction happens if the index of refraction of medium 1 is greater than that of medium 2. When the angle of incidence equals the critical angle, the angle of refraction is 90 degrees from a normal to the interface; in which case propagation occurs along the boundary between the mediums. Total reflection occurs when the angle of incidence is greater than the critical angle.

The critical angle is found using the following equation: $\theta_{critical} = \theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right)$

$n_1 =$ index of refraction of FC-40 between 300 MHz and 3 GHz = 1.89

$n_2 =$ index of refraction of free air $= 1.00$

Therefore $\theta c = \sin^{-1}(1.00/1.89) = 31.9$ degrees between 300 MHz and 3 GHz. This critical angle is what has to be shown via experiment using 300 MHz to 3 GHz EM waves.

2. EXPLANATION OF EXPERIMENTS

Near field measurements in a laboratory environment at Intel's Jones Farm facility were used to confirm this hypothesis. The near field measurements show the field intensity of the waves as they travel further and further from the source to the new medium. A picture of the near field scanner used is shown in Figure 1.

A signal generator creates a 300 MHz to 3 GHz RF sine wave at the point source. The signal generator used in this experiment was an Agilent E8257D, shown in Figure 2.

A monopole point source was used at the RF source because it is more straightforward to analyze the data than from a dipole source; a monopole point source is, for practical purposes, a half dipole. The monopole point source is mounted on a circuit board as show in the Figure 3 below.

The circuit board was mounted on a level plane and the equipment used to measure electric or magnetic field intensity was shock mounted to the same support as the circuit board. Shock mounts allow for a receiver pickup probe to travel in the X, Y and Z directions without losing bearing on the starting position of the point source that radiates RF fields. The medium [1] surrounding the point source was FC-40, a dielectric liquid that has an index of refraction of approximately 1.89 at 1 GHz. When H_2O is used as the medium surrounding the signal source, the critical angle is too small for high resolution [2] of refracted propagation as is shown in Figure 4.

A clear plastic container was sealed onto the surface of the circuit board prior to commencing with any of the experiments. A circular container was chosen in order to support consistent data gathering from any angle or position. The height of the FC-40 was determined by trial and error. As shown throughout the trials, the depth of FC-40 determined the ability of the RF radiation to interact with the surface of the FC-40. At 2 mm and 1 mm, the scan showed an intense field

Figure 1: Near field scanner.

Figure 2: Agilent E8257D.

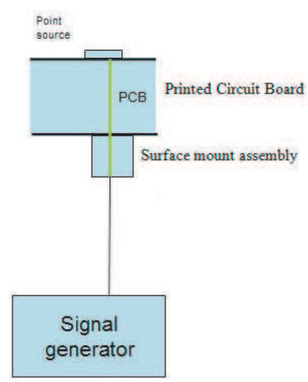
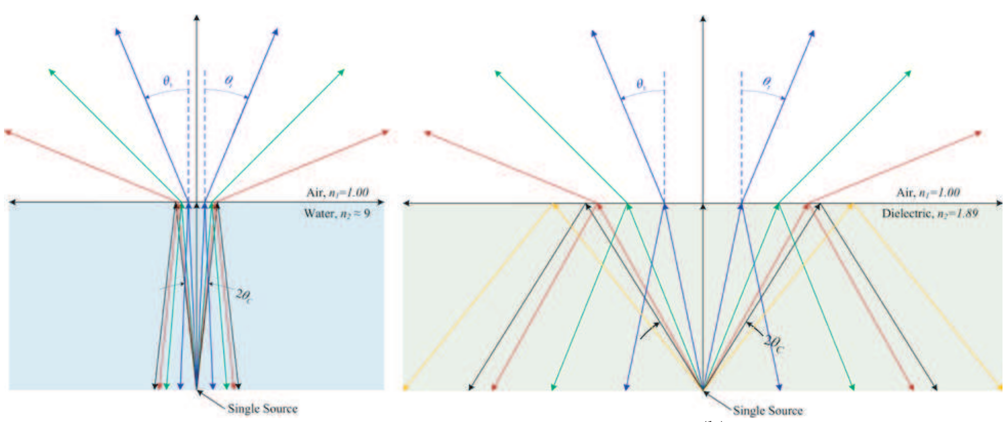


Figure 3: Circuit board setup diagram.



(a) (b)

Figure 4: Refraction of propagating EM waves at the interface between (a) water and air, and (b) FC-40 and air.

all the way from the point source to the surface; this result was neither ideal, nor easily analyzed since the diagram showed a maximum relative field strength all the way up to the surface. By comparison, 5 mm and 6 mm fluid depths showed the full range of the RF signal leaving the source. An electric or magnetic probe was attached to a near field scanner to measure field intensity as a function of distance from the source. Scanning consisted of using a probe to capture the RF signal by moving the probe back and forth relative to the point source while moving up in height on each pass. The pickup probe was connected to an oscilloscope for instant feedback during testing. The probe was also connected to motorized equipment that moved based on input parameters. The minimum size of the step that the probe took during its scans determined the resolution. The higher the resolution, the more detail that could be taken of the radiating signals. Nanometer resolution equipment was used in the research, design and manufacturing of circuit boards and microchips. Two probe options were available for these experiments: an E-probe for measuring electronic field intensity, and an H-probe for measuring magnetic field intensity. A diagram of the E-probe is shown in Figure 5 below.

16 trial sets (also referred to as scans) were measured, but only one data set will be presented here as shown in Figure 6. This scan reveals the signal propagating through the medium interface boundary with a color coded signal strength. Red signifies a strong signal strength and other colors represent a signal in various strengths from orange through yellow, green and blue. The four non-red colors represent high signal strength loss. As shown in Figure 6, the signal strength propagating in air above the FC-40 exhibits symmetrical RF signal penetration through the air/FC-40 interface relative to the point source consistent with Snell's law.

A wider scan above the liquid surface is shown in Figure 7.

The electromagnetic energy radiating from the point source *inside the FC-40* is shown in Figure 8. The E-field intensity lines also showed a symmetric pattern above the point source and there is some influence along the sides due to the cylindrical containment glass. This is due to the combination of the clean 3 GHz signal and 25 dBm of signal strength, as well as a balanced circuit board. The 2-D graph for the signals inside the liquid and in the air above the liquid are utilized in combination to form a complete picture of the signal as shown in Figure 9.

Figure 9 shows the combination of the scans inside and above the liquid. Note there is a change in the amplification of the signal strength between the air and liquid regions due to the sensitivity of the pickup probe in the two media. Figure 9 shows that a 3 GHz wave propagates through the air/FC-40 interface into the less dense air medium up to a critical angle of ± 31 degrees from the normal. This is approximately consistent with a Snell's Law calculation for the critical angle of

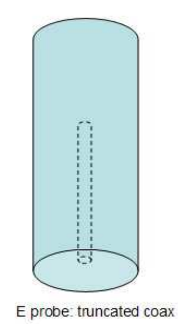


Figure 5: E-Probe diagram.

Scan number	Frequ	Strength (dBm)	Level of FC40 (mm)	Above or Below	Probe Type	Size of probe
14	3GHz	25	5	Above	E	.1mm

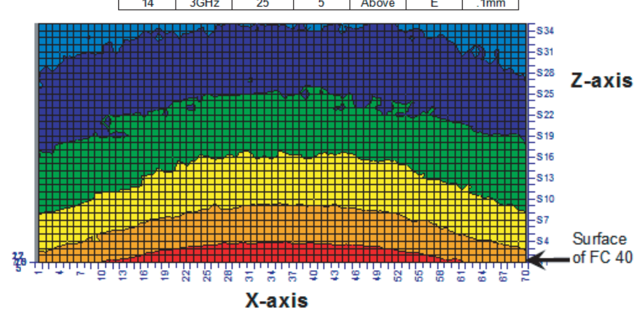


Figure 6: Electric field strength measured in air above the FC 40 surface interface.

Scan number	Frequ	Strength (dBm)	Level of FC40 (mm)	Above or Below	Probe Type	Size of probe
16	3GHz	25	5	Above	E	.1mm

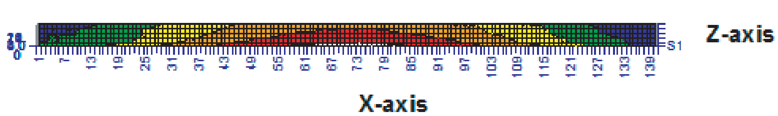


Figure 7: Wide scan of electric field strength measured in air above the FC 40 surface interface.

Scan number	Frequ	Strength (dBm)	Level of FC40 (mm)	Above or Below	Probe Type	Size of probe
15	3GHz	25	5	Below	E	.1mm

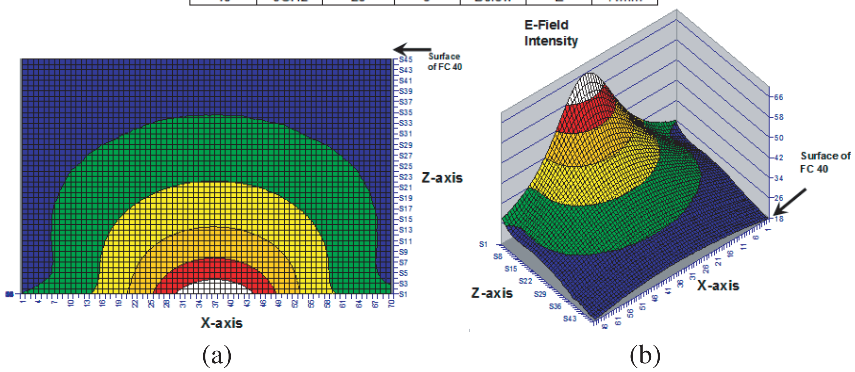


Figure 8: Electric field intensity strength: (a) side and (b) isometric view above the point source.

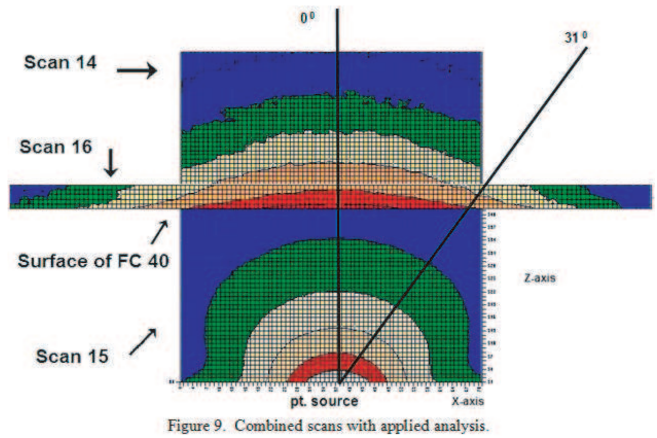


Figure 9: Combined scans with applied analysis.

31.9 degrees; i.e., Snell's Law predicts that electro-magnetic waves will refract up to the critical angle and beyond that angle it will exhibit only internal reflection into the FC-40 liquid medium from which it came. The absence of strong intensity field lines along the entire surface serves as verification of this law.

3. CONCLUSION

We conclude EM signals propagate from a point source in a FC-40 liquid through the air interface while experiencing refraction up to a critical angle. In this manner, Snell's Law has been experimentally verified for frequencies between 300 MHz to 3 GHz.

REFERENCES

1. Barnes, H., et al., "ATE interconnect performance to 43 Gbps using advanced PCB materials," *DesignCon 2008 Proceedings*, 2008.
2. Huray, P. G., *Maxwell's Equations*, Wiley-IEEE Press, 2009.

Backward Wave Modes of Partially Plasma Column Loaded Cylindrical Waveguide

E. Kelebekler and N. Yener
Technical Education Faculty, Kocaeli University, Turkey

Abstract— When two waves traveling in the same or opposite direction are interfered, they can attenuate or increase each other. The backward wave oscillators work with the same principal that interference between two waves traveling in opposite directions increases the amplitude along propagation direction. There are backward wave oscillators which use interaction of plasma-electron beam, in the literature. The Method of Moment (MoM) is a widely used technique for numerical simulation of propagation and scattering problems. In this study, backward wave modes of the plasma column loaded cylindrical waveguide have been investigated by using Method of Moment (MoM). In previous studies, the authors presented the validation of the method for gyro-resonance region. For the structure, the backward waves appear in the plasma resonance region. Unlike gyro-resonance region modes (called plasma modes), to obtain the plasma resonance region modes (called cyclotron modes), it is necessary to use very large dimensions for the linear algebraic equations system used in the MoM. Indeed in the implementation of the MoM in the plasma resonance region for the cyclotron modes, we have found that the method exhibits discrepancies with the exact solution which do not disappear unless very large numbers of expansion functions are used. The objective of this work is to report these discrepancies and suggest a possible numerical technique to overcome them. Also presented are the dispersion curves of the backward wave modes. This technique consists in using additional computer time and memory space. In particular it is demonstrated that the mean relative error is confined to less than 0.001 over the whole frequency band of interest which would be impossible if insufficient computing power was used.

1. INTRODUCTION

In this study, backward wave modes of the plasma column loaded cylindrical waveguide have been investigated by using Method of Moment (MoM). A wave is considered a backward (forward) wave if its group velocity, as indicated by slope of dispersion curve, is opposite (the same) in direction to the phase velocity. The phase velocity (υ_{ph}) and group velocity (υ_{gr}) are described in (1).

$$\upsilon_{ph} = \frac{\omega}{k}, \quad \upsilon_{gr} = \frac{d\omega}{dk} \tag{1}$$

where, ω is the operating frequency and jk is the wave propagation constant.

In this study, this definition, which is based on the dispersion properties of the wave propagation constant with respect to frequency, has been used for backward wave because the method used in the study directly gives the relation between the propagation constant and frequency. The variation of the electromagnetic fields is given below.

$$F(r, \phi, z) = F(r)\, e^{j(kz+n\phi-\omega t)} \tag{2}$$

where, n is the azimuthal variation number and r, ϕ, z are the cylindrical coordinates. The tensor permittivity of the plasma column is equal to below matrix.

$$\hat{\varepsilon} = \varepsilon_0 \begin{bmatrix} \varepsilon_1 & j\varepsilon_2 & 0 \\ -j\varepsilon_2 & \varepsilon_1 & 0 \\ 0 & 0 & \varepsilon_3 \end{bmatrix} \tag{3}$$

where, ε_0 is the permittivity of free space and the expressions of ε_1, ε_2, ε_3 in Equation (3) are given in (4). Their values change depending on the plasma frequency, the cyclotron frequency (so the material from which the plasma is produced) and the operating frequency.

$$\varepsilon_1 = 1 + \frac{1}{R^2 - \Omega^2}, \quad \varepsilon_2 = \frac{-R}{\Omega(R^2 - \Omega^2)}, \quad \varepsilon_3 = 1 - \frac{1}{\Omega^2} \tag{4}$$

where, R is the normalized cyclotron frequency (ω_c/ω_p) and Ω is the normalized operating frequency (ω/ω_p). Besides, ω_c is the cyclotron frequency and ω_p is the plasma frequency. The backward waves in plasma column loaded cylindrical waveguide are called the cyclotron modes and appear in the plasma resonance region [1] where the normalized operating frequency is between $\max(1, R)$ and Ω_u. Here, Ω_u is normalized upper hybrid frequency and described as below.

$$\Omega_u = \sqrt{1 + R^2} \tag{5}$$

The exact solution for the plasma column loaded cylindrical waveguide has been presented in different forms in the previous studies [1–5]. The exact solution of the structure has been used to check the accuracy of the solutions obtained from the MoM. The MoM is based on converting Maxwell's partial differential equations into linear algebraic equations that are written in matrix form. In the presented study, "Generalized Telegraphist's Equations" [6] or the transmission line equations have been utilized to generate linear algebraic equations system of the plasma column cylindrical waveguide and the validation of the method has been presented for the backward waves of the structure. Actually, the technique is a Fourier-Bessel series expansion technique and the number of expansion functions determines the dimension of the linear algebraic equations system.

In the previous studies, the authors presented the validation of the method for gyro-resonance region [7, 8]. Besides, they compared the method with two semi analytical methods, the quasi-static approximation and the asymptotic approximation, and they showed that the MoM is a better method than the other two methods for all frequencies of the gyro-resonance region [9]. Unlike gyro-resonance region modes (called plasma modes), to obtain the cyclotron modes, it is necessary to use very large dimensions for the linear algebraic equations system. Indeed in the implementation of the MoM in the plasma resonance region for the cyclotron modes, it has been observed that the method exhibits discrepancies with the exact solution which do not disappear unless very large numbers of expansion functions are used. In particular it is demonstrated that the mean relative error which equals to average of the relative error for whole frequency interval is confined to less than 0.001 over the whole frequency band of interest which would be impossible if insufficient computing power was used. Hence a numerical approach to solve this problem has been observed to be increasing the number of expansion functions.

In the study, the second section explains the MoM. In the third section, obtained results have been given. The last section constitutes the conclusion.

2. THE METHOD OF MOMENT

Maxwell's equations do not have an exact solution in the closed form for every physical structure. When the exact solution does not exist for any physical structure, a numerical solution, as the finite difference method or the finite element method, or a semi analytical method, as the transmission line method or the MoM, is investigated in order to obtain the solution. One of the best known semi-analytical methods for closed and sourceless waveguides was given by Schelkunoff in 1952 [6]. In his classical study, Schelkunoff derived the transverse field component from the potential and the stream functions for general structure of closed waveguides. The method transforms Maxwell's equations, consisting of partial differential equations, into an ordinary differential equations system containing differentiation with respect to propagation direction (z). The obtained system is also called the transmission line model of the structure. If Equation (2) is considered, the differentiation with respect to propagation direction is equivalent to multiplication with jk. Thus, the ordinary differential equations system is transformed into a linear algebraic equations system. The linear algebraic equations system for fully/partially gyro-electric or gyro-magnetic medium loaded waveguide is in the form of Equation (6).

$$jk(p) \begin{bmatrix} v(p) \\ i(p) \end{bmatrix} = \begin{bmatrix} 0 & Z(p) \\ Y(p) & 0 \end{bmatrix} \begin{bmatrix} v(p) \\ i(p) \end{bmatrix} \tag{6}$$

In this way, the problem of electromagnetic propagation in the gyro-electric medium loaded cylindrical waveguide is converted into an eigenvalue problem. In expression (6), p is the complex frequency, $jk(p)$ shows complex propagation constant, $v(p)$ and $i(p)$ are the unknown voltage and the current vectors, respectively. Besides, $Z(p)$ and $Y(p)$ are the complex impedance and admittance coefficient matrices per unit length, respectively.

The eigenvalues of the impedance-admittance coefficient matrix in (6) gives the propagation constants of the problem. The dimension of the system is determined from the number of known solutions of the empty waveguide used in the Fourier-Bessel expansion. The method is called as a semi-analytical method because of necessity of truncating an infinite summation of series at a point, while it uses known analytic solutions of the empty waveguide. This method is also called Galerkin version of the MoM [10].

3. THE BACKWARD WAVE MODES

The modes existing in the plasma resonance region are the backward wave modes. These modes have been presented as cyclotron modes [1], dynamic modes [2] and HE modes [3]. In the study, the plasma-resonance region's modes are called cyclotron modes as in [1] and symbolized with ${}^{R}_{\delta}C_{n,m}$. Here, δ is the normalized waveguide radius ($\omega_p a/c$) and m is the mode number. Also, a and c are the radius of the waveguide and the velocity of light in vacuum, respectively.

In the previous studies, the authors presented the plasma modes existing in gyro-resonance region by using the MoM. Unlike plasma modes, to obtain the cyclotron modes, it is necessary to use very large dimensions for the linear algebraic equations system used in the MoM. As a consequence of numerical computations for cyclotron modes, it is observed that the method exhibits discrepancies with the exact solution which do not disappear unless very large numbers of expansion functions (N) are used. These discrepancies occur at different frequency points for different number of expansion functions and disappear while the term number of expansion functions enlarges as shown Figure 1. At frequency points other than the discrepancy points, shown within ellipses in

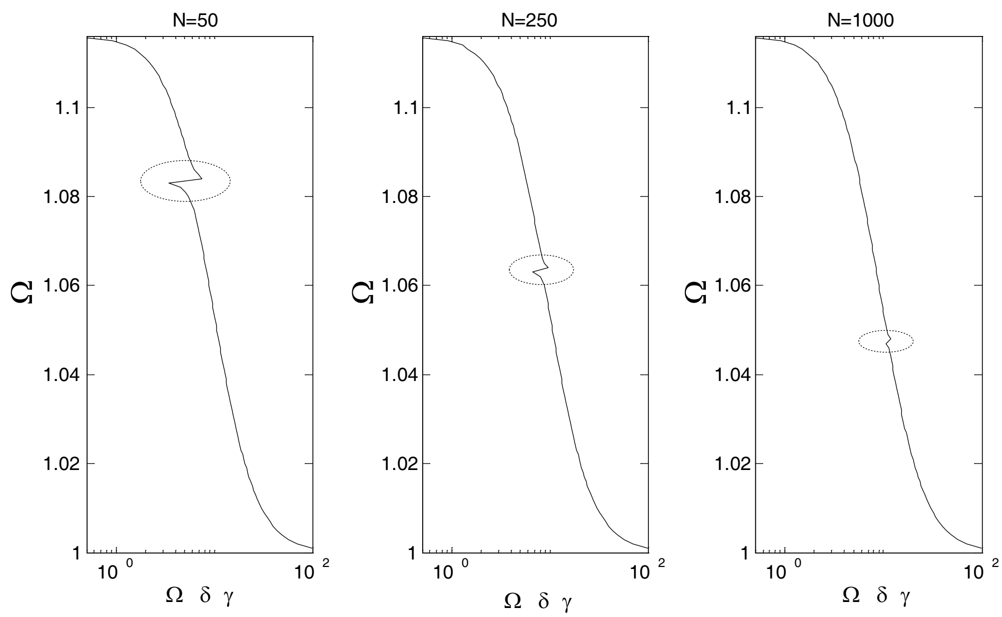


Figure 1: The dispersion curves obtained from the MoM for different number of expansion functions.

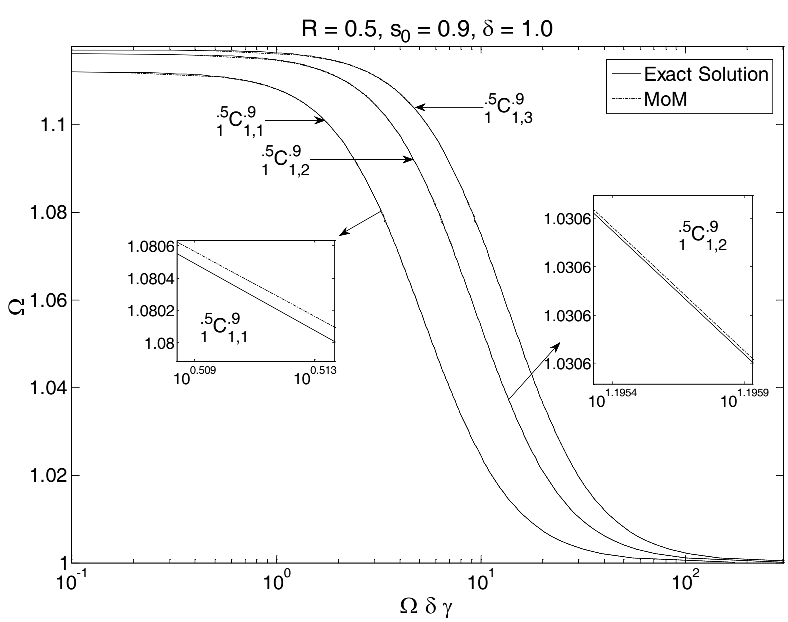


Figure 2: The dispersion curves obtained from the exact solution and the MoM for $R = 0.5$ and $s_0 = 0.9$.

the figure, the values obtained from the method conform with the exact solutions. In the figures, γ stands for k/k_0 where k_0 is the wave number in free space.

The dimension of the coefficients matrix is taken as $2N \times 2N$. In order to overcome the discrepancies, it is necessary to enlarge number of expansion functions (or augment dimension of the coefficient matrix), which means using additional computer time and memory space.

In numeral computations, first the number of expansion functions, which confines the mean relative error to less than 0.001 over the whole frequency band, has been determined for each structure. The numerical computations have been performed by using MATLAB R2009b on a computer which has Intel® Core™ i7 CPU 960 @ 3.20 GHz 3.19 GHz, 12 GB RAM and 64 bit operating system. As a consequence of computations, it is found that it is necessary to use bigger number of expansion functions in order to eliminate the discrepancies while the plasma column radius decreases in waveguide. The plasma to waveguide ratio is described as $s_0 = b/a$. Here b denotes the radius of the plasma column.

In this study, the first three degrees of the cyclotron modes ($m = 1, 2, 3$) have been investigated

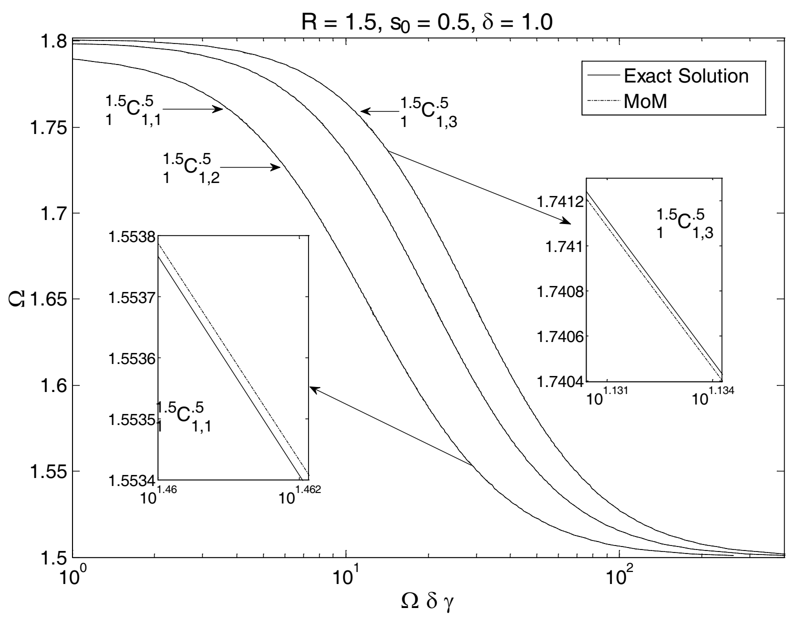


Figure 3: The dispersion curves obtained from the exact solution and the MoM for $R = 1.5$ and $s_0 = 0.5$.

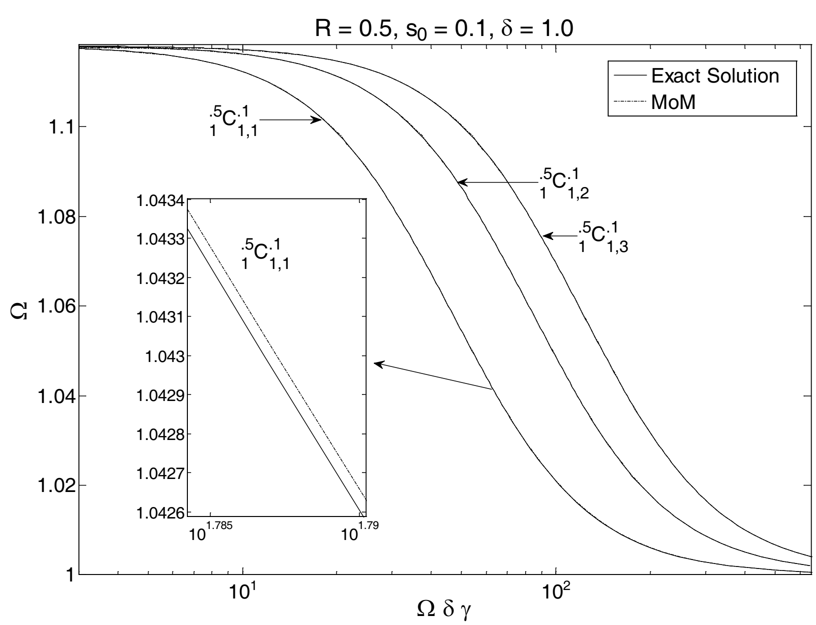


Figure 4: The dispersion curves obtained from the exact solution and the MoM for $R = 0.5$ and $s_0 = 0.1$.

using the MoM and compared with the exact solution given in [1]. For numerical computation, the waveguide radius, the normalized waveguide radius, the azimuthal variation and the plasma frequency have been taken fixed as $a = 3\,\mathrm{cm}$, $\delta = 1, n = 1$ and $\omega_p = 10^{10}\,\mathrm{rad/s}$, respectively. The numerical computations have been performed for two groups of the normalized cyclotron frequency. The first is for relatively weak magnetic field ($R = 0.5$) and the second is for the relatively strong magnetic field ($R = 1.5$) as described in [1]. The investigated structures have three different ratios of radii ($s_0 = 0.1, 0.5, 0.9$). As the result, the numbers of expansion functions which confine mean relative error to less than 0.001 have been determined as $N = 3000$ for $s_0 = 0.9$, $N = 4000$ for $s_0 = 0.5$ and $N = 5000$ for $s_0 = 0.1$ for both groups of the normalized cyclotron frequencies. Dispersion curves are given for all three ratios of radii while $R = 0.5$ or $R = 1.5$. The $R = 0.5$ and $s_0 = 0.9$ case, $R = 1.5$ and $s_0 = 0.5$ case, $R = 0.5$ and $s_0 = 0.1$ case are depicted respectively in Figures 2, 3, 4. The results obtained from the MoM conform with the exact solutions pretty well. In order to show differences of the results, the dispersion curves have been plotted by zooming at some frequency regions in the figures.

4. CONCLUSION

It is to be noted that for smaller dimensions of the coefficients matrix in the MoM the spike like formations observed in Figure 1 do not disappear. Hence one needs as long CPU times as several hours for each run for each frequency in the band $1 < \Omega < 1.12$ for $R = 0.5$ and the band $1.5 < \Omega < 1.8$ for $R = 1.5$. This problem has been encountered only for the frequency intervals above the gyro-resonance frequency region. Increasing the dimension of the coefficient matrix has eliminated these discrepancies in the structures reported above and the necessary minimum dimensions have been listed for each structure. Lastly dispersion curves of the backward wave modes have been presented comparatively with the exact solutions.

ACKNOWLEDGMENT

This work was supported by Kocaeli University Research Project No. 2010-034.

REFERENCES

1. Yip, G. L. and S. Le-Ngoc, "Dispersion characteristics of the dipolar modes in a waveguide partially filled with a magnetoplasma column," *Canadian Journal of Physics*, Vol. 53, 1163–1178, 1975.
2. Granatstein, V. L. and S. P. Schlesinger, "Electromagnetic waves in parameter space for finite magnetoplasmas," *Journal of Applied Physics*, Vol. 35, No. 10, 2866–2855, 1964.
3. Ivanov, S. T., E. G. Alexov, and P. N. Malinov, "Symmetrical electromagnetic waves in partially-filled plasma waveguide," *Plasma Physics and Controlled Fusion*, Vol. 31, No. 5, 941–953, 1989.
4. Bevc, V. and T. E. Everhart, "Fast-wave propagation in plasma-filled waveguides," *Journ. of Electron. Control*, Vol. 13, 185–212, 1962.
5. Bevc, V., "A new multipolar mode in bounded gyrotropic plasma," *J. App. Phy.*, Vol. 39, No. 3, 1492–1502, 1968.
6. Schelkunoff, S. A., "Generalized telegraphist's equation for waveguide," *Bell Syst. Tech. J.*, 785–801, 1952.
7. Kelebekler, E., "Investigation of the Gyro-resonance region modes by using the MoM for plasma column loaded cylindrical waveguide," *IEEE EDAPS Symposium*, 1–4, Hong Kong, December 2009.
8. Kelebekler, E. and N. Yener, "Farklı Yapıdaki Silindirik Dalga Kılavuzlarının Dispersiyon Egrilerinin Iletim Hattı Esdegerlikleri Yöntemi ile Elde Edilmesi," *13. Ulusal Kongresi*, 217–220, Ankara, Turkey, 2009.
9. Kelebekler, E., "Comparison of some semi analytic methods for plasma column loaded cylindrical waveguide," *IEEE Applied Electromagnetics Conference*, 1–4, Calcuta, India, December 2009.
10. Yener, N., "Necessary and sufficient conditions for the existence of backward waves in metallic waveguides," *Journal of Electromagnetic Waves and Applications*, Vol. 17, No. 12, 1713–1722, 2003.

Scattering Analysis of a Submerged Conducting Object in Lossy Media via Low Frequency EM

Amna Ajaz, Jia-Dong Xu, and Wei Yan
School of Electronics and Information, Northwestern Polytechnical University (NPU), Xi'an 710072, China

Abstract— This paper is an effort to study further the EM scattering from a conducting object submerged in a lossy media as has been investigated previously by King and later Marius. The applied simulation tool is Feko based on moment method. Models for the scattering in half-space and in layered media are defined in Feko. Also, analytic expression for scattering in half-space is computed with Matlab. Simulated scattered fields are in good agreement with the previous published results. Considering lateral wave propagation phenomenon in the presence of bounded media is observed to be useful in obtaining enhanced scattered field response from the submerged object.

1. INTRODUCTION

The emerging concerns to sensibly detect the presence of objects in hide outs of sea waters and to follow the activities of identified objects (civilian & military), asks for enhanced surveillance over the related offshore environments. The properties of the marine environment tend to privilege the use of acoustic waves in various underwater applications [1]. However, due inherent geometrical and back-ground noise limitations for acoustics in shallow sea, pushes the technological shift to non-acoustic means as well. One of the recent trends in monitoring the strategic littoral sites closely has started with the investment in the EM based wireless sensor network (WSN) technologies. Additionally, there is a need for analyzing the scattering scenarios from silent objects. As indicated in [2] such objects, in present times, are made quiet to an extent that their presence against the background could not be detected by only using conventional means of acoustics in strategic offshore zones. Also, due advancement in technology to control the fall out of the related EM fields from the objects, will further make them, unnoticed unless radiated by active emitters.

In present publication by using the EM simulation software, we will study the scattering scenario from a perfectly conducting cylindrical body submerged in a lossy environment. Many objects deployed for underwater operations employ the cylindrical shaped bodies, further justifies exploration and validation of its scattering response. Previously, Marius [3] has computed the example of remotely sensing a 3D PEC object, submerged fully in a lossy layered medium. He validated his simulation findings against the analytical results produced earlier by King in [4] for half-space case. Considering the 3D geometry of [3], the comparative scattering analysis was carried out by using the MoM based EM simulation software package Feko. It is learned that the propagation of lateral wave could not be ignored when considering bounded media with differing dielectric properties. Therefore, how this wave contributes to the scattering response from submerged body in sea medium is studied for its understanding in relation to finding silent objects.

The paper is constructed as follows: in Section 2, the geometrical and simulation models alongwith a brief outlook on the EM lateral waves is given. In Section 3, the comparative analysis of the simulation results is discussed followed by conclusions in Section 4.

2. DESCRIPTION OF THE MODELS

Previous outcome [(Fig. 1), 3] validated with analytical results of [4] and [(Fig. 2), 3] provides for comparison of the simulation results obtained from MoM based code Feko. Also, related analytical results in [4] are computed in Matlab and compared with the Feko result for half-space case. Here we attempted to study the simulation models with similar parameters as tested in [3] and based on integral equation method of moment technique. The geometry and numerical modeling of the problem under study are described as below:

2.1. Geometrical Illustrations

The geometry of the models as described in [3] for the cases of half-space and three-layered media are shown in Figs. 1 & 2 respectively. The observing point is placed far from the source in order to neglect the contributions of the direct and reflected fields; the electromagnetic phenomena observed

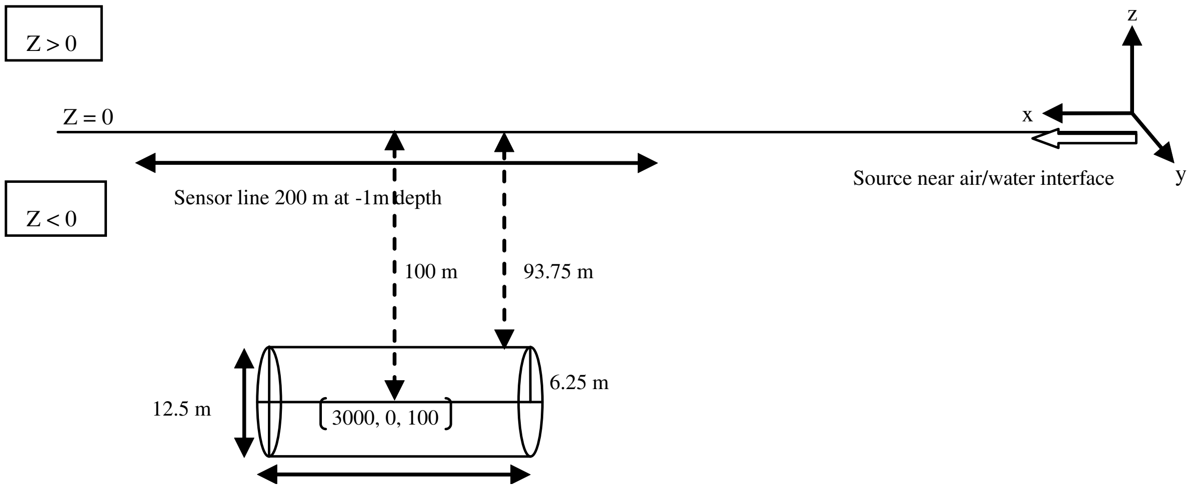


Figure 1: Geometrical configuration of the half-space model.

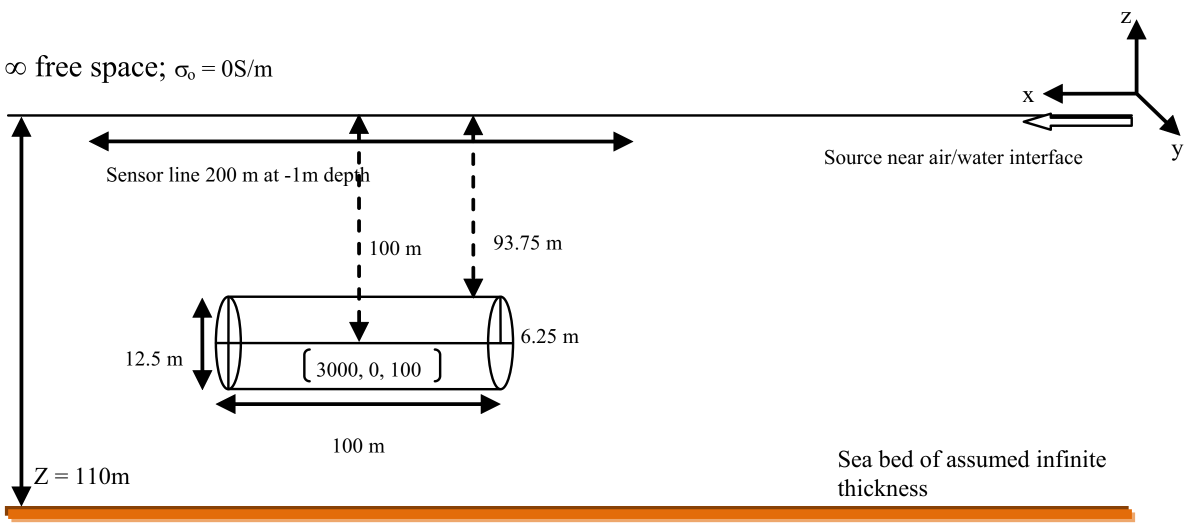


Figure 2: Geometrical configuration of the three-layered model.

in this portion of space are composed of lateral waves propagating on the sea surface [1] for the case of half-space and for the combined case of three layered environment.

In the conducting sea water, the corresponding wavelength of the selected excitation frequency of 25 Hz is found from,

$$\lambda_{seawater} = \frac{2\pi}{\sqrt{(\pi\mu_0\sigma f)}} = 316\,\mathrm{m} \tag{1}$$

where $\sigma = 4\,\mathrm{S/m}$; $\mu_0 = 4\pi \times 10^{-7}\,\mathrm{H/m}$ (same as that of free space).

The skin depth $\delta_{seawater}$ at a given frequency of 25 Hz, is

$$\delta_{seawater} = \sqrt{\frac{2}{\omega\mu\sigma}} = 50\,\mathrm{m} \tag{2}$$

with this skin depth the propagation of direct wave will be completely attenuated [3] and the signal approaching the sensor line will consist significantly of lateral wave travelled along the boundary of air-water media. It was learnt that all components of the lateral-wave field propagate along the surface of the sea without exponential attenuation until directly over the object and then travel vertically downward [5].

A three layered geometry in Fig. 2 is suggested to enhance the scattering response of the object by including lateral wave propagation along bounded media surrounding the submerged body.

2.1.1. Electromagnetic Lateral Waves

In literature, EM lateral waves have been discussed from time to time for over a century. To our understanding the defining characteristic of lateral waves is its appearance in dielectrically different bounded media. Presence of lateral waves is characterized by conical wavefronts below the surface [6]. The total field on or near the interface of air-dielectric is mainly expressed by lateral wave, where the amplitude of the field along the boundary is $1/\rho^2$ [7]. This wave is much less susceptible to the lossy nature of the earth, however its decay in lossless media is proportional to $r^{-3/2}$ rather than r^{-1} for the traditional space wave [8]. Also, from [6], we found that the excitation of lateral waves is related to the incident angle and magnitude of the incident fields on the surface. Lately Kai Li, has published comprehensive review on propagation of electromagnetic waves in stratified media [7]. His analytical derivations extended the expressions for lateral waves from already known two and three- layered environments to four-layered regions. Further, as the tangible systems based on lateral wave phenomenon are not developed therefore there is lack of physical interpretation about the underlying wave mechanism [6]. Recently, from applications stand point two references appeared in open literature incorporating use of lateral waves viz., in the domains of ground-based subsurface detection and shallow water wireless communication respectively [6, 9]. A concise description of lateral wave propagation can be found in [9], in addition, its utility was thought to be included while designing for wireless network prototype system for shallow water deployment. With the added effect of lateral wave propagation, a low data rate prototype system was proposed aiming at monitoring biological applications in shallow waters [9].

2.2. Simulation Models in Feko

For the simulation modeling of the problem, we utilized CADFEKO component [10] of the Feko software. Two different models as per the respective geometries (Figs. 1 &2) of the problems are defined alongwith the meshing. Also electromagnetic parameters and solution configurations are specified on the CADFEKO model. The core of Feko suite is implemented with MoM, however other frequency domain CEM methods and hybrid methods have also been incorporated for efficiently addressing variety of application areas [10]. Due presence of the single PEC scatterer we get the solutions using MoM option only [11]. The surface of the scatterer in our Feko models is meshed with 1524 patches. To include the planar boundaries, use is made of infinite planar layers by specifying following parameters:

i. Case-1 Half space numerical model: In rectangular coordinates, infinite homogeneous free space is selected for $z > 0$ while $z < 0$, an infinite homogeneous dielectric sea layer is specified. Sea water conductivity and relative permittivity, are selected as with $\sigma = 4\,\mathrm{S/m}$, $\varepsilon_r = 81$ respectively.

ii. Case-2 Three-layered numerical model: Infinite homogeneous free space for $z > 0$ while dielectric sea water with $\sigma = 4\,\mathrm{S/m}$, $\varepsilon_r = 81$ and finite depth of 110 m bounded by infinitely thick lower sea bed with $\sigma = 0.01\,\mathrm{S/m}$ as shown in Fig. 2.

The multilayered option in Feko is governed by special green's function. The simplest form of a Green's function is the free-space Green's function, which is used in the default MoM implementation. It is possible to use special Green's functions to incorporate features of the propagation space into the model. This means that properties of the structure are modeled implicitly [12]. The excitation frequency is set to 25 Hz as was used in previous works [3, 4]. The infinitesimal dipole in a homogeneous medium is set with electric moment of 1 A·m. For taking into account the same geometrical area as in [3], the set extent in the Feko models is selected as 5E+03.

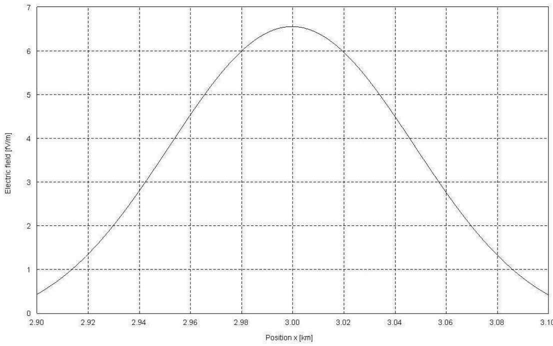
3. SIMULATION RESULTS

Simulation results obtained from Feko are compared with the already published computed and analytic outcomes in the literature (Refs. [3] & [4]). The simulated results from two model configurations are discussed in the following text:

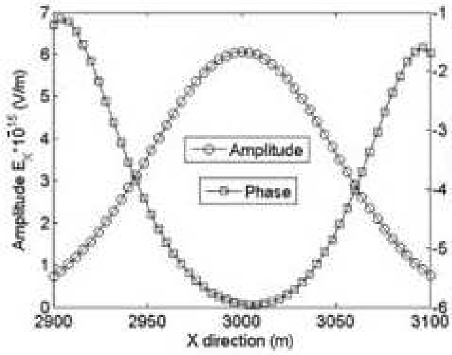
3.1. Case-1 Half Space Computations

For the two-layered case of Fig. 1, the corresponding graph obtained from Feko is illustrated in Fig. 3(a). It is seen to be in good fit with the secondary field profile as shown in Fig. 3(b). Also, the scattered field from Feko model is observed to follow the same range of amplitude $\times 10^{-15}$ as in Fig. 3(b).

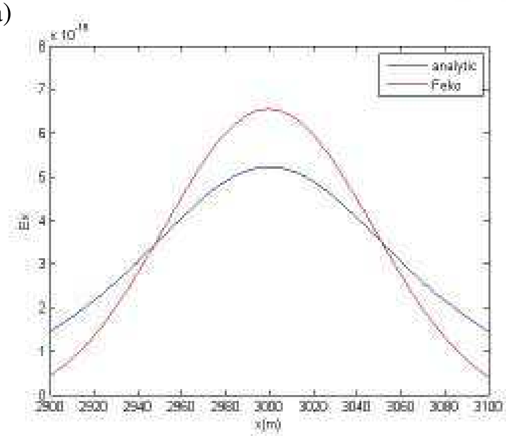
Additionally, the result from Feko simulation is compared with the graph obtained from analytic expression computations in Matlab as shown in Fig. 3(c). It is found that all three graphs falls nearly into the same range of maximum and minimum when the scatterer is a perfectly conducting



(a)

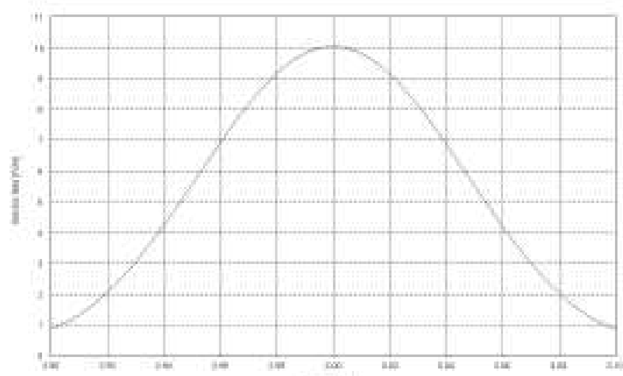


(b)

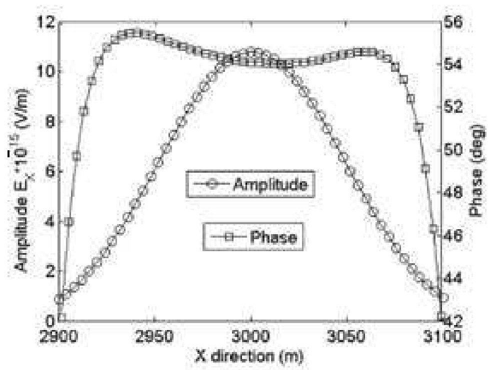


(c)

Figure 3: (a) Scattered electric field in Feko from PEC cylinder in half space. (b) Secondary electric field from the conducting cylinder in half space [3]. (c) Comparison of Feko and analytic scattered electric field.



(a)



(b)

Figure 4: (a) Scattered electric field in Feko from PEC cylinder in layered media. (b) Secondary electric field from the conducting cylinder in layered structure [3].

body. However, since Feko simulations also included full 3D geometry of the scatterer therefore the outcome is closer to 6.05×10^{-15} V/m of Fig. 3(b) above. (See Appendix A for Matlab editor command).

3.2. Case-2 Layered Medium Computations

For the three layered case of Fig. 2, which was suggested by Marius (2007) for considering the effect of lateral wave propagation from the water/sea-bed boundary as well. This effect can also be seen on the scattering amplitude of Feko graph as shown in Fig. 4(a). The amplitude has increased to 10×10^{-15} V/m above the centre of the 3D cylinder compared to half-space case of Feko simulations. Again there is a good fit between the secondary field profile (Fig. 4(b)) and Feko graph of Fig. 4(a).

4. CONCLUSION

This document serves the purpose to apply commercial EM analysis tool Feko for studying the scattering problem of PEC object surrounded by lossy medium of sea-water. On the basis of simulation results, it is learned to consider effect of lateral wave propagation for unconcealing the conducting body when submerged in layered structure. From literature review it was revealed that phenomenon of lateral wave propagation has yet to be matured with its incorporation in commercial embodiments. In future the effect of lateral wave for sea environment and in conjunction with the scattering objects will be investigated in more detail.

APPENDIX A: COMPUTATION IN MATLAB FOR ANALYTIC EXPRESSION

In addition to simulation model for half-space case in Feko, related analytic result of [4] is implemented with Matlab for comparison. Following commands are run from Matlab editor, for graphical presentation of King's expression for scattered field from the object:

$$x = -100 : 1 : 100;$$
$$z = 99;$$
$$d = \mathrm{sqrt}(x. \wedge 2 + z. \wedge 2);$$
$$\mathrm{omega} = 2 * \mathrm{pi} * 25;$$
$$\mathrm{miu} = \mathrm{pi} * 4e - 7;$$
$$E = i * \mathrm{omega} * \mathrm{miu} * 33.3 * 2.73 * 1e{-}9 * (\exp(i * 2.58)) * (1 + (i * \exp(-i * \mathrm{pi}/4)./(0.028 * d)) + (i./((0.028 * d). \wedge 2))).*\exp(-0.02 * d).*\exp(i * 0.02 * d)./(2 * \mathrm{pi} * d); x1 = 2900 : 1 : 3100;$$
$$\mathrm{plot}(x1, abs(E));$$

REFERENCES

1. Bruxelle, J. Y., "Electromagnetic propagation in conducting shallow water," *OCEANS '94, 'Oceans Engineering for Today's Technology and Tomorrow's Preservation', Proceedings*, III/522–III/527, Vol. 3, 1994, Digital Object Identifier: 10.1109/OCEANS.1994.364253.
2. Birsan, M., "Modelling the electromagnetic scattering by 3-D objects in a layered media, electromagnetic active detection of metallic targets," December 2004, url: http://handle.dtic.mil/100.2/ADA436222.
3. Birsan, M., "Remote sensing of 3-D conducting objects in a layered medium using electromagnetic surface waves," *IEEE Geoscience and Remote Sensing Letters*, Vol. 4, No. 4, 561–565, October 2007.
4. King, R. W. P., "Lateral electromagnetic waves from a horizontal antenna for remote sensing in the ocean," *IEEE Trans. Antennas Propag.*, Vol. 37, No. 10, 1250–1255, October 1989.
5. King, R. and M. F. Brown, "Lateral electromagnetic waves along plane boundaries: A summarizing approach," *Proceedings of IEEE*, Vol. 72, No. 5, 595–611, 1984.
6. Chen, C.-C., "Lateral wave radar system for forward detection," patent publication number WO 2010/042893 A2, Aprial 15, 2010.
7. Li, K., *Electromagnetic Fields in Stratified Media*, Zhejiang University Press, Hangzhou and Springer-Verlag GmbH Berlin Heidelberg, 2009.
8. Ortega, R., C.-C. Chen, and R. Lee, "FDTD modeling of scattering from distant buried objects," *IEEE Antennas and Propagation Society International Symposium*, 3309–3312, 2007.
9. Cella, U. M., R. Johnstone, and N. Shuley, "Electromagnetic wave wireless communication in shallow water coastal environment: Theoretical analysis and experimental results," *WUWNet'09*, 2009.
10. FEKO Suite 5.5, User's Manual.
11. Davidson, D. B., *Computational Electromagnetics for RF and Microwave Engineering*, Cambridge University Press, 2005.
12. Clarke, S. and U. Jakohus, "Dielectric material modeling in the mom-based code FEKO," *IEEE Antennas and Propagation Magazine*, Vol. 47, No. 5, 140–147, October 2005.

The Feature Selective Validation (FSV) as a Means of Formal Validation of Electromagnetic Data

A. P. Duffy and H. G. Sasse
De Montfort University, The Gateway, Leicester, UK

Abstract— The extensive use of computational electromagnetics is becoming a fundamental aspect of the design of many electronic and electrical systems, with the result that more reliance is being placed on those simulations, and the cost of getting it wrong in both time and money is becoming more significant. Clearly, a process for validating the codes and implementations is vital to expanding the use of computational electromagnetics and, at the heart of that validation, must be a robust method to quantify the quality of any comparisons used as part of the validation process. The Feature Selective Validation (FSV) method is the method of choice for providing that quantification in the first standard to set out the process of validation (IEEE Standard 1597.1). This paper provides an outline of the method, setting it in the context of the competing approaches to providing that data, and reviews some of the FSV developments in recent years and concludes with some of the key challenges to extending the reach of FSV.

1. INTRODUCTION

Custom-and-practice is not always the best way to achieve any goal. However, the evolution of the 'computational' branch of electromagnetics relied on the modellers looking at the results their solvers or model implementations had produced, comparing these against measurements or expected results and asking "is this good enough?" Clearly, a means of providing a quantified version of this comparison is desirable. This paper will review the state-of-the art in quantitative validation for computational electromagnetics as a timely and much needed assessment of current practice. The customary base technology for undertaking those comparisons has been visual inspection informed by experience and expectation. While this is acceptable and does allow the experience of the user to dominate, it does not take into account subjective differences between users. Statistical approaches to quantifying comparison using non-parametric tests such as the Kolmogarov-Smirnov test have found application in some areas, but statistical approaches generally fail for various reasons, such as failing to account for a likely group response of a number of users and failing to show whether that "good enough" is governed by broad agreement across much of the data or detailed agreement over all the data. This led to the development of the Feature Selective Validation (FSV) method in the late 1990s which has since been augmented and recently adopted into the IEEE standard 1597.1 [1]. This paper describes the origin of FSV based on studies in correlation and reliability functions as well as the set of guiding principles in its design, the introduction of categorisation and the use of the Visual Rating Scale to verify FSV and provide a set of natural language descriptors. It then reviews a range of FSV developments such as the use of FSV with complex valued signals, the introduction of the offset difference measure, studies into the impact of data density and research to account for the allowances normally made by users when visually assessing the comparisons. It concludes with a discussion on a number of the key challenges facing further FSV developments.

2. OUTLINE OF THE PROBLEM

While data of visually simple structure is straightforward to analyze using simple distance measures, it is common for data arising from computational electromagnetics, particularly when coming out of electromagnetic compatibility (EMC) studies, to be visually quite complex involving many changes of level and numerous resonant structures. Figure 1 presents typical data.

Historically, data such as that in Figure 1 would have been visually compared by the engineers involved in the comparison and a decision made as to whether such data is acceptable or whether the differences are such that the comparison is unacceptable. A number of individuals would view the data differently and a range of opinions would be obtained and from that a consensus would be formed. Any approach to quantifying such comparisons, with a view to formal validation, must account for this or the comparison method would be unlikely to find any adopters. A set of rules for quantification of validation comparisons was proposed at the outset of the research that gave rise to FSV [2]. These are:

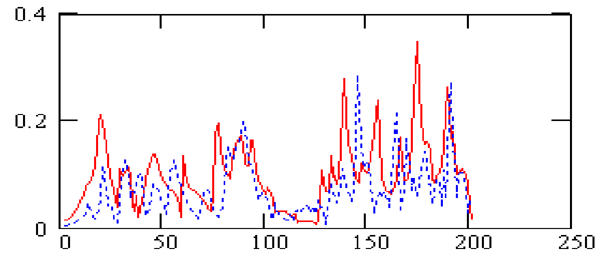


Figure 1: Typical data to be considered for comparison.

1. *Implementation of the validation technique should be simple.* i.e., no pre-requisites for special programming or mathematical knowledge.
2. *The technique should be computationally straightforward.* i.e., the computing power required should be relatively low.
3. *The technique should mirror human perceptions and should be largely intuitive.* This accounts for the previous discussion in this section.
4. *The method should not be limited to data from a single application area.*
5. *The technique should provide tiered diagnostic information.* Part of the purpose of this is to help improve the simulations and, therefore, feedback is vital.
6. *The comparison should be commutative.* i.e., the results should be the same irrespective of which data set is used as a 'reference'.

A number of approaches were initially considered and these provided a valuable foundation for the development of FSV. These are briefly discussed in the next section.

3. EARLY OPTIONS

Correlation is a process familiar to all engineers and this was first investigated as a means of quantifying data comparisons for validation purposes. It is computationally straightforward and does provide a 'global' goodness of fit. Extension to allow more detailed assessments to be made have been proposed [3] but not widely adopted, largely due to the difficulty in providing easily interpretable point-by-point analyses.

Reliability Factors were originally developed for surface crystallographers to assess discrepancies between experimental and modeled results and generally place a high weighting on regions of high rate of change of signals. Several have been proposed over the years [4, 5] and all have been heuristic, looking to generate a quantifiable signal out of the data. The method outlined in [5] is particularly interesting in that it is an R-factor composed of five individual components. Components A and E are given here and note that C is the fraction of energy range with slopes of different signs (B and D are omitted for the sake of space). It is particularly instructive to note the use of normalized differences and the signal derivatives in deriving this data.

$$A = \frac{\sum_{x_{\min}}^{x_{\max}} |I_{set1}(x) - cI_{set2}(x)|}{\sum_{x_{\min}}^{x_{\max}} |I_{set1}(x)|} \qquad E = \frac{\sum_{x_{\min}}^{x_{\max}} (I'_{set1}(x) - cI'_{set2}(x))^2}{\sum_{x_{\min}}^{x_{\max}} (I'_{set1}(x))^2} \tag{1}$$

None of the R-factors already in use satisfied the criteria for validation laid out above. However, their existence was a fundamental element of the research that led to FSV.

4. THE FEATURE SELECTIVE VALIDATION (FSV) METHOD AND NATURAL LANGUAGE DESCRIPTORS

One further element that helped drive the development of FSV, as it became, was the assertion that engineers in assessing data of this sort will consider the general amplitude/trend comparison and the areas of rapid rate-of-change more-or-less separately and decide whether one is more important or whether they are both equally weighted. This assertion received more support in [6] by the use of eye-tracking experiments. This gave rise to a heuristic that builds a global difference measure

(GDM) from an amplitude difference measure (ADM) and a feature difference measure (FDM) the latter two representing the comparison in general shape and in the rapidly changing features. The ADM was subsequently changed to include the absolute difference (offset) between the data levels [7] and this formulation was adopted in the development of the associated IEEE standard [1]. The data is filtered into three regions to help aid the analysis. The regions are:

- A 'DC' region which is low pass filtered to contain the DC-term in the FT and the first four data points. This data set is transformed back into the original domain of interest.
- A band-pass region that contains 40% of the remaining energy in the transform. This is transformed back to give 'Lo' data sets.
- A high-pass region that contains the remaining data points. Again, this is transformed back into the original domain to give 'Hi' data sets.

These data sets are then used to generate the two component difference measures, the Amplitude Difference Measure (ADM) and the Feature Difference Measure (FDM).

The resulting mathematics is given by (n is the nth data point and N is the total number of data points.)

$$\mathrm{ADM}(n) = \left|\frac{\alpha}{\beta}\right| + \left|\frac{\chi}{\delta}\right| \exp\left\{\left|\frac{\chi}{\delta}\right|\right\} \tag{2}$$

where $\alpha = (|Lo_1(n)| - |Lo_2(n)|)$, $\beta = \frac{1}{N}\sum_{i=1}^{N}((|Lo_1(i)| + |Lo_2(i)|))$, $\chi = (|\mathrm{DC}_1(n)| - |\mathrm{DC}_2(n)|)$, $\delta = \frac{1}{N}\sum_{i=1}^{N}((|\mathrm{DC}_1(i)| + |\mathrm{DC}_2(i)|))$.

The FDM is obtained from Equation (3)

$$\mathrm{FDM}(f) = 2\left(|\mathrm{FDM}_1(f) + \mathrm{FDM}_2(f) + \mathrm{FDM}_3(f)|\right) \tag{3}$$

where

$$\mathrm{FDM}_1(f) = \frac{|Lo_1'(f)| - |Lo_2'(f)|}{\frac{2}{N}\sum_{i=1}^{N}\left(\left(\left|Lo_1'(i)\right| + \left|Lo_2'(i)\right|\right)\right)} \qquad \mathrm{FDM}_2(f) = \frac{|Hi_1'(f)| - |Hi_2'(f)|}{\frac{6}{N}\sum_{i=1}^{N}\left(\left(\left|Hi_1'(i)\right| + \left|Hi_2'(i)\right|\right)\right)}$$

$$\mathrm{FDM}_3(f) = \frac{|Hi_1''(f)| - |Hi_2''(f)|}{\frac{7.2}{N}\sum_{i=1}^{N}\left(\left(\left|Hi_1''(i)\right| + \left|Hi_2''(i)\right|\right)\right)}$$

And the primes ($'$) and double primes ($''$) indicate first and second derivatives obtained from a simple central difference scheme. The overall quality measure, the Global Difference Measure (GDM) is obtained from the ADM and GDM by Equation (4).

$$\mathrm{GDM}(f) = \sqrt{\mathrm{ADM}(f)^2 + \mathrm{FDM}(f)^2} \tag{4}$$

The performance of the FSV routine was determined by comparison to group visual assessment involving 50 engineers with an EMC background assessing eight graphical comparisons [8] and using a visual ratings scale to improve intra-group offset effects and to provide a direct comparison to FSV [9]. This approach used probability density functions (confidence histograms) to do this. The GDM confidence histogram comparison for the data of Figure 1 is shown in Figure 2.

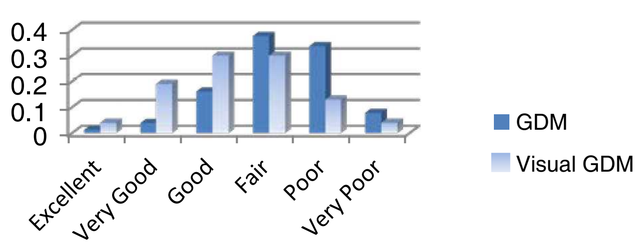


Figure 2: Confidence histogram (pdf) comparison of the GDM for the data of Figure 1.

5. ASSESSMENT OF FSV PERFORMANCE

There have been a number of additional studies to extend the reach of FSV or to assess its performance. For example:

- The effect of data density was studied [10] and it was concluded that there is no appreciable effect on the FSV output providing that any reduction in data density does not appreciably (visually) affect the appearance of the data.
- One mental process that appears to be underway when engineers assess the data is to visually 'shift' closely spaced features that may look to be the same or at least similar structures that have been translated or stretched between the two data sets (which could happen with different sized cavities between the measurements and the models due to discretization in the model). Studies have been undertaken to address this [11, 12].
- Various approaches have been proposed to use FSV with multiple signals, including real and imaginary [2] and many against a mean value [13].

6. CONCLUSIONS AND KEY CHALLENGES

FSV is still a technique in development and quantified validation for CEM is still a very young subject. It provides information that is essential for the formal validation of numerical modeling data in a way that appears to provide a good approximation to the group response of visual assessment. However, there are a number of pressing challenges to be overcome in order to extend the reach of FSV. These include a better mathematical representation and implementation of FSV, developing a better understanding of how humans approach the comparison of multiple dimension data, the effects of zero crossing data and an appreciation of the cumulative effects of numerical noise on the comparison [14].

REFERENCES

1. IEEE, "Standard for validation of computational electromagnetics computer modeling and simulation," IEEE Std 1597.1, Piscataway, NJ, 2009.
2. Martin, A. J. M., "Quantitative data validation (automated visual evaluations)," Ph.D. Thesis, De Montfort University, Leicester, 1999.
3. Duffy, A. P., M. S. Woolfson, and T. M. Benson, "The use of correlation functions to assist the experimental validation of numerical modeling techniques," *Microwave and Optical Technology Letters*, Vol. 7, No. 8, 361–364, 1994.
4. Zanazzi, E. and F. Jona, "Reliability factor for surface structure determination," *Surface Science*, Vol. 62, 61–80, 1977.
5. Van Hove, M. A., S. Y. Tong, and M. H. Elconin, "Surface structure refinements of 2H-MoS2, 2H-NbSe2 and W(100)p(2×1)-O via new reliability factors for surface crystallography," Vol. 64, No. 1, 85–95, 1977.
6. Coleby, D. E., "Assessment of techniques for electromagnetic modelling validation," Ph.D. Thesis, De Montfort University, Leicester, 2004.
7. Duffy, A., A. Orlandi, and H. Sasse, "Offset difference measure enhancement for the feature-selective validation method," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 50, No. 2, 413–415, May 2008.
8. Orlandi, A., A. P. Duffy, B. Archambeault, G. Antonini, D. E. Coleby, and S. Connor, "Feature selective validation (FSV) for validation of computational electromagnetics (CEM), Part II — Assessment of FSV performance," *Transactions on Electromagnetic Compatibility*, Vol. 48, No. 3, 760–767, August 2006.
9. Coelby, D. E. and A. P. Duffy, "A visual interpretation rating scale for validation of numerical models," *COMPEL: Int. J. for Computation and Mathematics in Electrical and Electronic Eng.*, Vol. 24, No. 4, 1087–1092, April 2005.
10. Duffy, A., A. Orlandi, and G. Antonini, "Testing the robustness of the feature selective validation (FSV) method for EMC/SI applications," *EMC Europe*, Barcelona, 2008.
11. Martin, A. J. M. and A. P. Duffy, "A compensation technique to aid the comparison of experimental and modeled data," *IEEE Int. Symp. on Electromagnetic Compatibility*, 475–480, 2002.
12. Edwards, R. S., "Uncertainty analyses in computational electromagnetism," Ph.D. Thesis, University of York, York, Uk, 2009.

13. Archambeault, B. and J. Diepenbrock, "Quantifying the quality of agreement between simulation and validation data for multiple data sets," *Asia-Pacific Symposium on Electromagnetic Compatibility*, 516–519, 2010.
14. Archambeault, B., A. P. Duffy, H. Sasse, X. K. Li, M. Scase, and M. Shafiullah, "Challenges in developing a multidimensional feature selective validation implementation," *IEEE International Symposium on Electromagnetic Compatibility*, Fort Lauderdale, 2010.

Comparing EMC-signatures by FSV as a Quality Assessment Tool

J. Knockaert[1], **D. Pissoort**[2], **and F. Vanhee**[2]

[1]University College of West-Flanders, Ghent University
Graaf Karel de Goedelaan 5, Kortrijk 8500, Belgium
[2]Catholic University College Bruges — Ostend, Catholic University Leuven, ESAT-MICAS
Zeedijk 101, Oostende 8400, Belgium

Abstract— Electromagnetic interference radiated by a consumer product is measured on a sample of similar devices under test. The EMC-signature of these devices can drift due to aging or deviations in manufacturing. It is proposed to use the Feature Selective Validation method to compare the emission from devices with a reference device in order to have a quality assessment tool.

1. INTRODUCTION

Electromagnetic interference radiated by a consumer product is measured on a sample of similar devices under test. This EMC-signature (Electromagnetic Compatibility signature) can drift due to several reasons, which can be grouped in changes during product lifetime and changes due to deviations at production. In the present paper the basic idea is to use the Feature Selective Validation (FSV) method to compare the EMC-signature of supposed identical devices. The result of FSV can give an idea of the equality or in other words quality of the devices. Obvious differences can give an alert to the quality controller. In this way FSV can be used as a quality check tool. The next section discusses the sources of the changing EMC-signature. The third section briefly summarizes FSV. Finally, the pitfalls and some examples are given.

2. SOURCES OF A CHANGING SIGNATURE

2.1. Changes During Product Life Time

The EMC-signature can change due to thermal stress. Thermal stress or an environmental temperature differing from the temperature at measurement time will influence the emitted spectrum. Measurements show drifts of both amplitude and frequency [1]. Drifts in frequency are due to a changing clock frequency and harmonics. Due to the temperature coefficient of passive components the initial spectrum can change up to 10 dB or more [2]. Changing passive components can also shift specific peaks and dips in the spectrum because of changing resonance frequencies. A second reason is aging. Aging has its influence on passive components and connections. During years, thermal cycles can degrade the used materials, changing the values of the components. During time the permittivity ε of capacitors will decrease and leakage currents will increase. As capacitors are a main component of filters it will influence both conducted and radiated emissions. Mechanical stress and vibrations are a third cause of degraded EMC performance. Mechanical stress and vibrations can influence connections. Especially capacitors are susceptible to mechanical resonances, where the connection (solder point) to the PCB can suffer. Hard environments for electronics are automotive and agricultural applications.

A way to have an idea of the influence of these parameters (temperature, mechanical stress and aging) on electronics is by using highly accelerated life time testing (HALT). In a HALT machine a stressful environment for electronics is created to indicate the weak points in the design. HALT test chambers force thermal conditions in the test environment from $-100°C$ to $180°C$ in several minutes. Besides this mechanical stress can be applied to the PCB up to 40 G.

2.2. Differences at Production Time

The EMC-signature of two supposed identical devices can be different. First reason is an anomaly at production time or assembly. Examples are a bad grounding connection or a slightly different length of used wires by the assembler. Different lengths results in different self- and mutual inductances, which change the electromagnetic emission of the device. Also, a different cable length will shift the frequency of the spectrum, depending on series and parallel resonances of the cable [3]. Second reason is the use of a different component. Components can become obsolete, while the replacing component can cause another EMC-signature. A third reason is tolerances on the components. This is again the case for passive components.

3. FSV

FSV is a method for validation of computational electromagnetics [4, 5], with applications in EMC and Signal Integrity. This method has shown its usefulness in the validation of EMC-models [6].

When comparing two datasets, normally measurements and simulations, FSV decomposes both datasets into two parts, trend and feature data. The trend data can be seen as the low frequency part, while the feature data or fast variations can be seen as the high-frequency part. Analysing the low-frequency part gives a measure of similarity of the trend (ADM or Amplitude Difference Measure). Analysing the high-frequency part of both datasets gives a measure of the similarity of the feature (FDM or Feature Difference Measure). These figures combine to a global goodness-of-fit value (GDM or Global Difference Measure). The strength of the FSV-method is the point-by-point comparison showing at which data points the comparison fails.

From previous parts, it is obvious that the EMC-signature can be used as a quality check tool. In this paper, it is proposed to use FSV (Feature Selective Validation Method) to compare two signatures. FSV is developed to compare two datasets, normally a measurement with a simulation result. Nevertheless, it was proven that FSV can also be used to compare two EMC-measurements [7]. Using FSV as a quality assessment is a proposed new application. FSV will be used to compare the EMC-signature of a reference device, with the signature of the device under test. The result of the comparison gives an idea if both devices are really similar. Besides this, FSV can be used to see if older devices are still as electromagnetic compatible as intended or if they have suffered from stress and aging.

4. PITFALLS

Before FSV can be used as a quality check tool, some problems have to be encountered. First problem is the noise on EMC-measurements. It is possible to compare two EMC-measurements, but the comparison suffers from the noise on the data which is typical for EMC peak measurements. Several solutions have been proposed [7]. Performing multiple peak measurements to be comparable with quasi peak measurements and performing quasi peak measurements are valuable solutions, but are time consuming. The only possible solution is performing peak measurements. The benefit is that not the full frequency range has to be measured. If particular resonance frequencies or clock frequencies are known, only a measurement at that frequency interval is necessary. This can decrease the measuring time drastically. Second problem is what validation will be used. Is the GDM (global difference measure) sufficient for evaluation or is a point by point comparison more appropriate? Measurements show that ADM, FDM and GDM give sufficient information to notice a problem. Third problem is choosing the boundary value to decide when quality is insufficient. Figure 1 shows two identical measurements, but the frequency is shifted. The FSV result is given in Figure 2. It shows that a shift in frequency is well noticed by FSV. Typical ADM and FDM values are 0.3 to 0.4, where 0.1 to 0.3 is normal for equal devices. A shift in amplitude gives similar results.

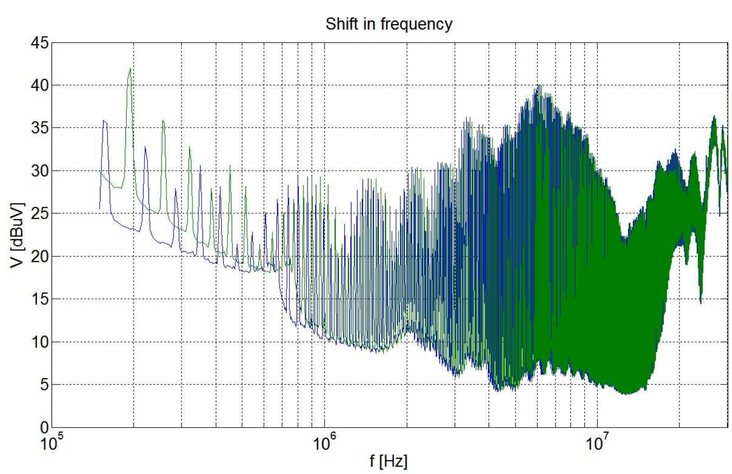


Figure 1: Frequency shift.

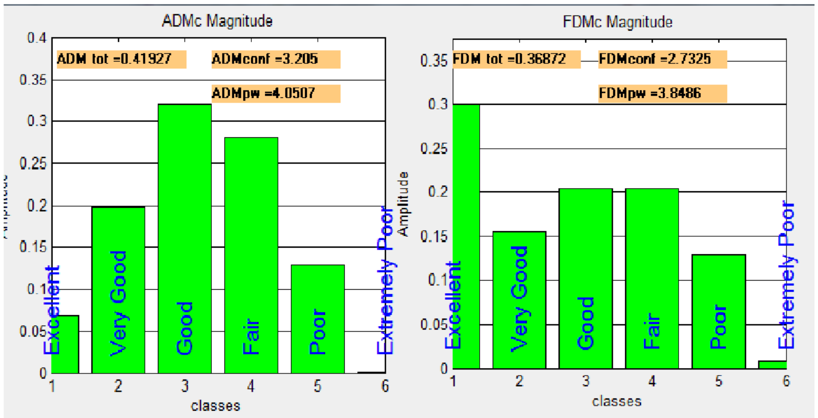


Figure 2: FSV results.

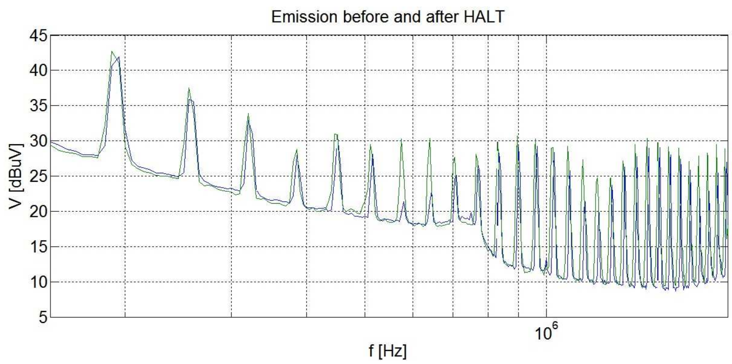


Figure 3: Conducted emission before and after HALT-test.

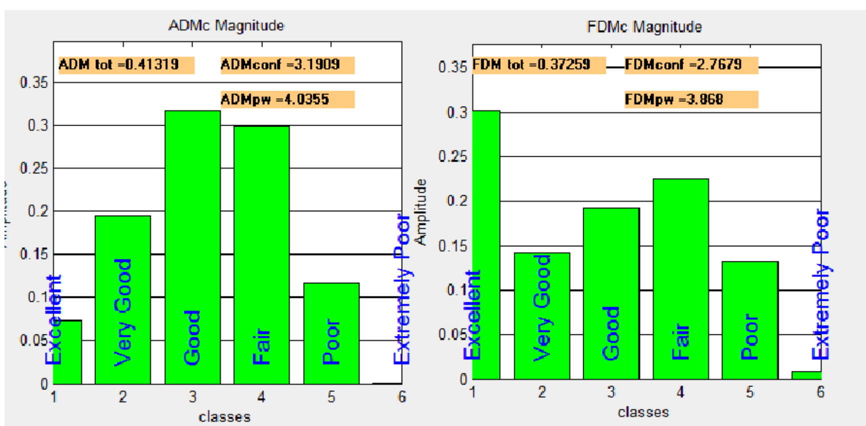


Figure 4: FSV results.

5. EXAMPLES

Figure 3 shows a conducted emission measurement of laptop, before and after a HALT test. During this test, the system was stressed with temperature cycles from -10°C up tot 60°C during several hours. The measurement is performed between 150 kHz and 30 MHz, but for the evaluation only the first 150 kHz to 2 MHz. As it is known from the reference device, the conducted emission is mainly determined by the switching power supply. This proves it is sufficient to focus only on the first harmonics. The measurements were performed in quasi-peak with a LISN and EMI-receiver. The result shows an obvious shift in frequency. The FSV-results are given in Figure 4. Comparing this with the result of two peak-measurements on the same device show that the frequency shift increases especially the ADM-value (from 0.10 to 0.37). The increase of the FDM-value is small (0.33 to 0.34).

The next two measurements (Figure 5) were performed on a frequency inverter, connected to

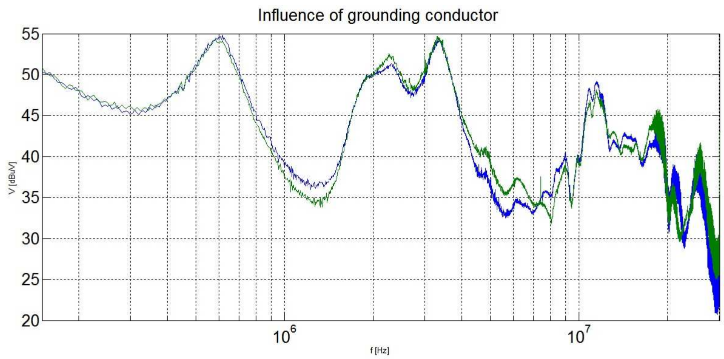


Figure 5: Conducted emission with two different PE-conducters.

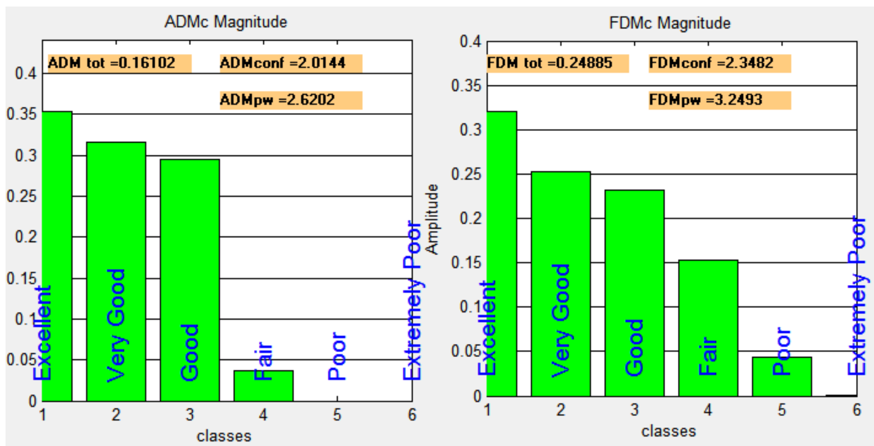


Figure 6: FSV results.

the mains through an EMI filter. The filter was connected to the reference plate by a 30 cm 2.5 mm^2 conductor. In the first measurement, the conductor was massive, in the second measurement the conductor was stranded. The result is a combined shift in amplitude and shift in frequency. As the grounding or the capacitors will be the problem in this setup, it is sufficient to look at the border frequency of the filter. An EMI filter for frequency inverters attenuates up to several MHz. A filter with bad grounding will show an increase at these frequencies. This means only a measurement up to 2 MHz will be sufficient.

6. CONCLUSION

Quality control can be done by measuring several parameters. One of the proposed parameters the EMC-signature. This device specific signature can be measured and compared to a reference device. A deviation of the signature can reveal a problem. It is proposed to use FSV as a tool to support the quality controller on making correct decisions. Some problems and examples were given. Further research will focus on how to choose an appropriate border value and on the practical implementation. Performing EMC measurements is time consuming, but the examples show that measurements in a limited well chosen frequency interval can make the method valuable.

ACKNOWLEDGMENT

Special thanks goes to T. Verbeerst for performing the measurements of Figure 5.

REFERENCES

1. Nyode, A. C., S. Ben Dhia, and A. Boyer, "Analysis of emission spectra evolution after accelerated aging of integrated circuits," *Proc. of the 7th International Caribbean Conference on Devices, Circuits and Systems ICCDCS2008*, 1–5, Mexico, Apr. 28–30, 2008.
2. Lafon, F., F. De Daran, L. Caves, M. Ramdani, M. Drissi, "Influence of aging and environnement conditions on EMC performances of electronic equipment — Influence of passive vs active components," *EMC Europe 2010*, 215–220, Wroclaw, Poland, 2010.

3. Knockaert, J., J. Catrysse, and R. Belmans, "Influence of inverter output cable and filters on conducted emission," *EPE2009*, 10, Barcelona, Spain, 2009.
4. Archambeault, B., S. Pratapneni, L. Zhang, D. C. Wittwer, and J. A. Chen, "Proposed set of specific standard EMC problems to help engineers evaluate EMC modeling tools," *IEEE International Symposium on Electromagnetic Compatibility*, Vol. 2, 1335–1340, 2001.
5. Antonini, G., S. A. Ciccomancini, A. Orlandi, C. Ritota, and A. Duffy, "Applications of FSV to EMC and SI data," *Proceedings of IEEE Int. Symposium on EMC*, Vol. 1, 278–283, Chicago, USA, August 8–12, 2005.
6. Duffy, A., A. Martin, G. Antonini, A. Orlandi, and C. Ritota, "The feature selective validation method (FSV) method," *Proc. of IEEE 2005 EMC Int. Symp.*, Chicago, USA, August 8–12, 2005.
7. Knockaert, J., J. Peuteman, J. Catrysse, and R. Belmans, "Using FSV to compare noisy datasets," *23rd Annual Review of Progress in Applied Computational Electromagnetics (ACES 2007)*, 538–543, Verona, Italy, March 19–23, 2007.

On the Psychological Processes of Decision Making in Displays of Electromagnetic Data

M. O. Scase[1], **M. Shafiullah**[1], **and A. P. Duffy**[2]

[1]Division of Psychology, De Montfort University, UK
[2]Faculty of Technology, De Montfort University, UK

Abstract— Feature selective validation (FSV) is an international standard for comparing visually complex 1D data. Initial verification of FSV was performed by presenting a visual rating chart accompanying comparative data. Eye-scanning data was analysed from observers as they compared graphically presented data. However, little is understood about how observers interpret displays of multidimensional datasets. We have been applying psychological principles of judgment and decision making to investigate the processes involved in this task. The goal has been to establish the main factors that influence observers when they make decisions when viewing displays. The interpretation of complex scenes by human observers to recognise or detect particular features is limited by the structure and performance of the visual system. Therefore to devise displays for observers to detect features from n-dimensional data sets one needs to appreciate the functional limitations of our perceptual system. First, observers must be screened for any visual dysfunction. Factors which can limit or affect processing include: attention; cognitive load; and differences between expert and novice observers data visualisation, interpretation and mental models. The psychological validation of multi-dimensional FSV models demands a range of experimental methods to collect data from human observers. This includes a) users responses (reaction times and errors) to a range of n-dimensional data with systematic manipulations of observer attributes, perceptual features and cognitive elements, b) real-time observational data collection using a talk-aloud protocol and c) eye-fixation data collection of observers. Outcomes of the psychological studies could then be fed into the development of formal validation models.

1. INTRODUCTION

Psychophysics is the study of the relationship between the physical attribute of a stimulus and its perceptual correlate. The term psychophysics was introduced by Fechner in 1860 [1] and there has been a long history of psychologists devising experimental techniques whereby the performance of human observers can be measured quantitatively whilst these observers are making perceptual judgements. Furthermore, psychologists have long investigated the more general processes involved in judgement and decision making [2]. This paper describes the input that psychologists can make using psychophysical techniques to help validate a technique used by engineers to investigate visually presented electromagnetic data. The goal is to provide principles that can aid in the human interpretation of multidimensional displays [3].

2. THE DEVELOPMENT OF FSV

IEEE Std. 1597.1 [4] describes the Feature Selective Validation (FSV) method as a "technique that combines an amplitude-based comparison with a feature-based comparison to give an overall better indication of the agreement between two data sets" [compared with simple difference-distance measures and correlation] it notes that it "has been calibrated to match human 'expert' comparison for decisions that are somewhat subjective". IEEE Std 1597.1 requires the use of FSV in order to comply with the standard. It specifies that a "minimum expectation is that the single value of Global Difference Measure (GDM) along with the Grade and Spread of the GDM will be reported". The GDM is a single value goodness-of-fit measure that describes the (dis)agreement between the two data-sets being considered based on the mathematical heuristics that give an overall measure of the difference in the envelopes for the two data sets and differences in the specific features of the two data sets. While simple difference based distance measures were found to be inaccurate for quantifying differences in the data being used for electromagnetic compatibility studies, the envelope measure, the Amplitude Difference Measure (ADM) and the Feature Difference Measure (FDM) were based on normalised difference measures of filtered portions of the data and of derivatives of filtered portions of the data [5]. The calibration, mentioned in the standard was carried out in a study involving 50 engineers [6]. The ADM and FDM were assumed to be sufficiently independent to be combined into the GDM by

$$\mathrm{GDM} = \sqrt{\mathrm{ADM}^2 + \mathrm{FDM}^2} \tag{1}$$

The assumption on which this was based was supported in the basic eye-tracking research [7]. This study noted that "any technique that does not specifically take into account what the engineers are actually comparing will never exactly emulate their visual comparisons" and went on to propose that engineers consider general shapes and fine detail as separate components in the analysis. While the study was only conducted with a handful of engineers, differences in approaches was noted. For example, one engineer took particular notice of peaks and another took equal notice of the peaks and the troughs in the data; in a further experiment, where the data became visually more complicated, one of the engineers initially concentrated on one trace, presumably to identify the overall shape. Another engineer, when faced with a different plot where there was a large offset between the two data sets, adopted a similar approach where one trace was considered first. A further engineer methodically compared each peak, in sequence, but ignored troughs and gradients.

This was a far from comprehensive study but it did add weight to the supposition that experience would determine how each engineer would approach the comparison as a visual exercise and this would dominate the (potentially broad) spread of opinions. It also leant weight to the assumption that envelopes and features were complementary aspects of the comparisons. This study, supporting the experiential feedback about the quality of FSV was necessary and sufficient to allow FSV to be used in IEEE Std 1597.1. However, further improvements to the method and, in particular, extensions to surface, volume or higher dimensionality data, require more robust and accurate assessment of how engineers do approach the task of visual data comparison, which, itself, requires knowledge and research methods specific to experimental psychology.

3. PSYCHOLOGICAL PRINCIPLES AND THE STRUCTURE AND PERFORMANCE OF THE VISUAL SYSTEM

The human visual system is able to extract the various attributes of the retinal image and process them via anatomically and functionally separate brain areas [8]. This modularity underpins modern cognitive neuropsychology [9]. Selective cognitive loss following head injury gives insight into the modular processing of sensory systems [10] and the mechanisms and limitations of how our visual system can interpret complex scenes.

Computational theoretical approaches to the study of visual processing arguably stem from the work of Marr [11] who devised algorithms based on physiological data to explain the stages in the process of seeing. Attentional mechanisms were incorporated into visual processing with the feature integration theory of vision [12] which built on earlier work [13] concerning cognitive bottleneck and the distribution of attentional processes [14–16].

The interpretation of complex scenes by human observers to recognise or detect particular features is limited by the structure and performance of the visual system. Therefore to devise displays for observers to detect features from n-dimensional data sets one needs to appreciate the functional limitations of our perceptual system.

There have been many attempts to produce visualisation of displays of complex data sets but almost all have come from a computer science or engineering viewpoint [17–26]. Some computer science approaches have drawn evidence from medical data [27] but have not gone as far as to link with psychological theory. Others have used multisensory approaches to combine signals stimuli from vision and audition [28] but it is not clear how our brains can equate signals from different modalities.

4. VISUAL SCREENING

Before presenting complex data to human observers for interpretation it is essential that observers be screened for visual dysfunction. For example if data presentation involves the use of colour then observers should have normal colour vision this can be tested with the Farnsworth-Munsell 100-hue test [29]. About 4% of the Caucasian male population are dichromats and require only two variables in a colour matching experiment and for whom there is a band of wavelengths in the spectrum that matches white [30]. Dichromacy is usually inherited, as a recessive, sex-linked characteristic, and occurs in less than 0.5% of women. This dichromacy is more commonly referred to as red-green colour blindness and individuals confuse reds, browns and greens. It is not only colour that needs to be screened. If presentation involves 3D displays then observers must have normal stereopsis. 3D vision can be assessed with the TNO test for stereoscopic vision [31]. Stereo vision sometimes does not develop normally because of reduced vision in one eye during childhood (amblyopia).

As well as congenital deficiencies it is important to realise that visual losses can be associated with disease (acquired deficiencies) [32]. Different diseases can produce changes at all stages of the visual pathway, from retina through optic nerve to the visual cortex. Examples of diseases that can result in visual losses include diabetes and multiple sclerosis as well as diseases more directly associated with visual disturbance such as glaucoma, retinitis pigmentosa and various optic neuropathies.

It should be appreciated that subtle visual changes (particularly with colour vision) are associated with drugs (cardiac agents, antimalarial, antituberculosis drugs and steroid ovulation inhibitors) [33] and chemicals (alcohol, tobacco, quinine) [34] so for precise visual discrimination tasks full medical details of participants should be noted.

5. LIMITING FACTORS

Techniques involving the combination of a number of visual attributes can be used to display data beyond 2D. A third dimension can come from the use of stereopsis. A fourth dimension can be added with motion and perhaps a fifth with colour. With added dimensions, particularly with multimodal displays, there are several factors that can confound experimental results. It might be difficult to equate one dimension with another. For example, to what extent is a displacement in 2D space equivalent to one in depth (3rd dimension) or motion (4th dimension)? As soon as colour is involved there are even more complicating factors. Displays with colour must be equiluminant or observers might be making decisions based on brightness differences rather than colour differences. Most display devices produce fairly low luminance short wavelength (blue) colours but much brighter medium wavelength (green) colours. Therefore a display involving greens and blues would have much brighter greens than blues unless the display is calibrated for equiluminance. Furthermore, with ever increasingly complex scenes there are issues of what has been called cognitive load and the brain being unable to process effectively all the information. To develop a model applicable to a range of data sets from varied contexts, issues related to observers skills and cognitive loads must also be examined. Cognitive load stems from the work of Miller [35] but the actual term was coined by Sweller [36]. It refers to the load on working memory and has been applied to conditions when human learning happens best. As tasks become more complex so the ability of participants to learn or perform the tasks decreases.

The study of data visualisation comes from a long history of information presentation beginning with work in the 19th century on 2D and 3D maps and sculptures [37–40] to computer simulation, information visualisation and surface and volume rendering [41]. Such techniques can give insight into the different methods that can be used for the visual presentation of multi-dimensional data.

6. EXPERT/NOVICE DIFFERENCES

It has been known since the mid-1960s that there are a cognitive processing differences between experts and novices [42]. Experts tend to remember meaningful combinations from their field of expertise [43, 44]. It has been found that for tasks involving visualisation and interpretation, the spatial transformation of external representation of the visually complex data set into internal representation is more global for expert observers [45]. The experts tend to develop goal-directed qualitative mental models of complex visualisation for interpretation and decision making [46].

7. VALIDATION METHODS

A range of psychological experimental methods could be used to help in the validation of multi-dimensional FSV models. Mental chronometry [47] or the study of reaction times can give an indication of the speed of processing. Error rates should be measures alongside reaction times to monitor any speed-accuracy tradeoff [48]. Reaction time tasks to a range of n-dimensional data with systematic manipulations of observer attributes, perceptual features and cognitive elements can give insight to the variety of processes involved in decision making.

Talk-aloud protocols [49] involve participants describing how they go about completing a task and this real-time observational data collection [50] can give objective feedback from participants.

Eye movement monitoring can give data on what parts in a display that observers look at, the duration and order of fixations and whether any features are re-examined or compared with other features before a decision is made Outcomes of the various psychological studies could then be fed into the development of formal validation models.

8. CONCLUSION

Established psychological principles can be used to help understand how participants make judgements on multidimensional displays. With appropriate screening for visual deficiencies coupled with careful experimental design the results of psychophysical tasks can give quantitative and qualitative data on the processes involved in decision making and help extend FSV validation to multi-dimensional datasets.

REFERENCES

1. Fechner, G. T., *Elemente der Psychophysik*, Breitkopf und Hartel, Leipzig, 1860.
2. Laming, D., *Human Judgment — The Eye of the Beholder*, Thomson Learning, London, 2004.
3. Archambeault, B., A. Duffy, H. Sasse, X. K. Li, M. Scase, M. Shafiullah, A. Orlandi, and D. di Febo, "Challenges in developing a multidimensional feature selective validation implementation," *IEEE Int. Symp. on Electromagnetic Compatibility*, Fort Lauderdale, USA, 2010.
4. IEEE Std 1597.1, "Standard for validation of computational electromagnetics computer modelling and simulation," Feb. 2009.
5. Duffy, A. P. and H. G. Sasse, "The feature selective validation (FSV) method as a means of formal validation of electromagnetic data," *PIERS Proceedings*, Marrakesh, Morocco, March 20–23, 2011.
6. Orlandi, A., A. P. Duffy, B. Archambeault, G. Antonini, D. E. Coleby, and S. Connor, "Feature selective validation (FSV) for validation of computational electromagnetics (CEM), Part II — Assessment of FSV performance," *IEEE Trans. EMC*, Vol. 48, 460–467, 2006.
7. Coleby, D. E., "Assessment of techniques for electromagnetic modelling validation," Ph.D. Thesis, De Montfort University, UK, 2004.
8. Zeki, S., *A Vision of the Brain*, Blackwell, Oxford, 1993.
9. Kolb, B. and I. Q. Whishaw, *Fundamentals of Human Neuropsychology*, 6th Edition, Worth Publishers, New York, 2009.
10. Andrewes, D., *Neuropsychology: from Theory to Practice*, Psychology Press, Hove, UK, 2001.
11. Marr, D., *Vision*, W. H. Freeman and Company, San Francisco, 1982.
12. Treisman, A., "Features and objects in visual processing," *Scientific American*, Vol. 255, 114B–125B, 1986.
13. Welford, A. T., "The psychological refractory period and the timing of high speed performance," *British Journal of Psychology*, Vol. 43, 2–19, 1952.
14. Broadbent, D. E., *Perception and Communication*, Pergamon, Oxford, 1958.
15. Deutsch, J. A. and D. Deutsch, "Attention: some theoretical considerations," *Psychological Review*, Vol. 93, 283–321, 1963.
16. Kahneman, D., *Attention and Effort*, Prentice Hall, Englewood Cliffs, NJ, 1973.
17. Birnholz, J. C. and E. E. Farrell, "Dataface: A suggested method for visual data reduction for pattern recognition," *Clin. Phys. Physiol. Meas.*, Vol. 2, 293–298, 1981.
18. Friel, S. N., F. R. Curcio, and G. W. Bright, "Making sense of graphs: Critical factors influencing comprehension and instructional implications," *Journal for Research in Mathematics Education*, Vol. 32, 124–158, 2001.
19. LeBlanc, J., M. O. Ward, and N. Wittels, "Exploring N-dimensional databases," *Proceedings of Visualization '90*, 230–237, IEEE Computer Science Press, 1990.
20. Mihalisin, T., T. Timlin, and J. Schwegler, "Visualizing multivariate functions, data, and distributions," *IEEE Computer Graphics and Applications*, Vol. 11, 28–35, 1991.
21. Najjar, L. J., *Using Color Effectively (Or Peacocks Can Fly) (IBM TR52.0018)*, IBM Corporation, Atlanta, GA, 1990.
22. Robertson, G. G., J. D. Mackinlay, and S. K. Card, "Cone trees: animated 3D visualizations of hierarchical information," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching through Technology*, 189–194, 1991.
23. Shovman, M. M., A. Szymkowiak, J. L. Bown, and K. C. Scott-Brown, "Changing the view: towards the theory of visualisation comprehension," *13th International Conference Information Visualisation*, 135–138, 2009.
24. Vessey, I., "Cognitive fit: A theory based analysis of the graph versus tables literature," *Decision Sciences*, Vol. 22, 219–240, 1990.
25. Winter, V. L., "Visualization and animation as a technique to assist in the construction of high assurance software," *CADE Workshop on Visual Reasoning*, 11–17, 1996.

26. Wolfe, R. H. and C. N. Liu, "Interactive visualization of 3D seismic data: A volumetric method," *IEEE Computer Graphics and Applications*, Vol. 8, 24–30, 1988.
27. Ogiela, L. and M. R. Ogiela, *Cognitive Techniques in Visual Data Interpretation*, Vol. 228, Studies in Computational Intelligence, Springer-Verlag, Berlin, 2009.
28. Rabenhorst, D. A., E. J. Farrell, D. H. Jameson, T. D. Linton, and J. A. Mandelman, "Complementary visualization and sonification of multidimensional data," *Proc. SPIE*, Vol. 1259, 147, 1990.
29. Farnsworth, D., *The Farnsworth-Munsell 100-hue Test for the Examination of Color Discrimination: Manual*, Munsell Color Company Inc., Baltimore, 1957.
30. Fletcher, R. and J. Voke, *Defective Colour Vision*, Adam Hilger, Bristol, 1985.
31. TNO, *TNO Test for Stereoscopic Visiont*, Lameris Ootech, Utrecht, 1972.
32. Foster, D. H., *Inherited and Acquired Colour Vision Deficiencies, Vision and Visual Dysfunction*, Vol. 7, Macmillan, Basingstoke, 1991.
33. Neubauer, A. S., K. Samari-Kermani, U. Schaller, U. Welge-Lüßen, G. Rudolph, and T. Berninger, "Detecting chloroquine retinopathy: electro-oculogram versus colour vision," *Br. J. Ophthalmol.*, Vol. 87, 902–908, 2003.
34. Merglera, D., L. Blaina, J. Lemairea, and F. Lalande, "Colour vision impairment and alcohol consumption," *Neurotoxicology and Teratology*, Vol. 10, 255–260, 1988.
35. Miller, G. A., "The magic number seven plus or minus two: Some limits on our capacity to process information," *Psychological Review*, Vol. 63, 81–97, 1956.
36. Sweller, J., "Cognitive load during problem solving: Effects on learning," *Cognitive Science*, Vol. 12, 257–285, 1988.
37. Maxwell, J. C. and P. M. Harman, *The Scientific Letters and Papers of James Clerk Maxwell*, Vol. 3, 1874–1879, Cambridge University Press, Cambridge, 2002.
38. Minard, C. J., "Carte figurative des pertes successives en hommes de l'armée qu'Annibal conduisit d'Espagne en Italie en traversant les Gaules (selon Polybe). Carte figurative des pertes successives en hommes de l'armée française dans la campagne de Russie," 1812–1813, 1869.
39. Nightingale, F., *Mortality of the British Army*, Harrison and Sons, London, 1857.
40. Snow, J., *On the Mode of Communication of Cholera*, (n.p.), London, 1855.
41. Rosenblum, L. J., *Scientific Visualization: Advances and Challenges*, Academic Press, London, 1994.
42. De Groot, A., *Thought and Choice in Chess*, Morton, The Hague, 1965.
43. Chase, W. and H. Simon, "Perception in chess," *Cognitive Psychology*, Vol. 4, 55–81, 1973.
44. Egan, D. and B. Schwartz, "Chunking in recall of symbolic drawings," *Memory and Cognition*, Vol. 7, 149–158, 1979.
45. Trafton, G. J., S. B. Trickett, and F. E. Mintz, "Connecting internal and external representations: Spatial transformations of scientific visualization," *Foundations of Science*, Vol. 10, 89–106, 2005.
46. Trafton, G. J., S. S. Kirschenbaum, T. L. Tsui, J. A. Ballas, and P. D. Raymond, "Turning pictures into numbers: Extracting and generating information from complex visualizations," *International Journal of Human-Computer Studies*, Vol. 53, 827–850, 2000.
47. Donders, F. C., "On the speed of mental processes," *Attention and Performance II, Acta Psychologica*, W. G. Koster, ed., Vol. 30, 412–431, 1868/1869.
48. MacKay, D. G., "The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior," *Psychological Review*, Vol. 89, 483–506, 1982.
49. Ericsson, K. and H. Simon, "Verbal reports as data," *Psychological Review*, Vol. 87, 215–251, 1980.
50. Dunbar, K., "How Scientists Think: on-line creativity and conceptual change in science," *Creative Thought: An Investigation of Conceptual Structures and Processes*, T. B. Ward and S. M. Smith, eds., 461–493, American Psychological Association, Washington, DC, 1997.

Numerical Noise Reduction in the Fourier Transform Component of Feature Selective Validation

Hugh G. Sasse and Alistair P. Duffy
De Montfort University, UK

Abstract— Abstract Feature Selective Validation (FSV) was developed as a means of making the subjective judgements about the quality of a comparison (a) repeatable, (b) consistent with the experience of a group of engineers and (c) amenable to automation. Over the past decade it has gained acceptance and is now the core technique in an IEEE standard. However, there is still scope for further research and refinements to the technique and its implementation are still possible. FSV, as defined in IEEE Std 1597.1, is built around the Fourier Transform, and the Fast Fourier Transform (FFT) is assumed to be the implementation of choice. When testing an implementation of the FFT it was found that the inverse of the transform differed from the original real signal by a small amount, and this was revealed by small imaginary components of the order of 10^{-15}. This was inferred to be numerical noise. The Standard does not discuss numerical noise, and for most signals such noise is irrelevant. However, when dealing with small signals or transients this becomes important, because there will be significant proportions of the data where the signal will be at or near zero. When the FSV components are calculated in dBs the constituent ratios will depend on two small and potentially pseudo-random numbers which then may lead to a false and potentially large component of FSV. Using the linearity property of the Fourier Transform, an approximation to this numerical error is obtained and this noise is then subtracted to provide a closer approximation to the ideal transform. This paper discusses the use of this approach for various classes of signal and the paper discusses implications for the reliability of the FSV results.

1. INTRODUCTION

It has been custom and practice in electromagnetic compatibility to perform comparisons between experiments, and between experiments and simulation, by eye. Given the noise floors involved on the one hand, and the experience of engineers on the other, this has been good enough, and much progress has been made under these conditions. However, there has been a desire to make this process more repeatable, and more rigorous. Disagreement which can only be settled by reference to one's experience is difficult to resolve satisfactorily, and with human decision makers one would expect some variability in performance. To this end, Feature Selective Validation (FSV) was developed as a means of providing a method which could be implemented easily, which would produce results which engineers could generally agree with, and, being automatic, was repeatable. It has now gained acceptance as a standard [1].

In terms of improvements to the method, it is now being applied to signals with significant regions of zeroes, such as transients, which is different from the original domains of application where there was significant signal strength. Recent experiments applying the method to two dimensional signals have revealed a noise floor of the order of 10^{-15} where zeroes were expected. This is after the inverse Fourier transform step. This paper analyzes and discusses this phenomenon.

2. NUMERICAL NOISE REDUCTION IN A LINEAR PROCESS.

The Fourier Transform is a linear process, by virtue of its being based on integrals [2]. The problems of error propagation in linear equations are well known, and a case commonly encountered in engineering is the solution of a system of linear equations. This may be expressed in matrix form as:

$$AX = B \tag{1}$$

When solving for for X a given A and B, rounding errors or truncation errors will occur in the computer arithmetic. One way to solve this problem is by computing an error term, as is done in [3]. Let X_0 be the first calculated value of X. Let B_0 be the result of multiplying A by X_0. If B and B_0 differ, as one would expect due to numerical errors, then one may solve the modified set of equations to obtain a better approximation to X. First, determine the error vector

$$\Gamma = B - B_0 \tag{2}$$

Then we can compute an error term, ϵ, for X:

$$A\epsilon = \Gamma \tag{3}$$

A new X_1 can be obtained using

$$X_1 = X_0 - \epsilon \tag{4}$$

This can then be used to generate B_1, a better approximation of B, and these steps are repeated until the ϵ terms are small enough.

This process relies purely on the linearity property of the matrix operations. In this case, the linear operation is the Fourier transform.

2.1. Numerical Noise Reduction in the Fourier Transform

The causes and characteristics of noise must be understood in order to be able to remove it. In an experimental setup there will be measurement noise, as well as numerical noise in the signal processing of the results. Understanding the different characteristics will aid in attributing the errors to the correct cause. Because of the linearity property of the Fourier transform, we know that

$$\mathcal{F}(a+b) = \mathcal{F}(a) + \mathcal{F}(B) \tag{5}$$

If noise is added at each stage of the transform, it is desirable to separate this noise out and remove it. The FSV method performs a transform, some filtering, and then the inverse transform. Just considering the operation without filtering on a single signal, performing a transform then an inverse, will introduce noise at each stage, this may be considered to be:

$$f_1(x) = \mathcal{F}^{-1}(\mathcal{F}(f(x) + \epsilon_f) + \epsilon_i) \tag{6}$$

As a result of this, the total noise added, ϵ, is

$$\epsilon \approx f_1(x) - f(x) \tag{7}$$

Unfortunately, this contains components from the inverse transform of the noise introduced by the forward transform. Whilst this makes the elimination of the noise difficult, it points to its impact on FSV. In the case of a transient signal, or one with a significant run of zeroes, then any noise picked up in the forward transform will be distributed by the inverse transform in a recognisable way. So, we can say that if the noise appears to be proportional to the signal, then there will be a component proportional to the Fourier transform of the signal on returning to the original domain of the signal. The consequence of this is that for a large transient, there will be a noise signal spread across the domain, possibly dominating any local signal.

An estimate of may be obtained by:

$$\epsilon_i \approx -\mathcal{F}(f_1(x)) - \mathcal{F}(f(x) + e_f) \tag{8}$$

The problem is that if ϵ_f is to be considered negligible in order to compute ϵ_i, then all the noise will be attributed to the inverse transform process. However, this may be sufficient to reduce the noise. Then

$$\epsilon_i \approx -\mathcal{F}(f_1(x) - f(x)) \tag{9}$$

This may then be removed from the intermediate signal.

3. METHOD

A fast fourier transform was implemented in Ruby, and its inverse was implemented using the swapping of real and imaginary parts applied to the input and results of the forward transform. This relationship is described in [4], and elsewhere.

The signal used initially is shown in Figure 1, which was subsequently zero padded to a length of 256 (next power of 2) for the Fast Fourier transform, and the resulting overall noise signal is shown in Figure 2.

The magnitude of the overall error signal ϵ is shown in Figure 3. The sizes are difficult to see so that statistics have been shown in Table 1. The numerical estimate of ϵ_i, obtained as described earlier, is shown in Figure 4, and its magnitude is shown in Figure 5. Once again the statistics for this information are shown, see Table 2.

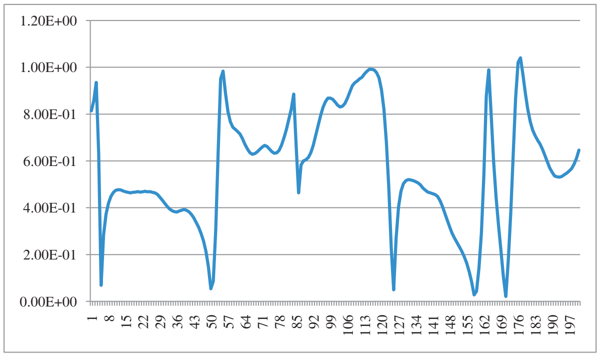


Figure 1: Chart 1 input signal.

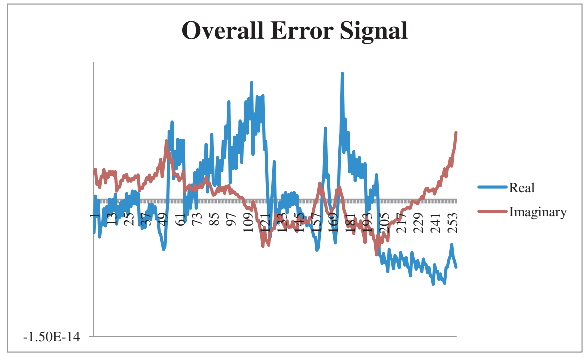


Figure 2: Chart, ϵ, real and imaginary parts.

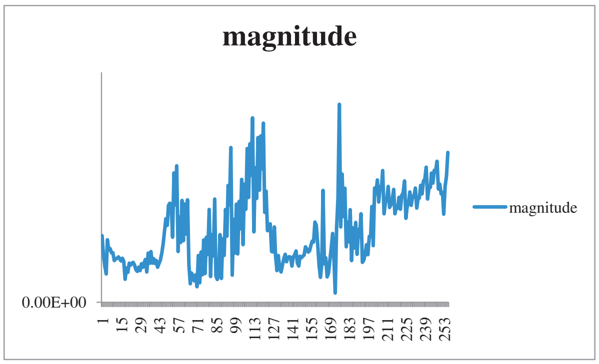


Figure 3: Magnitude of ϵ error signal.

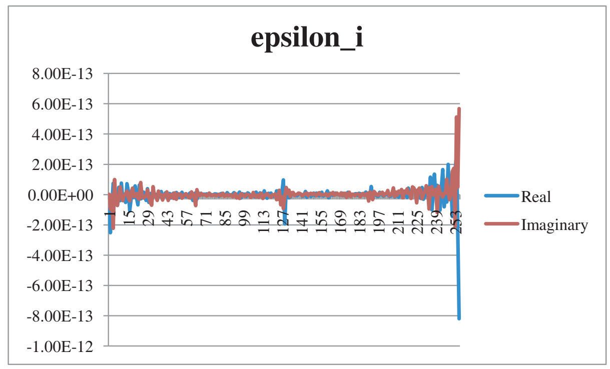


Figure 4: Real and imaginary parts of ϵ_i for Chart 1.

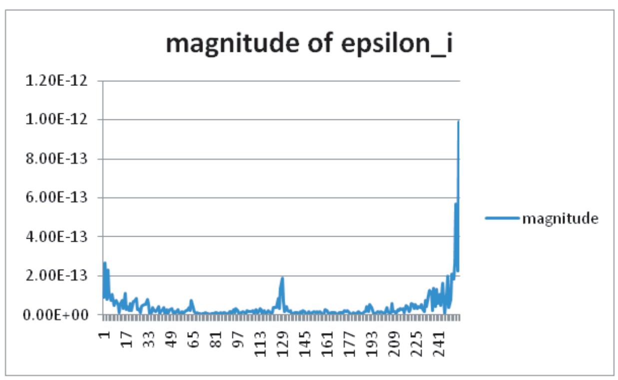


Figure 5: Magnitude of ϵ_i for chart 1.

Table 1: Statistical values for ϵ magnitude.

min	6.36e-16
max	1.38e-14
RMS	5.70e-15

Table 2: Statistical values for ϵ_i magnitude.

Min	3.10e-16
Max	9.97e-13
RMS	9.11e-14

4. DISCUSSION

The overall numerical noise is clearly small. In any given situation it may be dominated by measurement errors and this will be the topic of a further study. ϵ_i, the computed inverse transform noise, does look rather like the transform of a time reversed signal, much as one would expect, given the Fourier transform identities. Furthermore, the prediction that regions of zeroes in the input would be affected by the noise has been demonstrated in the zero-padded region of the input. At a very crude approximation, the noise does seem to track the magnitude of the input. Inputs with a large dynamic range may benefit from the extra inverse transform to deduce the numerical noise, so that it can be subtracted prior to filtering into DC, LO, and HI parts of the input. It may be desirable for some data to have FSV applied differently across different portions of the data, particularly where the data varies substantially across the range of the data, for example in transient simulations. The implication of this study is that a 'short period' FSV, where the denominator of the Difference Measures only spans a portion of the data, will take very careful thought.

REFERENCES

1. IEEE 1597.1, "Standard for validation of computational electromagnetics computer modeling and simulation," *IEEE*, 2009.
2. Stremler, F. G., *Introduction to Communication Systems*, Addison-Wesley, 1982, ISBN: 0-201-07259-9.
3. Dorn, W. S. and D. D. McCracken, *Numerical Methods with Fortran IV Case Studies*, John Wiley & Sons, 1972, ISBN:0-417-21919-3.
4. Duhamel, P., B. Piron, and J. M. Etcheto, "On computing the inverse DFT," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 36, No. 2, 285–286, February 1988.

Study of Transient Phenomena with Feature Selective Validation Method

R. Jauregui[1] **J. Rojas-Mora**[2], **and F. Silva**[1]

[1]Universitat Politècnica de Catalunya, Spain
[2]University of Barcelona, Spain

Abstract— In recent years, computational electromagnetism has had a great development thanks to the computational systems speed increase and their cost reduction. With those improvements the mathematical algorithms are able to work properly with more practical EMC issues. The problem that arises many times is to become confident with the results, in other words, to be able to quantitatively validate the results of the numerical simulation. In this paper we present a method to evaluate the difference between results obtained by visual expert opinion and those obtained from the FSV method.

1. INTRODUCTION

In recent years, computational electromagnetism has had a great development thanks to the computational systems speed increase and their cost reduction. These improvements have allowed us to analyze more complex systems and their EMC behavior.

One of the most interesting signals from the viewpoint of electromagnetic compatibility is impulsive noise, also known as transient phenomenon. Transient interference could be described as a signal that varies between two consecutive steady states during a short period of time compared to the time scale of interest. In other words, there must be a "momentary" change of the magnitude seen for a very short time in the sense that this short interval of time should be much less than one cycle of the signal affected [1].

The particular feature of them makes it difficult to analyze the effect of it in the different systems. For this reason, the numerical methods are a useful tool to analyze the effect of the transient signals in other systems [2, 3]. However, this new tool requires to ensure the quality of the results, in other words, to make a correct validation of the results.

The validation method most widely used today because of its versatility and simplicity in the field of EMC, is the Feature Select Validation method (FSV) [4, 5]. FSV, which is now an IEEE standard [6], has the advantage of analyzing the two major aspects that are considered in any validation, the magnitude levels and the shape of the graphs.

Despite all the benefits offered by the FSV in the validation processes, it has been observed in previous studies that there are some problems when trying to use this method in transient signals [7, 8]. To overcome these drawbacks, and to make a correct validation of transient signals, an original solution was proposed [8]: to mimic what an expert does when he analyzes a transient, dividing the signal in three different regions (pre-event, event, post-event) and assigning a specific weight to each of them. But not all transients are equal, so the main problem that arises in this new technique is the assignment of weights and the assessment of the size of each event for different types of transients.

This paper presents the analysis of three types of FSV methods to analyze some of the most representative transients in the EMC area. Specifically, eight transients were evaluated using the traditional FSV, the Weighted-FSV and the Manual-Weighted-FSV. All these methods were compared with expert opinion to help us make a more objective evaluation of each method.

2. FSV AND TRANSIENT PHONEMENA

The FSV method is based on the decomposition of the results into two groups; the first one discusses the difference in amplitude (Amplitude Difference Measure, ADM) and the second one the difference between the characteristics of the signal (Feature Difference Measure, FDM). The combination of these two indicators (ADM and FDM) is a measurement of the overall difference (Global Difference Measure, GDM).

All indicators ADM, FDM and GDM have the ability to be configured to perform a point-to-point analysis. The advantage of relying on point-to-point data is to know which areas of the data sets have the major differences. A subscript "i" is added to consider this point-by-point feature (ADMi, FDMi and GDMi). All of the indicators are calculated accordingly to the following equations ([6]):

$$ADM(n) = \frac{\alpha}{\beta} + ODM \cdot e^{ODM} \tag{1}$$

where:

$ODM = \frac{x}{\delta}$.

$x = (|DC_1(n)| + |DC_2(n)|)$.

$\delta = \frac{1}{N}\sum_{i=1}^{N}(|DC_1(i)| + |DC_2(n)|)$.

DC-data is the "very low" component of the original data.

Lo-data is the low-frequency component of the original data.

Hi-data is the hi-frequency component of the original data.

$$FDM(n) = 2(|FDM_1 + FDM_2 + FDM_3|) \tag{2}$$

where

$$FDM_1(n) = \frac{(|Lo_1'(n)| - |Lo_2'(n)|}{\frac{2}{N}\sum_{i=1}^{N}|(|Lo_1'(i)| + |Lo_2'(n)||}.$$

$$FDM_2(n) = \frac{(|Hi_1'(n)| - |Hi_2'(n)|}{\frac{6}{N}\sum_{i=1}^{N}|(|Hi_1'(i)| + |Hi_2'(n)||}.$$

$$FDM_3(n) = \frac{(|Hi_1''(n)| - Hi_2''(n)|}{\frac{7.2}{N}\sum_{i=1}^{N}|(|Hi_1''(i)| + |Hi_2''(i)|)|}.$$

$$GDMi(n) = \sqrt{ADM(n^2) + FDM(n)^2} \tag{3}$$

Another way to qualitatively analyse the FSV indicators is represented by a probability density function. This indicator is useful for a rapid and comprehensive analysis of the results. Histograms are sorted according to the quality of the results in excellent, very good, good, fair, poor and very poor.

This efficient and rapid method of analysis is ideal for most signals, but as mentioned above, some studies showed that the use of FSV for transient analysis is inadequate. For this reason, a variation of the method was created, **Weighted-FSV** [8]. This method divides the signal in three different sectors or regions and assigns a fixed weight to each of them. The first region (pre-event) is defined from $t = 0$ to the time the signal amplitude reached 5% of peak value. The pre-event represents 5% of the total weight on the transient analysis.

The second region (the event) is defined as the transition itself, ranging from the end of the pre-event up to a point that contains 65% of the signal energy (the largest region is chosen when the two sets of data match each other). Since this region is the most important of the entire disturbance, it is assigned a weight of 70% of the total.

Finally, the third region, the post-event, is defined from the end of the event to the end of the signal. Usually, this is the region of longer duration, but it is not the most important, for this reason a weight of 25% is assigned. However, despite this method has proven to be a good alternative for transient analysis, it has not been tested against a wide variety of them.

3. ANALYSIS OF THE TRANSIENTS

In this section we have analyzed eight typical EMC transients (Table 1). Figure 1 shows two short and very similar transients. In Figure 2, the transients analyzed have a longer settlement time and a larger ripple in each area of the event. The transients in Figures 3, 4 and 5 are more complex, with settlement times longer and with more ripples throughout the signal. Figure 6 shows two transients (Modulated Gaussian pulse) with equal amplitude but different modulation frequencies. Figure 7 shows two transient phenomena without pre-event and lengthy settlement times. In the last figure (Figure 8), shows two completely different transient signals.

The transients were carefully chosen and have a unique feature that will help us better understand the interpretation of the FSV, the "Weight-FSV" (W-FSV) and the "Dynamic-Region-FSV" (D-FSV). In the D-FSV method the weights are assigned in each zone manually, depending on the length and breadth that has the transient in that area in particular.

Finally, the results are compared with those obtained from a survey based on Likert scale questions [9], carried out with a sample of 18 experts in the area, who evaluated the transient based on previous studies [5]. To ease the comparison between the different methods and expert opinion, the mean square difference (see equation 4) is calculated for each indicator and each expert D-FSV method.

$$Diff(GDm_{expert}, GDMi_{FSV}) = \sqrt{\left(\frac{\sum_{i=1}^{n}(GDM_{expert(i)} - GDMi_{FSV(i)})^2}{n}\right)} \quad (4)$$

where: "i" is an element in the set of subjective values of the GDM indicator (from excellent to very poor).

4. DISCUSSION

Table 1 shows, as expected, that the worst results compared with the expert opinion are found in the FSV method. In the specific case of the transient signals, the FSV method has two problems. The first one is found in the ODM indicator (see Equation (1)), as it is severely affected by the offset levels. The reason can be found in the smaller offset in the signals when the value of "δ" decreases and the ODM grows. This error directly affects to the ADM indicator and, through it, to the GDM indicator. This is an important drawback for transients because most of them have offset

Table 1: Different type of transients analyzed.

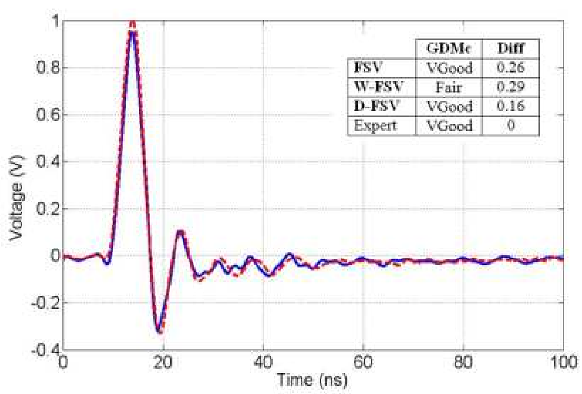


Figure 1: Weights used in D-FSV: pre=2%; event=95%; Post=3%.

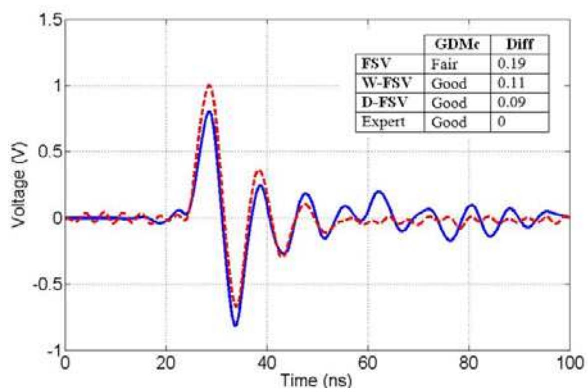


Figure 2: Weights used in D-FSV: pre=5%; event=70%; Post=25%.

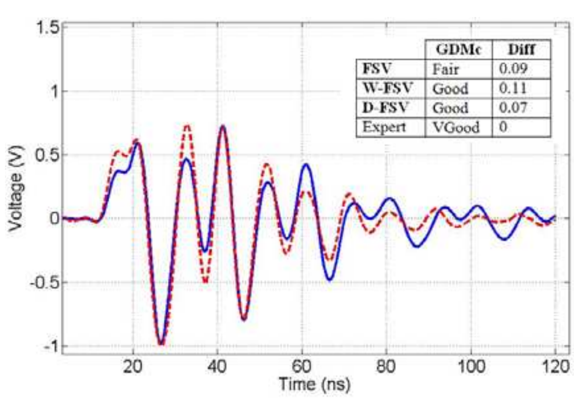


Figure 3: Weights used in D-FSV: pre=5%; event=85%; Post=10%.

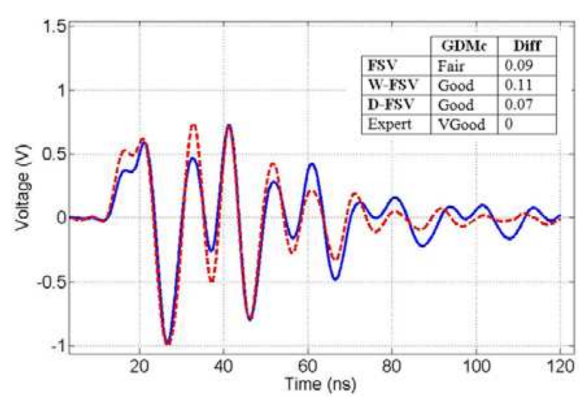


Figure 4: Weights used in D-FSV: pre=5%; event=80%; Post=15%.

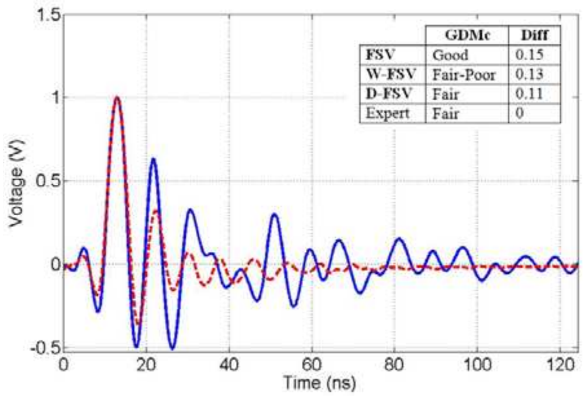


Figure 5: Weights used in D-FSV: pre=5%; event=70%; Post=25%.

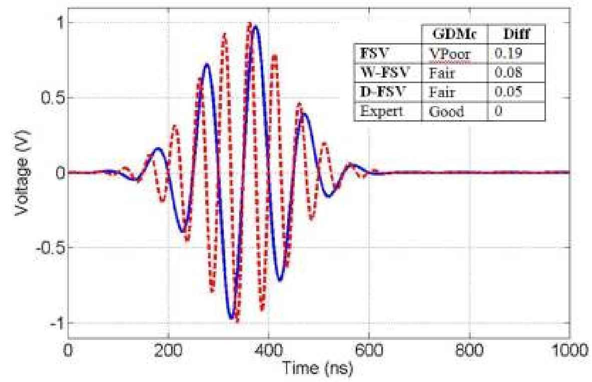


Figure 6: Weights used in D-FSV: pre=5%; event=85%; Post=10%.

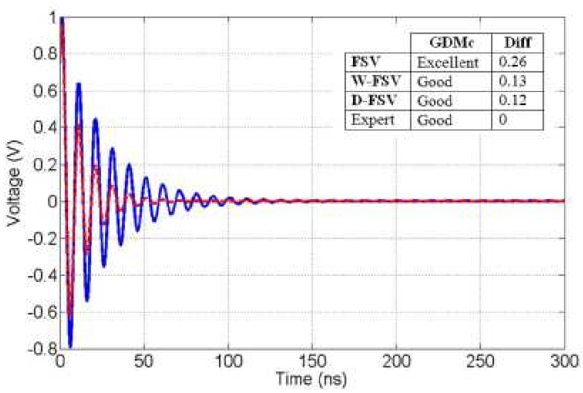


Figure 7: Weights used in D-FSV: pre=0%; event=80%; Post=20%.

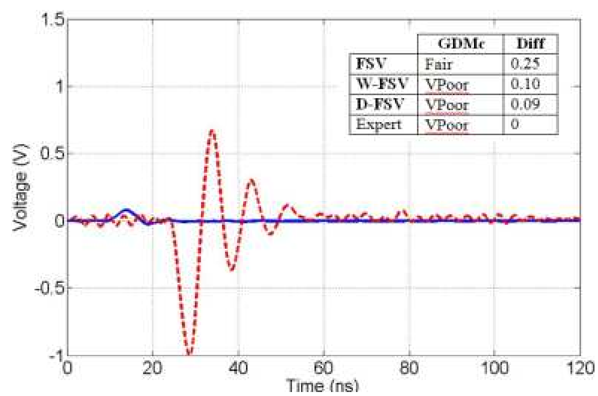


Figure 8: Weights used in D-FSV: pre=5%; event=80%; Post=15%.

levels close to zero. The second problem with FSV is, as stated above, that it gives equal weight to the three regions of the transient. This produces a grossly overestimating in the pre-event and the post event, skewing the final assessment away from the expert opinion.

The previous problems related with the FSV method were corrected with the W-FSV method. The ODM indicator deviations are solved by applying a large enough offset level to both signals to ensure an objective interpretation. On the other hand, the division of the transitional phenomenon into regions makes the results closer to the expert opinion in most of the cases.

However, this method still has some drawbacks. The main problem in W-FSV is that the weight remains fixed in all the proposed regions, regardless of the type of transient. Some good examples of this problem could be seen in the previous figures, where the importance of each event varies depending on the extent and duration of the event. An additional problem with this technique is that the possibility that the pre-event could be zero (Figure 7) or that the post-event could be extremely long (Figure 1) are not taken into account, making the results drift away from the assessments made by experts.

Finally, the results obtained by the method D-FSV presented in this paper, show an improvement over the W-FSV. The main reason for this improvement is that the weight assigned to each region of the event is selected manually in terms of magnitude and duration (Figures 1, 4, 6 and 7).

An important feature of this method is the procedure used to select the regions in the transient event. It was noted that if the power method (70% of total energy) is used to divide the event and post-event regions in the transient with very long or very short time of settlements (Figures 1, 4 and 6), some errors still appears in the interpretation of the transient. This division creates very short event and very long post-event regions in the transient, having a direct effect on the results of FSV. To solve this problem we use an approach based in the classical control theory [10] to divide the transient in the regions of interest. The theory defines the post-event region when the signal level decreases to 5% of the maximum. Conducting a study with the experts, the best value obtained for most of the transients was when the signal reached 10% of maximum.

5. CONCLUSIONS

This paper has addressed some problems related to validation of transient signals using different variants of the FSV method. The principal conclusion of this work is that, to use the FSV as a validation method for transient signals, it is not only necessary to divide each region of the transient carefully, but that a specific weighting for each region must also be carried out. The results obtained in this test show that it is necessary to assign weights to each region depending on both, its duration and amplitude.

Our results suggest that it is necessary to continue investigating ways of improving the analysis of transients in time domain through the FSV.

REFERENCES

1. Jauregui, R., "Electromagnetics analysis in time domain of complex structure using FDTD," Thesis, 2009.
2. Riu, P. J., R. Jauregui, S. Verdaguer, R. Santos, and F. Silva, "Transient radiated fields in 14/42 V powernet systems FDTD assessment," *EMC Europe*, Barcelona, Spain, 2006.
3. Portugues, I. E., P. J. Moore, I. A. Glover, and R. J. Watson, "A portable wideband impulsive noise location system," *IEEE Transactions on Instrumentation and Measurement*, IEEE Instrumentation and Measurement Society, Vol. 57, 2059–2066, 2008.
4. Duffy, A. P., A. J. M. Martin, A. Orlandi, G. Antonini, T. M. Benson, and M. S. Woolfson, "Feature selective validation (FSV) for validation of computational electromagnetics (CEM). Part I: The FSV method," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 48, 449–459, 2006.
5. Orlandi, A., A. P. Duffy, B. Archambeault, G. Antonini, D. E. Coleby, and S. Connor, "Feature selective validation (FSV) for validation of computational electromagnetics (CEM). Part II: Assessment of FSV performance," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 48, 460–467, 2006.
6. "IEEE 1597.1 Standard for validation of computational electromagnetics computer modeling and simulations," *The IEEE Electromagnetic Compatibility Society*, P.1 W.G., Jun. 2008.
7. Jauregui, R., P. J. Riu, and F. Silva, "Transient FDTD simulation validation," *IEEE International Symposium on Electromagnetic Compatibility*, 257–262, 2010.
8. Jauregui, R., F. Silva, A. Orlandi, H. Sasse, and A. Duffy, "Factors influencing the successful validation of transient phenomenon modelling," *Asia-Pacific Electromagnetic Compatibility Symposium and Technical Exhibition*, 2–5, Beijing, 2010.
9. Likert, R., "A technique for the measurement of attitudes," s.n., 1932.
10. Ogata, K., *Modern Control Engineering*, 2001.

Performance Improvement of FSV in a Special Situation

Gang Zhang[1], Lixin Wang[1], and Alistair Duffy[2]

[1]School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, China
[2]Department of Engineering, De Montfort University, Leicester, UK

Abstract— As the Feature Selective Validation (FSV) technique is becoming a dominant quantitative validation method of computational electromagnetic simulation results. Some studies have been undertaken, and subsequent enhancements proposed, to improve its performance. This paper illustrates a situation in which the FSV results do not agree with visual expectation: data sets consist of a high transient event and relative low pre-transient and post-transient regions. The reasons for this discrepancy are discussed in detail in this paper and shown to be related to the Offset Difference Measure (ODM) which is then modified in order to avoid this apparent error without any pre-processing to original data sets. It is finally shown that the modified FSV method could avoid the error and has no influence on the comparing of other data sets.

1. INTRODUCTION

The Feature Selective Validation (FSV) method [1] has been shown to be an effective tool in quantifying the agreement between measurements and simulations, especially within the field of Electromagnetic Compatibility (EMC). Therefore, it has been incorporated as the central technique of IEEE STD 1597.1 [2]. At present, research in the validation of computational electromagnetic is primarily focused on two issues: The first of these is the use of the FSV method, in particular, exploring its strengths and weaknesses [3]. And the other is the extension of FSV into the multidimensional FSV method for example.

There is a special situation that produces results that generally differ from user expectation: data sets consist of high transient event and relative low pre-transient and post-transient regions. The situation has been discussed in [4]. It was demonstrated that the performance of FSV could be effectively improved by offsetting the original data to one half plane and changing the weight of the transient event in the FSV results. However, this solution will increase computational cost, especially in the comparing of multidimensional data sets. Furthermore, this situation has the potential to arise more frequently in multidimensional data and may not be as obvious when it does occur. Therefore, further studies are still necessary to reveal the root of the issue and solve it thoroughly.

The paper first describes the FSV method in Section 2. After that, Section 3 illustrates the special situation mentioned above, and then Section 4 presents the modification to ODM, which successfully overcomes the apparent error and has no influence on the comparing of other data sets. Section 6 draws the conclusion.

2. THE FSV METHOD

The FSV method was established to mirror the decision-making process of a group of engineers when comparing data typical to EMC applications. Details of the FSV method can be found in [2]. To summarize the method, the original data sets are Fourier transformed first, and then the first 4 points can be grouped as DC component, the remainder are separated at the 40% of energy mark to give the low-frequency component and high-frequency component. These components can be used to compute the values of Amplitude Difference Measure (ADM) and Feature Difference Measure (FDM). The ADM is a measure of the agreement of the 'trend' information between the data sets. As part of ADM, the Offset Difference Measure (ODM) is introduced to reflect the level of DC difference between two sets of results [5]. The FDM is a measure of the agreement of the 'feature' information between the data sets. And the overall 'goodness of fit', which is called Global Difference Measure (GDM), is made up of the ADM and FDM.

The FSV can provide information in a number of different layers. The ADMi, FDMi, GDMi are calculated on a point-by-point basis for the data sets. Moreover, these values can be arranged into three probability density histograms in various agreement categories using the interpretation scale shown in Table 1. These histograms are denoted ADMc, FDMc, GDMc. Furthermore, the single value of ADM, FDM and GDM can be obtained by taking the means of ADMi, FDMi and GDMi, respectively.

Table 1: FSV interpretation scale [2].

FSV value (quantitative)	FSV interpretation (quanlitative)
Less than 0.1	Excellent
Between 0.1 and 0.2	Very Good
Between 0.2 and 0.4	Good
Between 0.4 and 0.8	Fair
Between 0.8 and 1.6	Poor
Greater than 1.6	Very Poor

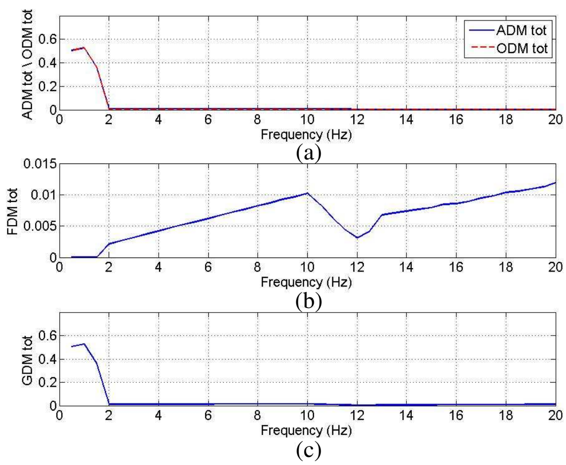


Figure 1: FSV results in different f.

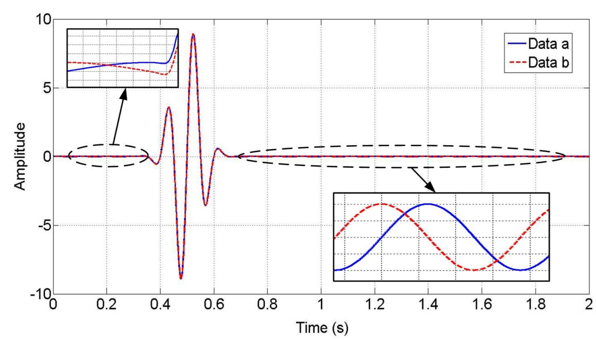


Figure 2: Data for comparison ($f = 1\,\text{Hz}$).

3. THE SPECIAL SITUATION

This section illustrates the special situation mentioned in Section 1 by means of a group of synthetic data sets. One of the parameters of the synthetic data sets is set to variable for the convenience of sensitivity analysis of the FSV method.

The situation could be demonstrated through the following data sets:

$$a(n) = 10\exp\left\{-\frac{(n-500)^2}{2\sigma^2}\right\}\sin(20\pi n\Delta t) + 0.01\sin(2\pi f n\Delta t)$$

$$b(n) = 10\exp\left\{-\frac{(n-500)^2}{2\sigma^2}\right\}\sin(20\pi n\Delta t) + 0.01\sin\left(2\pi f n\Delta t + \frac{\pi}{2}\right)$$

where $\sigma = 40$, $\Delta t = 0.001\,\text{s}$, f varies between 0.5 Hz and 20 Hz, and the length of time is set to 2 s. It can be seen that both of the data sets comprise two parts: transient part (Gaussian pulse) and non-transient part (sine). The discrepancy between the data sets solely exists in the latter part and its amplitude can be ignored compared with that of the former part.

It is well known that the FSV approach is to decompose the original pair of data sets into DC component, low-frequency component and high-frequency component. In this case study, these three components are the frequency below 2 Hz (DC), 2–12 Hz (low-frequency) and frequency above 12 Hz (high-frequency), respectively. With the increase in frequency, f, the non-transient part of the data sets a and b is identified as one of the three components above in order.

The values of ODM, ADM, FDM and GDM of the data sets a and b in different f are shown in Figure 1. It can be seen that there is disagreement between the values of GDM as may be reasonably expected from visual assessment in the DC frequency domain (0–2 Hz). For instance, the discrepancy between data sets a and b when f is 1 Hz can hardly be distinguished by engineers, as shown in Figure 2. So the FSV result should be 'Excellent', or at least 'Good'. However, the overall GDM of Figure 2 is 0.52 (Fair).

It can be found from Figure 1 that this discrepancy is caused by the abnormity of the ODM.

According to the FSV method, the ODM is given in Equation (1).

$$\mathrm{ODM}(n) = \left|\frac{\chi}{\delta}\right| \exp\left\{\left|\frac{\chi}{\delta}\right|\right\} \tag{1}$$

where

$$\chi = (|\mathrm{DC}_1(n)| - |\mathrm{DC}_2(n)|)$$
$$\delta = \frac{1}{N}\sum_{i=1}^{N}(|\mathrm{DC}_1(i)| + |\mathrm{DC}_2(i)|)$$

where n is the nth data point, $\mathrm{DC}_{\{1,2\}}$ are the DC components of datasets a and b. It is clear that the ODM is only related to the DC components. In this case study, the ODM solely reflect the discrepancy of non-transient part when f varies in the DC frequency domain (0–2 Hz), because there is little DC component in the transient part. While the FDM and the rest part of ADM only reflect the agreement of the transient part, so their values are close to zero. Consequently, both the ADM and GDM are dominated by the discrepancy of non-transient part, even though the agreement of the transient part is the uppermost feature of the data sets.

It can also be found that the GDM is not affected when the non-transient part is distinguished as low and high frequency component. Generally, low and high frequency component may represent the mainstream of the data sets, as they are extracted in the point view of energy. Therefore, the non-transient part with smaller energy may be negligible in this situation. In this frequency range, the calculation of FSV mirrors the process of visual assessment exactly.

However, the energy of DC component does not account for a definite proportion of the overall energy of data sets, because it is transformed back from the first 4 points of the spectrum. Therefore, the disagreement between the FSV results and visual assessment in this case study will reoccur when the relative discrepancy of DC components is significant and the energy of that is negligible compared with other components.

4. MODIFICATION TO THE FSV METHOD

The results and discussion above indicate that the visual assessment of DC discrepancy of data sets is vulnerable to the energy of other components, while ODM's algorithm does not take this into account and solely focuses on the discrepancy itself. That is the reason for the disagreement between the FSV results and visual assessment in the special situation. Taking into account the criteria given in [5], a modification is introduced to ODM to improve the performance of the FSV method, which is given in Equation (2). The low-frequency component is incorporated with DC component as the reference (δ_m) of DC discrepancy.

$$\delta_m = \frac{1}{N}\sum_{i=1}^{N}(|\mathrm{DC}_1(i) + \mathrm{Lo}_1(i)| + |\mathrm{DC}_2(i) + \mathrm{Lo}_2(i)|) \tag{2}$$

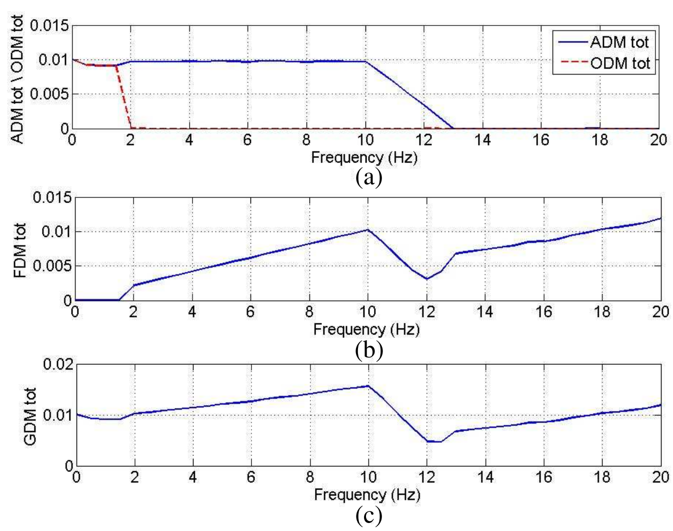


Figure 3: FSV results in different f.

It has been shown that the modified FSV method could improve the FSV results in the situation illustrated in the last section. The abnormity of ODM is eliminated in the region below 2 Hz as shown in Figure 3. Moreover, the range of the fluctuation of GDM is also limited to 'Excellent'. All of these changes make the FSV results become reasonable and in line with the visual assessment.

In order to evaluate the performance of the modification, the modified FSV method is also used to re-compare a set of pairs of data mentioned in [6]. The comparison of GDM confidence histograms obtained by the modified FSV method, original FSV method and visual assessment given by a number of engineers who had a strong background in EMC are shown in Figure 4. It appears that the modification in Equation (2) has slight influence on the FSV result when comparing other data sets.

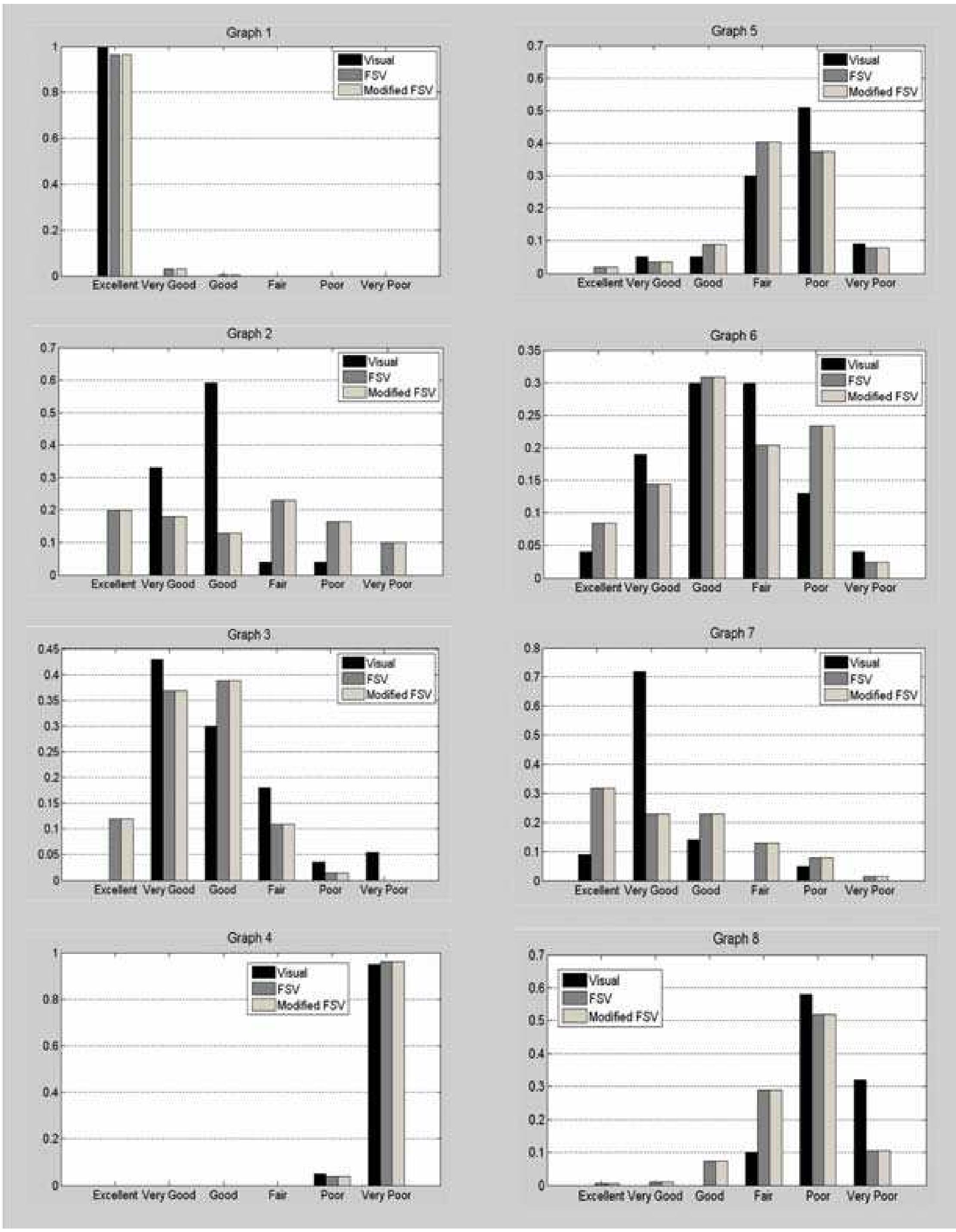


Figure 4: Comparison of the modified and original FSV method, referenced to visual assessment [6].

5. CONCLUSIONS

In this paper, a special situation has been discussed in which the algorithm of ODM is not in line with the process of visual assessment. A modification to ODM has been presented, which incorporates the low-frequency component as part of the reference of DC discrepancy. The study reported here suggests that this modification can eliminated the disagreement between FSV results and visual assessment in the special situation without any pre-processing to original data sets. It is also demonstrated that the change to the method has little effect on the assessment of other data sets.

The research undertaken for this paper is meaningful for the development of multidimensional FSV. Further research and discussion is still necessary, as the evidence of the good performance of the modification presented here is based on this preliminary study.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundational of China under Grant No. 5107-7016.

REFERENCES

1. Martin, A. J. M., "Quantitative data validation," Ph.D. dissertation, De Montfort Univ., Leicester, U.K., 1999.
2. 1597.1 IEEE Standard for Validation of Computational Electromagnetic Computer Modeling and Simulations, Standards Development Committee, USA, 2009.
3. Archambeault, B., A. Duffy, H. Sasse, X. Li, M. Scase, A. Orlandi, and A. Orlandi, "Challenges in developing a multidimensional feature selective validation implementation," *Proc. IEEE Int. Sym. EMC*, Florida, USA, July 2010.
4. Telleria, R. J., F. Silva, A. Orlandi, H. Sasse, and A. Duffy, "Factors influencing the successful validation of transient phenomenon modelling," *Asia-Pacific Sym. EMC*, 338–341, Beijing, China, April 2010,
5. Duffy, A., A. Orlandi, and H. Sasse, "Offset difference measure enhancement for the feature-selective validation method," *IEEE Trans. EMC*, Vol. 50, No. 2, 413–415, 2008.
6. Orlandi, A., A. Duffy, B. Archambeault, G. Antonini, D. E. Coleby, and S. Connor, "Feature selective validation (FSV) for validation of computational electromagnetics (CEM). Part II — Assessment of FSV performance," *IEEE Trans. EMC*, Vol. 48, No. 3, 460–467, 2006.

Supplement on the Non-constancy of Speed of Light in Vacuum for Different Galilean Reference Systems

Namik Yener
Technical Education Faculty, Kocaeli University
Umuttepe Campus, 41380 Izmit, Kocaeli, Turkey

Abstract— The objective of the paper is to make rigorous a development that was presented in a previous article. In particular we want to justify the steps of that development that led to negation of Special Relativity Theory. In that article, a simple medium with loss (medium (I), to which is attached a frame K) was considered interfaced by a perfect electric conductor filling a half space (medium (II), to which is attached a frame K') that is in uniform rectilinear motion with respect to medium (I). It is noted that the result obtained by taking the difference of dispersion relations for incident and reflected waves is actually a simultaneous solution of the two dispersion relations for frequency ω' and wave number k'_1. It is indicated that because of the loss in the medium, the solution set (ω', k'_1) is non-zero and a relation exists to falsify the Special Relativity Theory. When the loss is nullified, the only simultaneous solution of the two dispersion relations becomes the trivial solution $\omega' = k'_1 = 0$ and no relation exists to negate the Special Relativity Theory. The relation used to argue against the Special Relativity Theory involves frequency ω' measured from K', v the speed of K' with respect to K, speed of light in vacuum c and the constitutive parameters of medium (I); ε_1, μ_1, σ_1. It is thus shown that c depends on frequency and by rewriting the said relation employing an expression for ω' in terms of v, c and the incident wave parameters, it is shown at c depends on v too. Alternatively the solution for k'_1 is used in a similar manner and it is demonstrated that this expression can also be used to falsify the Special Relativity Theory.

1. INTRODUCTION

It has been established as the result of a series of articles by the author that when the most general homogeneous time invariant medium has loss, Special Relativity Theory [1] fails to account for this loss and inevitably one encounters a relationship which points to a dependence of the speed of light in vacuum c for a Galilean reference system on frequency or the speed v of Galilean reference system K' with respect to K, and hence on the Galilean reference systems [2–6]. Such relations are derivable using the Lorentz transformation, phase invariance principle, the particular boundary conditions of the problem and the dispersion relation for medium (I) attached to inertial frame K. As mentioned the key element here is the dissipative nature of medium (I). If medium (I) is lossless yet bianisotropic, homogeneous and time-invariant, no such relation emerges and the Special Relativity Theory cannot be contradicted [6].

The objective of this paper is to prove that for the medium (I) of [2], when it has loss, the step of division of the equation obtained through subtracting the dispersion relations for incident and reflected waves observed from K, by $\omega_i - \omega_r$, is justified (see Section 2.1). It was by this division that the relation that indicated a dependence of the speed of light in vacuum for K on the frequency was obtained. It will also be remarked that this relation can be interpreted to conclude that c depends directly on v, the relative speed of K' with respect to K and hence on the Galilean reference systems. This part of the work is contained in Section 2.2.

2. NOTES ON ARTICLE [2]

2.1. Justification of the Result of [2]

In article [2] the electromagnetic system considered was constituted of a medium (I) simple and lossy when observed from frame K attached to it, and a perfectly conducting medium (II) to which is attached frame K', where K' is in uniform rectilinear motion with speed v with respect to K along the Ox axis. The interface of the two media is an infinite plane perpendicular to the Ox axis. To obtain a common solution for the two dispersion equations for incident and reflected waves observed from K, first we transform the two equations into two equations in terms of frequency and wave numbers that are the measured from K' counterparts of the same quantities measured from K. To this end we observe from [2] that due to the phase invariance principle the boundary condition of vanishing tangential electric field component on the interface and under the Lorentz transformation

$$k_i \cos\theta = \alpha(k_1' \cos\theta' - \omega' r/c) \tag{1a}$$

$$k_r \cos\theta_1 = \alpha(k_1' \cos\theta_1' + \omega' r/c) \tag{1b}$$

$$\omega_i = \alpha(\omega' - k_1' rc \cos\theta') \tag{2a}$$

$$\omega_r = \alpha(\omega' + k_1' rc \cos\theta_1') \tag{2b}$$

$$k_i \sin\theta = k_1' \sin\theta' \tag{3a}$$

$$k_r \sin\theta_1 = k_1' \sin\theta_1' \tag{3b}$$

hold, where k and ω are the wave number and frequency θ and θ_1 are the angles of the incident and reflected waves with the Ox axis, $r = -v/c, \alpha = 1/\sqrt{1-r^2}$ with c equal to the speed of light in vacuum, primed and unprimed quantities refer to those measured from K' and K and subscripts i and r stand for incident and reflected waves.

From Snell's law $\theta' = \theta_1'$ follows.

When the dispersion relations for incident and reflected waves when observed from K are expressed in the way described above, i.e., in terms of quantities measured from K' or to put it in another way, if (1), (2) and (3) are utilized in the dispersion relations for incident and reflected waves when observed from K one gets,

$$\alpha^2(k_1' \cos\theta' - \omega' r/c)^2 + (k_1' \sin\theta')^2 = \mu_1\varepsilon_1\alpha^2(\omega' - k_1' rc\cos\theta')^2 + j\sigma_1\mu_1\alpha(\omega' - k_1' rc\cos\theta'), \tag{4a}$$

$$\alpha^2(k_1' \cos\theta' + \omega' r/c)^2 + (k_1' \sin\theta')^2 = \mu_1\varepsilon_1\alpha^2(\omega' + k_1' rc\cos\theta')^2 + j\sigma_1\mu_1\alpha(\omega' + k_1' rc\cos\theta'). \tag{4b}$$

For these two equations other than the trivial solution $\omega' = k_1' = 0$ we have another solution with $\omega' \neq 0$ and $k_1' \neq 0$ where

$$\omega' = \frac{j\sigma_1\mu_1 c^2}{2\alpha(1 - c^2\varepsilon_1\mu_1)}, \tag{5}$$

and

$$k_1' = \frac{j\sigma_1\mu_1 c}{2\alpha(1 - c^2\varepsilon_1\mu_1)}\sqrt{\frac{\alpha^2(1 - c^2\varepsilon_1\mu_1) + 1}{(\alpha^2 - 1)\cos^2\theta' + 1 - r^2c^2\varepsilon_1\mu_1\cos^2\theta'}}. \tag{6}$$

It is easy to verify that the values in (5) and (6) satisfy both equations in (4).

It is now clear that above value for ω' in (5) and the k' value in (6) constitute the simultaneous solution set for the dispersion equations of incident and reflected waves. Furthermore since $\omega_i - \omega_r = -2\alpha k_1' rc\cos\theta'$ by virtue of (2), this nonzero solution for k_1' justifies the division of the subtracted dispersion equations by $\omega_i - \omega_r$ in [2] to obtain (5) above.

If we now consider a lossless simple medium as medium (I), and modify (4), (5) and (6) by setting $\sigma_1 = 0$ we get the trivial solution $\omega' = k_1' = 0$ as the only set satisfying simultaneously the two dispersion relations for incident and reflected waves. Now neither of these equations incorporate loss and we cannot get a relationship as (5) (or as (40) of [2]) or as (6). Because now we have two identities in place of the two dispersion relations like the ones in (4), both of which have the form $0 = 0$ due to imposition of the condition $\omega' = k_1' = 0$. Thus with this electromagnetic system at hand we cannot argue the falsity of the Special Relativity Theory.

2.2. Another Interpretation of the Result of [2]

In [2] it is noted that because (5) has an infinite number of values ω' and the right hand side of (5) does not involve ω', c must necessarily vary with frequency, and the principle of the constancy of speed of light in vacuum must be put aside. Indeed if we use the inverse of (2a) written for ω' and in it substitute for k_i the value from the corresponding dispersion relation observed from K for medium (I), we obtain an expression for ω' in terms of ω_i. It is clear then that for an infinite number of values of ω_i, which is possible, we shall have an infinite number of ω' values as per this relation.

A similar procedure using (3a) and inverse of (1a) written for $k_1' \cos\theta'$ and the substitution for k_i the value from the corresponding dispersion relation observed from K for medium (I), we obtain another expression, this time for k_1' in terms of ω_i. It is clear then that for an infinite number of values of ω_i, which is possible, we shall have an infinite number of k_1' values as per this relation.

It can then be observed that these infinite numbers of (ω', k_1') value pairs obtained, satisfy also (4a) and (4b) simultaneously which are the dispersion relations for incident and reflected waves for medium (I) as observed from K. This statement follows from the fact that for incident and reflected waves, we have separate (ω_i, k_i) and (ω_r, k_r) pairs but a single (ω', k_1') pair. However this requirement entails satisfaction of (5) and (6) by these infinite numbers of (ω', k_1') value pairs which is a property contradicting the Special Relativity Theory, according to which for instance the right hand side of (5) must be a constant changing only with v due to the $1/\alpha$ factor. In other words c must change with changing ω'.

We now also point out that because of the inverse expression for ω', based on (2a) which reads

$$\omega' = \alpha(\omega_i + rck_i \cos\theta) = \alpha(\omega_i - k_i v \cos\theta), \tag{7}$$

(5) can be written also as

$$\frac{j\sigma_1\mu_1 c^2}{2\alpha(1 - c^2\varepsilon_1\mu_1)} = \alpha(\omega_i - k_i v \cos\theta). \tag{8a}$$

Now if we change v, by their definitions r, α will change assuming c is independent of v. Assume that ω_i which must be independent of v because it is a parameter of the incident wave, is also changed to preserve the right hand side of (8a) as a constant. But because $\varepsilon_1, \mu_1, \sigma_1$ are independent of v on the left hand side of (8a), c must change as a function of v while α keeps on changing again on the left hand side. Hence we conclude c must be a function of v in addition to being a function of ω' as stated in [2]. Therefore with this angle of view also we state c depends on the Galilean reference frames and the Special Relativity Theory is false in general.

It is worthy of note that just like Equation (5), (6) can also be used in a way similar to (5) to argue the falsity of Special Relativity Theory. Indeed, (6) can be written as

$$2\alpha(1-c^2\varepsilon_1\mu_1)\sqrt{r^2(\alpha^2-c^2\varepsilon_1\mu_1)+1+\tan^2\theta'}\alpha(k_i\cos\theta+\omega_i r/c)=j\sigma_1\mu_1 c\sqrt{\alpha^2(1-c^2\varepsilon_1\mu_1)+1} \tag{8b}$$

if one also notes that $k_1'\cos\theta' = \alpha(k_i\cos\theta + \omega_i r/c)$. Observing that as given in Appendix B of [2], $\tan\theta'$ can be expressed in terms of incident wave parameters, then the right hand side of (8b) will be independent of the incident wave parameters whereas the left hand side will be a function of these parameters and a function of ω_i and θ in particular. Therefore if ω_i is varied and if v is changed so that the effect of the variation of ω_i is nullified and the left hand side remains a constant, the right hand side can also remain constant only if c becomes a function of v. For otherwise because of the term α^2 on the right hand side, the right hand side cannot remain constant. Hence c has to be a function of v, and thus it has to depend on the Galilean reference systems.

3. CONCLUSION

The steps leading to a relationship (namely Equation (5)) which involves frequency ω', the relative speed of Galilean reference systems v and the speed of light in vacuum c but which does not involve the wave number, are made rigorous. These steps were originally reported in [2]. The fact that this relation has no wave number dependence makes it distinct from a dispersion relation, points out to an interdependence of frequency, c and v and hence negates the Special Relativity Theory.

ACKNOWLEDGMENT

The author is indebted to Dr. Celal Ceken for fruitful discussions.

REFERENCES

1. Einstein, A., "On the electrodynamics of moving bodies," *The Principle of Relativity*, Dover, New York, 1952.
2. Yener, N., "On the non-constancy of speed of light in vacuum for different Galilean reference systems," *Journal of Electromagnetic Waves and Applications*, Vol. 21, No. 15, 2241–2255, 2007.
3. Yener, N., "Determination of speeds of light in vacuum for different Galilean reference systems," *PIERS Proceedings*, 115–119, Beijing, China, March 23–27, 2009.
4. Yener, N., "The non-constancy of speed of light in vacuum for different Galilean reference systems (revisited)," *PIERS Proceedings*, 1631–1637, Xi'an, China, March 22–26, 2010.

5. Yener, N., "Non-constancy of speed of light in vacuum for different Galilean reference systems in case of an impulsive plane wave," *PIERS Proceedings*, 1638–1643, Xi'an, China, March 22–26, 2010.
6. Yener, N., "Homogeneous bianisotropic medium, dissipation and the non-constancy of speed of light in vacuum for different Galilean reference systems," *PIERS Online*, Vol. 6, No. 7, 664–668, 2010.

Permittivity of Vacuum and Speed of Light in Vacuum which Vary with Relative Speeds of Media in Uniform Rectilinear Motion with Respect to Each Other

Namik Yener
Technical Education Faculty, Umuttepe Campus, Kocaeli University
Izmit 41380, Kocaeli, Turkey

Abstract— Having determined in a series of articles that the principle of constancy of speed of light in vacuum of the Special Relativity Theory is false and has to be put aside because it cannot account for the loss in one of the two media which are in uniform rectilinear motion with respect to each other, it has become necessary to use different speeds of light in vacuum for different Galilean reference systems. In this paper this is demonstrated again by obtaining of a relation that relates the speed of light in vacuum c, constitutive parameters of the medium, frequency and the speeds of two media in uniform rectilinear motion with respect to each other. This in turn necessitates the revision of the concept of the permittivity of vacuum for different Galilean reference systems. Therefore first we determine the speed of light in vacuum for one of the two moving media, and define the permittivity of vacuum for that medium using the speed of light in vacuum found for that medium. We assume the permeability of vacuum is an invariant quantity given as $\mu_0 = 4\pi \times 10^{-7}$ [H/m]. The same procedure is repeated also for the second medium. The two media which are in relative motion with respect to each other and which are taken up in this paper, are a simple medium with loss and a perfectly conducting medium filling a half space. Their interface is an infinite plane perpendicular to the direction of the uniform rectilinear motion of the second medium which is constituted of the perfectly conducting half space. For simplicity we have assumed the incident wave that impinges on the interface has a wave vector that makes an angle of $\theta = \pi/2$ with the direction of the velocity of the moving medium and hence we can make use of the time dilation formula to transform the frequencies between the Galilean reference systems. The permittivity of vacuum and speed of light in vacuum results obtained are particular to this electromagnetic system.

1. INTRODUCTION

In a series of articles it has been established by the author that when there are two media in uniform rectilinear motion with respect to each other and one of which has dissipation the speeds of light in vacuum for the two media become dependent on the speeds of the media with respect to each other. This fact which negates the Special Relativity Theory has been proved and reported for the dissipative medium as general as a homogeneous and timeinvariant bianisotropic medium [1]. Therefore it has become necessary to redefine the concept of speed of light in vacuum for moving media. The objective of this paper is to achieve this and also redefine the concept of permittivity of vacuum as a parameter that depends on the reference frames in rectilinear and uniform motion while considering permeability of vacuum on the other hand as a universal constant equal to $\mu_0 = 4\pi \times 10^{-7}$ [H/m]. A simple dissipative medium at rest and a perfectly conducting medium in uniform and rectilinear motion constitute the electromagnetic system of this paper. Their interface is an infinite plane perpendicular to the velocity of the second medium [2]. For the two media the mentioned speeds of light in vacuum, and permittivities of vacuum are found in terms of the conductance, frequency, and permeability of the first medium and the relative speeds of the two media.

In this paper we assume that the incident wave vector makes an angle of $\theta = \pi/2$ with the direction of velocity of the rest medium (Ox axis). In this case the Doppler shift formula reduces to the time dilation formula between the two media and the Doppler shift is solely attributable to the factor of $\frac{c'}{c}\frac{1}{\sqrt{1-v_1^2/c^2}}$ that appears in the said formula [3]. In this case speed of light in vacuum and permittivity of vacuum results for medium (I) depend solely on measurements made from the laboratory frame. The same results for medium (II) however require measurement of frequency in the rest frame in addition to the above quantities measured from the laboratory frame.

It is generally desired to make the measurements from the laboratory frame rather than the rest frame. To arrive at results that involve only such quantities, the frequency as it appears from the rest frame must be transformed. This involves the angle the incident wave vector makes with the velocity of the rest frame [2]. I.e., there is implicit dependence of c on θ. By setting $\theta = \pi/2$ we decrease the number of the independent parameters by one. Furthermore if $\theta = \pi/2$ assumption was not made, instead of the linear equation in c^2 that we have in (3) below, we would have a cubic equation in c which would make the problem less amenable to analysis.

2. SPEED OF LIGHT IN VACUUM DEPENDENT ON GALILEAN REFERENCE FRAME

Taking $\varepsilon_c(\omega) = \varepsilon'(\omega) + j\varepsilon''(\omega)$ as the complex permittivity, we must have the condition that $\varepsilon''(\omega) > 0$ holds when $\omega > 0$. This is a physically necessary condition for the passivity of the medium. For reasons of causality $\varepsilon_c(\omega)$, where $\omega = \overline{\omega} + j\tilde{\omega}$ is true, must be analytic in the upper half plane. Indeed for metals at the lower frequencies up to infrared, one has $\varepsilon_c(\omega) = \varepsilon + j\sigma/\omega$ which is analytic in the whole complex ω plane except at $\omega = 0$ [4]. And it is this form of complex permittivity that we shall assume is true in this article. We shall use the nomenclature 'real part of permittivity' for the real part and 'conductivity' for σ.

Using Equations (21a) through (22b) of [2], and taking the difference of squares of the wave numbers for incident and reflected waves after setting $\theta = \pi/2$ one obtains: $k_i^2 - k_r^2 = (k_i \sin\theta)^2 - (k_r \cos\theta_1)^2 - (k_r \sin\theta_1)^2 = -\alpha^2(\omega' r/c' + \omega' r/c')^2 = 4\alpha^2\tilde{\omega}'^2(r/c')^2$ where $\omega' = j\tilde{\omega}'$. Now using time dilation relation $\omega'/\omega_i = \alpha(c'/c)$ in terms of the frequency [3], since we have transverse Doppler effect due to only the factor of $\frac{c'}{c}\frac{1}{\sqrt{1-v_1^2/c^2}}$ [5, 6] because we have set $\theta = \pi/2$, one obtains when the difference of dispersion relations for incident and reflected waves is taken, namely when Equations (26a) and (26b) of [2] are subtracted,

$$4\alpha^2\tilde{\omega}'^2\left(\frac{r}{c'}\right)^2 = \left\{\left(j\frac{\tilde{\omega}'}{\alpha}\frac{c}{c'}\right)^2\left[1-\alpha^4(1+r^2)^2\right]\right\}\varepsilon_1\mu_1 + j\sigma_1\mu_1\left\{\left(j\frac{\tilde{\omega}'}{\alpha}\frac{c}{c'}\right)\left[1-\alpha^2(1+r^2)\right]\right\}. \quad (1)$$

From (1)

$$2\frac{\tilde{\omega}'\alpha}{cc'}(1 - c^2\varepsilon_1\mu_1) = \sigma_1\mu_1 \quad (2)$$

follows which is now only in terms of real quantities as opposed to Equation (40) of [2] which includes imaginary quantities j and $\omega' = j\tilde{\omega}'$. This is in fact the equation used in [2] to negate the Special Relativity Theory. If ω_i is the pure imaginary incident wave frequency observed from K, with $\omega_i = j\tilde{\omega}$ and again using the time dilation relation for the frequency under the $\theta = \pi/2$ constraint, from (2) one gets for the speed of light in vacuum for K

$$c^2 = \frac{2\tilde{\omega}}{(\sigma_1 + 2\tilde{\omega}\varepsilon_1)\mu_1} + \frac{\sigma_1 v_1^2}{\sigma_1 + 2\tilde{\omega}\varepsilon_1}. \quad (3)$$

3. ε_0, PERMITTIVITY OF VACUUM FOR K THAT CHANGES WITH RELATIVE SPEEDS OF K AND K' AND SPEEDS OF LIGHT IN VACUUM FOR K AND K'

The relative real part of permittivity ε_{1r}, the relative magnetic permeability μ_{1r} and the conductivity of medium (I), are the quantities that can be obtained as the results of measurements made from K through comparisons with corresponding values for vacuum [7, 8]. Since medium (I) where these three quantities are measured is at rest with respect to K, they must not change with changing v_1 even though the real part of permittivity, ε_1 may change because of the change in c due to (3). Therefore since $\varepsilon_{1r}\mu_{1r}$ is a constant quantity, by definition

$$c^2\varepsilon_1\mu_1 = \varepsilon_{1r}\mu_{1r}, \quad (4)$$

must also not change with changing v_1. If in (3) we note that $\varepsilon_1 = \varepsilon_{1r}\varepsilon_0$, $\mu_1 = \mu_{1r}\mu_0$ where $\varepsilon_0\mu_0 = 1/c^2$, one can write down

$$\frac{2\tilde{\omega}}{\sigma_1\mu_{1r}\mu_0 + 2\tilde{\omega}\varepsilon_{1r}\varepsilon_0\mu_{1r}\mu_0} + \frac{v_1^2\sigma_1}{\sigma_1 + 2\tilde{\omega}\varepsilon_{1r}\varepsilon_0} = \frac{1}{\varepsilon_0\mu_0}. \quad (5)$$

One should observe that in (5), $\mu_0 = 4\pi \times 10^{-7}$ [H/m] is an invariant quantity and all other quantities in (5) excluding ε_0 are determinable by measurements made from K. Hence we are in the position to compute ε_0 from (5). Indeed (5) will yield:

$$\varepsilon_0 = \frac{\mu_{1r}\sigma_1}{2\tilde{\omega}(1 - \varepsilon_{1r}\mu_{1r}) + v_1^2\sigma_1\mu_{1r}\mu_0}. \quad (6)$$

ε_0 will be dependent on $\tilde{\omega}$ and v_1. Now we can formulate the expression for c^2 in terms of parameters that can all be determined as explained above:

$$c^2 = \frac{1}{\varepsilon_0\mu_0} = \frac{2\tilde{\omega}(1-\varepsilon_{1r}\mu_{1r}) + v_1^2\sigma_1\mu_{1r}\mu_0}{\sigma_1\mu_{1r}\mu_0}. \tag{7}$$

For the particular choice of frequency and medium (I) and hence of $\tilde{\omega}$, σ_1, μ_{1r} and ε_{1r}, (7) gives a hyperbola in terms of variables c and v_1.

Notice that the time dilation relation in terms of the frequency is $\frac{\omega'}{\omega_i} = \alpha(\frac{c'}{c})$. Here when medium (I) is simple but lossy, $c \neq c'$ has been proved in the previous series of articles and measurements that will yield ω' versus ω_i must necessarily accommodate a c'/c ratio that is different from unity. Using this formula for time dilation in terms of the frequency, one can establish the expression for c'^2 also utilizing (7).

Indeed one then has:

$$c'^2 = \left(\frac{\tilde{\omega}'}{\tilde{\omega}}\right)^2 (c^2 - v_1^2) = \left(\frac{\tilde{\omega}'}{\tilde{\omega}}\right)^2 \left[\frac{2\tilde{\omega}(1-\varepsilon_{1r}\mu_{1r})}{\sigma_1\mu_{1r}\mu_0}\right]. \tag{8}$$

In this expression all the quantities on the far right hand side are obtained by measurements. Except for $\tilde{\omega}'$ which is measured from K', all quantities are measured from K. Equation (8) can also be used to obtain v_2, which is the relative speed of K with respect to K', as follows:

$$v_2^2 = v_1^2 \left(\frac{\tilde{\omega}'}{\tilde{\omega}}\right)^2 \left[\frac{2\tilde{\omega}(1-\varepsilon_{1r}\mu_{1r})}{2\tilde{\omega}(1-\varepsilon_{1r}\mu_{1r}) + v_1^2\sigma_1\mu_{1r}\mu_0}\right]. \tag{9}$$

We can also find the expression for ε_0' permittivity of vacuum in frame K' using (8). Indeed if permeability of vacuum is again taken as $\mu_0' = 4\pi \times 10^{-7}$ [H/m] for frame K',

$$\varepsilon_0' = \left(\frac{\tilde{\omega}}{\tilde{\omega}'}\right)^2 \frac{\mu_{1r}\sigma_1}{2\tilde{\omega}(1-\varepsilon_{1r}\mu_{1r})} \tag{10}$$

will be found.

4. CONCLUSION

It has become evident that since the principle of constancy of speed of light for all Galilean reference systems has to be put aside, the concept of permittivity of vacuum and hence the concept of speed of light in vacuum have to be structured as per Equations (6) and (7). Thus we have to define a new permittivity of vacuum and hence a new speed of light in vacuum in order to make Maxwell's equations compatible with the principle of relativity. This problem has been attempted to be solved. For the case of an incident wave vector with an angle $\theta = \pi/2$ with the velocity direction of the uniform rectilinear motion of one of the moving media, the results have been obtained for both Galilean reference systems.

REFERENCES

1. Yener, N., "Homogeneous bianisotropic medium, dissipation and the non-constancy of speed of light in vacuum for different Galilean reference systems," *PIERS Online*, Vol. 6, No. 7, 664-668, 2010.
2. Yener, N., "On the non-constancy of speed of light in vacuum for different Galilean reference systems," *Journal of Electromagnenic Waves and Applications*, Vol. 21, No. 15, 2241–2255, 2007.
3. Yener, N., "Some consequences of the non-constancy of speed of light in vacuum for different Galilean reference systems," *PIERS Online*, Vol. 6, No. 7, 659–663, 2010.
4. Landau, L. D. and E. M. Lifschitz, *Electrodynamics of Continuous Media*, Pergamon, New York, 1984.
5. Jackson, J. D., *Classical Electrodynamics*, John Wiley & Sons, New York, 1975.
6. Yuan, Z. Z., *Special Relativity and Its Experimental Foundations*, World Scientific, Singapore, 1997.

7. Ocera, A., et al., "A four port structure for comparative measurement of complex permittivity and moisture," *IMTC*, 191–194, Sorrento, Italy, April 24–27, 2006.
8. Shenhui, J., et al., "Measurement of electromagnetic properties of materials using transmission/reflection method in coaxial line," *Asia-Pacific Conference on Environmental Electromagnetics*, Hangzhou, China, Nov. 4–7, 2003.

Reconstruction of Tumors in Human Livers by Magnetic Resonance Imaging

J. Mikulka[1], **E. Gescheidtova**[1], **P. Fiala**[1], **and K. Bartusek**[2]

[1]Department of Theoretical and Experimental Electrical Engineering, Brno University of Technology
Kolejni 2906/4, Brno 612 00, Czech Republic
[2]Institute of Scientific Instruments, Academy of Sciences of the Czech Republic
Kralovopolska 147, Brno 612 64, Czech Republic

Abstract— This article deals with three-dimensional reconstruction and visualization of segmented images of human liver with tumors. The source images were observed by standard hospital MR tomography equipment. There were acquired many slices of the liver area in all three planes. There were found the segmentation methods based on level set approach with many advantages. The resultant boundaries found by these methods are closed curves and bounds the whole tumor. The segmentation consists of automated processing of each slice. The found boundaries are consequently used for 3D image creation and the liver tumor is visualized in Matlab environment. The volume of the tumor is calculated and compared with the results of volume evaluation after manual boundary tracing. The algorithms are prepared for implementation to C++ application for medical usage.

1. INTRODUCTION

The aim of this work is 3D model creation of tumors in the human liver (Fig. 1). It is useful for diagnosis and decision about the treatment success. In traditional way it consists of many steps of MR image processing—preprocessing the images (noise suppression, contrast and edge enhancement), edge analysis and segmentation, segmented images processing. The goal of the proposed processing method is elimination of needful steps of processing and its simplification. The modern segmentation method was found and tested. The MR slices with visible tumors in human livers were segmented by edge-based level set segmentation method [1, 2]. This level set approach gives very good results in segmentation of noised MR images with low contrast and smooth edges so that it is not necessary to preprocess the image before the segmentation of image. The second step is calculation of the tumor volume from known pixel dimensions. The number of pixels in each bounded tumor will be numbered and multiplied by the slice thickness. We get the resultant volume of tumor by normalizing the number of all tumor subarea pixels with the one pixel dimensions. The calculated volume of tumor is compared with manual volume measurement.

This article follows the research of human liver segmentation described in [3], where the similar segmentation method was used but the whole boundaries of tumors havent been found. In this article such a method which can solve this problem is proposed and described.

The mentioned level set methods known as active contours are based on partial differential equation solution. The deformable model of active contour with initial shape and location in the image and by iterative solution of equation(s) the contour can change own shape, topology and location is defined. In the steady state the active contour bounds the found objects in their real boundaries. The problem of searching the tumor boundaries in the image is in unclosed area of the tumor. The tumors are sited in many cases on the edge of liver, so the boundary of tumor

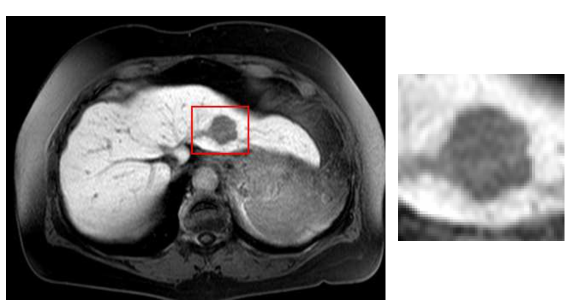


Figure 1: Example of one slice of the human liver (left), detail of the liver tumor (right).

disappears in the image background. The solution is just in use the edge-based level set approach, which can complete the real boundary of tumor, as well.

2. DEFORMABLE MODEL

The edge-based segmentation is described by this energy functional [1]:

$$F(\phi) = \lambda \int_{\Omega} g\delta(\phi)|\nabla\phi|\mathrm{d}x\mathrm{d}y + v \int_{\Omega} gH(-\phi)\mathrm{d}x\mathrm{d}y, \tag{1}$$

where the first term means the length of the zero level curve of Φ (level set distance function) and the second term is called weighted area of Ω_{Φ}^{-}. λ and v are the weighted coefficients of the mentioned terms, $\delta(\phi)$ is the Dirac function and H is the Heaviside function. The g function is the edge indicator defined by

$$g = \frac{1}{1 + |\nabla G_{\sigma} * I|^2}, \tag{2}$$

where I is the original image and G_{σ} is the Gaussian kernel with standard deviation σ. By calculus of variation, the first variation of the functional in (2) can be written as

$$\frac{d\phi}{dt} = \mu\left[\Delta\phi - \operatorname{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right] + \lambda\delta(\phi)\operatorname{div}\left(g\frac{\nabla\phi}{|\nabla\phi|}\right) + vg\delta(\phi). \tag{3}$$

This gradient flow is the evolution equation of the level set function Φ. The second and third term in the Equation (3) correspond to the length and area energy functional. The first term penalizes the deviation of the level set function from a signed distance function during its evolution.

3. SEGMENTATION RESULTS

The results of direct segmentation of images without any kind of the preprocessing are shown in Fig. 2. The slices of the liver were segmented by the edge-based level set method with polygonal initial contour inside the tumor slices. The segmentation process was done in about 5 seconds (AMD 1600XP, 768MB RAM).

4. VOLUMETRY RESULTS

All the images were segmented and the area of each tumor slices have been calculated. The comparison of volume assessment by manual tracing and automated segmentation is shown in the Table 1.

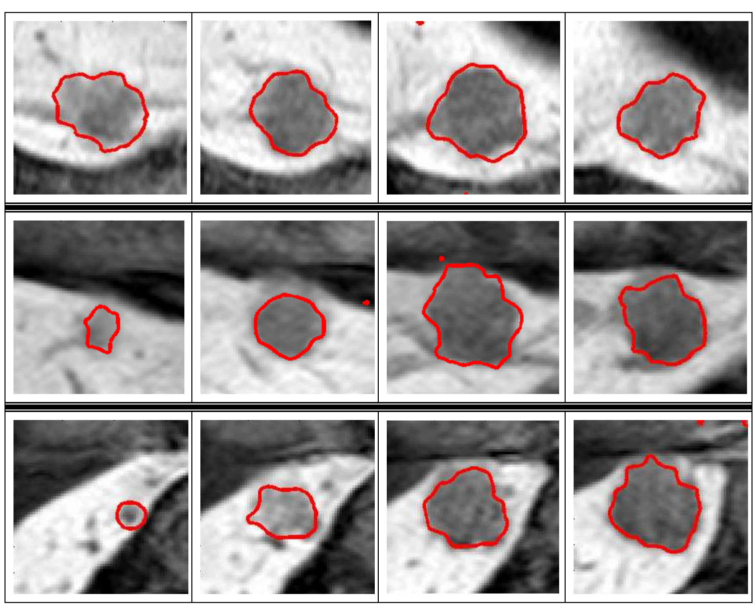


Figure 2: Segmentation results of the edge-based level set method — the red contours bound the liver tumor. From the top: sagittal plane, coronal plane, transverse plane.

Table 1: Segmentation results of the edge-based level set method.

		Plane		
		Sagittal	Coronal	Transverse
Sum of pixels [−]		3147	2631	3171
Pixel dimension [mm]		0.811	0.777	0.811
Total area [mm^2]		2069.848	1588.411	2085.633
Slice thickness [mm]		7	7	7
Tumor volume [cm^3]	Automatic segmentation	14.49	11.12	14.60
	Manual tracing	9.70	10.00	10.10
Approximate time of processing [s]	Automatic segmentation	20	25	30
	Manual tracing	80	100	120

The summary table is shown in Table 1 of the processing procedure. In the first step the sum of pixel in the tumor was calculated. The number of pixels was multiplied by pixel dimensions to obtain the total area of tumor. With known slice thickness we can calculate the volume of tumor. As we can see by the manual tracing the results of volume calculation are very similar in each body plane. The differences between volumes calculated in different body plane occur in automatic tumor segmentation but the time of automatic processing of tumor extraction is much faster.

5. CONCLUSIONS

The paper presents segmentation of human liver slices in three body planes. There were overall 15 MR slices with visible human liver and all the images were automatically successfully segmented by edge-based level set active contour method. The volume of tumor was calculated and it was compared with the manual tracing. The future work will be aimed at improving of the boundary of human liver tumors determination to get more precise results of volume. The algorithm will be implemented to open-source ImageJ software as a Java PlugIn for monitoring the development of tumor growth and three-dimensional reconstruction and visualization.

ACKNOWLEDGMENT

This work was supported within the framework of the research plan MSM 0021630513 and projects ME 10123 and GA 102/09/0314.

REFERENCES

1. Li, C., C. Xu, C. Gui, and M. D. Fox, "Level set evolution without re-initialization: A new variational formulation," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR05*, 430–436, D.C., U.S.A., 2005, ISBN 0-7695-2372-2.
2. Aubert, G. and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, 2nd Edition, Springer Science+Business Media, L.L.C., New York, 2006, ISBN 0-387-32200-0.
3. Mikulka, J., E. Gescheidtová, and K. Bartusek, "Processing of MR slices of human liver for volumetry," *PIERS Proceedings*, 202–204, Xi'an, China, March 22–26, 2010.

C-ring Metamaterial in Close Field

R. Kubásek, P. Drexler, P. Fiala, and K. Bartušek
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejní 2906/4, Brno 612 00, Czech Republic

Abstract— The article brings comparison of typical harmonic analysis with noise spectroscopy approach in metamaterials study. The measurement was performed in close electromagnetic field, where difficulty of harmonic analysis must be eliminated by shielding techniques. In other hand, analysis by noise spectroscopy is not deformed by stationary wave.

1. INTRODUCTION

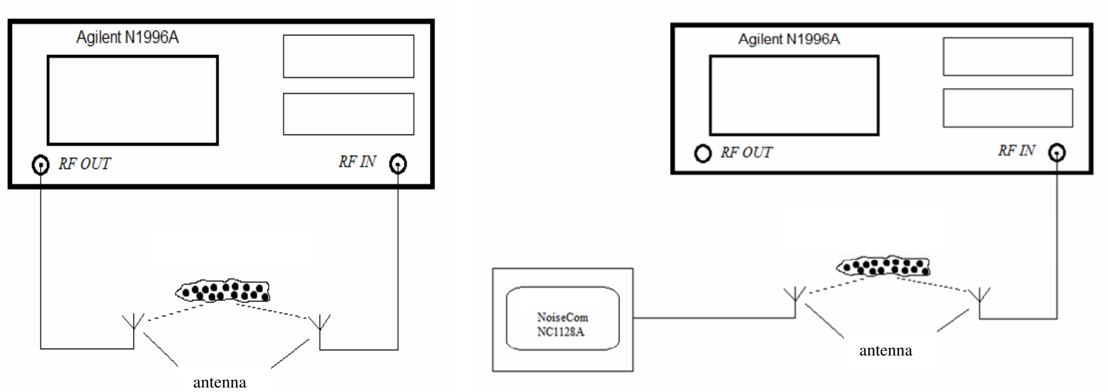
In the complex investigation of material structures for the micro-wave application (tensor and composite character), the properties of materials are studied by means of the typical single-frequency methods, which bring certain difficulties in the process [1]. In boundary changes with a size close to the wave-length wrong information concerning the examined objects occur [2, 3]. One of the possible ways of suppressing the negative sources of signals consists in the use of wide-band signals like white noise, and in the research into the problem of absorption in the examined material [4]. These methods require a source of noise, a receiving and a transmitting antenna, and A/D conversion featuring a large bandwidth; for our purposes, the bandwidth ranged between 0 Hz and 10 GHz.

2. METHOD OF MEASUREMENT

Two methods for frequency domain analysis are present, as shown in Fig. 1. The first is typical harmonic radiation consequently followed by heterodyne spectral analysis. The second approach consists of noise radiation of examined material performed continuously. Spectral analysis is performed any time. The stochastic parameters of noise signal must be satisfied.

The purpose of the transmitting antenna connected to the noise generator output is to form an electromagnetic wave. As a matter of fact, in the field of noise we indeed have to consider a whole spectrum of electromagnetic waves, and it is not possible to define the antenna proximity area. In addition to this, most principles or rules related to the configuration of antennas have to be regarded as void here. The electromagnetic wave is let to impinge on the investigated material and the reflected or partially absorbed wave is then received through the receiving antenna, to the output of which an oscilloscope has been connected. This type of measurement configuration can be seen in Fig. 2(a).

Both antennas ought to feature a large bandwidth with, if possible, constant amplitude and defined radiaton pattern. In this respect, let us mention the fact that there exist approaches to the design of antennas that come close to the broadband requirements of noise spectroscopy. Suitable solutions include, for example, the spiral fractal antenna or the planar log-periodic antenna. The designed planar log-periodic antenna is applicable for transmission within the frequency range of



(a) (b)

Figure 1: (a) Harmonic analysis, (b) noise examination approach.

between100 MHz and 10 GHz; its real characteristics or qualities depend heavily on the quality of the design practical implementation. Fig. 2(b) shows the realization of a planar log-periodic antenna. The numerical design was performed for currently available materials and its evaluation exhibited the undulating module frequency characteristics. The antenna realization experiments using the PCB showed, above all, troubled transmission at higher frequencies from 2 GHz and problematic modification of the feeder. Other antenna designs are directed towards applying the fractal spiral version in Fig. 2(c).

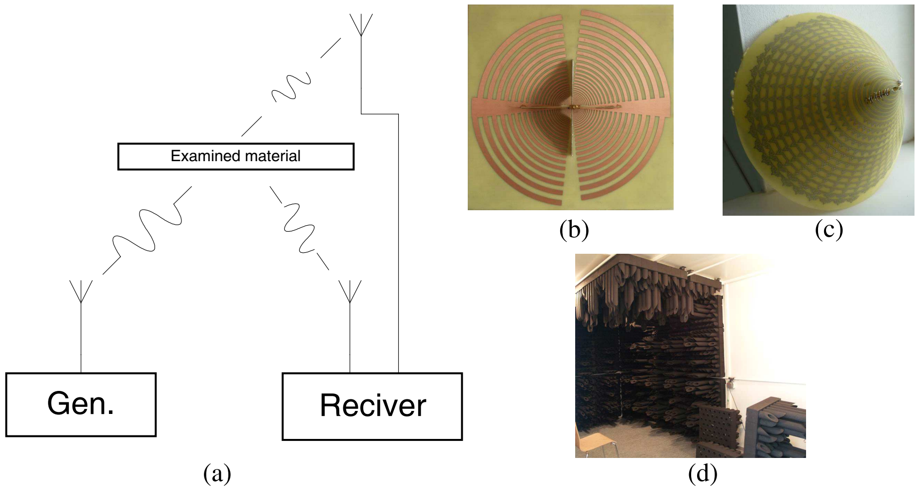


Figure 2: (a) The experiment configuration, (b) planar log-periodic antenna, (c) spiral fratal antenna, (d) Faraday cage.

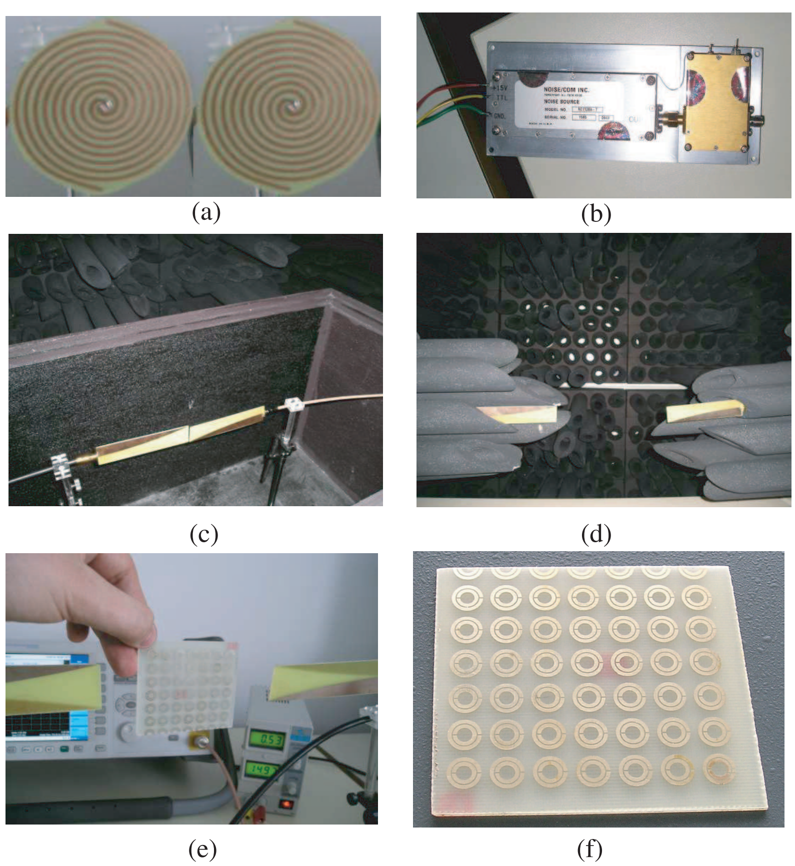


Figure 3: (a) Spiral antenna, (b) noise generator, (c) Vivaldi antenna surrounded by shielding material, (d) Vivaldi antenna in shield box, (e) noise measurement outside the faraday cage, (f) metamaterial sample.

The experiments and spectroscopy tests were performed in an anechoic laboratory. We have selected a system of compelementing the Faraday cage shielding with absorbers of electromagnetic waves. The absorbers were designed for the range of 100 MHz -10 GHz with damping below 35 dB. Thanks to the shielded and separated chamber, the external environment should not affect the internal part of measurement, and the complemented absorbers enable us to lower foreign signals to a level below -60 dB/mW. Thus, the measurement were not be affected by the outside environment of external electromagnetic sources like mobile phone and Wifi network signals or stationary waves and reflection within the Faraday cage.

Figure 3(a) shows spiral antennas, whose frequency characteristics is shown in Fig. 4(a). Spiral antenna works approximately from 2 GHz above. Fig. 3(b) shows noise generator NOISECOM NC1128A (10 MHz to 10 GHz, $P_{out} = 0$ dBm). Vivaldi antenna with shielding material can be seen in Figs. 3(c) and 3(d). The measurement using noise signal is static wave free and can be perform without shielding, Fig. 3(e). This statement was proofed in experimental part. Fig. 3(f) shows C-ring metamaterial sample resonating on 3.70 GHz, also other sample resonating on 3.59 GHz was used.

3. EXPERIMENTS RESULTS

Firs experiment results for harmonic radiation, see Fig. 1(a), are shown in Fig. 4. Figs. 4(a) and 4(b) show spectral characteristic of spiral and Vivaldi antennas. Vivaldi antennas has wider bandwidth and higher gain, but also higher ripple in useful frequency band. Figs. 4(c) and 4(d) show spectra with metamaterial in direct way of radiation between antennas. The higher attenuation of examined metamaterial is reached at its resonating frequency in 10 MHz accuracy. The bandwidth of stopband corresponds to quality of manufactured metamaterial.

Second measurement was performed with noise radiation of metamaterial, see Fig. 1(a). The spectral characteristic of spiral and Vivaldi antennas are shown in Fig. 5(a) and 5(b). Spectral characteristic are equivalent with the previous one. The noise source has $P_{out} = 0$ dBm, that corresponds to approximately -30 dbm/Hz in power spectral density. Therefore spectral characteristic

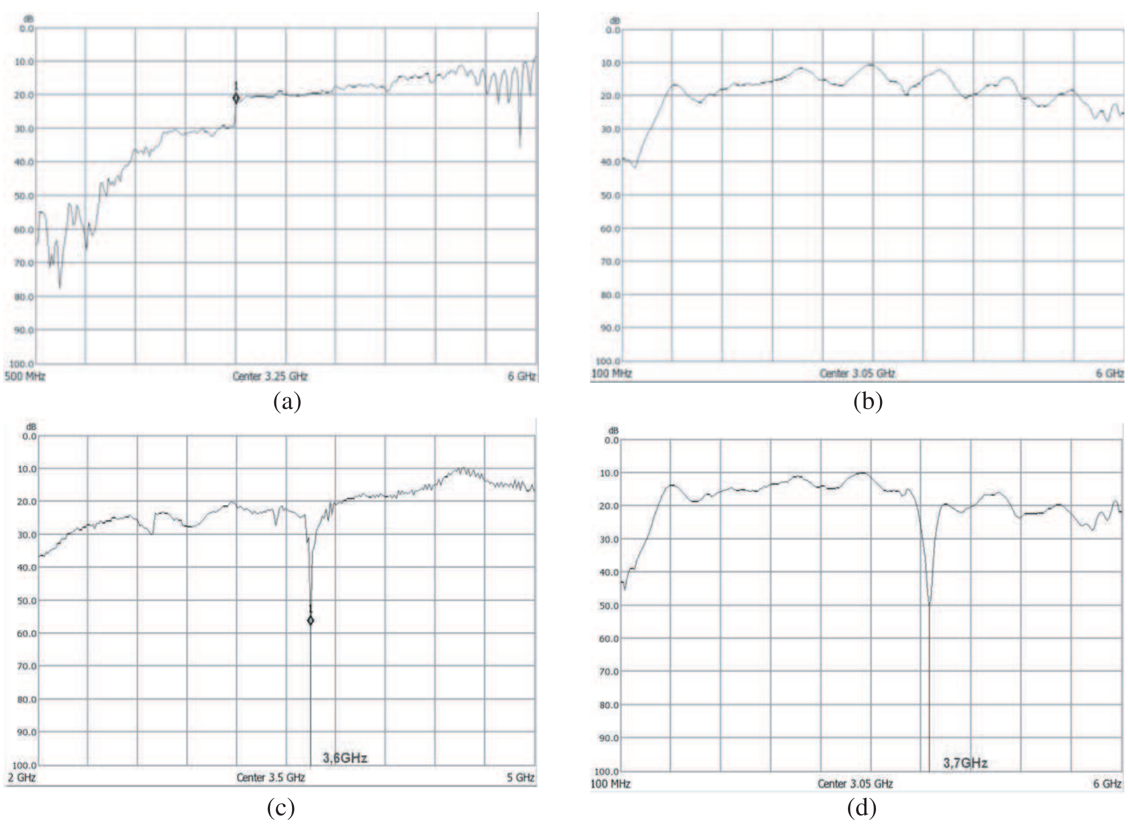
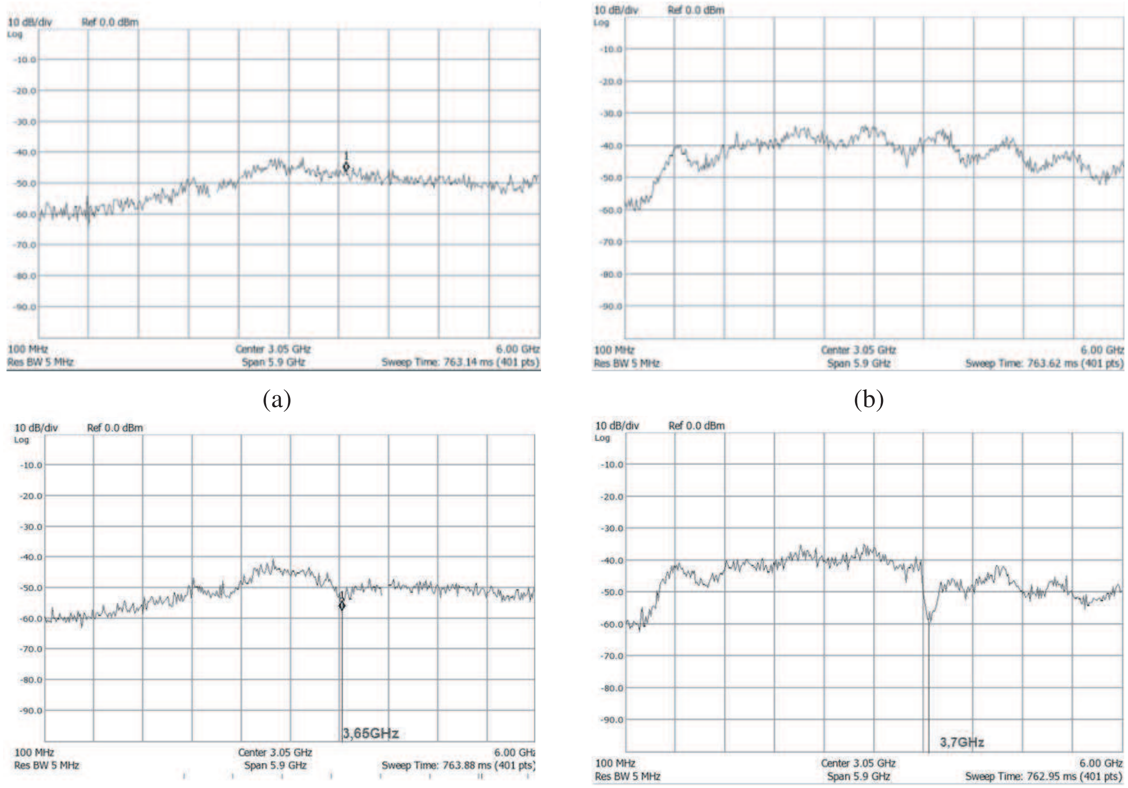


Figure 4: (a) Spectral characteristic of spiral antennas, (b) spectral characteristic of Vivaldi antennas, (c) frequency response with metamaterial and spiral antennas, (d) frequency response with metamaterial and spiral antennas.



(a) (b) (c) (d)

Figure 5: (a) Spectral characteristic of spiral antennas, (b) spectral characteristic of Vivaldi antennas, (c) frequency response with metamaterial and spiral antennas, (d) frequency response with metamaterial and spiral antennas.

by noise source has 30 dB lower gain. The spectra with metamaterial in direct way of radiation between antennas are shown in Figs. 5(c) and 5(d). The noise background of spectral analyzer is to close to be seen the same attenuation at resonance frequency. This is the disadvantage of using noise source radiation against the harmonic analysis. This disadvantage can be suppressed by using additional amplifiers at receiver point. Theoretically, in noise analysis, the static wave can be made. All reflections are adding to original signal in stochastic way. The result is again stochastic noise with higher power. Paradoxically, any reflection can increase quality of measurement. We tested several measurement setups, with and without shielding. The spectral characteristic was very similar. That proofed our hypothesis.

4. CONCLUSION

Article presents and compares measurements with harmonic and noise source for radiation of c-ring metamaterial. Spectral characteristic can be considered as comparable. The higher advantage of noise approach is absence of stationary waves. Lower dynamic of the measurement is disadvantage, mainly for time characteristic measurement. The dynamic of oscilloscopes is too small for wide band noise measurement.

ACKNOWLEDGMENT

This work was supported by the project of the Grant Agency of the Czech Republic No. 102/09/0314 and ELCOM, No. MSM 0021630513 and project of the BUT Grant Agency FEKT-S-10-13.

REFERENCES

1. Maslovski, S., S. Tretyakov, and P. Alitalo, "Near-field enhancement and imaging in double planar polariton-resonant structures," *J. Appl. Phys.*, Vol. 96, 1293, 2004.
2. Freire, M. and R. Marques, "Near-field imaging in the megahertz range by strongly coupled magnetoinductive surfaces: Experiment and ab initio analysis," *J. Appl. Phys.*, Vol. 100, 63105, 2006.

3. Machac, J., P. Protiva, and J. Zehentner, "Isotropic epsilon-negative particles," *2007 IEEE MTT-S Int. Microwave Symp. Dig.*, TH4D-03, Honolulu, USA, June 2007.
4. Fiala, P. and E. Gescheidtova, "Tuned structures for special THz applications," *Proceedings of Progress In Electromagnetics Research Symposium*, 151–155, Beijing, China, March 23–27, 2009.
5. Protiva, P., J. Mrkvica, and J. Machác, "Universal generator of ultra-wideband pulses," *Radioengineering*, Vol. 17, No. 4, 2008.

Sensors and Experimental Model Development for PD Localization in HV Transformers

P. Drexler, **P. Fiala**, **M. Friedl**, **P. Marcon**, **M. Steinbauer**, **and Z. Szabo**
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejni 2906/4, Brno 612 00, Czech Republic

Abstract— For experimental verification of partial discharge localization method a measurement model setup has been built. The model setup is equipped with antenna sensors, RF signal processing circuitry and data acquisition unit. The model setup is described in the paper as a following step of the research. The result of simulated discharge measurement which has been performed up to the present are presented. The comparison to partial discharge time-domain and spectral characteristic is given also.

1. INTRODUCTION

One of the problematic phenomena in the field of high-voltage technology, is the occurrence of partial discharge (PD). Several other effects have combined with this notion over time [1]. In consequence of these effects there emerge short electromagnetic pulses with a defined and measurable spectrum in the characteristic frequency band [2]. The group of end products attributable to the emergence of interfering signals involves, for example, displacement current in a dielectric, pulse current on the interface between dielectrics, or the dielectric/metal interface owing to high electric field intensity and structure of the dielectric.

In HV and VHV transformers the dielectric is mineral or synthetic oil. Due to the increase in PD activity in transformer oil free atoms of carbon, hydrogen and oxygen develop from hydrocarbons, and there also generates a certain percentage of water and other organic compounds. All of these elements decrease the quality of the dielectric. In addition to that, rapid increase in pulse activity may cause the formation of a hazardous explosive compound of oxygen and hydrogen. Then, this situation may result in explosion and damage to the device.

In order to prevent the transformer failure the observation of pulse activity is necessary. The occurrence of discharge with substantial charge transport level can be localized in critical areas of the transformers. Having a possibility to localize the increased discharge activity in some of the critical areas allows us to undertake precautions in order to avoid the critical transformer failure. The discharge activity localization can be determined on the basis of processing of signals from suitable installed sensors. The evaluation of the discharge location utilizes the model of wave propagation inside the transformer, which has been presented in early work [3].

2. MEASUREMENT MODEL SETUP

For the evaluation of PD detection and localization the measurement model has been designed and built. The model is based on the oil filled metal tank which is box-shaped. The tank contains 230 l of transformer oil. Four inserts are assembled on the wall of the tank. The inserts are in the form of cylinder cavity and they allows to put the antennas inside in order to be able to collect PD radiated RF energy. The oil tank is shown in the left of the Figure 1. In the right of the Figure 1 the antennas positions in the tank with their labels are shown also. The tank is equipped with the metal cover which can be removed to insert additional objects with specified material properties $\varepsilon_{r,n}$, $\mu_{r,n}$ and $\gamma_{r,n}$ inside the tank. The insertion of additional objects allows to observe changes in signal time relations which are caused by the additional reflections on the walls and different propagation velocities in the objects and at their boundaries.

The presence of objects with properties different to the oil environment (middle of the Figure 1) causes the propagation of waves with different wave numbers $\bar{k}_n$ in compare to the primary wave number $\bar{k}_1$.

$$\bar{k}_1 = \sqrt{-j\omega_1\mu_1(\gamma + j\omega_1\varepsilon_1)}; \quad \bar{k}_n = \sqrt{-j\omega_n\mu_n(\gamma + j\omega_n\varepsilon_n)}. \tag{1}$$

Considering the known configuration of the model (or the HV transformer) there is a possibility to determine the probability areas of PD occurence by means of the numerical model evaluation. The example of probabilities $P1, \ldots, Pn$ areas are shown in the right part of Figure 1 for $P1$ and $P2$.

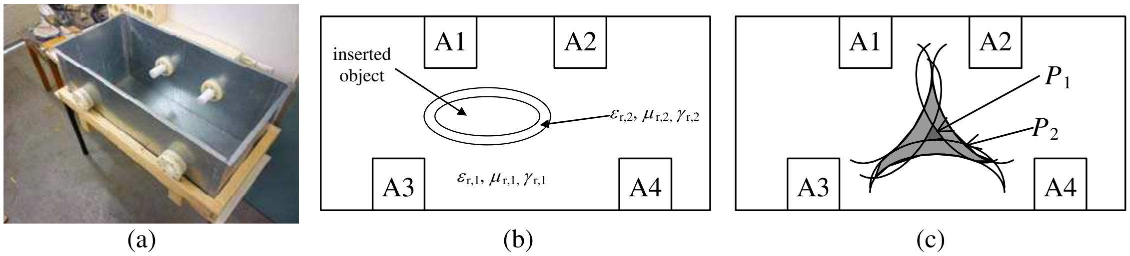


Figure 1. (a) Tank with inserts for antennas assmb, (b) position of antennas with labels, (c) PD occurrence areas.

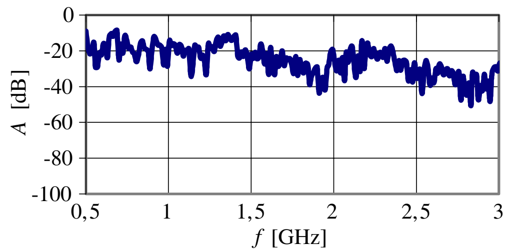


Figure 2. The example of insertion loss frequency characteristic of the measurement model.

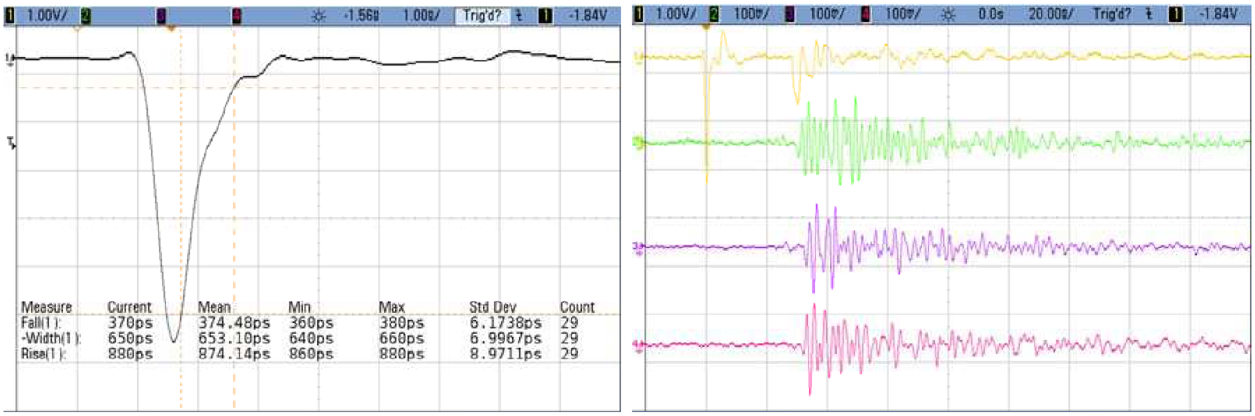
The basic properties of proposed antenna sensor have been presented in [3]. From these sensors the cone-type antenna and the Vivaldi-type antenna has been chosen for further experiments in oil tank. In order to get an idea about the wave propagation inside the oil filled tank with regard to spectral characteristic of this model the measurement of insertion loss has been performed. The measurement has been performed for various position of transmitting (TX) antenna and the receiving (RX) antenna in the inserts. For example, the TX antenna was placed in insert A1 and the insertion loss has been measured for the position of RX antenna in insert A2, A3 and A4. In the next step the TX antenna has been moved into the insert A2 etc.. Special attention was given to the Vivaldi antenna. The Vivaldi-type antenna belongs to linear polarization antennas. Hence a various mutual positions of antennas polarization planes were examined. The polarization planes have been placing to horizontal and to vertical positions during the experiment. The results of this measurement show that there is no distinctive effect of the TX and RX antennas position, eventually the mutual polarizations orientation, on spectral characteristic of the tank oil environment insertion loss. The character of the results is probably caused by the presence of multiple wave reflections on the tank walls. The example of the insertion loss frequency characteristic for the TX cone-type antenna in position A1 and the RX cone-type antenna in position A4 is shown in Figure 2. It should be noted that the insertion loss characteristic contains the frequency characteristics of both antennas also.

3. TIME DOMAIN MEASUREMENT

We can localize the area of the PD activity by evaluating the time relation of the signal received on antennas. The verification of the time-of-arrival difference in different receiving channels has been made by means of short testing pulse. The testing pulse was applied on the TX antenna and the signals on the remaining RX antennas were captured. The positions of the antennas in the measurement model have been designed in the way that the distances of three RX antennas to the TX antenna are mutually different. TX antenna has been placed in the position A1. RX antennas have occupied positions A2, A3 and A4 in Figure 1.

A negative going short pulse of electrical voltage with the peak level of approx. $-6\,\mathrm{V}$ was fed in TX antenna in position A1. The pulse was supplied by the pulse generator based on the avalanche effect in the RF transistor in the transition from saturation into the off-state. The pulse fall time is $t_f = 370\,\mathrm{ps}$, rise time is $t_r = 880\,\mathrm{ps}$ and the pulse width is $t_w = 650\,\mathrm{ps}$. The pulse waveform is shown in the left part of the Figure 3.

In order to preserve the time-of-arrival differences the cables between the RX antennas and the acquiring device were of the same length. The signals on the receiving antennas were captured by the digital storage oscilloscope with the sampling rate 2 GSa/s and with the analog bandwidth of 1 GHz. The captured waveforms are shown in the right part of the Figure 4.



(a) (b)

Figure 3. The pulse waveform for the time-of-arrival difference verification (a) and captured waveforms of signals at the RX antennas (b).

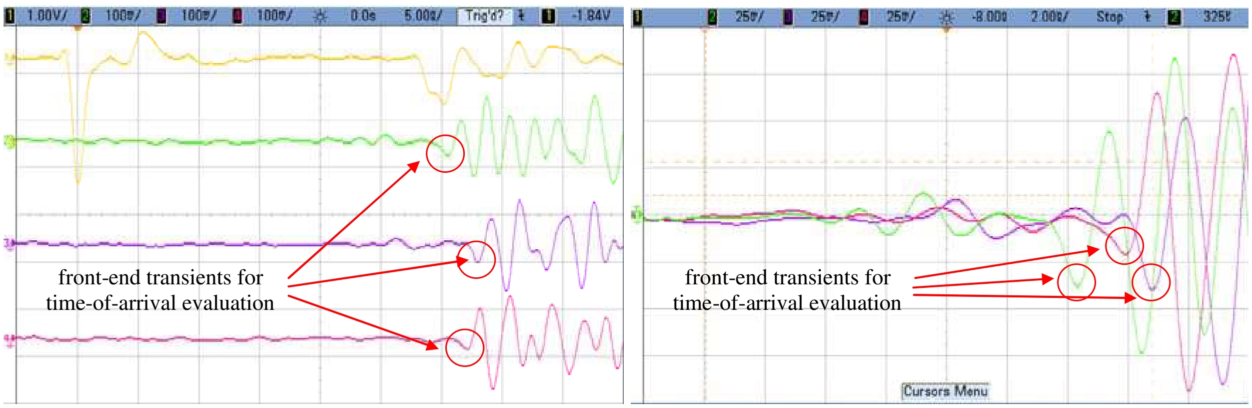


Figure 4. Detail of front-end transients for time-of-arrival evaluation.

Figure 5.

In the channel 1 (the top waveform) a waveform of excitation pulse is shown. The signals on the RX antennas are shown in the rest of the channels. Due to the propagation in cavity with conducting walls a considerable amount of multiple reflections can be observed in the RX channels. However, for the signals time relations evaluation the occurrence of the first remarkable transient has to be watched. The details of the front-end transients in the RX channels are shown in Figure 5. The received pulse signal is represented by the negative going transients, whose time of occurrence difference is in order of nanoseconds. The time difference is a function of the propagation velocity which can be calculated for the oil environment with defined dielectric constant. The sequence of the front-end transients is in agreement with the distances between the TX antenna and RX antennas.

It should be noted the even in the case of a small dimensions of the measurement model the time differences between the signals are measurable. In the target application the dimensions of the transformer to be watched are much greater which results in greater time differences also. A substantial is the fact the in the real application the direct visibility between the area of PD occurrence and the receiving point is not guaranteed. This fundamental issue causes that remarkable transient can be received after a number of reflections. This can introduce a several delay even if the PD occurrence area is close to the receiving antenna. The solution of the PD occurrence location can be evaluated by means of the numerical model which has been presented in previous work [3]. This effect can be modeled by inserting of additional walls into the oil tank which is the goal of further research.

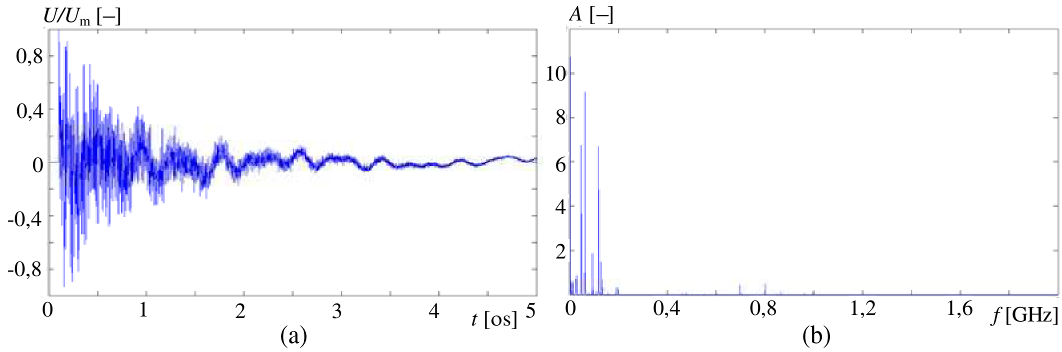


Figure 6. Waveform (a) and spectrum (b) of the PD radiated signal.

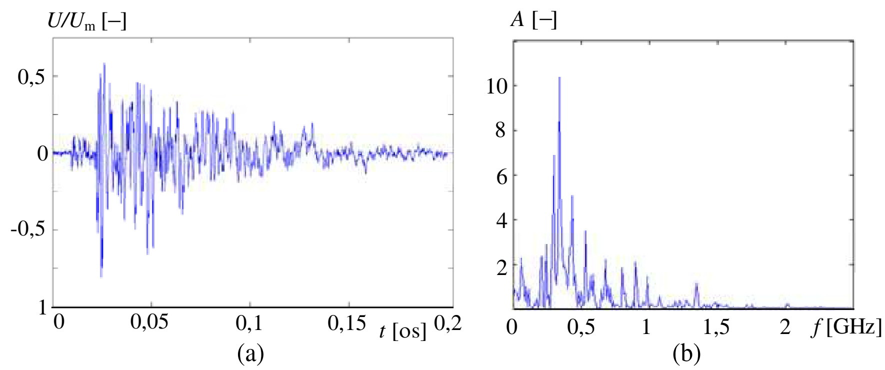


Figure 7. Waveform (a) and spectrum (b) of the spark-gap discharge radiated signal.

4. MEASUREMENT OF SIMULATED PARTIAL DISCHARGE

In order to test the designed detection system in the measurement model a need for suitable discharge source occurs. A spark-gap with low breakdown voltage has been proposed. Due to the low level of charge transport during the PD activity in order of 10^{-10}C the low breakdown voltage is required in order to get the similar charge transport level. The spark-gap consists in two tungsten electrodes of sharp-tip shape to maximize the field intensity in the gap. The gap distance has been set to 0.1 mm. The spark-gap has been placed in the oil tank and the electromagnetic radiation of the discharge has been captured by means of four antenna sensors. The waveform and the spectrum of the spark-gap discharge should approximate the characteristic of the partial discharge. The waveform and the spectrum of the PD radiation is shown in Figure 6. The signal has been obtained by measurement on real high voltage transformer during the PD.

The characteristic obtained by measurement of the spark-gap discharge in the oil tank are shown in Figure 7. The differences of the characteristics in Figure 6 and Figure 7 are obvious. The PD waveform has longer time duration in compare to spark-gap discharge waveform. The next difference is the content of harmonic components. The most of the power is carried by the harmonic components on frequencies in range 10–200 MHz. While the waveform of the spark-gap discharge has shorter time duration and the frequency content is also different. The most of the signal power is carried by frequencies of hundreds of MHz. There components with frequencies up to 2 GHz are present also.

5. CONCLUSIONS

The description of the measurement model for PD detection and localization method evaluation is presented in the paper. The spectral transmission characteristic of the model equipped with antenna sensor is shown further. In order to demonstrate the ability of transients time-of-arrival resolving the time domain measurement has been performed. Simulated discharge is compared to the real PD acquisition. Regarding to differences another discharge source should be intended which is a part of further research. The designed circuitry for RF signal processing is a point of subsequent publication.

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, grant of Czech ministry of industry and trade no. FR-TI1/001, GACR 102/09/0314 and project of the BUT Grant Agency FEKT-S-10-13.

REFERENCES

1. Fiala, P., "Transformer partial discharge modeling, minimal breakdown value set in a critical parts of transformer design, Research report," Laboratory of modeling and optimization fields in electromagnetic systems, FEI VUT and ABB EJF a.s. Brno No. 2/99, 18.3.1999, Brno, Czech Republic, 1999.
2. Sarathi, R., A. J. Reid, and M. D. Judd, "Partial discharge study in transformer oil due to particle movement under DC voltage using the UHF technique," *Electric Power Systems Research*, Vol. 78, 1819–1825, Elsevier, 2008.
3. Fiala, P., T. Jirku, P. Drexler, and P. Dohnal, "Detection of partial discharge inside of HV transformer, modeling, sensors and measurement," *PIERS Proceedings*, 1013–1016, Cambridge, USA, July 5–8, 2010.

Comparison of Different Methods for Measurement of Shielding Fabrics Properties

Z. Szabó
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejní 2906/4, Brno 612 00, Czech Republic

Abstract— Shielding is a very popular method of ensuring of electromagnetic compatibility and for protecting of electronic and electrical equipment and humans against radiated electromagnetic energy. The suitable alternative to the classical shielding materials can be special shielding-fabrics. The main advantages of these fabrics are their low weight, flexibility and their easy processability. The knowledge of shielding effects of different types of material represents a basic prerequisite for further development and implementation of shielding devices. Measuring of shielding and absorption properties of fabric materials is relatively difficult. There were developed many measurement methods that are used in various laboratories to solve these problems. A lot of producers who develop this type of fabrics are searching for relatively simple, not time-consuming and reliable measurement methods for shielding and absorption properties measurement which operate in a wide frequency range. This paper presents a comparison results from two different methods for measurement of shielding and absorption properties of shielding fabrics.

1. INTRODUCTION

Shielding is a very popular method to ensure the electromagnetic compatibility, protection of electronic equipments and human beings against radiated electromagnetic energy. Decrease of radiation disturbances and increase of immunity to electromagnetic fields is obtained with grounded shielding eventually in combination with other suppression components. The shielding is also used to isolate some places from the external source of electromagnetic interference or to prevent radiation electromagnetic disturbance from the internal shielded source. There were used metallic materials with well known electromagnetic qualities for the shielding.

There are more and more used plastic materials for shielding with a conductive coating or with embedded conductors which ensure the shielding flexibility. Recently the researchers attention is focused on even lighter and more flexible materials which are fabrics coated with absorptive film. These materials due to their flexibility and low cost price are promising candidates for instruments and human protection against unwanted electromagnetic radiation effect.

Measuring shielding and absorption qualities of fabrics and plastic materials is relatively difficult. There were developed many measuring methods that are used in various laboratories to solve these problems. A lot of producers who develop this type of fabrics are searching for relatively simple, not time-consuming and reliable measuring methods to measure shielding and absorption properties which operate in a wide frequency range.

2. ELECTROMAGNETIC WAVE ABSORBTION

Attenuation of the electromagnetic energy may be characterised by the shielding effectiveness, SE and the insertion loss IL. The ability of a shield to screen out electromagnetic fields has been quantitatively defined in MIL-STD-285 (June 25, 1956) [1] as an attenuation or the ratio (expressed in dB) of electromagnetic field strengths E_0/E_1 measured without and with the tested material when it separates field source and receptor respectively:

$$SE_{\mathrm{dB}} = 20 \cdot \log \frac{E_0}{E_1}\,[\mathrm{dB}] \tag{1}$$

where E_0 and E_1 are the electric field strength without and with the shield in position, respectively. Also it can be explained as the ratio of transmitted power (p_t) to incident power (p_i) measured in decibels (dB) as follows:

$$SE_{\mathrm{dB}} = 10 \cdot \log \frac{p_t}{p_i}\,[\mathrm{dB}] \tag{2}$$

Alternatively magnetic field strength can be used. It is also known that electric and magnetic field measurements give different results because they depend not only on material performance but also on the wave impedance.

The known methods of measuring the shielding effectiveness of materials are all based on the definition above, no matter which technique is used to generate incident fields or to measure transmitted fields.

Shielding effectiveness measurement results of a practical enclosure are largely influenced by the shape of the enclosure, incidence of the fields and by the functional apertures. Nevertheless, the knowledge of the enclosure material performance is the point of departure for the shielding design.

More recently, other methods have been introduced that use coaxial transmission lines supporting TEM mode propagation: the sample is placed as a shunt across the line and in this way it is exposed to a guided plane wave. Ffor this reason they are also called plane wave measurement methods. Also, using transmission lines instead of antennas makes the simulation of near field, electric and magnetic possible. This method makes the use of two TEM cells coupled with an aperture and will be discussed in the rest of this paper [2, 4].

2.1. Coaxial Holder Method

The SE of the composites was analyzed by the method using coaxial cable which is designed according to ASTM D4935 Standard [3]. This coaxial method is applicable to near-field measurement. The measuring instrument was composed of two parts: spectral analyser (Agilent CSA Spectrum Analyzer N1996A-506 network analyzer) and sample holder (Figure 1). This technique involves determining the shielding effectiveness of a base material (flat panel or coupon) using an insertion-loss method. The flat, thin sample is irradiated with an electromagnetic wave over the frequency range of interest. The method uses a coaxial transmission line with an interrupted inner conductor and a flanged outer conductor. The sample is placed between the flanges in the middle of the cell. A schematic of this instrumental set-up is shown Figure 1 [4, 6].

2.2. Dual Transverse Electromagnetic Cell

The DTEM test consists of interposing a 100×100 mm sheet of the sample-under-test (SUT) in the aperture between two TEM cells, configured so that one cell acts as the transmitter while the other serves as the receiver as shown in Figure 2. Through proper calibration and by alternating the transmit/receive function of both cells, in theory this test allows for separation of "electric" and "magnetic" field insertion loss of the sample between the cells. Small aperture theory is used to explain the coupling between the two cells constituting the DTEM cell. If the aperture is electrically small, then the scattering effect is essentially equivalent to that produced by an appropriate set of dipole moments. These dipole moments may be used to predict the scattered fields, which in the case of the dual TEM cell give a description of the expected aperture coupling. The dipole moments depend on the incident (exciting) fields and the shape, size, and orientation of the aperture. These latter effects are summarized in a quantity termed the aperture polarizability [2, 6].

3. PROJECT AND REALIZATION

The main measurement method in the controlled medium was the coaxial holder method. This method uses a holder transmission line and a vector network analyzer, as is shown in Figure 1. The sample material is placed and fixed in the flanged circular coaxial transmission line holder. Generally, the maximum operating frequency is around 2 GHz. The increasing of the maximum operating frequency determines the decreasing of the flanged coaxial line dimensions and, of course, the sample dimensions. This measuring system is compact and allows automation and data proceeding by computer control. The difficulty of this measurement method arises from the sample preparation. Thus, the dimensions of the sample must be small, especially for higher frequency measurements and the influence of the contact resistance between the sample and the coaxial holder is necessary

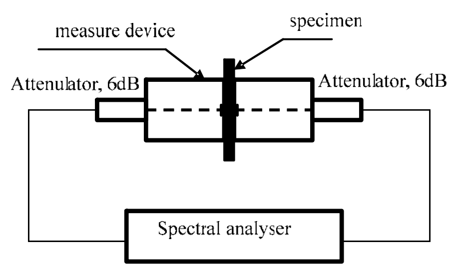


Figure 1: SE measurement device.

to be considered. In this method we used the above mentioned theoretical descriptions. A device for coaxial holder method was made in the laboratory at DTEEE. This device consists of two cylindrical electrodes which are matched to 50 Ω transmission line. To generate and measure the SE the spectral analyzer N1996A-506 from Agilent was used. Before the measurement the device was calibrated to eliminate the device insertion loss and unwanted signals. The calibration sets the device insertion loss to zero. After the sample was inserted between the measuring electrodes on spectral analyzer, the level of insertion loss has changed. Signals were transmitted by coaxial cable. The frequency was scanned within the range of 10 MHz∼1 GHz with 400 data points in both reflection and transmission. The samples SE was calculated from the measurement results [6].

Measurements of SE can be also based on the use of a dual-TEM cell (see Figure 3). A typical TEM (transverse electromagnetic) cell consists of a section of rectangular coaxial transmission line tapered down at each end to match ordinary 50 Ω coaxial line. The TEM cell is well established as a device that creates a known broad-band isolated test field. A dual-TEM cell is then simply a pair of TEM cells with the added feature of an aperture in a shared wall. The aperture transfers power from the driving cell fed at the Port 1 to the receiving cell. The insertion loss provided by putting a sample on the aperture gives an evaluation of the shielding effectiveness of the material (tested at Port 2 or 4) Figure 2.

The measuring place consists of an Agilent spectral analyzer and a DTEM cell. This cell has been designed for measurement with a 50 Ω line and for measurement of thin samples with maximum dimensions of 100 × 100 mm. For electromagnetic field generation and analysis of the resulting signal a spectral analyser Agilent CSA Spectrum Analyzer N1996A-506 (from 100 kHz to 6 GHz) has been used. The apparatus incorporates both the signal generator and the spectrum analyzer. Signals

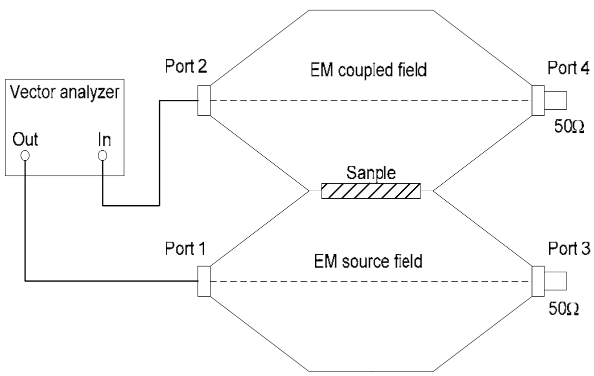


Figure 2: Block diagram of dual-TEM cell. Samples are placed between the two cells.

Figure 3: Dual-TEM cell for SE measurement.

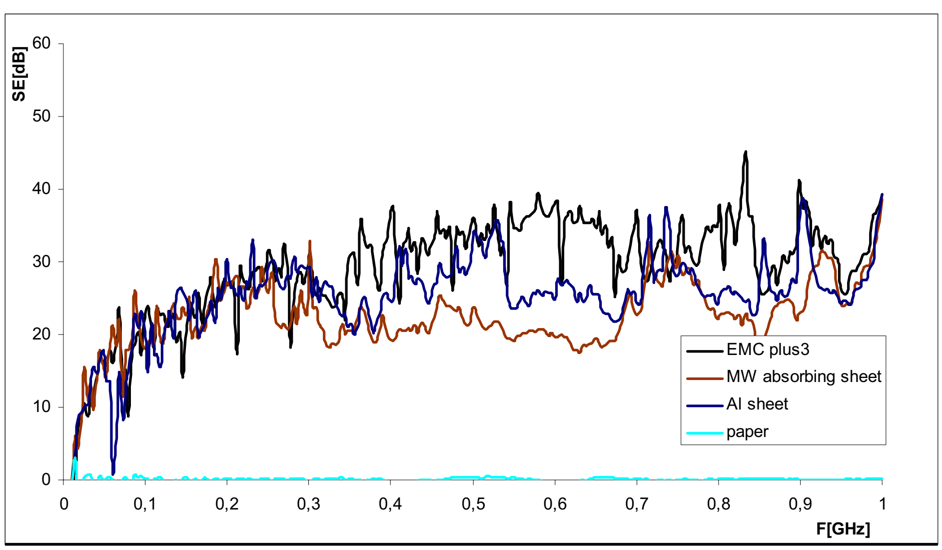


Figure 4: Shielding effectiveness for FlecTron fabric.

were transmitted by coaxial cable. The frequency was scanned within the range of 10 MHz∼1 GHz with 260 data points in both reflection and transmission. The samples SE was calculated from the measurement results.

4. EXPERIMENT RESULTS

The following figures (Figures 5–8) show the measurement results of the shielding effectiveness of various shielding fabrics which were measured by the above described methods. The Figure 4 illustrates SE measured with the coaxial holder method and the Figure 5 illustrates SE measured with the DTEM cell. On both figures we can see that the results from the SE measurement are approximately the same. There are shown few results from the tested materials, e.g., EMC plus3 thin woven shielding fabric with the surface resistance $0.01\,\Omega/\mathrm{m}^2$. The highest attenuation level of this fabric is at 0.5 GHz which is 35 dB in average. If we compare these measurement results with a non-woven fabric made from carbon fibers with a surface conductance of $10\,\Omega/\mathrm{m}^2$ then the EMC plus3 type fabric has slightly better shielding qualities. The shielding effectiveness of fabric made from carbon fibers is in average 15 dB in range from 0.2 to 1 GHz.

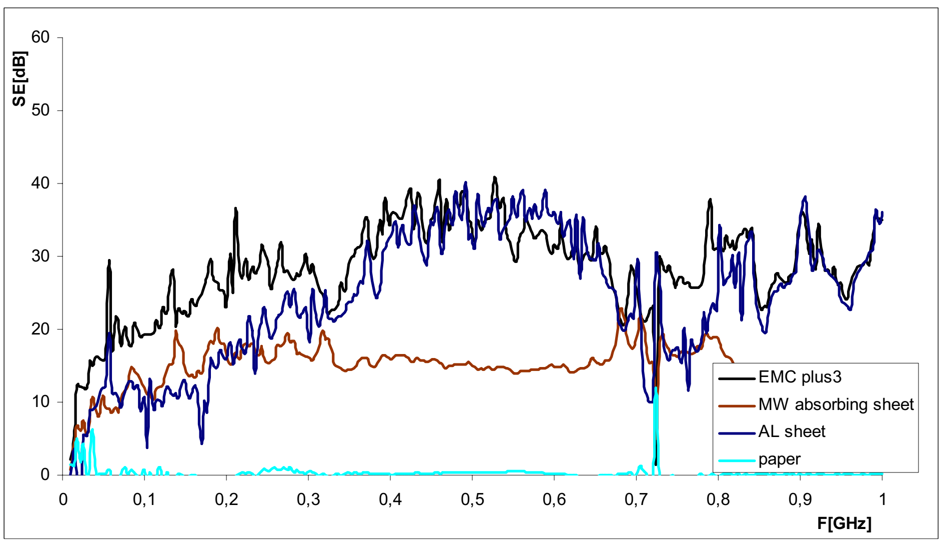


Figure 5: Insertion loss for FlecTron fabric type.

5. CONCLUSION

This research work preformed various measurement methods to compare the difference of shielding effectiveness between chamber structure and construction. The various chambers have different types of electromagnetic field and propagation of EM wave. There were proposed two different methods to describe and measure the fabric absorption quality. To verify the functionality of these methods an experimental measurement on few different types of samples has been made. The results of these experiments were demonstrated on graphic diagrams and they were described.

ACKNOWLEDGMENT

The research described in the paper was financially supported by the research program MSM 0021630516 and research plan MSM 0021630513, Ministry of Defence of the CR, Ministry of Industry and Trade of the CR (Diagnostics of Superfast Objects for Safety Testing, FR-TI1/368), Czech Science Foundation (102/09/0314) and project of the BUT Grant Agency FEKT-S-10-13.

REFERENCES

1. MIL-STD-285, "Method of attenuation measurement for enclosures, electromagnetic shielding, for electronic test purposes," June 25, 1956.
2. Manara, A., "Measurement of material shielding effectiveness using a dual TEM cell and vector network analyzer," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 38, No. 3, Aug. 1996.

3. ASTM, D4935 Standard Test Method for Measuring the Electromagnetic Shielding Effectiveness of Planar Materials, ASTM, 1999, IMEKO, 2007.
4. Valeriu, D., E. Vremera, A. Salceanu, I. Nica, and O. Baltag, "On the characterization of electromagnetic shielding effectiveness of materials".
5. Wilson, P. F. and M. T. Ma, "Techniques for measuring the electromagnetic shielding effectiveness of materials: Part 11-Near-field source simulation," *IEEE Transactions on Electromagnetic Compatibility*, 251–259, Vol. 30, No. 3, Part 2, Aug. 1988.
6. Szabó, Z. and P. Fiala, "Characterisation and testing shielding fabrics," *PIERS Online*, Vol. 5, No. 7, 609–612, 2009.
7. Avloni, J., L. Florio, A. R. Henn, R. Lau, M. Ouyang, and A. Sparavigna, "Electromagnetic shielding with polypyrrole coated fabrics," http://arxiv.org/.

Propagation of Electromagnetic Wave in Layered Heterogeneous Medium

R. Kadlec, E. Kroutilová, D. Nespŏr, and P. Fiala
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejní 2906/4, Brno 612 00, Czech Republic

Abstract— The paper presents the problem of numerical modelling of high frequency electromagnetic (EM) waves propagation in inhomogeneous materials. For this method, a numerical model was prepared. For a layered heterogeneous medium, an algorithm for the reflection on several layers was prepared in MatLab program environment. Reflection and refraction in many directions on heterogenous material are solved by means of the numerical method. Central in this respect are the refractions and reflections on the boundary of materials with different properties. This method is suitable for the design application of metamaterials. The deduced algorithm was projected for the wide spectrum of EM waves.

1. INTRODUCTION

Generally, inhomogeneities and regions with different parameters appear even in the cleanest materials. During the electromagnetic wave passage through a material there occur an amplitude decrease and a wave phase shift, owing to the material characteristics such as conductivity, permittivity, or permeability. If a wave impinges on an inhomogeneity, a change of its propagation there occurs. This change materializes in two forms, namely in reflection and refraction. In addition to this process, polarization and interference may appear in the waves.

2. ELECTROMAGNETIC WAVES IN ISOTROPIC DIELECTRICS MATERIALS

Algorithms were generated in the Matlab program environment that simulates reflection and refraction in a lossy environment on the interface between two dielectrics. This section of the paper is linked to the previous modelling of light applying the related geometrical laws. The reflection and refraction is in accordance with Snell's law for electromagnetic waves as shown in Fig. 1. The form of Snell's law is

$$\frac{\sin\theta_0}{\sin\theta_2}=\frac{k_2}{k_1}=\frac{\sqrt{\mathrm{j}\omega\mu_2\cdot(\gamma_2+j\omega\varepsilon_2)}}{\sqrt{\mathrm{j}\omega\mu_1\cdot(\gamma_1+j\omega\varepsilon_1)}}, \tag{1}$$

where k is the wave number, γ is the conductivity, ε the permittivity and μ the permeability. Relation (1) is defining for the boundary line between the dielectrics medium. Generally, k_1 and k_2 are complex; then angle θ_2 is also complex. An electromagnetic wave is understood as the electric field strength and the magnetic field strength. The electric component incident wave according to Fig. 1 follows the formula

$$\mathbf{E}_i=\mathbf{E}_0 e^{-j\,k_1\,u_{n0}\times r}, \tag{2}$$

where $\mathbf{E}_0$ is the amplitude electric field strength on the boundary line, r is the positional vector, and un_0 is the unit vector of propagation direction.

The intensity of reflection beams and the intensity of refraction beams are expressed according to the formula

$$\mathbf{E}_r=\mathbf{E}_l e^{-jk_1\,u_{n1}\times r},\quad \mathbf{E}_t=\mathbf{E}_2 e^{-jk_2\,u_{n2}\times r}, \tag{3}$$

where $\mathbf{E}_1$ is calculated from the intensity on boundary line $\mathbf{E}_0$ and reflection coefficient ρ_E ,and $\mathbf{E}_2$ is calculated from the intensity on boundary line $\boldsymbol{E}_0$ and transmission factor τ_E:

$$\mathbf{E}_1=\rho_E\cdot\mathbf{E_0},\quad \mathbf{E_2}=\tau_E\cdot\mathbf{E_0}. \tag{4}$$

The calculation of reflection coefficient ρ_E and transmission factor τ_E with utilization of wave impedance Z_v is according to these relations:

$$\rho_E=\frac{\mathbf{E_1}}{\mathbf{E_0}}=\frac{Z_{v2}\cos\theta_1-Z_{v1}\cos\theta_2}{Z_{v2}\cos\theta_1+Z_{v1}\cos\theta_2},\quad \tau_E=\frac{\mathbf{E_2}}{\mathbf{E_0}}=\frac{2Z_{v2}\cos\theta_1}{Z_{v2}\cos\theta_1+Z_{v1}\cos\theta_2}. \tag{5}$$

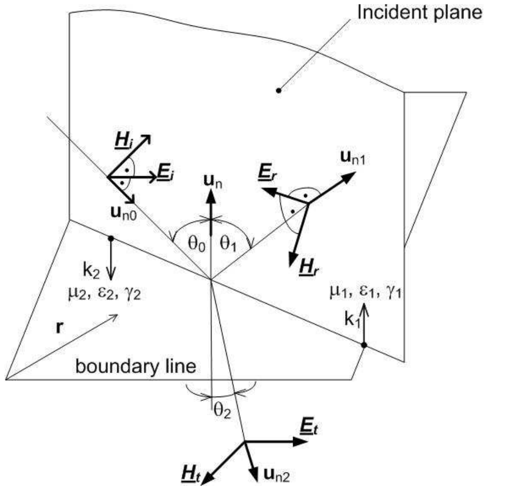


Figure 1: Reflection and refraction of light [2].

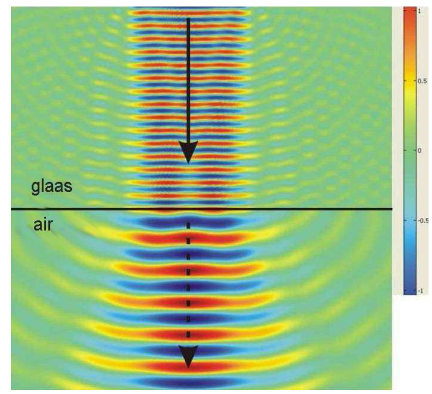


Figure 2: Reflection and refraction on a planar boundary line [5].

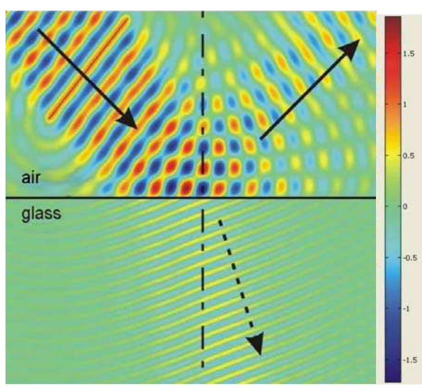


Figure 3: Reflection and refraction on a planar boundary line [5].

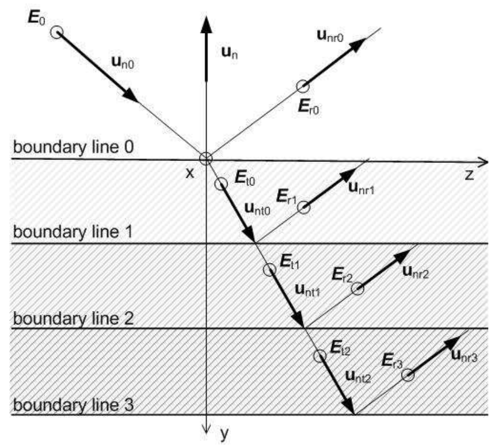


Figure 4: Reflection and refraction on a layered heterogenous material.

For numerical modelling, there is a suitable relation in the form of

$$\begin{aligned}\mathbf{E_r} &= \frac{\mu_2 k_1 \cos\theta_0 - \mu_1\sqrt{k_2^2 - k_1^2 \sin^2\theta_0}}{\mu_2 k_1 \cos\theta_0 + \mu_1\sqrt{k_2^2 - k_1^2 \sin^2\theta_0}}\mathbf{E_0} \cdot e^{-j\,k_1\,u_{n1}\times r},\\ \mathbf{E_t} &= \frac{2\mu_2 k_1 \cos\theta_0}{\mu_2 k_1 \cos\theta_0 + \mu_1\sqrt{k_2^2 - k_1^2 \sin^2\theta_0}}\mathbf{E_0} \cdot e^{-j\,k_2\,u_{n2}\times r}.\end{aligned} \tag{6}$$

These relations are calculated from the basic variable and they facilitate an acceleration of the calculation process.

Interpretation of the Fresnel equations and Snell's laws is simple in the case of the refraction on boundary line between the dielectrics medium. In case of refraction in a lossy medium, angle θ_2 is complex. According to relation (1), angle θ_2 depends on wave numbers k_1 and k_2, which are generally complex; then, in medium 2 an inhomogeneous wave is propagated.

Example of reflection and refraction on a planar boundary line in COMSOL program is shown in Fig. 2 (at perpendicular incidence of the wave on the interface) and Fig. 3 (at the incidence of the wave on the interface at an angle of 45°).

For a layered heterogeneous medium, an algorithm is deduced for the reflection on several layers. The reflection and refraction on a heterogonous material is solved by the help of the numerical method. The reflection on a layered material on n layers generates n primary electromagnetic waves, according to Fig. 4. The interpretation of propagation of electromagnetic waves on a

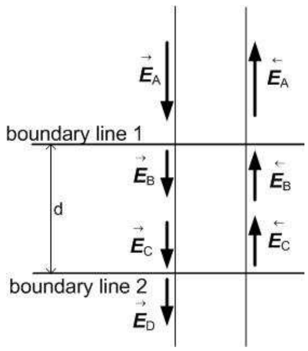


Figure 5: Perpendicular incidence of the wave on layered material.

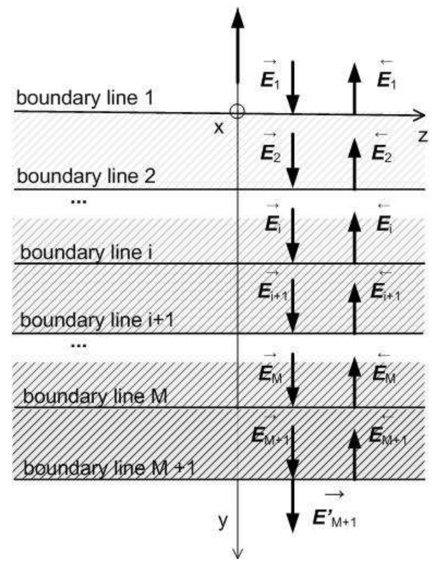


Figure 6: Multilayer heterogenous material.

layered heterogeneous medium is according to relation

$$\mathbf{E}_{rl} = \mathbf{E}_{il}\rho_{El} \cdot e^{-j\,k_{(l+1)}\,u_{nrl}\times r_l}, \quad \mathbf{E}_{tl} = \mathbf{E}_{il}\tau_{El} \cdot e^{-j\,k_{(l+2)}\,u_{ntl}\times r_l}, \tag{7}$$

where $\mathbf{E}_{rl}$ and $\mathbf{E}_{tl}$ are the reflection and refraction electromagnetic waves on the boundary line $(l = 0, \ldots, \max)$ according to Fig. 2, $\mathbf{E}_{il}$ is the amplitude electric field strength on boundary line l, ρ_{El} and τ_{El} are the reflection coefficient and transmission factor on boundary line l, k_l is the wave number of layer, r_l is the electromagnetic wave positional vector on boundary line l, u_{ntl} and u_{nrl} are the unit vectors of propagation direction.

Special case is perpendicular incidence of the electromagnetic wave on the interface according to Fig. 5. The interpretation of perpendicular incidence for incident wave $\tilde{\mathbf{E}}_{\mathbf{A}}$ and reflection wave $\overleftarrow{\mathbf{E}}_{\mathbf{A}}$ is according to relation

$$\vec{\mathbf{E}}_A = \frac{e^{j\,k_2\cdot u_{nA}\cdot r} + \rho_{12}\rho_{23}\cdot e^{-j\,k_2\cdot u_{nA}\times r}}{\tau_{12}\tau_{23}}\vec{\mathbf{E}}_D, \quad \overleftarrow{\mathbf{E}}_A = \frac{\rho_{12}e^{j\,k_2\cdot u_{nA}\cdot r} + \rho_{23}e^{-j\,k_2\cdot u_{nA}\times r}}{\tau_{12}\tau_{23}}\vec{\mathbf{E}}_D, \tag{8}$$

where ρ_{12} is reflection coefficient of wave in external medium which is reflective on boundary line 1, ρ_{21} is reflection coefficient of wave in internal medium which is reflective on boundary line 1. Transmission factors τ are indexed analogically. k_2 is the wave number internal medium and $u_{nA}\times r$ is distance between boundary lines.

The general case of arbitrary number of dielectric layers is shown in Fig. 6 [3]. There are M layers and $M+1$ interface and $M+2$ dielectric media, including the external media. The incident and reflected waves are considered on the top of each interface. The overall reflection response $\rho_1 = \tilde{\mathbf{E}}_{\mathbf{1}}/\overleftarrow{\mathbf{E}}_{\mathbf{1}}$, can be obtained recursively in a variety of ways, such as by the propagation matrices, the propagation of the impedances at the interfaces, or the propagation of the reflection responses. The forward and backward waves on the top of interface i are related to those on the top of interface $i+1$ by

$$\begin{bmatrix} \tilde{\mathbf{E}}_i \\ \overleftarrow{\mathbf{E}}_i \end{bmatrix} = \frac{1}{\tau_i}\begin{bmatrix} \mathrm{e}^{\mathrm{j}k_i\cdot u_{ni}\times r} & \rho_i\mathrm{e}^{-\mathrm{j}k_i\cdot u_{ni}\times r} \\ \rho_i\mathrm{e}^{\mathrm{j}k_i\cdot u_{ni}\times r} & \mathrm{e}^{-\mathrm{j}k_i\cdot u_{ni}\times r} \end{bmatrix}\begin{bmatrix} \tilde{\mathbf{E}}_i+1 \\ \overleftarrow{\mathbf{E}}_i+1 \end{bmatrix}, \ i = M,\ M-1, \ldots, 1, \tag{9}$$

where $\tau_i = 1 + \rho_i$ and $k_i \cdot u_{nA} \cdot r$ is the phase thickness of the i layer. Assuming no backward waves in the most bottom medium, these recursions are initialized at the $(M+1)$ interface as follows

$$\begin{bmatrix} \tilde{\mathbf{E}}_{M+1} \\ \overleftarrow{\mathbf{E}}_{M+1} \end{bmatrix} = \frac{1}{\tau_{M+1}}\begin{bmatrix} 1 & \rho_{M+1} \\ \rho_{M+1} & 1 \end{bmatrix}\begin{bmatrix} \tilde{\mathbf{E}}'_{M+1} \\ 0 \end{bmatrix} = \frac{1}{\tau_{M+1}}\begin{bmatrix} 1 \\ \rho_{M+1} \end{bmatrix}\tilde{\mathbf{E}}'_{M+1}, \tag{10}$$

Similarly the recursions for the total electric and magnetic fields, which are continuous across each interface, are given by

$$\begin{bmatrix} \mathbf{E}_i \\ \mathbf{H}_i \end{bmatrix} = \begin{bmatrix} \cos(k_i \cdot u_{ni} \times r) & \mathrm{j} \cdot Z_{vi} \sin(k_i \cdot u_{ni} \times r) \\ \mathrm{j} \cdot Z_{vi}^{-1} \sin(k_i \cdot u_{ni} \times r) & \cos(k_i \cdot u_{ni} \times r) \end{bmatrix} \begin{bmatrix} E_{i+1} \\ H_{i+1} \end{bmatrix}, \; i = M,\; M-1, \ldots,\; 1, \tag{11}$$

3. CONCLUSION

Numerical modelling of wideband electromagnetic signals on field of multilayer and periodic structure optical materials in Matlab program is very time demanding. This method is suitable for specific purposes of detail analysis.

For simple cases (such as a planar interface), the behaviour of an impinging wave can be calculated analytically by the help of Snell's refraction/reflection law and the Fresnel equations. However, it is difficult (and often infeasible) in more complex structures to perform an analytical calculation. Therefore, numerical methods are applied to facilitate the calculation process, and a wide range of programs like ANSYS, Comsol, or Matlab can be utilized in the realization of numerical modelling.

Algorithms created in the Matlab environment are verified by the help of programs based on the finite element method, namely programs such as Comsol and ANSYS. The paper includes a theoretical analysis and references to the generated algorithms, which are verified using numerical models.

ACKNOWLEDGMENT

The research described in the paper was financially supported by the research program MSM 0021630516 and research plan MSM 0021630513, Ministry of Defence of the CR, Ministry of Industry and Trade of the CR (Diagnostics of Superfast Objects for Safety Testing, FR-TI1/368), Czech Science Foundation (102/09/0314) and project of the BUT Grant Agency FEKT-S-10-13.

REFERENCES

1. Stratton, J. A., *Teorie Elektromagnetického Pole*, STNL, Praha, 1961.
2. Dedek, L. and J. Dedková, *Elektromagnetismus*, Vol. 2, 232, Vitium, Brno, 2000, ISBN 80-214-1548-7.
3. Orfanidis, S., *Electromagnetic Waves and Antennas*, 1031, 2008, Available from WWW: <www.ece.rutgers.edu/∼orfanidi/ewa>.
4. Moss, C., *Numerical Metods for Electromagnetic Wave Propagation and Scattering in Comlex Media*, 240, 2004, Available from WWW: <http://portal.acm.org/citation.cfm?id=1023429>.
5. Nešpor, D., "Electromagnetic wave propagation study in heterogeneous structures," Ph.D. thesis, 2009.
6. Jin, G. A., "Survey of radiosity and ray-tracing methods in global illumination," Institute for Computer Graphics, 2000, Available from WWW: <http://www.icg.seas.gwu.edu/cs367/global.pdf>.
7. 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.

Measurement of Concentration and Water Flow

M. Friedl, P. Fiala, P. Marcon, and R. Kubasek
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejni 2906/4, Brno 612 00, Czech Republic

Abstract— This article deals with the non-destructive measuring of water concentration and movement in the tissues of tree structures. The related measurement method is based upon monitoring impedance changes of the sensor or, alternatively, upon observing changes of the sensor resonance frequency in a porous material. Water flow and concentration affect the sensor permittivity and, thus, change its impedance. The final selection of the sensor was based on the corresponding experimental measurement in the DTEE anechoic chamber. The following analyses aimed at measuring water flow in the applied tube. The sensor was placed on the outside of the tube and connected with a vector analyzer. Then, we measured the impedance of the sensor in the range of 100 Hz to 40 MHz. The actual measurement process was realized at the DTEE outdoor facilities. For the measurement, we fixed the tube to one location and, subsequently, recorded the sensor impedance value during the through-flow of water. Another instance of experimental water flow measurement was performed in relation to the added magnetic field. We had created the magnetic field by the help of two super magnets. Magnetic fields both vertical and horizontal in relation to the sensor plane were tested. The last measurement task consisted in determining the dependence of the sensor impedance on the concentration of water in porous materials.

1. INTRODUCTION

The measurement of fluid concentration and its flow through tree structure tissues constitutes a rather difficult task that cannot be realized in a simple manner if the planned measurement process is conceived as non-destructive. In this respect, the presented report contains the analysis of an experimental method designed to support the basic evaluation of measurement utilizing electromagnetic sensors [1] to determine the characteristics of fluid flow through different strata of plants. The measurement method is based upon the monitoring of impedance changes that occur within the sensor environment, namely in the sensor resonance area above the porous material [2–4].

2. SENSORS

Four basic concepts have been theoretically designed and practically produced to facilitate the relevant research activities. Sensor No. 1 shown in Fig. 1 was realized during the preparations as a simple coil and operated in resonance mode; sensor No. 2, though operated similarly in resonance mode, consisted in a modified coil with capacitive coupling of the conducting plane (Fig. 1, No. 2). Sensor No. 3 was created in the form of an equivalent capacitor and operated in resonance mode, Fig. 1, No. 3. The last of these sensors was designed as a parallel configuration of capacity and induction of a magnetic field whose components are invariably out-of-parallel in relation to the fluid movement vector, Fig. 1, No. 4. During the actual process of measurement, we applied a vector analyzer in order to sample the fluid movement change at a low speed as well as to scan the fluid volume. Within this context, we also evaluated the resonance phase change of the sensor — measured sample system depending upon the fluid velocity and volume flow.

3. EXPERIMENTAL MEASUREMENT OF WATER FLOW

The measurement method is based upon monitoring impedance changes of the sensor or, alternatively, upon observing changes of the sensor resonant frequency depending on the through-flowing water. The sensor was located in the immediate vicinity of the water through-flow and the subsequent processing of the measured values materialized through the OMICRON LAB BODE 100 vector analyzer. The actual experimental measurement took place in the DTEE anechoic chamber. The sensors were fixed on the outside of a hollow cylinder (tube) and interconnected with the vector analyzer; then, in the course of the water through-flowing, measurement was performed of the sensor impedance within the range of 100 Hz to 40 MHz.

The resulting instances of impedance behaviour for sensor No. 1 are provided in Fig. 2. The red curve corresponds to the module of the measured impedance, while the blue curve represents the impedance phase. Further, the dashed line shows the initial condition during which there is no flow of water through the cylinder and the continuous line indicates the period of water through-flow. It

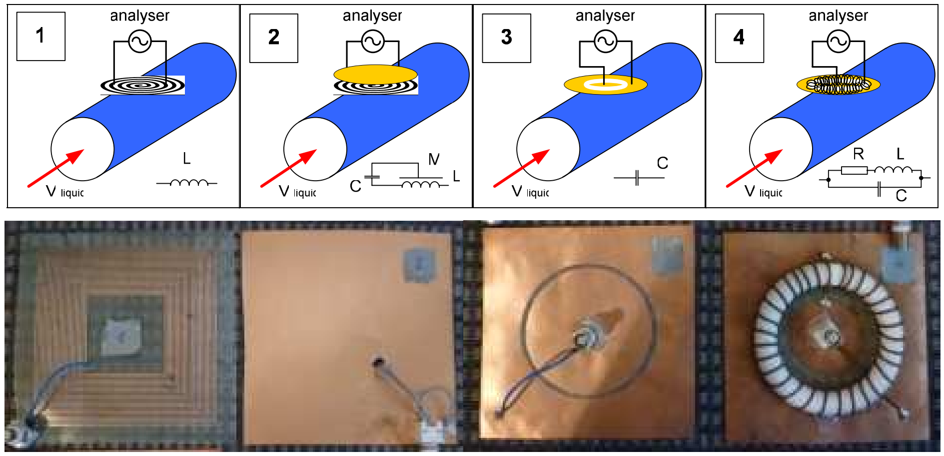


Figure 1: Sensors 1, 2, 3 and 4.

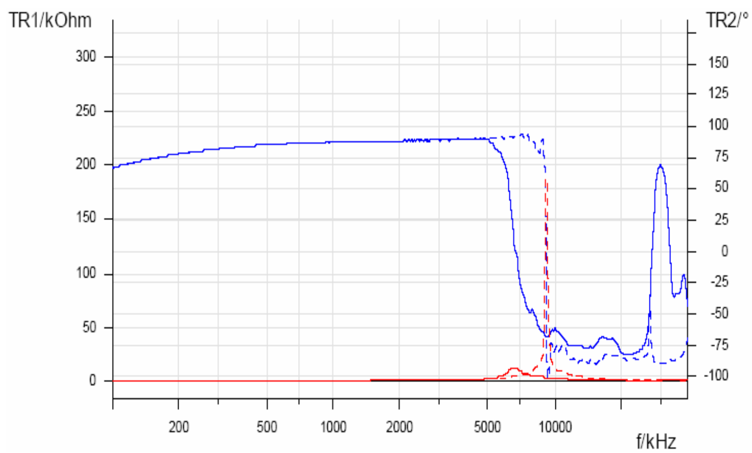


Figure 2: Sensor 1 impedance behaviour at water through-flow.

follows from the measured values that, in general terms, the volume of through-flowing water can be measured by means of the sensor impedance change. A suitable method for the evaluation of water through-flow consists in measuring the change of the sensor impedance phase at the selected frequency; the feasibility of this procedure generally rests upon the fact that the phase characteristic of the sensor impedance is sufficiently steep. The resulting impedance behaviour related to the other sensors is not mentioned in this report owing to space limitations.

4. MEASUREMENT WITH MAGNETIC FIELD

The experimental measurement using B-field was realized at the outdoor facilities of the Department of Theoretical and Experimental Electrical Engineering. For the measurement, both vertical and horizontal applications of B-field were used; in this respect, the primary obtained results implied that the vertical B-field has a positive influence upon the sensitivity of the given system. However, it is necessary to perform other measurements in order to prove more reliably the effectiveness of the applied B-field for any higher-sensitivity analysis of the sensor impedance.

5. MEASUREMENT IN A FOREST

Utilizing the knowledge and observation obtained within the experimental stage of our work, we performed further measurements materialized in a forest through the medium of a tree and four sensors which, for the purposes of this research activity, were based upon a different (destructive) principle. Sensors numbered 1, 2, 3 and 4 were fixed to the tree (Fig. 4), with the measuring conductors connected to the analyzer.

Then, we calibrated the measuring system and started the actual procedure. The experimental measurement proceeded through the resonance frequency and phase reading for the period of three hours. In order to attain results capturing changes of through-flow in the tree, it is necessary

Figure 3: The Scots Pine measurement photos.

Figure 4: Tree fixation. Details of the A, B, C and D sensors.

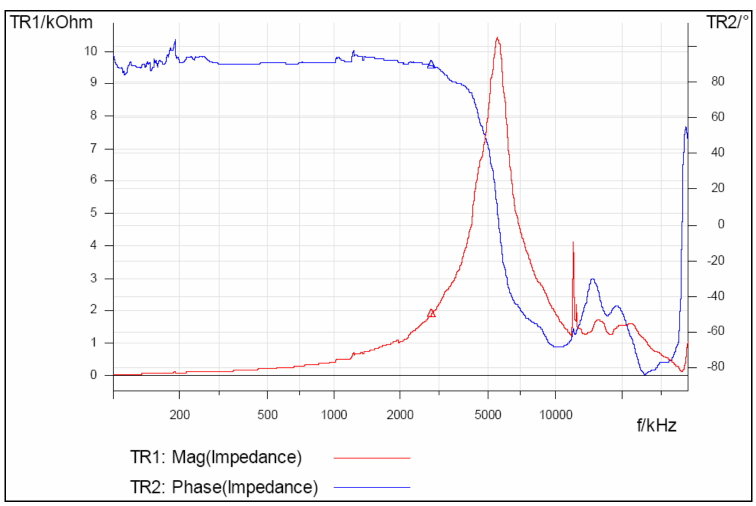


Figure 5: Impedance behaviour of sensor No. 1 fixed to the tree.

to repeat the daylong measurement procedure within the given season of the year. Our activities, however, aimed at verifying the measurement method based on laboratory experiments. When fully processed, the obtained results demonstrated the functionality of the sensors from the theoretical electro-magneto-hydrodynamic point of view.

The forest measurement took place at the grounds of Mendel University in Brno (the Mendel university of agriculture and forestry) and utilized a coniferous tree, the Scots Pine, as the test object (Fig. 3). The same tree has been used in the process of measuring fluid motion within a tree structure by the help of a different (destructive) method. The results obtained from both measuring approaches introduced herein have been subject to analysis.

Figures 5 and 6 show the results obtained from sensors numbered 1 and 4. These results correspond to impedance values of the individual sensors following their fixation to the structure

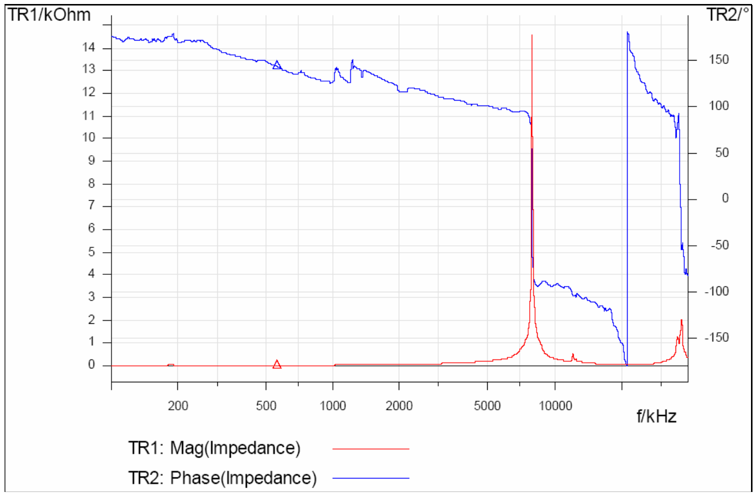


Figure 6: Impedance behaviour of sensor No. 4 fixed to the tree.

of the tree. The red curves represent the module of the measured impedance, while the blue curves correspond to the impedance phase. It is obvious from the measured results that sensor 4, which combines the inductive and the capacitive principles, attains the highest quality Q of all the sensors.

6. CONCLUSION

The current trend in the process of given quantities measurement is markedly directed towards applying non-destructive methods. This report presents the description of a non-destructive method utilizable for measuring water through-flow and concentration. The measurement method is based on the monitoring of impedance changes of a sensor or, alternatively, on the observation of changes of the sensor resonant frequency in a porous material.

Four sensors have been experimentally designed and practically analyzed in order to perform the necessary measurements within this task. The primarily realized instances of measurement confirmed the applicability of the basic idea that the constructed sensors could be utilized for non-destructive measurement of fluid movement in porous materials. Another task consisted in measuring the given characteristics depending on the use of B-field; in this respect, the results imply that the B-field perpendicular to the fluid movement has a positive influence upon the sensitivity of the entire measurement system. However, it is necessary to perform further research in order to prove more reliably the effectiveness of the applied B-field. Here, an increase in the system sensitivity is most likely to provide significant support for the given measurement as well as a means to obtain more relevant results.

The last of the tasks realized was a practical forest measurement. While it is true that the measurement time period would have to be substantially longer in order to provide results with a higher degree of accuracy, we also have to point out that our activities were aimed mainly at verifying the measurement method based on laboratory experiments. The obtained results have been duly analyzed to sustain further development of the described method. It is obvious from these results that sensor 4, which combines the inductive and the capacitive principles, attains the highest quality Q of all the sensors applied. Therefore, the sensor is most likely to be prioritized within the planned non-destructive measurements. The reading of the sensor resonant frequency by an automatic reading system will be made more accurate. By means of long-term measurement of a given quantity, more relevant results will be secured for further comparison with the effects of the destructive method.

ACKNOWLEDGMENT

The research described in the paper was financially supported by FRVS (a fund of university development), by research plans No. MSM 0021630513, No. MSM 0021630516, by grant of the Ministry of Industry and Trade of the Czech Republic No. FR-TI1/001, FR-TI1/368, GACR 102/09/0314, and by a project of the BUT Grant Agency coded FEKT-S-10-13.

REFERENCES

1. Fiala, P., "Model indukčního prutokoměru," *ANSYS Users Meeting*, s.6–15, Česká republika a Slovensko, Třeš' Sekce C 1–28, 6, SVS FEM s.r.o, Brno, September 25–26, 2001.
2. Deng, X., Y. Chen, Z. Yan, and Z. Wei, "AC impedance model of array electrodes in multisensor fusion system for two-phase flow measurement," *IEEE Transactions on Instrumentation and Measurement*, Vol. 59, No. 6, 1722–1726, June 2010.
3. Szczepanik, Z. and Z. Rucki, "Field analysis and electrical models of multi-electrode impedance sensors," *Sens. Actuators A: Phys.*, Vol. 133, 13, 2007.
4. Bera, S. C. and B. Chakraborty, "A novel technique of flow measurement for a conducting liquid," *IEEE Transactions on Instrumentation and Measurement*, Vol. 58, No. 8, 2512–2517, August 2009.

Using Metamaterials as Electromagnetic Lens for MR Tomograph

D. Nespor, P. Drexler, P. Fiala, and K. Bartusek
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejni 2906/4, Brno 612 00, Czech Republic

Abstract— It is possible to manipulate electromagnetic field by means of dielectric materials. Electromagnetic field will be affected due to values of dielectric constant ε_r and permeability μ_r. Viktor Veselago has defined theoreticaly the usage of material with negative parameters (metamaterial) as an electromagnetic lens. If we put metamaterial between transceiver and receiver, we can see an increase in detected energy. It appears that the metamaterial behaves like an electromagnetic lens. The subject of this work is construction and optimization of electromagnetic lens for mid-field MR tomograph. Principles of lens construction are in [1–3]. Paper contains analytic calculations of basic design of the lens and numerical calculations of the optimized lens.

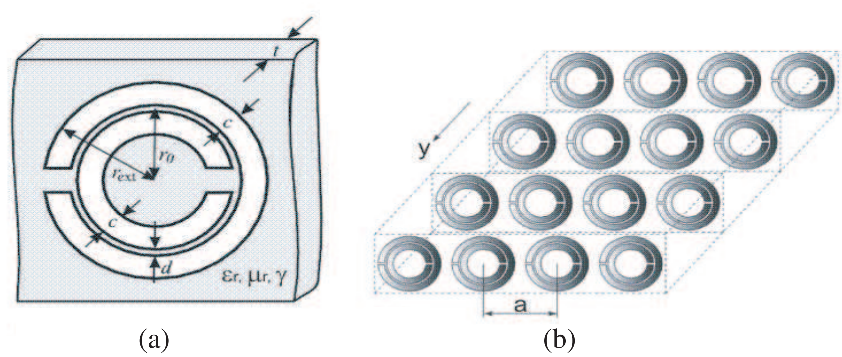
1. INTRODUCTION

Metamaterials are composite structures which have electrical and magnetic parameters (permittivity and permeability) which don't exist in the nature. Metamaterials are composed of small segments. For example, they compose of long parallel conductors or planar coils. In fact, metamaterials aren't homogenous. But if we use an electromagnetic wave with wavelength much larger than the size of metamaterial segment, we can consider the metamaterial as homogenous. Generally, metamaterials with negative parameters represent the media which has negative value effective permittivity ε_{ef} or permeability μ_{ef} [1]. These materials are known since 1960s [5, 6] but only theoretically.

Negative permeability can increase magnetic coil range. It means that metamaterial with negative permeability behave as an electromagnetic lens [3]. Negative effective permeability can be created by Split Ring Resonators (SRR) [1]. The resonator is constructed as a concentric metal ring with interspaces. These rings are deposited as regular matrix on dielectric substrate. Figure 1(a) shows one Double Split Ring resonator. Figure 1(b) shows matrix of Double Split Ring resonators. On these periodic structure will be created negative permeability.

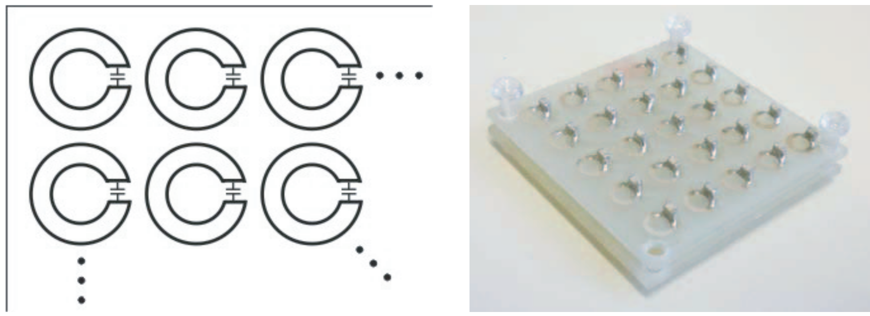
2. ELECTROMAGNETIC LENS CONSTRUCTION

For construction of electromagnetic lens a Single Split Ring Resonators (SSRR) is used. Resonators have outer diameter 5 mm and inner diameter 3.5 mm. These resonators are putted as regular matrix on the dielectric substrate FR4 as shows Figure 2(a). The substrate thickness is 2 mm and relative permittivity is 4.5. On the substrate are five resonators in a row and five in a column. Numeric analysis of this structure was performed with Comsol software. If planar electromagnetic wave with varying frequency is fed on the structure it resonates on frequency over 6 GHz. The frequency characteristic of S_{21} parameter is shown in Figure 3(a). The frequency of the MR system is 198.75 MHz. But single split resonators for that frequency would be too big. That is why the constructed resonators with resonant frequency about 6 GHz has been modified by means of chip



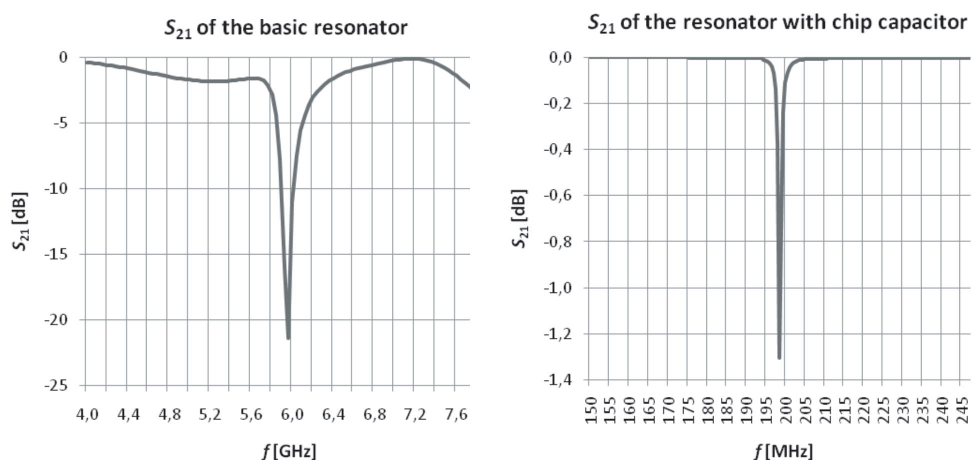
(a) (b)

Figure 1: Double split ring resonators [1].



(a) (b)

Figure 2: (a) Lens structure design, (b) lens realization.



(a) (b)

Figure 3: Results of numerical analysis of S_{21} for: (a) basic resonators, (b) resonators with chip capacitors.

capacitors. There ceramic chip capacitors has been assembled over the interspaces [1, 4, 7]. Capacity calculation was made analytically and then numerically with Comsol software. Parameter S_{21} of numeric solution is shown in Figure 3(b). Capacity value is 112 pF. This capacity was approximated as parallel combination three chip capacitors: $C_1 = 100\,\text{pF}$, $C_2 = 10\,\text{pF}$, $C_3 = 1,2\,\text{pF}$. Lens realization is shown Figure 2(b). The measured resonance frequency of the lens is 199 MHz.

3. LENS OPTIMALIZATION

The solution described above is inconvenient in many aspects. The main inconveniency is the inaccuracy of capacitors and soldered pads. Next inconveniency is the need of the surface mount. That's why we have designed the spiral resonator. This design was created with Comsol software. Spiral resonators are possible to manufacture as a planar types without the external component assembly. That is the main advantage of these resonators. Diameter of this resonator is 7 mm, width of ring strip is 0.1 mm and number of the turns is eight. The rings are etched on the substrate CER10. Relative permittivity of this substrate is 10 and substrate thickness is 1 mm. Spiral resonator is shown in Figure 4(a). The calculated S_{21} parameter of the spiral resonator is shown in Figure 5(a).

There are many technology problems of manufacturing. Therefore the resonators are manufactured with resonant frequency lower than desired and then the resonators have to be trimmed. Trimmed spiral resonator is shown in Figure 4(b). Parameter S_{21} of the trimmmed spiral resonator is shown in Figure 5(b).

The main requirement for metamaterials application is that the wavelength of electromagnetic wave is much higher than the segments of the metamaterial. Therefore the next design is double-sided spiral resonator. A resonator has square shape and in the middle is via. Via connected both sides of the resonator. For the double sided resonator the CER10 substrate is used again. Size of the resonator is 5 mm. Width of the one ring strip is 0.1 mm and number of turns is five. Double-sided resonator is shown in Figure 6. Parameter S_{21} in shown in Figure 7.

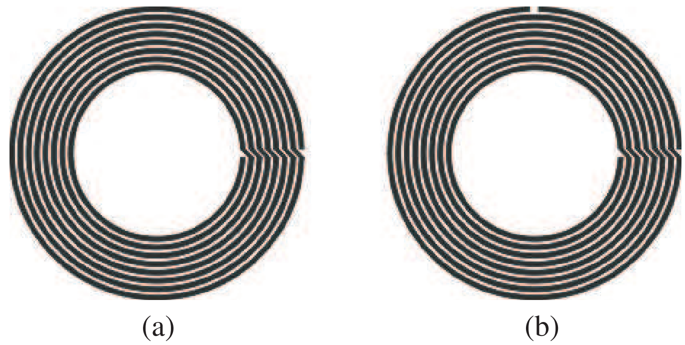


Figure 4: (a) spiral resonator design, (b) trimmed spiral resonator design.

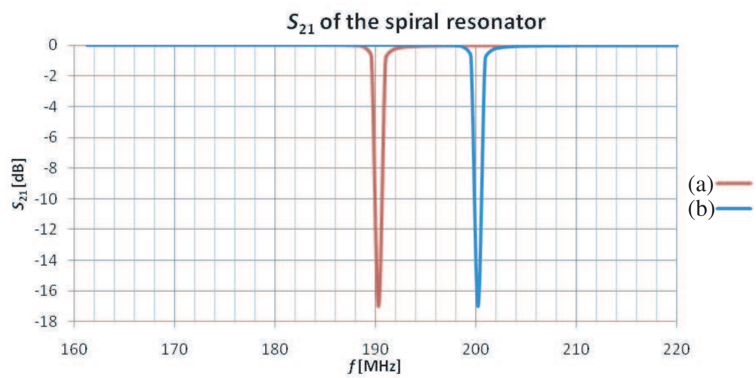


Figure 5: (a) S_{21} of the spiral resonator, (b) parameter S_{21} of trimmed spiral resonator.

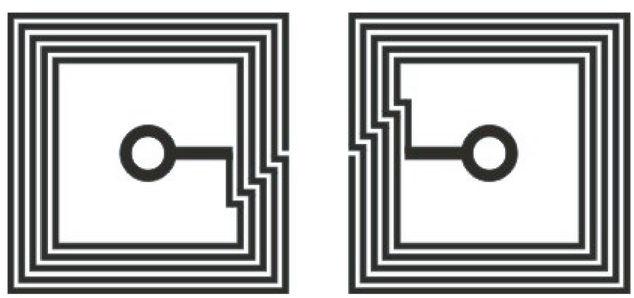


Figure 6: S_{21} of the spiral double-sided resonator.

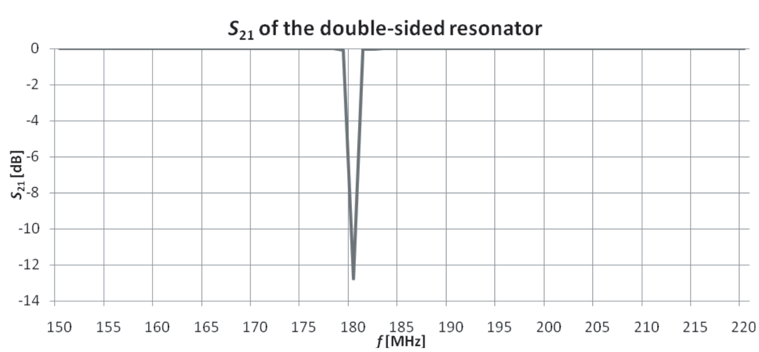


Figure 7: S_{21} of the spiral double-sided resonator.

4. CONCLUSION

The properties of the first prototype of the designed lens have been examined. The chip capacitors assembly brings a lot of issues, which has to be solved to obtain a correct behaving device. Hence, the new resonator structure has been proposed as a planar resonator with distributed capacitance in the form of spiral layout. Two designs have been described — the single sided spiral resonator and the double sided spiral resonator. In this configuration the resonant frequency can be trimmed in order to achieve the desired value.

The results of applied research of metamaterial structures for low frequencies ($f \sim 200\,\mathrm{MHz}$) are discussed in the paper. In order to achieve the possibility of real application in MR systems a lot of theoretical issues has to be solved. The other area to research is the technology of these structures fabrication and its properties measurement.

ACKNOWLEDGMENT

The work described in the paper was financially supported by the research project GA102/09/0314, research plan MSM 0021630513 and project of the BUT Grant Agency FEKT-S-10-13 and grant of Czech ministry of industry and trade No. FR-TI1/001.

REFERENCES

1. Marqués, R., F. Martín, and M. Sorolla, *Metamaterials with Negative Parameters*, John Wiley & Sons, Inc., Hoboken, New Jersey, Rok, 2007, ISBN 2007017343.
2. Caloz, C. and T. Itoh, *Electromagnetic Metamaterials: Transmission Line Theory and Microwave Applications*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2006, ISBN 2005048976.
3. Manuel, J., L. Jelinek, and R. Marques, "On the application of $\mu = -1$ metamaterial lences for magnetic resonance imaging," Departamento de Electronica y Electromagnetismo, Universidad de Sevilla, 2008.
4. Bartusek, K, P. Drexler, P. Fiala, R. Kadlec, and R. Kubasek, "Magnetoinductive lens for experimental mid-field MR tomography," *PIERS Online*, Vol. 6, No. 7, 621–624, 2010.
5. Cui, T. J. and J. A. Kong, "Time-domain electromagnetic energy in a frequency-dispersive left-handed medium," *Phys. Rev. B*, Vol. 70, 2004, ISBN 205106.
6. Huangfu, J., L. X. Ran, H. S. Chen, X. M. Zhang, K. S. Chen, T. M. Grzegorczyk, and J. A. Kong, "Experimental confirmation of negative refractive index of a metamaterial composed of Omega-like metallic patterns," *App. Phys. Lett.*, Vol. 84, 2004, ISBN 1537-1539.
7. Kadlec, R., P. Fiala, and P. Drexler, "A new approach to periodical structure analysis," *Progress In Electromagnetics Research Symposium*, 323–326, Cambridge, USA, July 5–8, 2010.

The Design of High-impedance and High-voltage Input Amplifier for Measurement of Electropotentials on Solid-liquid Phase Boundary

Z. Roubal[1], Z. Szabó[1], M. Steinbauer[1], D. Heger[2], and R. Kubásek[1]

[1]Department of Theoretical and Experimental Electrical Engineering, Brno University of Technology
Kolejní 2906/4, Brno 612 00, Czech Republic
[2]Department of Chemistry, Faculty of Science, Masaryk University
Kamenice 5/A8, Brno 625 00, Czech Republic

Abstract— On the interface of solid-liquid phases of water solutions, certain electric potential occurs during the freezing process. This is caused uneqal distribution of ions between the solid and liquid phases. As freezing is often used for the preservation of biological samples, the influence of the electric field induced by this process upon biological samples is the subject of investigation. For this purpose, we constructed specialized measuring devices to facilitate the measurement of this potential.

In this paper, a design will be analyzed of an electrometric amplifier for the measurement of voltage in the order of hundreds volts. Because the measured source shows a very high inner resistance and low capacity, the amplifier input resistance must be greater than $10^{14}\,\Omega$ with a negligible parallel capacity. For the same reason, using an input voltage divider is a problematic step. Such a high input impedance can be achieved only when applying a special input amplifier. A standard operational amplifier shows a measurement range of about $\pm 10\,\mathrm{V}$, for the expansion of this range excluding the input voltage divider, it is necessary to use voltage shifting for the operational amplifier power supply. A circuit with floating supplies is susceptible to oscillation unless supported by the right frequency corrections. The proposed electrometric amplifier with a high voltage input range will be simulated using Pspice.

1. INTRODUCTION

In the field of research into biochemical substances there has emerged the need to measure the electric potential of phase changes upon solidification of aqueous solutions. This potential is referred to as freezing potential or Workam-Reynolds phenomenon [1–5]. At this point, it is necessary to note that the results obtained by the related researchers differed in respect to the applied measuring method, measurement system configuration, and methodology. When performing a comparison of these results, readers of the herein mentioned reports can identify within the measured potentials a scatter that ranges between 100 mV and hundreds of volts for a chemically identical sample. The principle of freezing potential consists in the generation of an electric charge on the interface of fluid and solid [2]; the described process is typical of water or liquid solutions! A separation of charge occurs between ice and a sulution. This separation results from differences in the partition coefficients of anions and cations and it generates an electric potential, which is known as freezing potential. The mobility of ions is changed upon freezing. A model of freezing potential was first created by Lefebre [2]; the proposed approach includes charge generation, redistribution and neutralization. This model was later perfected by Bronshteyn and Chernov, who included a charge redistribution in ice due to ionic diffusion and H^+ ion flow driven by the electrical field in the crystal.

2. THE MEASURING APPARATUS

The basic research was materialized in laboratory conditions showing a lower degree of repeatability. For this reason, we designed and built a measuring apparatus (Fig. 1); the illustration of its primary internal structure is provided in Fig. 2. In the described measurement device concept, cooling has been preset in the direction from the bottom to the upper sections of the apparatus. Thus, we can attain repeated generation and measurement of an electric potential on the interface of the sample solid phase. At the moment when ice reaches the inner electrode, discharge occurs and the measured electric voltage will drop to zero.

The entire measuring device is positioned in a thermally insulated vessel where liquid nitrogen will be produced (nitrogen boiling temperature equals to $-195,80°\mathrm{C}$). The lower section of the vessel shows a shape and configuration enabling high-quality accumulation and transfer of heat (with cooling realized by means of liquid nitrogen); simultaneously, however, the vessel facilitates

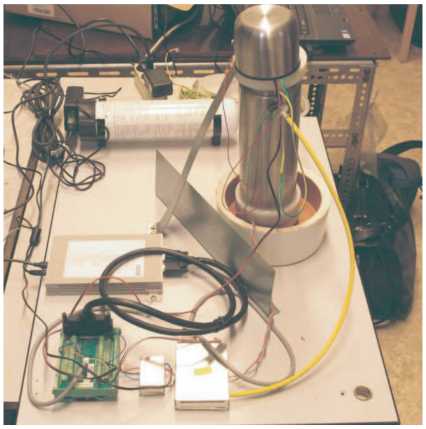


Figure 1: The assembled apparatus.

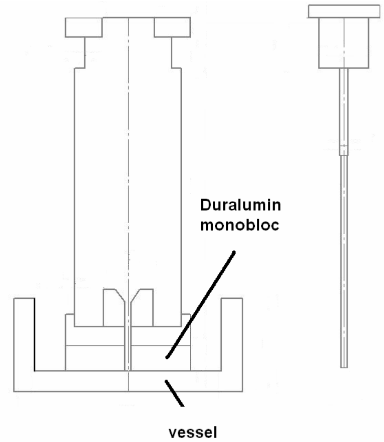


Figure 2: Internal structure of the measuring apparatus.

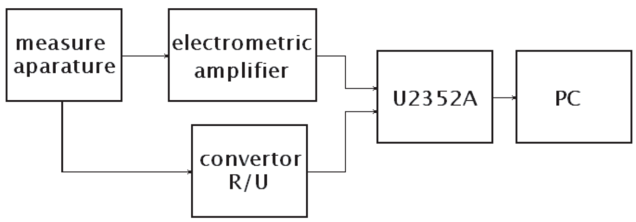


Figure 3: The measuring apparatus overall diagram.

the elimination of problems resulting from the change in linear expansion. In the shielded vessel having a hot and a cold section there exists free, gas-filled space that prevents the occurrence of air humidity freezing. At the very initial stage of measurement, the head housing a capillary as well as the sample to be tested is inserted in an overcooled duralumin monobloc; thus, a repeatable refrigeration process starts. At the moment of the sample insertion in the overcooled space, the tested sample phase begins to change and the fluid-solid phase interface progressively moves upwards; now, freezing potential is measured. Following the phase change reach of the other electrode, freezing potential will dischare itself.

3. THE PROPOSED MEASURING STRING

An electric potential on the interface between two phases of the sample behaves as a source of potential with a high differential resistance. The measurement must be realized using a system with an electrometric amplifier at the input. The duralumin monobloc temperature is measured by the help of a PT100-type metal resistive sensor.

The supply of the sensor materializes through the source of constant current 1mA; in addition, voltage scanning is realized on the sensing device (element). The digitization of the related two voltages takes place through an Agilent U2352A data acquisition measurement module. Further, the data measured are processed by a PC using the Agilent VEE environment. The temperature provided by the PT100-type platinum sensor is evaluated by means of solving the quadratic equation. The measurement result consists in the time behaviour of freezing potential in the time domain.

4. THE ELECTROMETRIC AMPLIFIER

In the process of designing an electrometric amplifier there may occur a certain technical problem concerning high input voltage. As a consequence, the measured voltage value can range within several hundreds of volts. A standard solution consists in applying a resistor divider at the input, Fig. 4. In electrobiology, however, we can not use this type of solution as the signal source contains capacity in orders of pF; even when special high-ohm $100\,\mathrm{G\Omega}$ resistors are used, the discharge time constant of the circuit ranges within orders of tenths of seconds. it is therefore obvious from the description that the discussed solution does not help us to meet the desired target.

One of the proposed methods of solution to the problem lies in the application of an electrometric amplifier not equipped with any input divider; in this type of amplifier, then, we assume floating power supply in relation to the input voltage. Fig. 8 shows the diagram of such an amplifier. Here, the input voltage is amplified by an electrometric amplifier supplied from a floating source. The related output is connected to a high-voltage amplifier supplied by a raising voltage changer. This output manages a high-voltage straight line source, which shifts voltage levels of the electrometric amplifier. The input voltage is read at the high-voltage source output. The internal resistance of this configuration is defined only by the electrometric amplifier volume resistivity and may reach up to $10^{14}\,\Omega$. The discharge time constant is, with inner capacity of the signal source, approximately 1000 s, which will not affect the measured values of freezing potential.

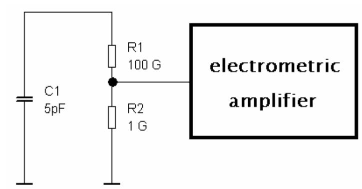


Figure 4: Resistor divider at the input.

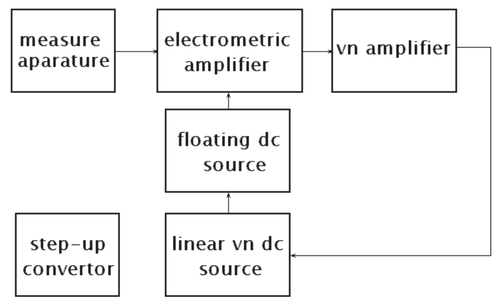


Figure 5: The block diagram of an electrometric amplifier with floating power supply.

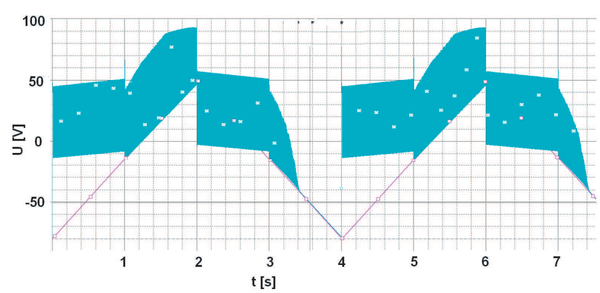


Figure 6: A floating supply electrometric amplifier: an instance of oscillation.

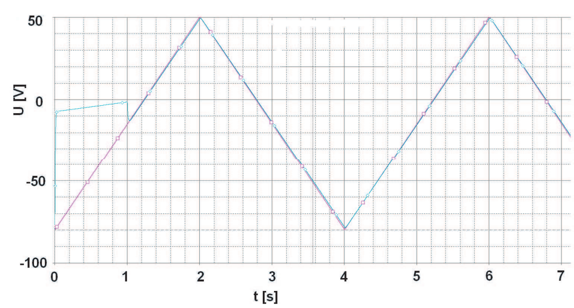


Figure 7: A frequency compensated amplifier.

A diagram of this type of electrometric amplifier has been designed and simulated using Spice. In the design, an OZ LMC6041 was utilized as an electrometric amplifier; typically, its input current is 2 fA. Fig. 6 presents the overall diagram. Operational amplifiers having the input voltage of 300 V are generally not available, therefore we built a high-voltage amplifier based on discrete hv transistors. Owing to the connection sensitivity to oscillation, it is necessary to use correct values of capacitor C_1. With respect to the maximum input resistance, the amplifier does not have input protection. This problem is solved through the application of RC filter(s) R_{10} and C_2. The filter restrains voltage spikes at the input and the amplifier is capable of monitoring the changes occurring at its own input.

Figure 7 provides an example of possible oscillations: the input voltage is shown as the violet course, while the output voltage pertains to the dark green course. Fig. 8 illustrates the situation following compensation. At the beginning of the measurement, the amplifier input must be short-circuited in order to facilitate stabilization of the initial conditions. Simulation in the Pspice environment indicates input resistance at 10^{14} ohm. The amplifier input current is markedly represented by the charging of capacitor C_2 during the input voltage changes.

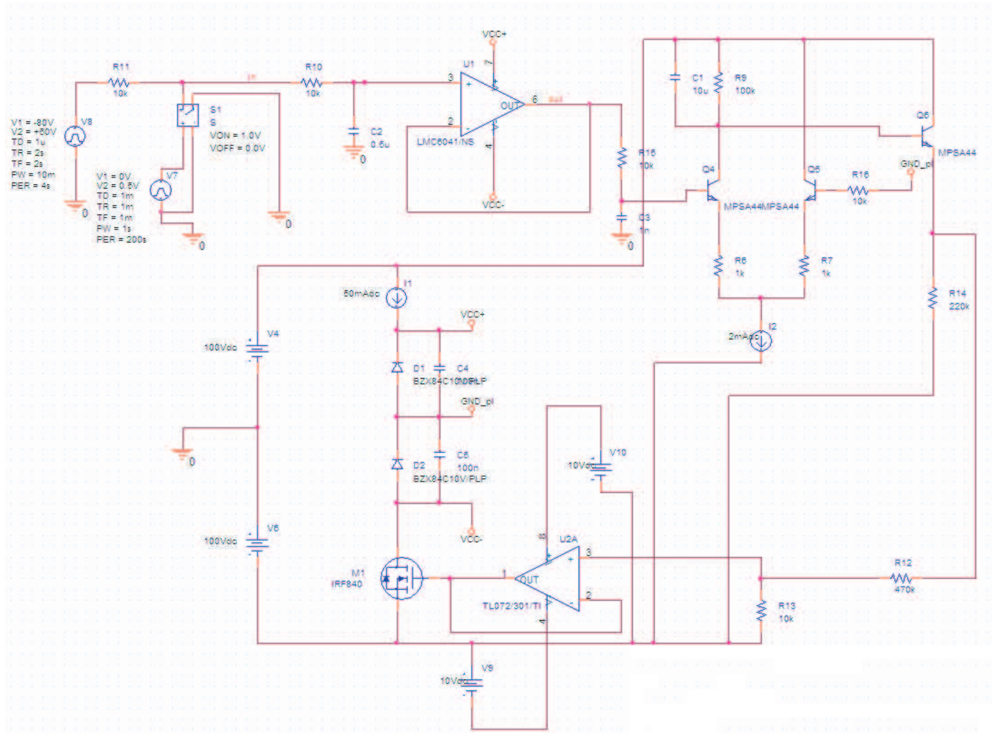


Figure 8: Concrete diagram of a floating supply electrometric amplifier (after compensation).

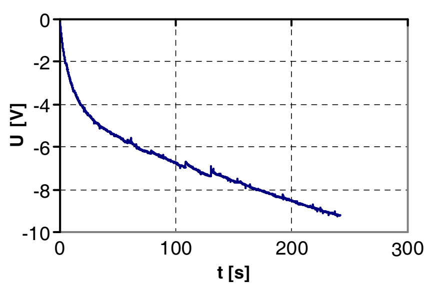


Figure 9: Only the charging of spurious capacities.

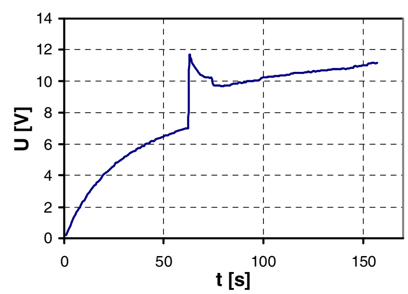


Figure 10: The freezing potential with a parasitic exponential.

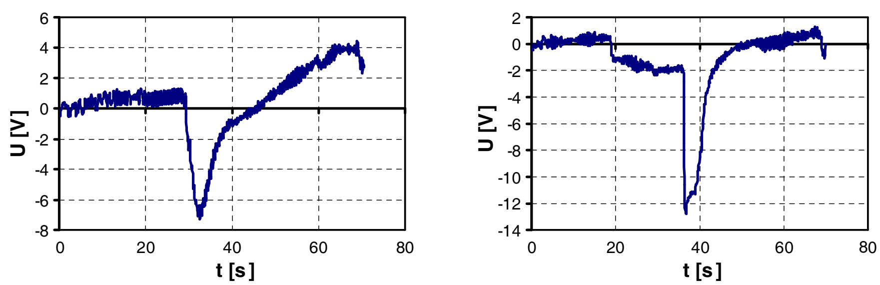


Figure 11: Freezing potential with the elimination of spurious capacities charging for two instances of measurement (the second measurement of refrigeration commencement at sec. 30).

5. THE MEASURED DATA

Using a special electrometric amplifier as well as measurement apparatuses, we measured the potentials of chemical solutions. At the initial stage of the experiments, the measurement was degraded by an electric charge in certain parts of the measuring apparatus. The effects on the concerned parts manifested themselves adversely during the experiment evaluation. Voltage surge caused by the freezing of the solution occurred non-repestedly and its amplitude showed different characteristics. The situation is described in Fig. 9, which shows the charging of spurious capacities. Fig. 10 shows the freezing potential added to an erroneous signal from the electric charge of structural parts of the apparatus.

6. CONCLUSION

A special electrometric amplifier has been designed and frequency-compensated. The amplifier is capable of performing measurement in the order of hundreds of volts with a large input resistivity. By the help of the measuring apparatus, we measured potential at the interface of the sample phase change. The measured data show that this potential can be repeatedly measured using a structurally modified apparatus.

ACKNOWLEDGMENT

The work described in the paper was financially supported by the research project GA102/09/0314, GACR 203/09/P445, research plan MSM 0021630513, and project of the BUT Grant Agency FEKT-S-10-13.

REFERENCES

1. Sola, M. I. and H. R. Corti, "Freezing included electrical potentials and Ph charges in aqueous electrolytes," *An. Asoc. Quím Argent.*, Vol. 81, No. 6, 483–498, 1993.
2. Lefebre, V., "The freezing potential effect," *J. Colloid Interfacing Sci.*, Vol. 25, No. 2, 263–269, 1967.
3. Bronshteyn, V. L. and A. A. Chernov, "Freezing potentials arising on solidification of dilute aqueous solutions of electrolytes," *J. Crystal Growth*, Vol. 112, No. 1, 129–145, 1991.
4. Robinson, C., C. S. Boxe, M. I. Guzmán, A. J. Colussi, and M. R. Hoffmann, "Acidity of frozen electrolyte solutions," *J. Phys. Chem. B*, Vol. 110, No. 15, 7613–7616, 2006.
5. Parameswaran, V. R., C. R. Burn, A. Profir, and Q. Ngo, "A note on electrical freezing and shorting potentials," *Cold Regions Science and Technology*, Vol. 41, No. 2, 83–89, 2005.
6. Keithley, *Low Level Measurements Handbook: Precision DC Current, Voltage, and Resistance Measurements*, 6th Edition, Keithley Instruments, Inc., Cleveland, Ohio, 2004.

Analytical Expressions of Diffraction' Free Beams Obtained by Diffraction on an Opaque Disk

Q. Huang[1, 2], **S. Coëtmellec**[1], **A. Louis**[3], **F. Duval**[3], **H. Leblond**[4], **and M. Brunel**[1]

[1]UMR 6614 CORIA CNRS, Université de Rouen, France
[2]Xidian University, Xi'an, China
[3]IRSEEM, ESIGELEC, Saint-Etienne du Rouvray, France
[4]LΦA EA 4464, Université d'Angers, France

Abstract— We establish approached analytical expressions that show that the beam produced after diffraction of a gaussian beam by an opaque disk and collimation by a lens can be described by a diffraction-free J_0 Bessel function. We further show that a similar analytical expression can be established in the case of femtosecond pulses.

1. INTRODUCTION

Diffraction-free beams have attracted much attention since pioneer work of Durnin and co-workers [1]. Different beam-shaping techniques exist, based on the use of an annular aperture, a computer generated-hologram or an axicon [2]. Although an annular aperture induces many losses, the technique based on this element has recovered much interest in view of recent studies involving sub-wavelength apertures [3]. They could indeed allow the design of systems in sub-wavelength optics.

We have recently proposed another method based on the occultation of an incident beam. We could indeed demonstrate that diffraction of a Gaussian beam by an opaque disk leads to the generation of a diverging Bessel'like beam, that can then be collimated with a lens into a diffraction-compensated beam [4, 5]. Numerical developments could corroborate experimental results. They showed that the beam diffracted by the opaque disk can be expressed as a sum of approximately 20 Bessel functions of odd orders. Unfortunately, these expressions are so complex that they do not allow to establish any simple expression of the diffraction-compensated beam after the collimating lens. Numerical integration is required to obtain correct simulations. We present in this communication a simplest formulation. After some simplifications that are detailed expressly, the beam is shown to be correctly described by a zeroth-order Bessel diffraction-free beam. We further show that this expression can be generalized to the case of 100 fs incident pulses.

2. DIFFRACTION BY AN OPAQUE DISK

We consider the diffraction of a Gaussian beam by an opaque disk, as described in Figure 1. A CW laser beam is first filtered through a pinhole. The filtered beam, which is a quasi-gaussian beam, is then focused with a lens L_1 An opaque disk whose diameter D is 300 µm is positioned just before the focus point. It is well centered on the optical axis, but behind the focus point. A lens L_2 collimates the beam diffracted by the opaque disk. The opaque disk is exactly positioned at the focus point of the lens L_2. The diffracted pattern is observed with a CCD camera.

According to the paraxial approximation, the expression of the scalar electric field diffracted by the opaque disk is given at distance z after the disk by:

$$u(x,y,z)=-\iint\limits_{disk\ screen}\frac{2\,i\,k}{4\pi(z-z_q)}e^{-\frac{x_q^2+y_q^2}{w(z_q)^2}}e^{-ik\frac{x_q^2+y_q^2}{2R(z_q)}}e^{\left(ik(z-z_q)+\frac{(x-x_q)^2+(y-y_q)^2}{2(z-z_q)}\right)}\,dx_q\,dy_q$$

$w(z_q)$ and $R(z_q)$ represent the waist and the beam radius of the incident Gaussian beam, in the plane of the opaque disk. After some developments, this integral can be written

$$u(x,y,z)=-\frac{2ikD^2e^{ik(z-z_q)}e^{-a(z)}}{8(z-z_q)}e^{i\frac{k\rho^2}{2(z-z_q)}}\sum_{m=0}^{+\infty}(-1)^m\left(\frac{1}{2a(z)}\right)^{m+1}\left(\frac{k\rho D}{2(z-z_q)}\right)^m J_m\left(\frac{k\rho D}{2(z-z_q)}\right)\quad(1)$$

with: $a(z)=\frac{D^2}{4}(\frac{1}{w(z_q)^2}-ik(\frac{1}{2(z-z_q)}-\frac{1}{2\,R(z_q)}))$.

Figure 2 shows the comparison between an experimental profile, the theoretical development of reference [4] and the present expansion considering only the coefficient $m=0$ (first-order approximation). For these experimental results, the laser source is a CW He-Ne laser emitting at

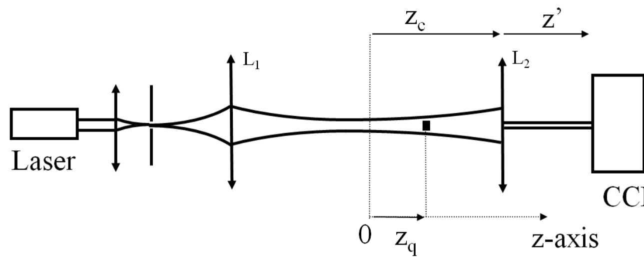


Figure 1: Experimental set-up.

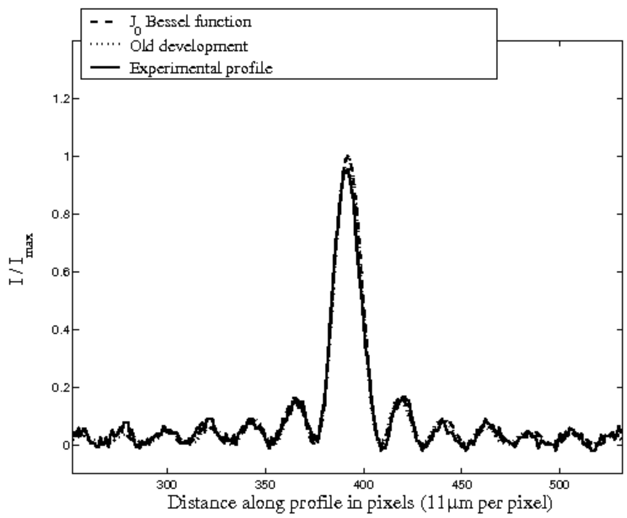


Figure 2: Comparison between the experimental profile, the new development limited to the sole J_0 Bessel function, and the old development of reference [4] which required the sum of 20 Bessel functions of odd orders.

632.8 nm. The focal length of L_1 is 10 cm. Focus point is 15 cm after the lens L_1. The diameter of the focus point is 70 µm at $1/e$. The opaque disk is 3.5 mm behind the focus point. The diffracted pattern is observed with a CCD camera positioned 10.5 cm after the opaque disk (Lens L_2 is not present in this first experiment). We can see that the first-order approximation using only a J_0 Bessel function allows a good fit of experimental results. Expression (1) can thus be reasonably simplified by the zeroth-order Bessel function in paraxial approximation. So, we get

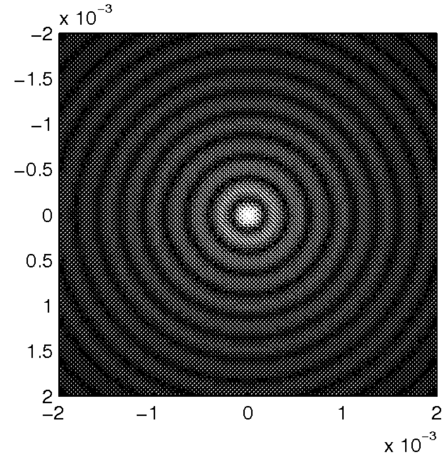
$$u(x,y,z) = -\frac{2ikD^2 e^{ik(z-z_q)}}{8(z-z_q)} e^{i\frac{k\rho^2}{2(z-z_q)}} \frac{e^{-a(z)}}{2a(z)} J_0\left(\frac{k\rho D}{2(z-z_q)}\right) \tag{2}$$

with $\rho = \sqrt{x^2+y^2}$.

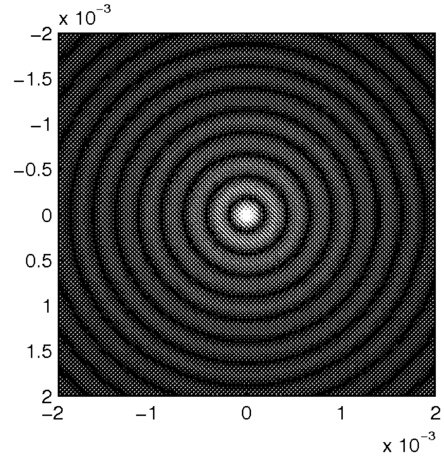
Note that the expansion limited to the sole J_0 Bessel function does never diverge. For comparison, expansions of higher orders would be more precise in their domain of convergence, but they would be limited to a domain of convergence given by $\rho < 4|a|(z-z_q)/(kD)$. In this case, the limit is $\rho < 2.5$ mm.

3. DIFFRACTION FREE BEAMS AFTER COLLIMATION BY A LENS

let us now consider collimation by the lens as detailed in Figure 1. On the plane $z = z_c$ where the collimating lens is located, the scalar field is reduced to the zeroth-order Bessel function in paraxial approximation. Propagation through the lens is then described by a transmitting phase factor, while propagation in free space after the lens can be expressed with a Fresnel transform. Assuming that the radius R of the lens L_2 is very large, these operations lead to a very simple result. If the opaque disk is located just on the focus point of the lens L_2, that is $f_{\{L2\}} = z_c - z_q$, The amplitude of the scalar field on the plane z' after the collimating lens L_2 can indeed be written:



(a)



(b)

Figure 3: Comparison between (a) the old development of reference [4] which required the sum of 20 Bessel functions of odd orders and (b) the new development limited to the sole J_0 Bessel function.

$$u(x', y', z') = -\frac{2ikD^2 e^{ik(z'-z_q)} e^{-a(z_c)}}{16a(z_c)(z_c - z_q)} e^{-i\frac{k(z'-z_c)D^2}{8(z_c-z_q)^2}} J_0\left(\frac{k\rho' D}{2(z_c - z_q)}\right) \text{ with } \rho' = ({x'}^2 + {y'}^2)^{1/2''} \quad (3)$$

To illustrate this, Figure 3 shows a comparison between the beam patterns predicted 30 cm after the collimating lens using the old expansion of reference [4] over 20 Bessel functions of odd orders (Figure (a)), and the pattern predicted using the first-order J_0 Bessel function approximation (Figure (b)). Other parameters are the opaque disk diameter $D = 300\,\mu\text{m}$, $z_q = 3.5\,\text{mm}$, and $f_{\{L2\}} = 25\,\text{mm}$. Although they are not reported here, transverse profiles show a good quantitative accordance between experimental profiles and both theoretical developments. The J_0 Bessel approximation (3) is thus a good approached expression of the nondiffracting beam and is theoretically justified.

4. FEMTOSECOND DIFFRACTION FREE PULSES

Femtosecond pulses exhibit a large spectrum and the previous model cannot be used anymore. We consider ultrashort pulses with constant waist width under the paraxial approximation. Their propagation can be described by a combination of Fresnel diffraction for each spectral component, and a temporal filter for proper superposition of the components [6]. In the temporal Fourier domain, we can write $\tilde{E}(\vec{r}, \omega) = \tilde{U}(\vec{r}, \omega)\tilde{G}(\omega)$ where $\tilde{E}(\vec{r}, \omega)$ is the Fourier transform of the diffracted field $E(\vec{r}, t)$, and $\tilde{G}(\omega) = \exp(-(\omega - \omega_0)^2/\Delta\omega^2)$ is the spectrum of the pulses. By using the Parseval theorem, we obtain the diffraction intensity:

$$I(\rho, z) = \frac{C}{16(z - z_q)^2} e^{-\frac{D^2}{2w(z_q)^2}} \int_{-\infty}^{+\infty} \left|\tilde{U}(\vec{r}, \omega)\right|^2 \left|\tilde{G}(\omega)\right|^2 d\omega$$

where C is a normalization constant. It gives:

$$I(\rho, z) = \frac{Cw(z_q)^4}{16(z - z_q)^2} e^{-\frac{D^2}{2w(z_q)^2}} \int_{-\infty}^{+\infty} \frac{1}{1 + \tau^2\omega^2} e^{-\left(\frac{\omega - \omega_0}{\Delta\omega}\right)^2} J_0^2(\alpha\omega) d\omega$$

with $\alpha = \rho D/(2c(z - z_q))$ and $\tau = w(z_q)^2/(2c)((z - z_q)^{-1} - R(z_q)^{-1})$. Typical values of the parameters are $\alpha \approx 10^{-14}\,\text{s}$, $\tau \approx 10^{-16}\,\text{s}$, $\omega_0 \approx 3000\,\text{THz}$, $\Delta\omega \approx 30\,\text{THz}$. Introducing the new variable $x = (\omega - \omega_0)/\Delta\omega$, it is thus possible to expand the integral of previous relation in a series of powers of $\Delta\omega$. Simulations are presented in Figure 4. They show simulations using first-order, third-order and fifth-order Taylor–type expressions and the old integration of reference [6] in the case of 100 fs pulses. All simulations deliver very similar results in our case, because the spectrum width is relatively small. It appears that the first-order expansion is sufficient. Assuming then that the lens radius is very large, the expression of the intensity after the collimating lens can be evaluated and after some steps, it can be approached by:

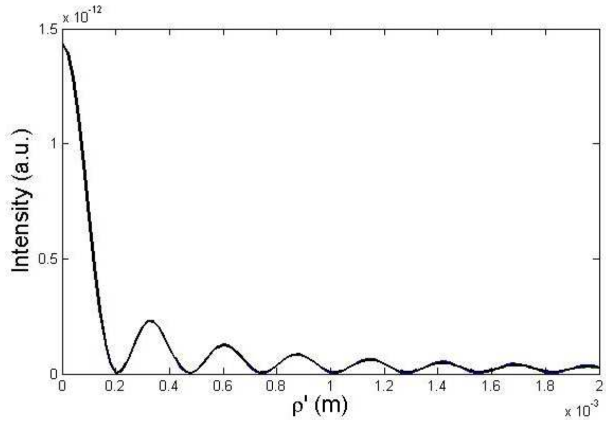


Figure 4: Comparison between first-order, third-order and fifth-order Taylor-type expressions and old numerical treatment of reference [6]. All curves are superposed.

$$I(\rho', z') = \frac{Cw(z_q)^4}{16(z_c - z_q)^2} e^{-\frac{D^2}{2w(z_q)^2}} \frac{\Delta\omega\sqrt{\pi} J_0^2(\beta\omega)}{1 + \tau^2\omega_0^2} \quad \text{with} \quad \beta = \rho' D/(2c(z_c - z_q)).$$

This relation shows the diffraction free nature of the beam. This expression represents a very important simplification of previous procedures which required long numerical developments as detailed in reference [6].

5. CONCLUSION

In conclusion, we have established approached analytical expressions that show that the beam produced after diffraction of a gaussian beam by an opaque disk and collimation by a lens can be described by a diffraction-free J_0 Bessel function. We have further showed that a similar analytical expression can be established in the case of femtosecond pulses. Those simple analytical relations allow the design of systems in different wavelength regions, particularly in the microwave regime, while very long computing times were necessary using previous numerical developments.

ACKNOWLEDGMENT

This work is supported by Carnot ESP.

REFERENCES

1. Durnin, J., J. J. Miceli, and J. H. Eberly, "Diffraction free beams," *Phys. Rev. Lett.*, Vol. 58, 1499–1501, 1987.
2. McGloin, D. and K. Dholakia, "Bessel beams: Diffraction in a new light," *Contemp. Phys.*, Vol. 46, 15–28, 2005.
3. Li, Z., K. Boratay Alici, H. Caglayan, and E. Ozbay, "Generation of an axially asymmetric Bessel-Like beam from a metallic subwavelength aperture," *Phys. Rev. Lett.*, Vol. 102, 143901, 2009.
4. Brunel, M. and S. Coëtmellec, "Generation of non diffracting beam through an opaque disk," *J. Opt. Soc. Am. A*, Vol. 24, 3753–3761, 2007.
5. Brunel, M., D. Mgharaz, and S. Coëtmellec, "Generation of diffraction-compensated beams with a pigtailed laser diode," *IEEE Photon. Technol. Lett.*, Vol. 20, 742–744, 2008.
6. Brunel, M., D. Mgharaz, and S. Coëtmellec, "Generation of femtosecond diffraction-compensated beam through an opaque disk," *Opt. Express*, Vol. 16, No. 14, 10390–10397, 2008.

Rescaled Range Analysis of ELF Natural Electromagnetic Noise from Antarctica

A. Salinas[1], **S. Toledo-Redondo**[1], **J. A. Morente**[2], **J. A. Portí**[2], **E. A. Navarro**[3], **A. Méndez**[1], **J. F. Fornieles**[1], **M. Rodríguez-Sola**[2], **N. Novas**[4], **J. A. Gázquez**[4], **and R. García-Salvador**[4]

[1]Department of Electromagnetism and Matter Physics, University of Granada, Granada 18071, Spain
[2]Department of Applied Physics, University of Granada, Granada 18071, Spain
[3]Department of Applied Physics, University of Valencia, Burjassot, Valencia 46100, Spain
[4]Department of Computer Architecture and Electronics, University of Almería
La Cañada de San Urbano 04120, Almería, Spain

Abstract— We present the results of a statistical rescaled range (R/S) analysis of natural electromagnetic noise in the extremely low frequency (ELF) band. The Hurst exponent was derived from the records of two horizontal magnetic antennas taken in Antarctica with a sample frequency of 512 Hz, to guarantee enough resolution with the small number of samples in the analysis. The study shows that ELF radio noise is a persistent random process, maintaining its value between 0.75 and 0.85 for the different series analyzed. A numerical simulation was performed, in which each discharge was modeled by the radiated field of a wire antenna excited by a Gaussian pulse, demonstrating a way of inferring lightning rates from R/S analysis.

1. INTRODUCTION

Fractals are uncommon mathematical objects defined in an imprecise way that shows auto-resemblance, i.e., one part is similar to the whole. The auto-resemblance concept is equivalent to scale invariance. Fractals are interesting because many natural phenomena have auto-resemblance. Traditional geometry does not account for the properties of these shapes. The mathematician Benoît Mandelbrot, known for his work with fractals, defines a fractal set as a collection having a fractal dimension higher than its topological dimension. The Hurst exponent, which is involved in R/S calculations, relates to the fractal dimension [1].

Natural electromagnetic noise in the ELF band produced by electrical storms on Earth can be studied in terms of the Hurst exponent by employing R/S analysis. This technique was developed by the engineer Harold Edwin Hurst in order to study the Nile floods, with the aim of optimizing water retention. The method is fully described in [2, 3].

In this work we present the R/S analysis of three records of both (North-South and East-West) horizontal components of the magnetic field, taken between January and February 2008 in Antarctica (where the anthropogenic noise should be low), employing the magnetotelluric method [4].

2. R/S ANALYSIS

R/S analysis starts from a digital series of 2^N samples, which in our case is a magnetic field time series. For each value of the index i ($1 \leq i \leq N$) we can define a partition of the series, which consists of splitting it into data blocks of $M = 2^i$ length. The average is then calculated from each partition i for each data block j:

$$\langle x\rangle_{i,j} = \frac{1}{M}\sum_{l=1}^{M} x_{(j-1)\cdot M+l}, \tag{1}$$

to obtain the sum (integral) of the time series with the average value taken

$$y_j(k) = \sum_{l=1}^{k} (x_j(l) - \langle x\rangle_{i,j}). \tag{2}$$

The range and standard deviation is also calculated from each partition i for each data block j:

$$R_{i,j} = \max_{l\leq k\leq M}(y_j(k)) - \min_{l\leq k\leq M}(y_j(k)) \tag{3}$$

$$S_{i,j} = \sqrt{\frac{1}{M}\sum_{l=1}^{M}(x_j(l) - \langle x\rangle_{i,j})^2}. \tag{4}$$

The average value of $R_{i,j}$ and $S_{i,j}$ can be calculated for each partition i, and $Z[M]$ is defined as the quotient between them:

$$Z[M] \equiv \frac{R_M}{S_M} = \frac{\langle R_{i,j}\rangle}{\langle S_{i,j}\rangle}. \tag{5}$$

The Hurst exponent is defined from the empiric relation between the coefficient Z and the number of samples in the data block by following a power relation which indicates scale invariance:

$$Z = \left(\frac{M}{2}\right)^{Hu} \tag{6}$$

Hu being the Hurst exponent (also called K by H. E. Hurst and later H by B. Mandelbrot). R/S analysis tries to find a power relation between Z and the number of samples, as illustrated in Equation (6). By representing $\log_2(Z(M))$ against $I = \log_2(M)$ we obtain a straight line whose slope determines the Hurst exponent. The Hurst exponent takes values in the interval (0.7, 0.8) for many geophysical processes [5, 6]. Due to the complexity of natural electromagnetic phenomena and of those involving different scales in general [7], it is convenient to define different Hurst exponents for different time windows, i.e., for different numbers of samples.

We can think of a time series as the sum of discontinuities, tendencies, periodic components and a stochastic component. The last of these would account for all the effects not included in those listed previously. One important fact to ascertain about a time series is whether its behavior is persistent, random or anti-persistent. If neighboring samples are not correlated then the stochastic component is random and the Hurst exponent takes values of $Hu = 0.5$. If the correlation is positive then the local fluctuations are lower than the average fluctuations and the behavior is persistent. If the correlation is negative then the neighbor values tend to be far apart and the behavior is anti-persistent [3]. In summary, if $0.5 < Hu < 1$ then the series is persistent and if $0 < Hu < 0.5$ then the series is anti-persistent.

The time series employed in this study corresponds to a sample frequency of 512 Hz and the number of samples is $2^{18} = 262144$, which corresponds to roughly 9 minutes of recording. The measurements were taken at three different sites, called S_{22}, S_{23}, and S_{24}. We obtained the Z coefficient as a function of the number of samples in the data block, as can be seen in Figure 1.

The Z coefficient shows a common behavior for the six series at values of $I \leq 6$. At around $I = 6$ we can observe a change in the slope and a major difference in the Z values. In order to obtain the Hurst exponent in different time scales a lineal adjustment has been made, by using earlier and later points for each adjustment. The results can be seen in Figure 2. For the interval $I =$ 3–5, the Hurst exponent falls between 0.75–0.85, depending on the component and the site. For a time scale of $I =$ 6–9, the exponents take values around 0.5, which correspond to random

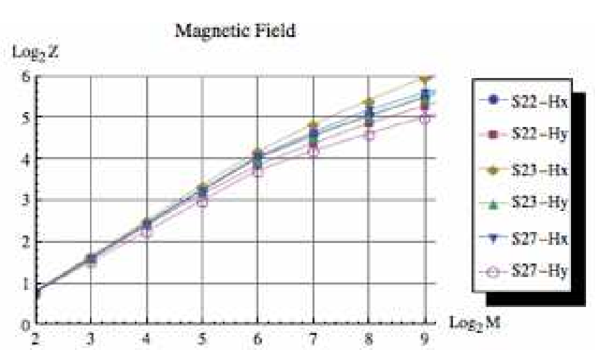


Figure 1: $\log(Z)$ against $I = \log(M)$ representation for the horizontal components of the magnetic field at three different sites.

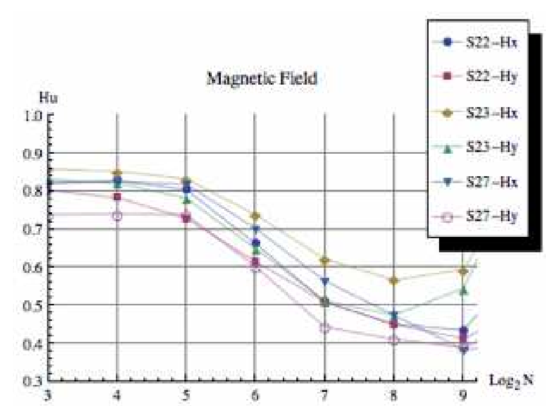


Figure 2: Hu against $I = \log(M)$ representation for the horizontal components of the magnetic field at three different sites.

behavior. When the number of samples increases, i.e., for $I > 9$, the Hurst exponent goes beyond unity (not shown in the figure). This happens because as the number of samples increases we have fewer data blocks to average.

R/S analysis is sensitive to the internal structure of the noise and could therefore be a tool for finding its different sources and the specific processes involved in its generation. The statistical treatment of the time series is not hard to implement; the difficulty here is in establishing the connection between the results and the underlying physical processes. In [5], a possible relation between global lightning activity (i.e., the global lightning rate) and the number of samples in which Hu abandons its persistent behavior (see Figure 2) is indicated. This relation can be expressed as:

$$r = \frac{f_s}{m_k} \tag{7}$$

where r is the average number of lightning strokes per second, f_s is the sample frequency and m_k is the number of samples in which Hu starts to decrease towards 0.5 (knee of the curve). From the results depicted in Figure 2 we can establish a global lightning rate of 16–32 strokes per second, because the sample frequency is 512 Hz. This result is in agreement with other observations [8].

In order to corroborate the empirical law expressed in Equation (7), a numerical simulation has been performed. The field generated by a stroke has been substituted by the response of a wire dipole antenna to a Gaussian pulse excitation. Hallén's equation has been solved for the spectral content by using the Method of Moments (MoM) to obtain the distribution of charge and current in the antenna. From these distributions we can obtain the temporal fields by using the Inverse Fourier Transform, and by applying Jefimenko's expressions we can obtain the electric and magnetic fields at any point in the space [9]. We have taken the electric field measured in the equatorial plane at a distance of five times the length of the antenna. This is a resonant structure and the radiated field presents resonances (see Figure 3), like the electromagnetic noise in the Earth-ionosphere cavity in the ELF band.

Each "stroke" of our numerical simulation is composed of 774 samples and the "natural electromagnetic noise" is obtained by superposing strokes with a fixed time interval in between, giving a final signal composed of 2^{16} samples. The lightning rate is selected according to the number of samples inserted between two discharges. For this study, three different signals were created, with a distance between discharges of 16, 64 and 128 samples, which correspond to 1/16, 1/64 and 1/128 strokes per second respectively (we are taking $f_s = 1\,\text{Hz}$).

R/S analysis has been applied to these signals. In Figure 4, we can observe the behavior of $\log_2 Z$ as a function of $I = \log_2 N$. The legend of the figure indicates the number of samples inserted between strokes and represents, as previously stated, the inverse of the lightning rate. So the curve labeled as 16 presents a constant value for Hu until $I = 4$, where the knee appears and its value changes abruptly to 0. As stated in Equation (7), the lightning rate must be 1/16, which coincides with its theoretical value. The same happens for the other two synthetic signals, where the knee appears at $I = 6$ and $I = 7$, which means rates of 1/64 and 1/128 lightning per second respectively.

The signals recorded in Antarctica present similar lightning rates and we cannot differentiate between them regarding the location of the knee. It would be necessary to record signals at maximum and minimum times of lightning activity in order to see the differences with the method presented.

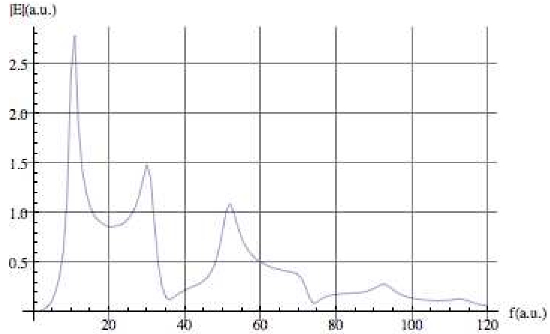


Figure 3: Electric field spectrum radiated by a wire antenna, measured at its equatorial plane.

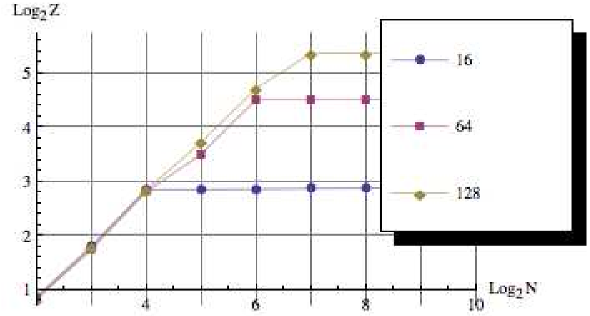


Figure 4: R/S analysis applied to the numerical simulation.

3. CONCLUSION

In this study, the R/S method has been applied to ELF-band (512 Hz sample frequency) electromagnetic noise, measured in Antarctica during January and February 2008. The Hurst exponent obtained shows the persistent behavior of the process for the scale between 2^2 and 2^5, and random behavior between 2^6 and 2^9. The number of samples where the transition occurs seems to be related to the lightning rate in the atmosphere, as shown in a numerical simulation.

ACKNOWLEDGMENT

This work was partially supported by the Ministerio de Educación y Ciencia of Spain and Consejería de Innovación, Ciencia y Empresa of the Andalusian Government under projects with references FIS2007-63293, PO7-FQM-03280, POL2006.13836.C03.01, and CGL2009-07721, co-financed with FEDER funds of the European Union.

REFERENCES

1. Feder, J., *Fractals*, Plenum Press, New York, 1988.
2. Hurst, H. E., R. P. Black, and Y. M. Simaika, *Long-term Storage: An Experimental Study*, Constable, London, 1965.
3. Turcotte, D. L., *Fractals and Chaos in Geology and Geophysics*, Cambridge University Press, New York, 1997.
4. Cagniard, L., "Basic theory of the magnetotelluric method of geophysical prospecting," *Geophysics*, Vol. 18, No. 3, 605–635, 1953.
5. Nickolaenko, A. P., C. Price, and D. D. Iudin, "Hurst exponent derived for natural terrestrial radio noise in Schumann resonance band," *Geophysics. Res. Lett.*, Vol. 27, No. 19, 3185–3188, 2000.
6. Nickolalenko, A. P. and M. Hayakawa, *Resonances in the Earth-ionosphere Cavity*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
7. Mandelbrot, B. B. and J. R. Wallis, "Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence," *Water Resour. Res.*, Vol. 5, 967–988, 1969.
8. Christian, H. J., R. J. Blackslee, D. J. Boccippio, W. L. Boeck, D. E. Buechler, K. T. Driscoll, S. J. Goopdman, J. M. Hall, W. J. Koshakl, D. M. Mach, and M. F. Stewart, "Global frequency and distribution of lightning as observed by the optical transient detector (OTD)," *Proc. of 11th Conference on Atmospheric Electricity*, 726, Guntersville, Alabama, USA, June 7–11, 1999.
9. Jackson, J. D., *Classical Electrodynamics*, 3rd Edition, Wiley University Press, New York, 1999.

Hybrid Method to Compute the Magnetic Field in Bird Cage Coil for a Magnetic Resonance Imaging System

N. Benyahia[1], **and M. E. Latreche**[2]

[1]Department of Electrical Engineering, University of Biskra, Algeria
[2]Department of Electro technical, University of Constantine, Algeria

Abstract— In this paper we present a simple method to solve the Biot-Savart formula by calculation the magnetic field created at each point (M) in the three dimension (3D) mesh of the study area by a hybrid method.

1. INTRODUCTION

Radiofrequency (RF) coils are key components in a magnetic resonance imaging (MRI) system. They serve two proposes. The first is to generated RF pulse at the Larmor frequency to excite the nuclei in the object to be imaged. The second is to pick RF signals emitted by the nuclei at the same frequency. To obtained high quality MRI images, RF coils must able to produce a homogenous field in the volume of the interest at the Larmor frequency, so that the nuclei can be excited uniformly [1].

Over two decade, ago, birdcages resonators are most popular, because they generate a very homogenous field over a large volume within the coils.

There are many computational methods which have been published for calculating the magnetic field of the birdcage resonator [2, 3]. These cases have been treated by evaluating the magnetic field analytically in terms of functions such as complete or incomplete elliptic integrals of first and second types [4].

So in this work we purpose to compute the magnetic field which the sum of magnetic fields produced by the end rings and legs of RF resonator at each point in the 3D mesh study area.

2. FORMULATION

To compute the magnetic field produced by current I passing through the surface bounded by C, we can apply Biot-Savart's law specifically the magnetic field at $M(x, y, z)$ is given by:

$$\vec{B}(M) = \int_C \frac{\mu_0 I}{4\pi} \frac{\vec{dl} \times \vec{R}}{\left|\vec{R}\right|^3} \tag{1}$$

with:

$$\vec{R} = \vec{r} - \vec{r_0} = R_x\vec{i} + R_y\vec{j} + R_z\vec{k} \tag{2}$$

where μ_0 is the free space permeability, r_0 and r are position vectors of the source point $P(x_0, y_0, z_0)$ and field point $M(x, y, z)$, respectively, $\vec{dl}$ is a vector differential line element along the birdcage coils conductors and is given by:

$$\vec{dl} = dlx_0\vec{i} + dly_0\vec{j} + dlz_0\vec{k} \tag{3}$$

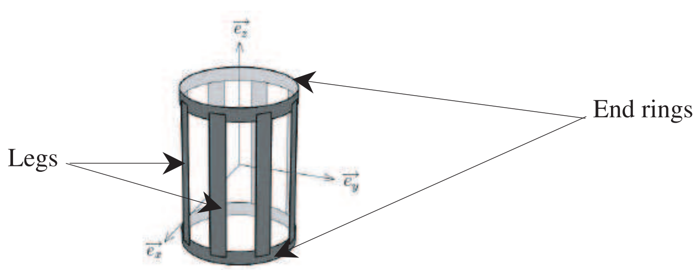


Figure 1: Illustration of Birdcage coil.

And $(\vec{i}, \vec{j}, \vec{k})$ are the unit vectors of the Cartesian coordinate.

After, development we obtain the magnetic field components from (1), (2) and (3):

$$\begin{cases} B_x(x,y,z) = \dfrac{\mu_0 I}{4\pi} \sum \left(\dfrac{R_z dl_y - R_y dl_z}{\left|\vec{R}\right|^3} \right) \\ B_y(x,y,z) = \dfrac{\mu_0 I}{4\pi} \sum \left(\dfrac{R_x dl_z - R_z dl_x}{\left|\vec{R}\right|^3} \right) \\ B_z(x,y,z) = \dfrac{\mu_0 I}{4\pi} \sum \left(\dfrac{R_y dl_x - R_x dl_y}{\left|\vec{R}\right|^3} \right) \end{cases} \tag{4}$$

The current in legs was distributed according to Equation (5) which gives a highly homogeneous magnetic field pattern inside the coil [4, 5]:

$$I_j = I_0 \exp\left(ik\theta_j\right) \tag{5}$$

Here I_0 is the current amplitude, $k = 1, \ldots, N-1$, and $\forall 1 \leq j \leq N$, $\theta_j = 2\pi(j-1)/N$.

Using expression (4), (5) and geometry proprieties we can calculate:

$$\begin{cases} \vec{B_{ah}}(x,y,z) = B_{ahx}\vec{i} + B_{ahy}\vec{j} + B_{ahz}\vec{k} \\ \vec{B_{ab}}(x,y,z) = B_{abx}\vec{i} + B_{aby}\vec{j} + B_{abz}\vec{k} \\ \vec{B_{br}}(x,y,z) = B_{brx}\vec{i} + B_{bry}\vec{j} + B_{brz}\vec{k} \end{cases} \tag{6}$$

where: B_{ah}, B_{ab} and B_{br} are the magnetic field produced by the high end ring, law end ring and N legs, respectively.

So the magnetic field generated by a birdcage resonator is the sum of all magnetic field produced by end rings and legs in 3D mesh study area:

$$\vec{B_T}(x,y,z) = (B_{ahx} + B_{abx} + B_{brx})\vec{i} + (B_{ahy} + B_{aby} + B_{bry})\vec{j} + (B_{ahz} + B_{abz} + B_{brz})\vec{k} \tag{7}$$

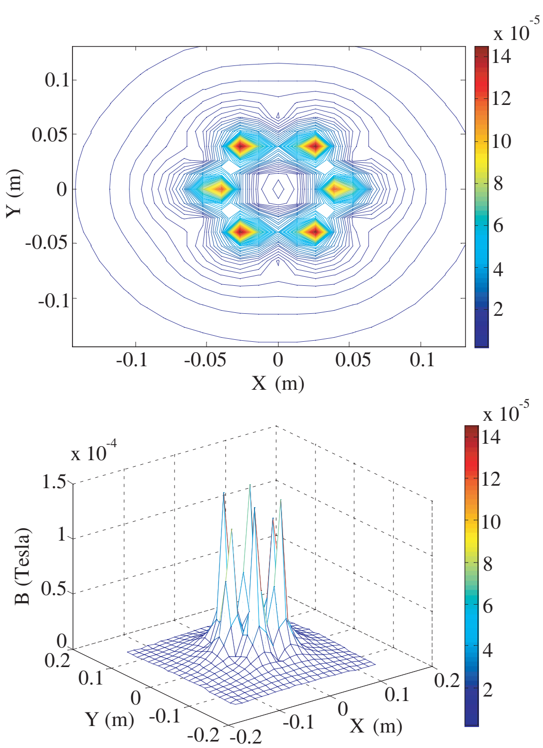


Figure 2: Contour and 3D plots of magnetic field for mode 1.

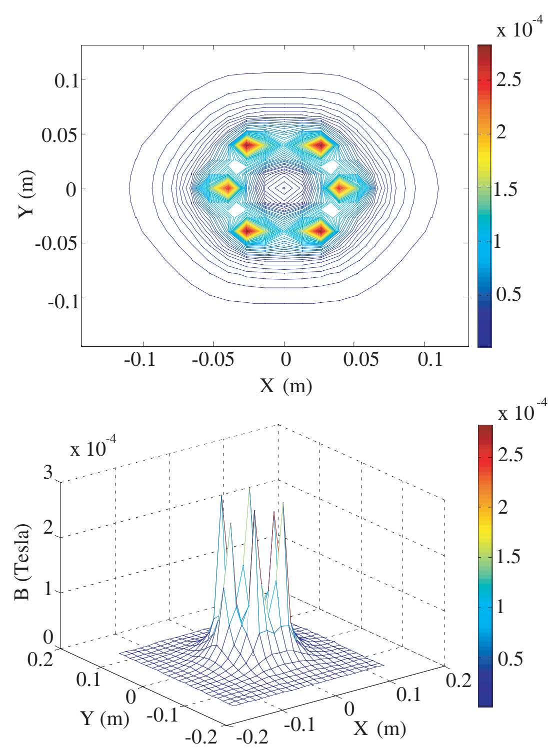
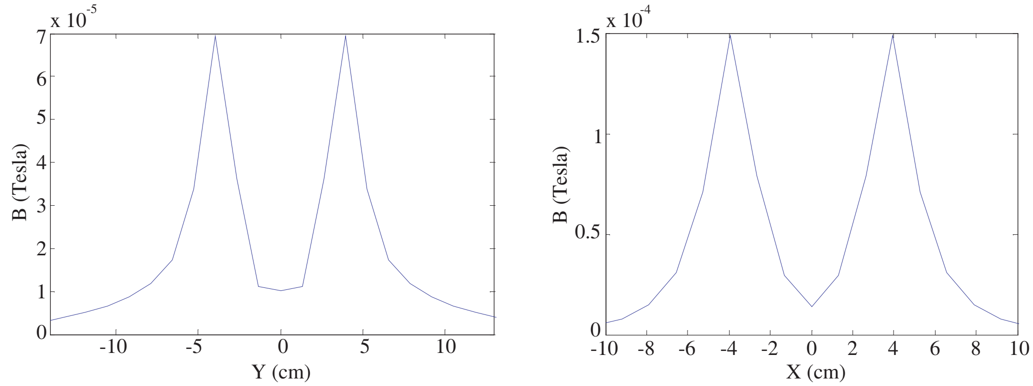


Figure 3: Contour and 3D plots of magnetic field for mode 3.



(a) (b)

Figure 4: Magnetic field for (a) mode 1 and (b) mode 3.

So:

$$B_T(x, y, z) = \sqrt{B_{Tx}^2 + B_{Ty}^2 + B_{Tz}^2} \tag{8}$$

3. RUSULTS AND DISCUSSION

The diagrams in Figs. 2, 3, and 4 refer to the magnetic field patterns relevant to the birdcage coil with $N = 6$ legs, length $= 12.8$ cm, diameter $= 8.9$ cm.

The diagrams in Figs. 2 and 3 refer to the magnetic field patterns in the plan perpendicular to the birdcage axis in its middle.

In order to verify the magnetic field homogeneity produced inside the coil, we plot the mode 1 field and the mode 3 along the y axis on Fig. 4 we observe that only the field associated to the mode 1 is not zero at the center of the birdcage coil. It is the only one with energy and homogeneity property in the center. That is why it is in particular the one used in MRI.

4. CONCLUSIONS

The present article describes a simple method to solve the analytical expression (Biot-Savart) using a Numerical simulation to compute the magnetic field produced by an infinitesimal element located on the high end ring, law end ring and legs in each point in the 3D mesh study area, so the magnetic field inside the birdcage coil is the sum of all this magnetic fields.

The method developed can be coupled with numerical methods (FEM, FDTD, and MVF) to compute the RF field inside a simpler phantom.

ACKNOWLEDGMENT

One of authors (N. B) thanks Mr H. Chennoufi (University of Constantine, Algeria), Mr A. Cheriat and L. Alloui (University of Biskra, Algeria) for providing helpful and encouragement.

REFERENCES

1. Jin, J., *Electromagnetic Analysis and Design in Magnetic Resonance Imaging*, CRC Press, New York, 1999.
2. Hayes, C., W. Edelstein, J. Schenck, O. Muller, and M. Eash, "An efficient, highly homogeneous radiofrequency coil for whole-body nmr imaging at 1.5 T," *J. Magn. Reson.*, Vol. 63, 622–628, 1985.
3. Chu, Y., "Numerical calculation for the magnetic field in current-carrying circular arc filament," *IEEE Transactions on Magnetics*, Vol. 34, No. 2, 502–504, 1998.
4. Boissoles, P. and G. Caloz, "Accurate calculation of mutual inductance and magnetic fields in birdcage coil," *Preprint IRMAR*, France, February 2006.
5. Eriksson, O. and P. Kjäll, "Interaction between the RF-field and a stereotactic frame during MRI, a 3D study," *Proceedings of COMSOL Multiphysics User's,* Stockholm, Danemark, 2005.

Coils and Magnets: 3D Analytical Models

R. Ravaud[1], **G. Lemarquand**[1], **and V. Lemarquand**[2]

[1]Universite du Maine, LAUM, UMR CNRS 6613, Le Mans, France
[2]Universite de Toulouse, LAPLACE, UMR CNRS 5213, IUT Figeac, France

Abstract— This paper deals with analytical models for computing the magnetic fields created by both kinds of magnetic sources: coils and permanent magnets. It shows that the same analytical model can be used for modeling a permanent magnet or a coil and highlights and explains the mathematical difficulties encountered in solving problems with cylindrical properties. As a remark, each type of source, coil or permanent magnet, can be modeled by both approaches too: the coulombian or the amperian one. The presented three-dimensional mathematical model is accurate whatever the magnetic source dimensions, so for the near field and the far field.

1. INTRODUCTION

Various electromagnetic applications are composed of two thick coils that form a loosely coupled transformer. Then, the first coil produces a magnetic field which is partly picked up by a secondary coil. A good criterion for evaluating the efficiency of such a device is the mutual inductance of the coupled transformer. The calculation of the mutual inductance between circular coils has been studied by many authors. On one hand, the approaches used are based on the application of Maxwell's formula, Neumann's formula and the Biot Savart law. These approaches lead to expressions based on elliptic integrals of the first, second and third kind, the Heummann's Lambda function or the Bessel functions. On the other hand, authors generally use the finite element method, the boundary element method or numerical algorithms which evaluate the Laplace's equation in all points in space. Then, this paper presents a method for calculating the mutual inductance between two thick coils which leads to an interesting analytical formulation too. As said previously, both models of a source can be used. Thus, this paper presents general expressions that have been obtained with the coulombian and the amperian models of a permanent magnet. But the point is that some configurations may be modeled indifferently with the two previous models where others should be modeled either with the coulombian model or the amperian current model in order to obtain interesting formulations for the result. For instance, the use of the coulombian model for a thin coil will lead to useful formulations whereas the amperian current model only will give good results in the case of a thick coil. In this paper, a thick coil is replaced by a toroidal conductor carrying a unifom current volume density for calculating both the force exerted between two thick coils and their mutual inductance. The Lorentz Force is used for calculating the exact axial force exerted between two thick coils. Then, the potential vector and the Stocke's theorem are used for calculating the mutual inductance of two thick coils in air. The obtained formulation is interesting as it is fully based on elliptic functions whose numerical calculation is fast and robust.

1.1. Radial Component

The magnetic field produced by a cylindrical permanent magnet can be determined with the same analytical formulation as the one used for a cylindrical thin coil. However, the magnetic field produced by a permanent magnet is not of the same order of magnitude as the one produced by a thin coil. Besides, two kinds of models are used: the coulombian approach and the amperian current one. We can say that the choice of the model used does not depend on the magnetic source nature. Indeed, in the coulombian approach, a cylindrical magnet axially magnetized can be replaced by two charged planes which are located on the lower and upper faces of this cylinder. In the same way, a thin coil carrying uniform current density can also be represented by two charged planes. However, in the case of permanent magnet whose polarization is J in Tesla, its equivalent current must satisfy the following equation:

$$\vec{k} = \frac{\vec{J} \times \vec{n}}{\mu_0} \tag{1}$$

where $\vec{n}$ is the normal unit and μ_0 is the permeability of the vacuum. This implies that a cylindrical permanent magnet can be replaced by a thin coil whose current surface density is $\vec{k} = k\vec{u}_\theta$. In the case of a thin coil of N loops carrying a current I and whose height is h, its equivalent coulombian model satisfy the following equation:

$$J = \mu_0 \frac{NI}{h} \tag{2}$$

Finally, these two previous equations lead to the following conclusion: the magnetic field produced by a permanent magnet or by a thin coil has the same topology. From this point of view, as shown in Fig. 1, we can say that a cylindrical permanent magnet and a cylindrical thin coil own the same fictitious charge distribution. The authors generally use the amperian current model [1] or the coulombian approach [2] for calculating the magnetic field produced by permanent magnets. However, from a calculation point of view, these approaches do not lead to the same analytical expressions [3]. We assume, for this configuration, the amperian current model is more interesting. Therefore, the calculation of the magnetic field created by its constitutive elements, magnets or coils, is generally the first step for studying the characteristics of magnetomechanical devices [4–6]. Eventually, it is to be noted that various electromagnetic applications are composed of two coils that form a loosely coupled transformer. Basically, in such configurations, the first coil generates a magnetic field in all points in space. Then, this magnetic field is partly picked up by the secondary coil. However, a decrease in power transfer efficiency can be caused by a lower mutual inductance between two coils [7]. In other words, it is very useful to know the accurate value of the mutual inductance between two coils [8]. The calculation of the mutual inductance for cylindrical coils was studied by many authors [9–11]. These papers are generally based on the application of the fundamental laws of magnetostatics (Biot Savart law, Lorentz Force, Maxwell's equations). By using these approaches, the mutual inductance of circular coils can be expressed in terms of analytical and semi-analytical functions, as the elliptic integrals of the first, second and third kind or the Bessel functions [12].

2. NOTATION AND GEOMETRY

We present first the notation and the geometry which will be used throughout this paper. For this purpose, let us first consider Fig. 2 in which we have represented two cylinders axially decentred. As stated previously, these cylinders can represent either permanent magnets or thin coils. The lower cylinder (1) has the following characteristics: r_1 radius, $h_1 = z_2 - z_1$ height, and the upper cylinder (2): r_2 radius, $h_2 = z_4 - z_3$ height.

Furthermore, for the third and fourth parts of this paper, the axial distance between the two cylinders is denoted d. Therefore, the parameters defined in this section will be used throughout this paper.

$J_1(J_2)$ polarization in Tesla, the cylinder 1, respectively 2, considered as a permanent magnet.

$k_1(k_2)$ current surface density [A/m], $N_1(N_2)$ number of loops, $I_1(I_2)$ current [A], the cylinder 1, respectively 2, considered as a thin coil.

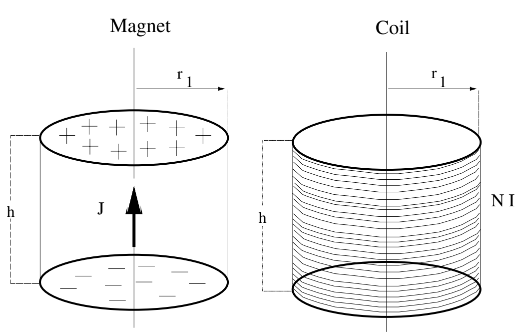


Figure 1: Representation of a cylindrical permanent magnet whose polarization is J [T] and a thin coil having N loops in which a current I circulates: these cylindrical topologies have the same radius r_1 and the same height h.

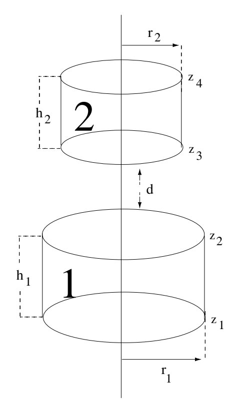


Figure 2: Geometry considered for the rest of this paper: two coaxial cylinders radially centered: they can represent a thin coil or a cylindrical permanent magnet.

3. MAGNETIC FIELD PRODUCED BY PERMANENT MAGNETS AND THIN COILS

Let us consider the cylinder **1** in Fig. 2. This cylinder is supposed to be a thin coil having a current surface density k_1 that equals $\frac{N_1 I_1}{h_1}$. By using the Biot-Savart Law, the magnetic induction field $\vec{B}_1$ created by this thin coil is expressed as follows:

$$\vec{B}_1(\tilde{\tilde{r}}, \tilde{\tilde{z}}) = \frac{\mu_0}{4\pi} \iint_S \vec{k_1} d\tilde{s}_1 \times \left\{ -\nabla G(\vec{\tilde{\tilde{r}}}, \vec{\tilde{r}}) \right\} \tag{3}$$

where $G(\vec{\tilde{\tilde{r}}}, \vec{\tilde{r}})$ is the Green's function that is defined as follows:

$$G(\vec{\tilde{\tilde{r}}}, \vec{\tilde{r}}) = \frac{1}{\sqrt{\tilde{\tilde{r}}^2 + \tilde{r}^2 - 2\tilde{r}\tilde{\tilde{r}}\cos(\tilde{\theta}) + (\tilde{\tilde{z}} - \tilde{z})^2}} \tag{4}$$

Furthermore, the parameters defined with the $\tilde{r}$ notation correspond to the points located on the lower cylinder whereas the parameters defined with the $\tilde{\tilde{r}}$ notation correspond to the points located on the upper one. After having integrated with respect to $\tilde{r}$ and $\tilde{z}$, we obtain the two following magnetic induction field components $B_r(r, z)$ and $B_z(r, z)$:

$$B_r(r, z) = \frac{\mu_0 k_1}{2\pi} \mathbf{C_a} = \frac{\mu_0 N_1 I_1}{2\pi h_1} \mathbf{C_a} \qquad \text{and} \qquad B_z(r, z) = \frac{\mu_0 k_1}{2\pi} \mathbf{C_b} = \frac{\mu_0 N_1 I_1}{2\pi h_1} \mathbf{C_b} \tag{5}$$

where $\mathbf{C_a}$ and $\mathbf{C_b}$ are scalar coefficients that depend on the source topology:

$$\mathbf{C_a} = \sum_{i=1}^{2} (-1)^i \left\{ \frac{a_i}{r\sqrt{\alpha_i}} \mathbf{K}\left[\frac{-2b}{\alpha_i}\right] - \frac{\sqrt{\alpha_i}}{r} \mathbf{E}\left[\frac{-2b}{\alpha_i}\right] \right\} \tag{6}$$

$$\mathbf{C_b} = \sum_{i=1}^{2} (-1)^i \epsilon_{4,i} \left\{ (cr - br_1) \left(\mathbf{\Pi}\left[\epsilon_{1,i}, \epsilon_{3,i}, \epsilon_{2,i}\right] + \mathbf{\Pi}\left[\epsilon_{1,i}, \epsilon_{2,i}\right] \right) + (br_1 - a_i r) \left(\mathbf{F}\left[\epsilon_{3,i}, \epsilon_{2,i}\right] + \mathbf{K}\left[\epsilon_{2,i}\right] \right) \right\} \tag{7}$$

where the parameters a_i, α_i, b, c, $\epsilon_{1,i}$, $\epsilon_{2,i}$, $\epsilon_{3,i}$, $\epsilon_{4,i}$ are defined in Table 1. We illustrate our previous analytical expressions in Fig. 3 where we have represented the radial and axial components of the magnetic induction field created by a thin coil. The parameter values are $h_1 = 25\,\text{mm}$, $r_1 = 25\,\text{mm}$ and $k_1 = 32000\,\text{A/m}$, this values correspond to a current of $I_1 = 4\,\text{A}$ and $N_1 = 200$. Let us still consider the cylinder **1** in Fig. 2. If this cylinder is a cylindrical permanent magnet having an axial polarization J_1. By using the equivalence between the coulombian approach and the amperian current model, the magnetic induction field created by a cylindrical permanent magnet can be expressed as follows:

$$B_r(r, z) = \frac{J_1}{2\pi} \mathbf{C_a} \qquad \text{and} \qquad B_z(r, z) = \frac{J_1}{2\pi} \mathbf{C_b} \tag{8}$$

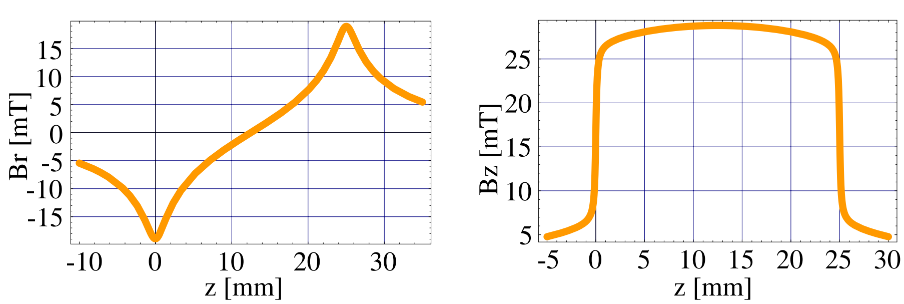


Figure 3: Representation of the radial and axial components of the magnetic induction field versus the radial and axial distances created by a thin coil in air; $h_1 = 25\,\text{mm}$, $r_1 = 25\,\text{mm}$, $k_1 = 32000\,\text{A/m}$; left representation: $r = 24\,\text{mm}$, right representation: $r = 24.9\,\text{mm}$.

Table 1: Definition of parameters used in this paper.

Parameter	Definition
α_i	$(r-r_1)^2+(z-z_i)^2$
$\epsilon_{1,i}$	$\frac{(a-c)}{(a+b)}$
$\epsilon_{2,i}$	$\frac{(a-b)}{(a+b)}$
$\epsilon_{3,i}$	$-\arcsin\left[\frac{a+b}{a-b}\right]$
$\epsilon_{4,i}$	$\frac{2r_1(z-z_i)}{b(a-c)\sqrt{-a-b}}$
a_i	$r^2+r_1^2+(z-z_i)^2$
b	$2rr_1$
c	$r^2+r_1^2$
ϖ	z_i-z_j
ψ	$\frac{\tau-\epsilon}{\tau+\epsilon}$
β	$\frac{\epsilon}{\epsilon-\tau}$
γ	$\varpi^2-\tau+\epsilon$
δ	$\frac{\epsilon}{\tau+\epsilon-\varpi^2}$
κ	$\epsilon\sqrt{\tau-\epsilon}$
ν	$-\varpi^4+2\varpi^2\tau-\tau^2+\epsilon^2$
ι	$\frac{\epsilon}{\tau+\epsilon}$
τ	$r_1^2+r_2^2+(z_i-z_j)^2$
ϵ	$-2r_1r_2$
μ	$(r_1-r_2)^2+(z_i-z_j)^2$

4. FORCE EXERTED BETWEEN TWO PERMANENT MAGNETS OR TWO THIN COILS

This section presents analytical models for calculating the axial force exerted between thin coils or cylindrical permanent magnets. The analytical expression of the axial force is expressed as follows:

$$F_z=\frac{\mu_0k_1k_2}{4\pi}\iint_{S_1}\iint_{S_2}\frac{\left(\tilde{\tilde{z}}-\tilde{z}\right)\cos(\tilde{\theta})\tilde{r}\tilde{\tilde{r}}d\tilde{\theta}d\tilde{z}d\tilde{\tilde{\theta}}d\tilde{\tilde{z}}}{\left(\tilde{\tilde{r}}^2+\tilde{r}^2-2\tilde{r}\tilde{\tilde{r}}\cos(\tilde{\theta})+(\tilde{\tilde{z}}-\tilde{z})^2\right)^{\frac{3}{2}}}\tag{9}$$

After integrating with respect to $\tilde{z}$, $\tilde{\tilde{z}}$, $\tilde{\theta}$ and $\tilde{\tilde{\theta}}$, the final analytical expression of the force F_z exerted between two thin coils in air is expressed as follows:

$$F_z=\frac{\mu_0k_1k_2}{2}\mathbf{C_d}\tag{10}$$

where A_z is the fully analytical part and f_z, the analytical part based on elliptic functions:

$$\mathbf{C_d}=\sum_{i=1}^{2}\sum_{j=3}^{4}(-1)^{i+j}(r_1r_2)\left(A_z+f_z\right)\tag{11}$$

with

$$\begin{aligned}A_z &= \frac{(\varpi^2-\tau)\pi-2\sqrt{\nu}}{2\epsilon}\ln\left[\frac{-4\epsilon^2}{\nu^{\frac{3}{2}}}\right]-\frac{(\tau-\varpi^2)\pi-2\sqrt{\nu}}{2\epsilon}\ln\left[\frac{4\epsilon^2}{\nu^{\frac{3}{2}}}\right]\\ f_z &= \frac{2\mathbf{i}\varpi}{\epsilon\sqrt{\tau+\epsilon}}\left\{(\tau+\epsilon)\mathbf{E}\left[\arcsin\sqrt{\frac{1}{\psi}},\psi\right]-\epsilon\mathbf{F}\left[\arcsin\sqrt{\frac{1}{\psi}},\psi\right]\right\}\\ &\quad+\frac{2\varpi}{\kappa\sqrt{\beta}}\left\{\frac{\epsilon}{\iota}\mathbf{E}\left[\psi\right]-\epsilon\sqrt{\iota}\mathbf{K}\left[\psi\right]+\sqrt{\beta\psi}\left((\tau-\varpi^2)\mathbf{K}\left[2\iota\right]+\gamma\mathbf{\Pi}\left[2\delta,2\iota\right]\right)\right\}\end{aligned}\tag{12}$$

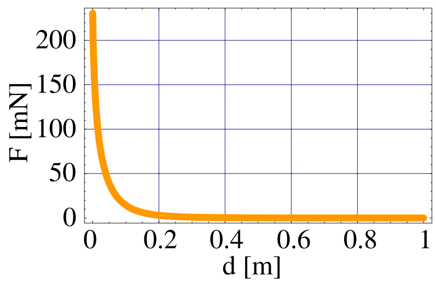


Figure 4: Representation of the force exerted between two thin coils versus the axial distance d. We take the following parameters: $r_1 = 0.0875\,\mathrm{m}$, $r_2 = 0.0875\,\mathrm{m}$, $h_1 = 0.025\,\mathrm{m}$, $h_2 = 0.025\,\mathrm{m}$, $k_1 = k_2 = 8000\,\mathrm{A/m}$.

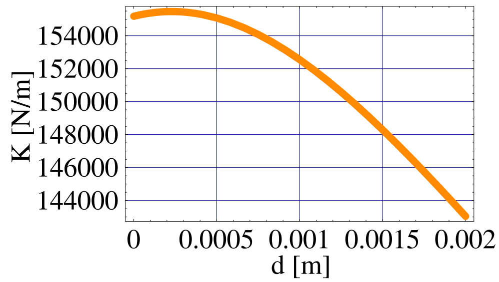


Figure 5: Representation of the axial stiffness exerted between cylindrical permanent magnets in air versus the axial distance d; $h_1 = 25\,\mathrm{mm}$, $h_2 = 20\,\mathrm{mm}$, $r_1 = 87.5\,\mathrm{mm}$, $r_2 = 85\,\mathrm{mm}$, $J_1 = J_2 = 1\,\mathrm{T}$.

where **i** verifies $i^2 = -1$ and the other parameters can depend on i and j (see Table 1). We have represented in Fig. 4 the axial force exerted between two thin coils in air. It is emphasized here that this curve had been verified with the filament method in a previous paper [13]. Let us consider the cylinders **1** and **2** in Fig. 2. These cylinders are supposed now to be cylindrical permanent magnets axially magnetized. By using the analytical expression of the force exerted between two thin coils in air, we can directly write that the axial force exerted between two cylindrical permanent magnets:

$$F_z = \frac{J_1 J_2}{2\mu_0}\mathbf{C_d} \tag{13}$$

5. STIFFNESS EXERTED BETWEEN PERMANENT MAGNETS AND THIN COILS

Let us consider the cylinders **1** and **2** in Fig. 2. These cylinders are supposed to be cylindrical permanent magnets axially magnetized. The axial stiffness can be deducted from the following equation:

$$K_z = -\frac{d}{dz}F_z \tag{14}$$

where F_z is the axial force exerted between the two permanent magnets. The axial stiffness exerted between two cylindrical permanent magnets is expressed as follows:

$$K_z = \frac{J_1 J_2}{\mu_0}\mathbf{C_f} \tag{15}$$

where $\mathbf{C_f}$ is a scalar coefficient that depends on the source distributions.

$$\mathbf{C_f} = \sum_{i=1}^{2}\sum_{j=3}^{4}(-1)^{i+j}\left\{-\frac{\tau}{\sqrt{\mu}}\mathbf{K}\left[\frac{2\epsilon}{\mu}\right] + \sqrt{\mu}\mathbf{E}\left[\frac{2\epsilon}{\mu}\right]\right\} \tag{16}$$

It is emphasized here that it is simple to calculate the radial stiffness K_r as it satisfies the following equation [14] for ironless structure:

$$2K_r + K_z = 0 \tag{17}$$

Let us still consider the cylinders **1** and **2** in Fig. 2. These cylinders are supposed now to be thin coils in air. The axial stiffness exerted between two thin coils in air can be deducted directly from the equivalence between the coulombian approach and the amperian current model:

$$K_z = \mu_0 k_1 k_2 \mathbf{C_f} \tag{18}$$

6. MUTUAL INDUCTANCE BETWEEN TWO THIN COILS

6.1. Analytical Calculation

The mutual inductance between two thin coils can be derived from the following equation:

$$M = \frac{\mathbf{\Phi_{1\to 2}}}{I_1} \tag{19}$$

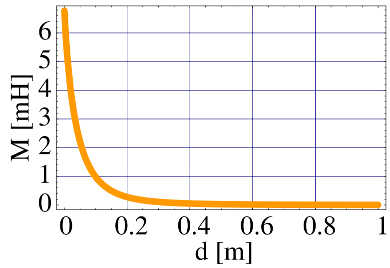


Figure 6: Representation of the mutual inductance between two thin coils in air versus the axial distance d;$r_1 = 0.0875$ m, $r_2 = 0.085$ m, $h_1 = 0.025$ m, $h_2 = 0.025$ m, $N_1 = N_2 = 200$.

where $\mathbf{\Phi_{1\to 2}}$ is the flux created by the thin coil **1** in the thin coil **2**. By using here previous integral formulations and after mathematical manipulations, the reduced form of the mutual inductance can be expressed as follows:

$$M = \frac{\mu_0 N_1 N_2}{(z_4 - z_3)(z_2 - z_1)} \mathbf{C_g}. \tag{20}$$

where $\mathbf{C_g}$ is a scalar depending on the two source distributions.

$$\mathbf{C_g} = \frac{r_1 r_2}{2} \sum_{i=1}^{2} \sum_{j=3}^{4} (-1)^{i+j} \left(m_1(r_1^2 + r_2^2, 2r_1 r_2) + \frac{z_i - z_j}{2r_1 r_2} m_2(r_1^2 + r_2^2, 2r_1 r_2, z_i - z_j) \right) \tag{21}$$

with

$$m_1(a, b) = \frac{2}{3\sqrt{\xi_1}} \left(-2b\mathbf{K}\left[\frac{-2b}{\xi_1}\right] - \frac{\xi_2}{b} \left(2\xi_1 \mathbf{E}\left[\frac{-2b}{\xi_1}\right] - 2\xi_2 \mathbf{K}\left[\frac{-2b}{\xi_1}\right] \right) \right) \tag{22}$$

where $\xi_1 = a - b + c^2$ and $\xi_2 = a + c^2$ and

$$\begin{aligned} m_2(a, b, c) \quad = \quad & a\pi + 2c\sqrt{\frac{-1}{\xi_4}} \left(-\xi_4 \mathbf{E}\left[\frac{\xi_3}{\xi_4}\right] - b\mathbf{K}\left[\frac{\xi_3}{\xi_1}\right] \right) - 2ci\xi_1 \mathbf{E}\left[\arcsin\sqrt{\frac{\xi_1}{\xi_3}}, \frac{\xi_3}{\xi_1}\right] \\ & - \frac{2bci}{\sqrt{\xi_1}} \mathbf{F}\left[\arcsin\sqrt{\frac{\xi_1}{\xi_3}}, \frac{\xi_3}{\xi_1}\right] - \frac{2c}{\sqrt{\xi_3}} \left(a\mathbf{K}\left[\frac{2b}{\xi_3}\right] + (b - a)\mathbf{\Pi}\left[\frac{2b}{a+b}, \frac{2b}{\xi_3}\right] \right) \\ & + 2c\sqrt{b^2 - a^2} \left(\ln\left[\frac{4b^2}{(b^2 - a^2)^{\frac{3}{2}}}\right] - \ln\left[\frac{-4b^2}{(b^2 - a^2)^{\frac{3}{2}}}\right] \right) \end{aligned} \tag{23}$$

where $\xi_3 = a + b + c^2$. We have represented in Fig. 6 the mutual inductance of two coils versus the axial distance between them. The computational cost required for representing this curve is about 0.8 s.

7. CONCLUSION

This paper has presented a synthesis about analytical calculations of cylindrical magnetic sources. We have presented the analytical expressions of the magnetic induction field created by cylindrical permanent magnets axially magnetized and thin coils in air. These expressions have been confirmed with the finite element method. Then, we have presented analytical models allowing us to calculate the axial force exerted between two cylindrical permanent magnets axially magnetized and two thin coils in air. Moreover, we have presented the analytical calculation of the stiffness exerted between these magnetic source distributions. Eventually, we have presented an analytical calculation of the mutual inductance between two thin coils in air. More generally, we can say that the equivalence between the coulombian model of a magnet and the amperian current model is a previous element of information for the design of magnetomechanical devices composed of thin coils or cylindrical permanent magnets.

REFERENCES

1. Furlani, E. P., S. Reznik, and A. Kroll, "A three-dimensonal field solution for radially polarized cylinders," *IEEE Trans. Magn.*, Vol. 31, No. 1, 844–851, 1995.
2. Babic, S. I. and C. Akyel, "Magnetic force calculation between thin coaxial circular coils in air," *IEEE Trans. Magn.*, Vol. 44, No. 4, 445–452, 2008.
3. Ravaud, R. and G. Lemarquand, "Comparison of the coulombian and amperian current models for calculating the magnetic field produced by radially magnetized arc-shaped permanent magnets," *Progress In Electromagnetics Research*, Vol. 95, 309–327, 2009.
4. Ravaud, R. and G. Lemarquand, "Modelling an ironless loudspeaker by using three-dimensional analytical approaches," *Progress In Electromagnetics Research*, Vol. 91, 53–68, 2009.
5. Ravaud, R., G. Lemarquand, and V. Lemarquand, "Force and stiffness of passive magnetic bearings using permanent magnets. part 1: Axial magnetization," *IEEE Trans. Magn.*, Vol. 45, No. 7, 2996–3002, 2009.
6. Ravaud, R., G. Lemarquand, V. Lemarquand, S. Babic, and C. Akyel, "Calculation of the magnetic field created by a thick coil," *Journal of Electromagnenic Waves and Applications*, Vol. 24, No. 10, 1405–1418, 2010.
7. Costa, E. M. M., "Responses in transformers built with planar coils inner ring coils excited by square waves," *Progress In Electromagnetics Research B*, Vol. 18, 43–58, 2009.
8. Grover, F. W., *Inductance Calculations*, Dover, New York, 1964.
9. Kamon, M., M. J. Tsuk, and J. White, "Fast-henry: A multiple-accelerated 3d inductance extraction program," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 1750–1758, 1994.
10. Koledintseva, M. Y., J. L. Drewniak, T. P. van Doren, D. J. Pommerenke, M. Cocchini, and D. M. Hockanson, "Method of edge currents for calculating mutual external inductance in a microstrip structure," *Progress In Electromagnetics Research*, Vol. 80, 197–224, 2008.
11. Babic, S. I. and C. Akyel, "Calculating mutual inductance between circular coils with inclined axes in air," *IEEE Trans. Magn.*, Vol. 44, No. 7, 1743–1750, 2008.
12. Conway, J., "Inductance calculations for noncoaxial coils using bessel functions," *IEEE Trans. Magn.*, Vol. 43, No. 3, 1023–1034, 2007.
13. Akyel, C., S. I. Babic, and M. M. Mahmoudi, "Mutual inductance calculation for non-coaxial circular air coils with parallel axes," *Progress In Electromagnetics Research*, Vol. 91, 287–301, 2009.
14. Elies, P. and G. Lemarquand, "Analytical study of radial stability of permanent magnet synchronous couplings," *IEEE Trans. Magn.*, Vol. 35, No. 4, 2133–2136, 1999.

Discussion on the Magnetic Pole Volume Density in Analytical Models of Permanent Magnets

R. Ravaud[1], **G. Lemarquand**[1], **and V. Lemarquand**[2]

[1]Universite du Maine, LAUM, UMR CNRS 6613, Le Mans, France
[2]Universite de Toulouse, LAPLACE, UMR CNRS 5213, IUT Figeac, France

Abstract— This paper deals with the study of the magnetic field created by ring permanent magnets whose polarization is radial. The influence of the magnetic pole volume density on the magnetic field is studied. For this purpose, the expressions of the magnetic field components created by the magnetic pole volume density of a ring permanent magnet whose polarization is radial are determined semi-analytically in a useful form. To do so, calculations are obtained by using the Coulombian model of a magnet. Different configurations of radially polarized magnets are studied in this paper. The ring permanent magnets dimensions are varied, especially their inner and outer radii, to consider radially thin or thick magnets as well as smoothly curved or not ones.

1. INTRODUCTION

Many electrical devices and engineering applications use ring permanent magnets. Indeed, the magnetic field they create can be used to design magnetic bearings, stepping motors, sensors. As a consequence, the precise determination of the magnetic field created by such configurations is required. Many authors are interested in the modeling of the magnetic field created by ring permanent magnets whose polarization is either radial or axial or coils [1–7]. Consequently, several interesting approaches are commonly used in order to obtain semi-analytical expressions which are valid for all points in space [8–10]. Moreover, authors also use numerical means or 2 D analytical calculations for the determination of both the magnetic field and the magnetic forces created between arc-shaped permanent magnets [11]. The expressions of the contribution of the magnetic pole volume density are determined semi-analytically. The total computationnal time, though far shorter than finite element calculations, is increased when the volume pole density is taken into account. So, it is interesting to study the influence of this volume density on the result in order to discuss the necessity of its taking into account with regard to the result accuracy. Indeed, taking all the pole densities leads to the exact value of the magnetic field components, but what is the difference if the volume densities are neglected?

To give some answers, different configurations of radially polarized magnets are studied in this paper. The ring permanent magnets dimensions are varied, especially their inner and outer radii, to consider radially thin or thick magnets as well as smoothly curved or not ones.

This paper has two objectives. The expression of the radial component is presented in terms of elliptic integrals and one term must be solved numerically though the axial component expression is presented in a useful form based on one numerical integration. Then, we discuss the utility of taking into account the magnetic pole volume density on the evaluation of the magnetic field created by thin or massive disks. Moreover, the length influence of their radius of curvature is discussed in each case.

2. CALCULATION OF THE MAGNETIC FIELD COMPONENTS CREATED BY THE MAGNETIC POLE VOLUME DENSITY OF A RING PERMANENT MAGNET WHOSE POLARIZATION IS RADIAL

2.1. Notation and Geometry

The geometry considered is shown in Fig. 1. The ring inner radius is r_1 and the outer one is r_2. The lower altitude is z_1 and the upper one is z_2. Moreover, this ring permanent magnet has a radial polarization which is denoted J. It is noted that the coulombian model of a permanent magnet is used. For ring permanent magnets whose polarization is radial, the magnetic pole density can be modeled as both a magnetic pole volume density σ_v and two magnetic pole surface densities $-\sigma_s$ and $+\sigma_s$ which are located on the inner and outer faces of the ring permanent magnet. The magnetic field created by the two magnetic pole surface densities has been determined in a previous paper. In this paper, we look at the contribution of the magnetic pole volume density on the magnetic

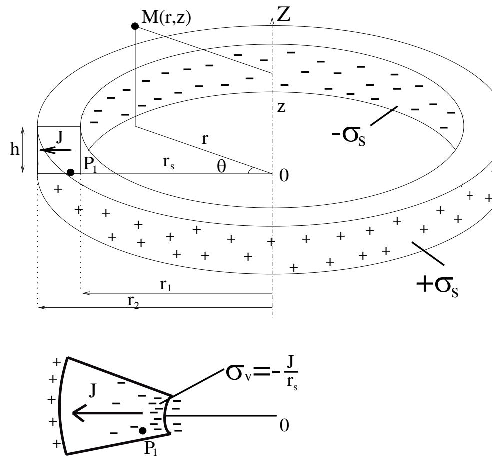


Figure 1: Representation of the studied geometry: ring permanent magnet whose polarization is radial. The inner radius is r_1, the outer one is r_2, the lower altitude is z_1 and the upper one is z_2.

field created by a ring permanent magnet whose polarization is radial. For ring permanent magnets whose polarization is radial, the magnetic pole volume density σ_v in cylindrical coordinate system is given by (1).

$$\sigma_v = -\frac{J}{r_s} \tag{1}$$

where r_s is the radial abscissa. We can say that the magnetic pole volume density allows us to have a magnetic charge equilibrium of the ring permanent magnet. Indeed, for a ring permanent magnet whose polarization is radial, the magnetic pole volume density σ_v must verifies

$$-S_{ext}\sigma_s + S_{int}\sigma_s + V\sigma_v = 0 \tag{2}$$

where S_{ext} and S_{int} are the outer and inner faces of the ring, V is its volume. As S_{ext} is higher than S_{int}, a magnetic pole volume density σ_v appears in the ring magnet (Eq. (1)). The magnetic field components created by the magnetic pole volume density are denoted $H_r^{(v)}$ and $H_z^{(v)}$. The magnetic field components created by the magnetic pole surface densities are denoted $H_r^{(s)}$ and $H_z^{(s)}$. The exact magnetic field components creating by the ring permanent magnets are denoted $H_r^{(total)}$ and $H_z^{(total)}$. It can be noted that the last case corresponds to the total magnetic field created by a ring permanent magnet whose polarization is radial. As a result, the magnetic field created by the magnetic pole volume density is given by (3).

$$\vec{H}^{(v)} = \iiint_{(V)} \frac{\sigma_v}{4\pi\mu_0} \frac{\overrightarrow{P_1M}}{\left|\overrightarrow{P_1M}\right|^3} r_s dr_s d\theta dz_s \tag{3}$$

2.2. Radial Component

The radial component expression can be integrated analytically two times. The last integration can be decomposed in two parts so as to have only one term which must be determined numerically. As a consequence, we obtain an expression with an analytical part and a numerical part. Such an expression is given by (4).

$$H_r^{(v)} = \frac{J}{2\pi\mu_0}\left(h_r - \int_{u_1}^{u_2}(1-u^2)\sum_{i=1}^{2}\sum_{j=1}^{2}\arctan\left[\frac{(r_i - ru)(z-z_j)}{\sqrt{-r^2(-1+u^2)\sqrt{r^2+r_i^2-2rr_iu+(z-z_j)^2}}}\right]du\right)$$

where

$$\begin{aligned} h_r = & f\left(z-z_1, r^2+r_1^2+(z-z_1)^2, 2rr_1, u_2\right) - f\left(z-z_1, r^2+r_1^2+(z-z_1)^2, 2rr_1, u_1\right) \\ & + f\left(z-z_1, r^2+r_2^2+(z-z_1)^2, 2rr_2, u_2\right) - f\left(z-z_1, r^2+r_2^2+(z-z_1)^2, 2rr_2, u_1\right) \\ & + f\left(z-z_2, r^2+r_1^2+(z-z_2)^2, 2rr_1, u_2\right) - f\left(z-z_2, r^2+r_1^2+(z-z_2)^2, 2rr_1, u_1\right) \\ & + f\left(z-z_2, r^2+r_2^2+(z-z_2)^2, 2rr_2, u_2\right) - f\left(z-z_2, r^2+r_2^2+(z-z_2)^2, 2rr_2, u_1\right) \end{aligned} \tag{4}$$

and

$$f(a_1, b_1, c_1, u) = -\eta a_1 \sqrt{\frac{b_1 - c_1 u}{b_1 + c_1}} \left(\sqrt{1-u^2} + \frac{(a_1^2 + b_1)\arcsin[u]}{c_1} - \frac{\sqrt{\frac{b_1+c_1}{c_1(1+u)}}}{c_1\sqrt{b_1 - c_1 u}\sqrt{1-u^2}} \right) \tag{5}$$

and

$$\begin{aligned} \eta = & (b_1 - c_1)\sqrt{\frac{c_1(-1+u)}{b_1 - c_1}}(1+u)\,\mathbf{E}^*\left[\arcsin\left[\sqrt{\frac{b_1 - c_1 u}{b_1 + c_1}}, \frac{b_1+c_1}{b_1-c_1}\right]\right] \\ & (-a_1^2 + b_1)\sqrt{\frac{c_1(1+u)}{b_1+c_1}}\sqrt{1-u^2}\,\mathbf{F}^*\left[\arcsin\left[\sqrt{\frac{1+u}{2}}\right], \frac{2c_1}{b_1+c_1}\right] \\ & + c_1(1+u)\sqrt{\frac{c_1(-1+u)}{b_1-c_1}}\,\mathbf{F}^*\left[\arcsin\left[\sqrt{\frac{b_1 - c_1 u}{b_1 + c_1}}\right], \frac{b_1+c_1}{b_1-c_1}\right] \\ & + (b_1 - a_1^2 - c_1)\sqrt{\frac{c_1(1+u)}{b_1+c_1}}\sqrt{1-u^2}\,\mathbf{\Pi}^*\left[\frac{2c_1}{-a_1^2+b_1+c_1}, \arcsin\left[\sqrt{\frac{1+u}{2}}\right], \frac{2c_1}{b_1+c_1}\right] \\ & - 2\sqrt{1-u^2}\log\left[a_1 + \sqrt{b_1 - c_1 u}\right] \\ & - \frac{\sqrt{\beta}}{c_1}\log\left[\frac{4c_1^2\left(c_1 + a_1^2 u - b_1 u + \sqrt{\beta}\sqrt{1-u^2}\right)}{\beta^{\frac{3}{2}}(a_1^2 - b_1 + c_1 u)}\right] \end{aligned} \tag{6}$$

$$\beta = -a_1^4 + 2a_1^2 b_1 - b_1^2 + c_1 \tag{7}$$

$\mathbf{F}^*[\Phi, m]$ is the elliptic integral of the first kind given by (8).

$$\mathbf{F}^*[\Phi, m] = \int_0^\Phi \frac{1}{\sqrt{1 - m\sin(\theta)^2}} d\theta \tag{8}$$

$\mathbf{E}^*[m]$ is the complete elliptic integral of the second kind given by (9).

$$\mathbf{E}^*[m] = \int_0^{\frac{\pi}{2}} \sqrt{1 - m\sin(\theta)^2} d\theta \tag{9}$$

$\mathbf{\Pi}^*[n, \Phi, m]$ is the incomplete elliptic integral of the third kind given by (10).

$$\mathbf{\Pi}^*[n, \Phi, m] = \int_0^\Phi \frac{1}{(1 - n\sin(\theta)^2)} \frac{1}{\sqrt{1 - m\sin(\theta)^2}} d\theta \tag{10}$$

The radial component of the magnetic field created by the magnetic pole volume density is shown in Fig. 2. The values taken are $r_1 = 0.025$ m, $r_2 = 0.028$ m, $z_1 = 0$ m, $z_2 = 0.003$ m, $z = 0.001$ m and $\sigma^* = 1$ T. It can be noted that this semi-analytical expression has a low computational cost compared to a fully numerical integration of (3). With Mathematica, the time necessary to determine the radial component at a given observation point is 0.03 s with our semi-analytical calculation whereas this time is 1.21 s with the fully numerical integration of (3).

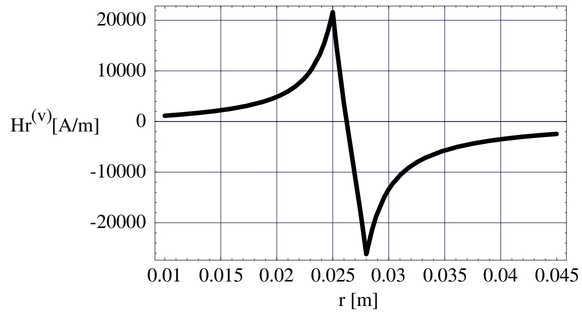


Figure 2: Representation of the radial component $H_r^{(v)}$ [A/m] versus the radial distance r [m]; $r_1 = 0.025$ m, $r_2 = 0.028$ m, $z_1 = 0$ m, $z_2 = 0.003$ m, $z = 0.001$ m and $\sigma^* = 1$ T.

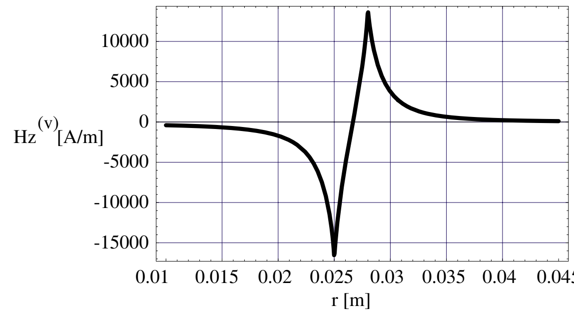


Figure 3: Representation of the radial component $H_z^{(v)}$ [A/m] versus the radial distance r [m]; $r_1 = 0.025$ m, $r_2 = 0.028$ m, $z_1 = 0$ m, $z_2 = 0.003$ m, $z = 0.001$ m and $\sigma^* = 1$ T.

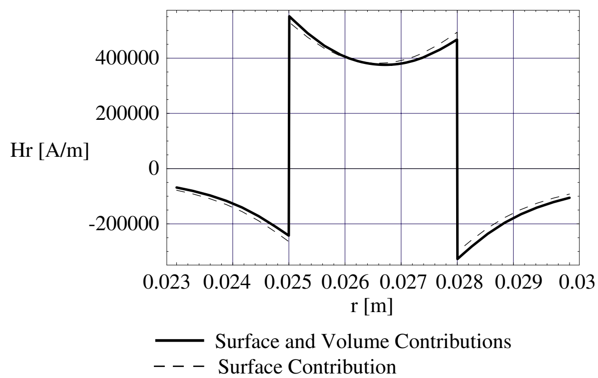


Figure 4: Representation of the radial component of the magnetic field created by either the surface contribution of the two curved planes or both the surface contribution and the volume contribution of the magnetic poles; $r_1 = 0.025$ m, $r_2 = 0.028$ m, $z_1 = 0$ m, $z_2 = 0.003$ m, $z = 0.001$ m, $\sigma^* = 1$ T.

2.3. Axial Component

The expression of the axial component can be integrated analytically two times and can be presented in a very useful form (11).

$$H_z^{(v)} = \frac{J}{4\pi\mu_0} \sum_{i=1}^{2} \sum_{j=1}^{2} -(-1)^{i+j} \int_0^{2\pi} \tanh^{-1}\left[\frac{\sqrt{r^2 + r_i^2 + (z - z_j)^2 - 2rr_i\cos(\theta)}}{r_i - r\cos(\theta)}\right] d\theta \qquad (11)$$

The axial component of the magnetic field created by the magnetic pole volume density is shown in Fig. 3. The values taken are $r_1 = 0.025$ m, $r_2 = 0.028$ m, $z_1 = 0$ m, $z_2 = 0.003$ m, $z = 0.001$ m and $\sigma^* = 1$ T. It can be noted that this semi-analytical expression has a low computational cost compared to a fully numerical integration of (3). With Mathematica, the time necessary to determine the axial component at a given observation point is 0.05 s with our semi-analytical calculation whereas this time is 0.55 s with the fully numerical integration of (3).

3. INFLUENCE OF THE MAGNETIC POLE VOLUME DENSITY ON THE MAGNETIC FIELD CREATED BY THE RING PERMANENT MAGNETS WHOSE POLARIZATION IS RADIAL

The aim of this section is to determine when the magnetic pole volume density must be taken into account in the determination of the magnetic field created by a ring permanent magnet whose polarization is radial. The two components of the magnetic field are singly studied.

3.1. Study of the Magnetic Field Radial Component

The expression of the radial component of the magnetic field created by the magnetic pole volume contribution of a ring permanent magnet whose polarization is radial has been determined with (4). First, we compare the representation of the radial component by taking into account either only the surface magnetic pole density or both the surface and volume magnetic pole densities. Such a

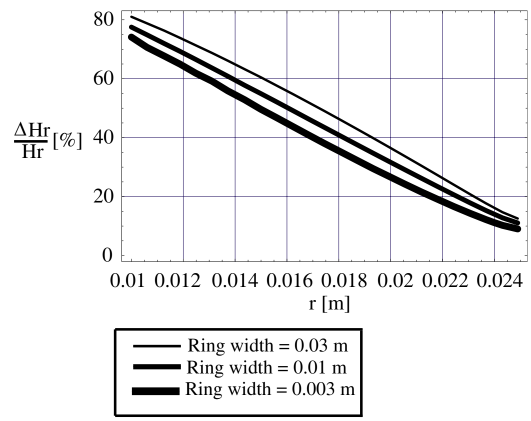


Figure 5: Representation of the relative radial difference $\frac{\Delta H_r}{H_r}$ versus the radial distance $r\,[\mathrm{m}]$; three radial widths are studied: ($r_2 - r_1 = 0.003\,\mathrm{m}$, $r_2 - r_1 = 0.01\,\mathrm{m}$ and $r_2 - r_1 = 0.03\,\mathrm{m}$), $r_1 = 0.025\,\mathrm{m}$, $z_1 = 0\,\mathrm{m}$, $z_2 = 0.003\,\mathrm{m}$, $z = 0.001\,\mathrm{m}$ and $\sigma^* = 1\,\mathrm{T}$.

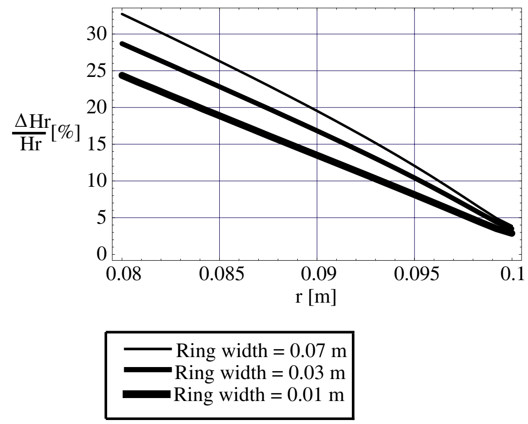


Figure 6: Representation of the relative radial difference $\frac{\Delta H_r}{H_r}$ versus the radial distance $r\,[\mathrm{m}]$; three radial widths are studied: ($r_2 - r_1 = 0.01\,\mathrm{m}$, $r_2 - r_1 = 0.03\,\mathrm{m}$ and $r_2 - r_1 = 0.07\,\mathrm{m}$) , $r_1 = 0.1\,\mathrm{m}$, $z_1 = 0\,\mathrm{m}$, $z_2 = 0.003\,\mathrm{m}$, $z = 0.001\,\mathrm{m}$ and $\sigma^* = 1\,\mathrm{T}$.

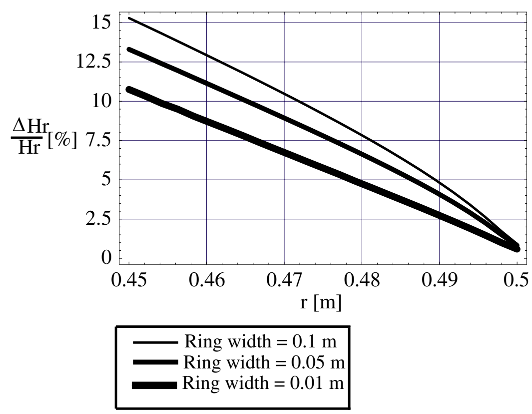


Figure 7: Representation of the relative radial difference $\frac{\Delta H_r}{H_r}$ versus the radial distance $r\,[\mathrm{m}]$; three radial widths are studied: ($r_2 - r_1 = 0.01\,\mathrm{m}$, $r_2 - r_1 = 0.05\,\mathrm{m}$ and $r_2 - r_1 = 0.1\,\mathrm{m}$), $r_1 = 0.5\,\mathrm{m}$, $z_1 = 0\,\mathrm{m}$, $z_2 = 0.003\,\mathrm{m}$, $z = 0.001\,\mathrm{m}$ and $\sigma^* = 1\,\mathrm{T}$.

comparison can be made with Fig. 4. Indeed, Fig. 4 shows that the magnetic field radial component is different when the magnetic pole volume density is taken into account. This figure is interesting because it also shows that this difference varies versus the radial distance r. Consequently, different configurations must be compared. First, we study the radial relative difference between the exact magnetic field created by the ring permanent magnet and the one created by the magnetic pole surface density located on the ring curved planes. The radial relative difference is given by (12).

$$\frac{\Delta H_r}{H_r} = \frac{H_r^{(s)} - H_r^{(total)}}{H_r^{(total)}} \tag{12}$$

This relative radial difference is determined for different inner radius and for different radial widths of the ring (Figs. 5, 6 and 7). The influence of the magnetic pole volume density is first studied in front of the ring magnets. In short, Figs. 5, 6 and 7 show several important things. First, if the magnetic pole volume density is omitted, the relative radial difference becomes more and more important when the magnetic field radial component is determined far from the ring magnets. When the inner radius is $0.025\,\mathrm{m}$, we incur an error of at least 10% in calculating the magnetic field radial component either close to the ring magnets whereas we incur an error of at least 50% in calculating this one when the distance in length between the ring magnet and the observation point equals $0.01\,\mathrm{m}$. Moreover, we see that this error is more important for the same observation point when the ring width is larger. For instance, for an observation point which equals $0.018\,\mathrm{m}$,

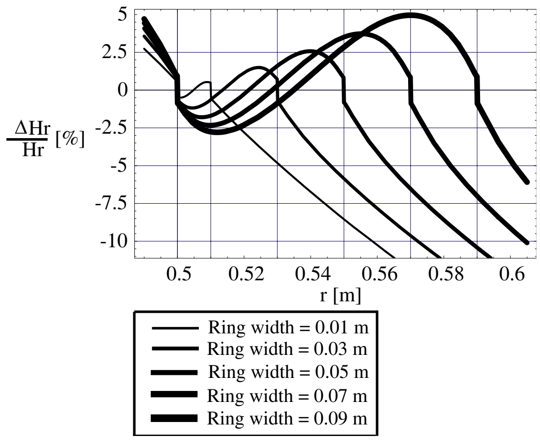


Figure 8: Representation of the relative radial difference $\frac{\Delta H_r}{H_r}$ versus the radial distance r [m] inside the ring magnet for different ring widths ($r_2 - r_1 = 0.01\,\mathrm{m}$, $r_2 - r_1 = 0.03\,\mathrm{m}$, $r_2 - r_1 = 0.05\,\mathrm{m}$, $r_2 - r_1 = 0.07\,\mathrm{m}$ and $r_2 - r_1 = 0.09\,\mathrm{m}$); $r_1 = 0.5\,\mathrm{m}$, $z_1 = 0\,\mathrm{m}$, $z_2 = 0.003\,\mathrm{m}$, $z = 0.001\,\mathrm{m}$ and $\sigma^* = 1\,\mathrm{T}$.

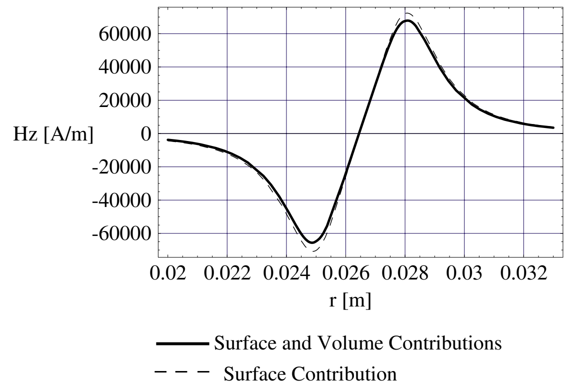


Figure 9: Representation of the radial component of the magnetic field created by either the surface contribution of the two curved planes or both the surface contribution and the volume contribution of the magnetic poles.

there is a difference of about 10% in the determination of the magnetic field radial component between the thin ring and the massive ring. When the inner radius is 0.01 m, the error we incur is less important than the one made in the previous case. Indeed, we incur an error of at least 3% in calculating the magnetic field radial component either close to the ring magnets. Moreover, Fig. 7 clearly shows that this error becomes very small when the inner radius becomes important (here $r_1 = 0.5\,\mathrm{m}$). Therefore, we deduce that the magnetic pole volume density must be taken into account when the inner radius is small. In the three configurations, we see that the ring width has an important influence which becomes more and more visible when the magnetic field is determined far from the ring magnets. Let us consider now the magnetic field inside the ring permanent magnet which corresponds to the demagnetization field. For this purpose, the relative radial difference has been represented in Fig. 8 for different ring widths inside the ring magnet. Fig. 8 shows several important things. First, when the ring width is small (for instance $r_2 - r_1 = 0.01\,\mathrm{m}$), the relative radial difference is small compared to the one determined when the ring width becomes larger (for instance $r_2 - r_1 = 0.09\,\mathrm{m}$). Second, for each configuration, it always exists an observation point where the magnetic pole volume density has not any influence on the magnetic field created by the ring magnets. Third, we see that the relative radial difference is less important inside the ring magnet than in front of the ring magnet. Consequently, we deduce that the magnetic field radial component can be determined inside the ring magnet without taking into account the magnetic pole volume density. However, if the magnetic field radial component is determined near the outer curved planes, the relative radial difference becomes more important. In any case, we can conclude that the magnetic pole volume density must be taken into account when the magnetic field radial component is determined outside of the ring magnets if the distance in length between the ring magnet and the observation point becomes important. Inside the ring magnet, the relative radial difference is less important and we deduce that we can only use the analytical expressions of the magnetic pole surface density to determine the magnetic field radial component.

3.2. Study of the Axial Component

The expression of the axial component of the magnetic field created by the magnetic pole volume contribution of a ring permanent magnet whose polarization is radial has been determined with (11). First, we compare the representation of the axial component by taking into account either only the surface magnetic pole density or both the surface and volume magnetic pole densities. Such a comparison can be made with Fig. 9. Indeed, Fig. 9 shows that the magnetic field axial component is different when the magnetic pole volume density is taken into account. This figure is interesting because it also shows that this difference varies versus the radial distance r. Consequently, different configurations must be compared. First, we study the axial relative difference between the exact magnetic field created by the ring permanent magnet and the one created by the magnetic pole

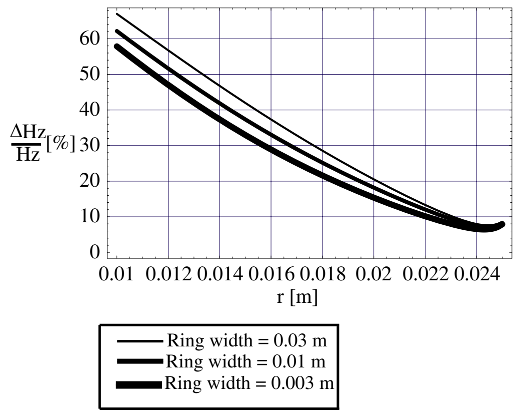


Figure 10: Representation of the relative axial difference $\frac{\Delta H_z}{H_z}$ versus the radial distance r [m]; three radial widths are studied: ($r_2 - r_1 = 0.003$ m, $r_2 - r_1 = 0.01$ m and $r_2 - r_1 = 0.03$ m), $r_1 = 0.025$ m, $z_1 = 0$ m, $z_2 = 0.003$ m, $z = 0.001$ m and $\sigma^* = 1$ T.

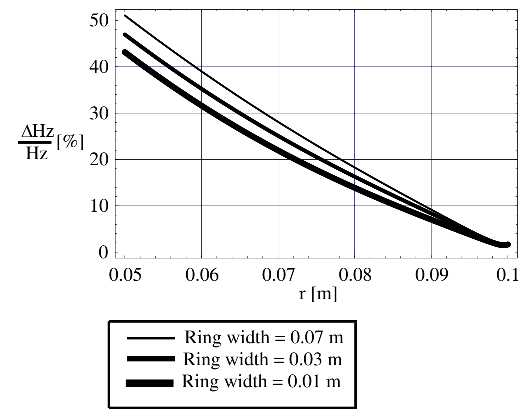


Figure 11: Representation of the relative axial difference $\frac{\Delta H_z}{H_z}$ versus the radial distance r [m]; three radial widths are studied: ($r_2 - r_1 = 0.01$ m, $r_2 - r_1 = 0.03$ m and $r_2 - r_1 = 0.07$ m), $r_1 = 0.1$ m, $z_1 = 0$ m, $z_2 = 0.003$ m, $z = 0.001$ m and $\sigma^* = 1$ T.

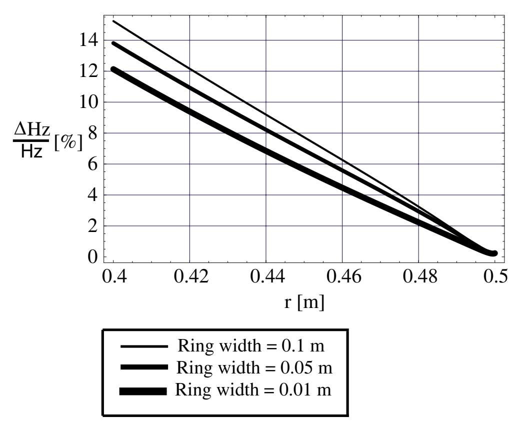


Figure 12: Representation of the relative axial difference $\frac{\Delta H_z}{H_z}$ versus the radial distance r [m] inside the ring magnet for different ring widths ($r_2 - r_1 = 0.01$ m, $r_2 - r_1 = 0.05$ m, and $r_2 - r_1 = 0.1$ m); $r_1 = 0.5$ m, $z_1 = 0$ m, $z_2 = 0.003$ m, $z = 0.001$ m and $\sigma^* = 1$ T.

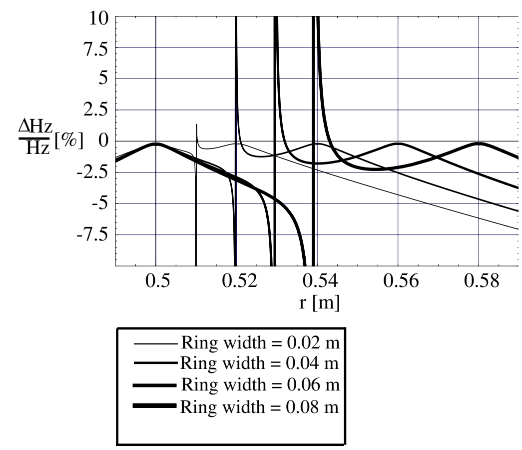


Figure 13: Representation of the relative axial difference $\frac{\Delta H_z}{H_z}$ versus the radial distance r [m] inside the ring magnet for different ring widths ($r_2 - r_1 = 0.02$ m, $r_2 - r_1 = 0.04$ m, $r_2 - r_1 = 0.06$ m and $r_2 - r_1 = 0.08$ m); $r_1 = 0.5$ m, $z_1 = 0$ m, $z_2 = 0.003$ m, $z = 0.001$ m and $\sigma^* = 1$ T.

surface density located on the ring curved planes. The axial relative difference is given by (13).

$$\frac{\Delta H_z}{H_z} = \frac{H_z^{(s)} - H_z^{(total)}}{H_z^{(total)}} \tag{13}$$

This relative axial difference is determined for different inner radius and for different radial width of the ring (Figs. 10, 11 and 12). The influence of the magnetic pole volume density is first studied in front of the ring magnets. In short, Figs. 10, 11 and 12 show several important things. First, if the magnetic pole volume density is omitted, the relative axial difference becomes more and more important when the magnetic field axial component is determined far from the ring magnets. When the inner radius is 0.025 m, we incur an error of at least 8% in calculating the magnetic field axial component either close to the ring magnets whereas we incur an error of at least 30% in calculating this one when the distance in length between the ring magnet and the observation point equals 0.01 m. As said previously in the study of the radial component, we see that this error is more important for the same observation point when the ring width is larger. For instance, for an observation point which equals 0.018 m, there is a difference of about 8% in the determination of the magnetic field axial component between the thin ring and the massive ring. When the inner radius is 0.01 m, the error we incur is less important than the one made in the previous case. Indeed, we incur an error of at least 1% in calculating the magnetic field radial component either close to the ring magnets. Moreover, Fig. 12 clearly shows that this error becomes very small either close to the ring magnets (about 0.2%) when the inner radius becomes important (here $r_1 = 0.5\,\mathrm{m}$). Therefore, we deduce that the magnetic pole volume density must be taken into account when the inner radius is small. In the three configurations, we see that the ring width has an important influence which becomes more and more visible when the magnetic field axial components is determined far from the ring magnets. Let us consider now the magnetic field axial component inside the ring permanent magnet which corresponds to the demagnetization field. For this purpose, the relative axial difference has been represented in Fig. 13 for different ring widths inside the ring magnet. Figure 13 shows several important things. First, we see that for each configuration, we incur an important error if we determine the magnetic field axial component in the middle of the ring permanent magnet without taking into account the magnetic pole volume density. If we do not determine the magnetic field axial component in the middle of the ring magnet, the magnetic pole volume density can be omitted only for thin ring magnets (for instance, $r_2 - r_1 = 0.02\,\mathrm{m}$ or $r_2 - r_1 = 0.04\,\mathrm{m}$). For ring magnets whose width becomes larger, the relative axial difference becomes important either close to the ring inner face. Consequently, we deduce that the magnetic pole volume density can be omitted only for thin ring magnets. However, we see that for each configuration, the relative axial difference is not very important either close to the ring outer face and consequently, the magnetic pole volume density can be omitted.

4. CONCLUSION

This paper has discussed the utility of taking into account the magnetic pole volume density in the calculation of the magnetic field components produced by ring permanent magnets. It has been shown that for small thin rings and for ring magnets whose inner radius is higher than 0.1 m, it is not necessary to take into account the magnetic pole volume density if the magnetic field components are determined either close to the ring inner face. Moreover, we have shown that for thin rings, the magnetic pole volume density can be omitted if the magnetic field radial component is determined inside the ring magnet.

REFERENCES

1. Ravaud, R., G. Lemarquand, and V. Lemarquand, "Magnetic field created by tile permanent magnets," *IEEE Trans. Magn.*, Vol. 45, No. 7, 2920–2926, 2009.
2. Hilton, J. and S. McMurry, "Halbach cylinders with improved field homogeneity and tailored gradient fields," *IEEE Trans. Magn.*, Vol. 43, No. 5, 1898–1902, 2007.
3. Ravaud, R., G. Lemarquand, V. Lemarquand, S. Babic, and C. Akyel, "Calculation of the magnetic field created by a thick coil," *Journal of Electromagnetic Waves and Applications*, Vol. 24, No. 10, 1405–1418, 2010.
4. Perigo, E., R. Faria, and C. Motta, "General expressions for the magnetic flux density produced by axially magnetized toroidal permanent magnets," *IEEE Trans. Magn.*, Vol. 43, No. 10, 3826–3832, 2007.
5. Selvaggi, J. P., S. Salon, O. M. Kwon, and M. V. K. Chari, "Computation of the three-dimensional magnetic field from solid permanent-magnet bipolar cylinders by employing toroidal harmonics," *IEEE Trans. Magn.*, Vol. 43, No. 10, 3833–3839, 2007.
6. Selvaggi, J. P., S. Salon, O. M. Kwon, M. V. K. Chari, and M. DeBortoli, "Computation of the external magnetic field, near-field or far-field from a circular cylindrical magnetic source using toroidal functions," *IEEE Trans. Magn.*, Vol. 43, No. 4, 1153–1156, 2007.
7. Ravaud, R., G. Lemarquand, V. Lemarquand, and C. Depollier, "Permanent magnet couplings: Field and torque three-dimensional expressions based on the coulombian model," *IEEE Trans. Magn.*, Vol. 45, No. 4, 1950–1958, 2009.
8. Furlani, E. P., "A two-dimensional analysis for the coupling of magnetic gears," *IEEE Trans. Magn.*, Vol. 33, No. 3, 2317–2321, 1997.
9. Furlani, E. P., "Field analysis and optimization of ndfeb axial field permanent magnet motors," *IEEE Trans. Magn.*, Vol. 33, No. 5, 3883–3885, 1997.
10. Ravaud, R. and G. Lemarquand, "Modelling an ironless loudspeaker by using three-dimensional analytical approaches," *Progress In Electromagnetics Research*, Vol. 91, 53–68, 2009.
11. Ravaud, R. and G. Lemarquand, "Comparison of the Coulombian and Amperian current models for calculating the magnetic field produced by radially magnetized arc-shaped permanent magnets," *Progress In Electromagnetics Research*, Vol. 95, 309–327, 2009.

A Peak to Average Power Ratio Reduction of Multicarrier CDMA System Using Error Control Selective Mapping

Sajjad A. Memon, **A. W. Umrani**, **F. A. Umrani**, **and A. K. Baloch**
Institute of Information and Communication Technologies
Mehran University of Engineering & Technology
Jamshoro 76062, Pakistan

Abstract— Multicarrier Code Division Multiple Access (MC-CDMA) is one of the most promising techniques for high bit rate and high capacity transmission for future broadband mobile services. One major drawback associated with multicarrier systems that significantly degrades power efficiency and makes it less preferred by the industry is Peak to Average Power Ratio (PAPR). One method for reducing PAPR in multicarrier CDMA systems is to use Selective Mapping. We aim to analyze the system performance of multicarrier CDMA system and reduce PAPR using selective mapping that employs convolutional codes to avoid the need of side information. Simulation results show that the error control selective mapping is more effective than the selective mapping. Results also show that the increasing number of PAPR control bits results in much reduced PAPR value but also increases the system complexity.

1. INTRODUCTION

Multicarrier code division multiple access (MC-CDMA) is a promising technique for high-bit-rate and high capacity transmission in mobile communications. It can achieve high-bit- rate transmission with protection against both frequency selective fading and time dispersion channel while at the same time offers a spectrum efficient multiple access strategy [1, 2]. However, a major disadvantage is the high Peak-to-Average Power Ratio (PAPR) that results in nonlinear distortion at a high power amplifier (HPA) and the degradation of the bit error rate (BER). Among many solutions proposed to combat the PAPR issue in MC-CDMA systems, the distortion less methods are very attractive, since the information in transmitted signals is undistorted [3]. The Partial Transmit Sequences (PTS) [4] and the Selective Mapping (SLM) [1, 3, 4] are two of the typical distortion less methods. A modified SLM technique is proposed [5] by integrating selective mapping with error control in the OFDM systems to overcome the need for the transmission of explicit side information. However the effects of error control SLM (EC-SLM) for MC-CDMA systems have not been clarified. In this paper we investigate, the effects of EC-SLM for MC-CDMA using convolutional codes. We explore the EC-SLM which is similar to SLM method [3], but includes robust EC ability and eliminates the need of explicit information to be transmitted and errors in the decoder.

The rest of the paper is organized as follows: Section 2 details the problem of PAPR in MC-CDMA system and proposes a technique to reduce it. Some simulation results are shown in Section 3. The paper is concluded in Section 4.

2. ERROR CONTROL SELECTIVE MAPPING

In selective mapping, parallel data symbols are multiplied by the phase sequences before inverse fast Fourier transform (IFFT) and then symbol sequence with the lowest PAPR is selected and transmitted. To recover the source data at the receiver, information is needed regarding which channel representation of the block of source data has been transmitted. In general, this information is transmitted separately and is called side information [1]. Figure 1 shows the functional block diagram of MC-CDMA system with error-control selective mapping. The complex-valued data symbol $d(k)$ is multiplied with the user specific spreading code $c^{(k)} = (c_0^{(k)}, c_1^{(k)}, \ldots, c_{L-1}^{(k)})^T$ of spreading factor L. The complex-valued sequence obtained after spreading is given in vector notations as, $s(k) = d(k)c(k) = (S_0^{(k)}, S_1^{(k)}, \ldots, S_{L-1}^{(k)},)^T$. A multi-carrier spread spectrum signal is obtained after modulating the components $S_l^{(k)}$, $l = 0, \ldots, L-1$ in parallel onto L sub-carriers. Let us assume that K users are actively transmitting data. The spread data symbols of K users are summed, and then input to the IFFT of size $N = 1 \times L$. The resultant baseband transmission signal for one MC-CDMA symbol, $0 \le t \le T_{s'}$ is expressed as

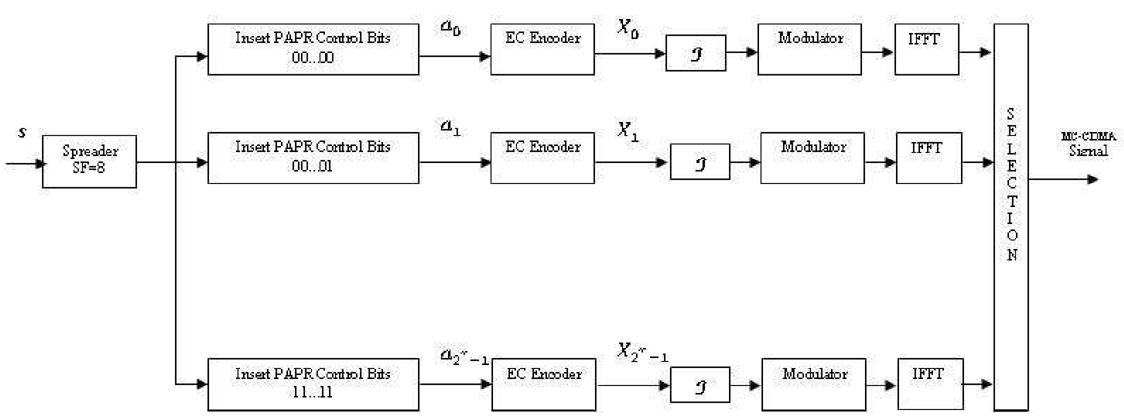
$$s(t) = \sum\nolimits_{l=1}^{L} \sum\nolimits_{k=1}^{K} d_1(k) c_l(k) e^{j2\pi\{(l-1)\}t/T_s} \tag{1}$$

The PAPR of the transmitted signal in (1) can be written as,

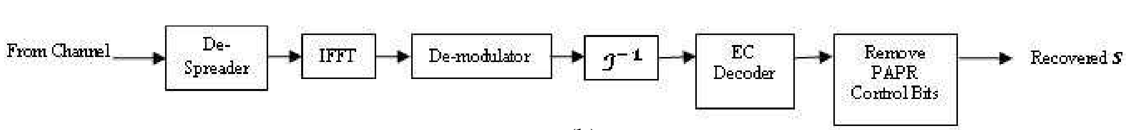
$$\text{PAPR} = \frac{\max |s(t)|^2}{E[|s(t)|^2]} \tag{2}$$

where $\max |s(t)|^2$ denotes the peak power and $E[|s(t)|^2]$ denotes the average power. If the number of subcarriers becomes large, the PAPR of transmitted signal becomes large and thus the signal is distorted by the nonlinear amplifier. To reduce the nonlinear distortion of nonlinear amplifier, we must reduce the PAPR of transmit signal. The basic idea of SLM is to generate several OFDM symbols as candidates then select the one with the minimum PAPR for transmission. The information on phase sequences used for the transmitted signal must be conveyed to the receiver so that receiver can use the side information to tell which candidate was selected for transmission. In the conventional SLM scheme, this information is transmitted as explicit side information [1]. However, due to noise, error free decoding the signal may not be possible. Errors in the reverse mapping would result in the data of whole symbols being lost. In EC-SLM coding as shown in Figure 1(a), during each encoding interval r PAPR control bit(s) are appropriately inserted into the source symbol block s to generate a set of encoder input symbol blocks $\{a_0, a_1, \ldots, a_{2^r} - 1\}$, and each of these blocks is then encoded which is followed by an interleaver to increase the randomness of coded symbols. The coded symbol block with the lowest PAPR is selected to represent the source block. To decode the received words an optional de-interleaver may be used to perform de-interleaving and standard decoding is followed by removing the PAPR control bits. No side information is required by the EC-SLM decoder.

Now let the constraint length of a convolutional code is M and the number of states in the trellis of this code is $2^{(M-1)}$. Assume that for each encoding interval, PAPR control bits are inserted prior to the source block s. In contrast to conventional convolutional encoding where there is only one initial state for the trellis, in a convolutional-SLM encoder (either recursive or non-recursive in form) the r PAPR control bits result in $U = 2^r$ different initial states, and therefore result in Udifferent paths through the trellis where all U different coded blocks are associated with s. However, with non-recursive convolutional encoders, the U different paths starting from the U different states soon merge into the trellis, resulting in coded blocks in the selection set which are strongly correlated with each other. Therefore, non-recursive convolutional coding is unsuitable for use in EC-SLM coding. Instead we use recursive convolutional codes. These codes can be either recursive systematic convolutional (RSC) codes or recursive non-systematic convolutional (NSC) codes. RSC encoders with different initial states result in paths that do not merge in the trellis, and the U different coded blocks which represent the same source block s are relatively random to



(a)



(b)

Figure 1: Error control selective mapping. (a) Encoder. (b) Decoder.

one another. To further improve PAPR, we inserted an interleaver after encoding to improve the randomness of each coded block.

3. SIMULATION RESULTS

Figures 2 and 3 shows the complementary cumulative distribution function (CCDF) plot of the PAPR using recursive systematic convolutional codes selective mapping for MC-CDMA system with $N = 128$ subcarriers. There are $r = 6$ PAPR control bits which results in $U = 64$. The

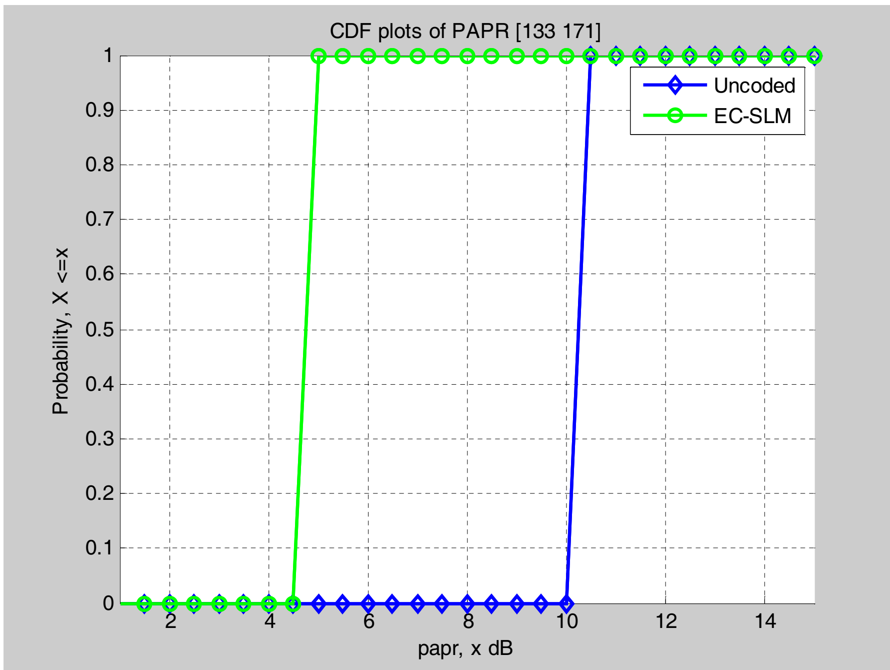


Figure 2: Complementary cumulative distribution function of the PAPR of recursive systematic convolutional codes-selective mapping.

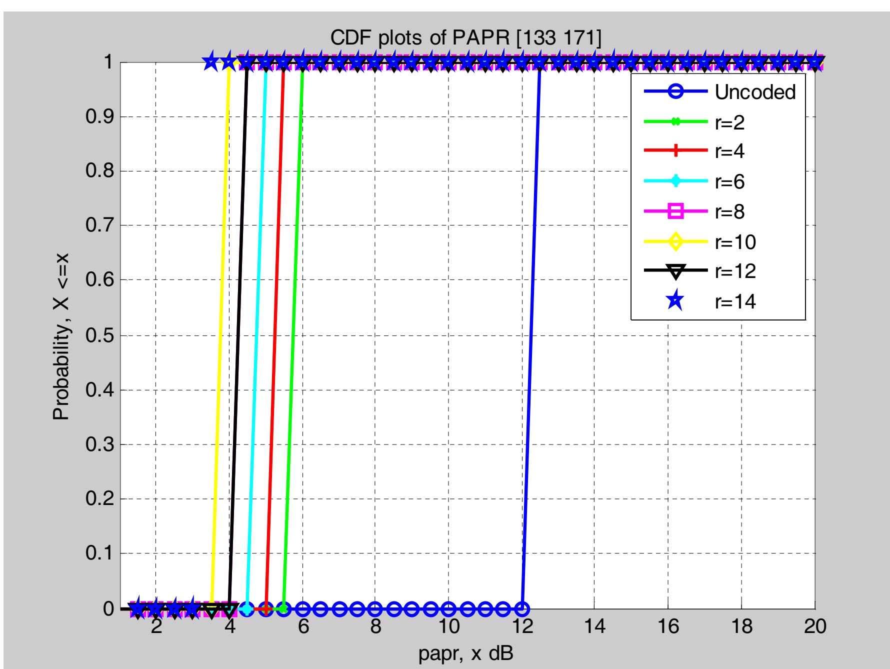


Figure 3: Complementary cumulative distribution function of the PAPR of recursive systematic convolutional codes-selective mapping.

performance is evaluated in AWGN channels with BPSK modulation. The EC codes used conforms to the industry standard with code rate 1/2, constraint length $M = 7$, and RSC code with the generator polynomials $G(133,171)_8$ expressed in octal.

Figure 2 plots the cumulative probability of PAPR for MC-CDMA system against the value of PAPR in dB. As can be seen, the observed PAPR for un-coded MC-CDMA system is distributed from approximately 10 dB to 10.5 dB. Where the same distribution for coded MC-CDMA (with EC-SLM), is from 4.5 dB to 5 dB. It demonstrates that RSC-SLM is effective method to improve the statistical PAPR performance of MC-CDMA signals. With $r = 6$ and after encoding, the PAPR reduction is of 5.5 dB compared to un-coded MC-CDMA system as shown in Figure 2. Figure 3 shows the CCDF plot for several values of PAPR control bits. It demonstrates that as the number of PAPR control bits increases the PAPR reduction also increases. However this improvement is at the cost of system complexity and it further degrades the error rate performance. Since there are no errors after decoding, the coded BER performance of SLM is very close to the BER performance of Convolutional codes (not shown here). The degradation in the performance is merely due to the induction of redundant bits for PAPR control. This results in a tradeoff between choosing number of PAPR control bits and improving BER performance.

4. CONCLUSIONS

We have extended the previously proposed selective mapping method for PAPR control in MC-CDMA to error-control selective mapping coding (EC-SLM) for both effective PAPR reduction and error control. Several well known EC codes such as convolutional codes, turbo codes and LDPC codes can be adopted in our new EC-SLM approach. Simulation results have shown that, the EC-SLM is more effective than the SLM in reducing the PAPR of MC-CDMA and avoid the need for the transmission of explicit side information. Simulation results also showed that the increasing number of PAPR control bits results in much reduced PAPR value but also increases the system complexity. By observation of our simulation results indicate that the optimum value of PAPR control bits, which results in significant PAPR reduction, is one less than the constraint length of convolutional codes.

REFERENCES

1. Ohkubo, N. and T. Ohtsuki, "Design criteria for phase sequences in selected mapping," *IEICE Trans. Commun.*, Vol. E86-B, No. 9, 2628–2636, Sept. 2003.
2. Yang, L., M. You, and J. Li, "Optimized spreading code reallocation technique for PAPR reduction in MC-CDMA systems," *IEEE Global Telecom. Conf.*, 1–6, Nov. 2009.
3. Liu, Y. and G. Lu, "Enhanced selected mapping via candidates combination for PAPR reduction in OFDM," *IEEE 5th WiCom Conf.*, 1–4, Oct. 2009.
4. Ruangsurat, N. and R. M. Rajatheva, "An investigation of peak to average power ratio in MC-CDMA combined with partial transmit sequence," *IEEE VTC'01 — Spring*, Vol. 1, 761–765, Rhodes, May 2001.
5. Yan, X. F., "Error-control selective mapping coding for PAPR reduction in OFDM systems," *IEEE VTC'04*, Vol. 1, 583–587, Apr. 2005.

Detection of Singularities by Wavelet Technique for Extracting Leaky Waves in Piezoelectric Material

D. Benatia[1], T. Fortaki[1], and M. Benslama[2]

[1]Département d'Electronique, Faculté des Sciences de l'Ingénieur, Université de Batna, Algeria
[2]Département d'Electronique, Faculté des Sciences de l'Ingénieur, Université de Constantine, Algeria

Abstract— In this paper, we propose a new numerical method for leaky wave detection of an acoustic microwave signal during the propagation of acoustic microwaves in a piezoelectric substrate. Moreover, we know that the Fourier transform presents a global spectral study of signal, this is not interesting if we want to study a signal locally and know its features in a more precise manner. By the use of wavelet transform, we can reduce this drawback. The originality of the wavelet transform consists of the local analysis of signal singularities where abrupt events appear and hence access to hidden information by using the scale of this transform as up scaling parameters. These singularities (correspond to abrupt variations) inform us of presence of leaky waves in piezoelectric materials.

Furthermore, this transform proved its efficiency in many applications, such as signal processing and the analysis of waves in microstrip structures. Hence, it can play an important role in the modeling of singularities in acoustoelectronic.

1. INTRODUCTION

The investigation of bibliography in micro-acoustic area permits us to point the state of the art. Two major works can be mentioned, the first one is Greeb's paper [1], which explore the interaction using effective permittivity concept, another work of Lakin [2] which elaborate a perturbation theory to explain the interaction phenomena. Following the work of Greeb, Milson [3] in 1977 elaborates a relation based of the charge density, and develops a formalism based on the Fourier Transform. In spite of it, this method doesn't permit to distinguish easily the different modes of propagation by a numerical methods based on the inverse Fourier transform. These works were taken by Junjhunwalla [4] that granted a particular attention to the SSBW (Surface Skimming Bulk Waves). Yashiro and Goto [5] introduce the method of the stationary phase of Lightill [6] to calculate the singularities that inform us on the presence of the pseudo-waves, particularly the leaky waves.

In our case, we propose another approach for the modelling of the acoustic microwaves with a complementary vision to the literature mentioned above. In this approach we interested especially in the detection of singularities by the use of a wavelet transform as detection tool [7, 8] in order to mark the mode of a leaky waves [9, 10].

2. PHENOMENOLOGICAL TENSORIAL PIEZOELECTRIC EQUATIONS

The signal to be treated will be applied to the electrodes of the transducer that generate the compression and dilatation, so a piezoelectric wave is generated and propagated in the X direction (Figure 1).

We consider the space coordinates: $X_1 = X$, $X_2 = Y$, $X_3 = Z$.

The mechanical state of the medium is defined by two magnitudes of tensorial type, the stress T_{ij} and the mechanical deformation (Strain) S_{ij} $(i, j = 1, 2, 3)$. The electric state of the medium is defined by two vectors, the electric field E_k and the electric induction D_i. The stress tensor and the electric induction are given by:

$$T_{ij} = C_{ijkl} \cdot S_{kl} - e_{kij} \cdot E_k \tag{1}$$

$$D_i = e_{ikl} \cdot S_{kl} + \varepsilon_{ik} \cdot E_k \tag{2}$$

with $i, j, k, l = 1, 2, 3$, where ε_{ik}: permittivity tensor (F/m), e_{jkl}: piezoelectric tensor (c/m), C_{ijkl}: elastic tensor (N/m^2).

The strain is bound to the relative displacements of the particles of the material environment is defined by:

$$S_{ij} = \frac{1}{2}\left(\frac{\partial U_i}{\partial X_j} + \frac{\partial U_j}{\partial X_i}\right) \tag{3}$$

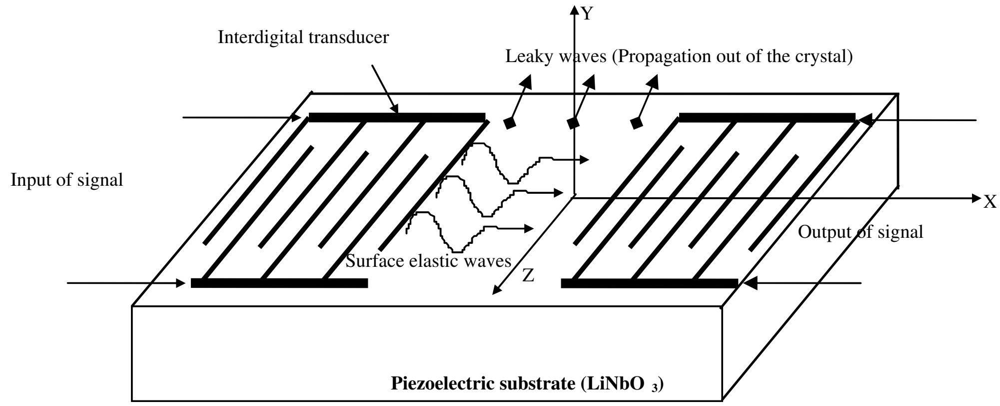


Figure 1: $LiNbO_3$ Crystal excited by transducer.

where U_i represents the elastic displacement of the particle ($i = 1, 2, 3$).

Note that in the quasi-static approximation, we can define an electric field of components:

$$E_i = -\frac{\partial U_4}{\partial X_i} \tag{4}$$

where U_4 is the electric potential (with $i = 1, 2, 3$).

In the quasi-static approximation, the Maxwell's equation amount to the Poisson's equation:

$$\text{div} \cdot \vec{D} = \frac{\partial D_i}{\partial X_i} = 0 \tag{5}$$

The movement of the particles under the action of stress (constraints), is described by the following:

$$\nabla T = \frac{\partial T_{ij}}{\partial X_i} = \rho \cdot \frac{\partial^2 U_j}{\partial t^2} \tag{6}$$

where ρ is the mass density of medium.

Replacing (3) and (4) in (1) and (2), we obtain:

$$T_{ij} = C_{ijkl} \cdot \frac{1}{2}\left(\frac{\partial U_k}{\partial X_l} + \frac{\partial U_l}{\partial X_k}\right) + e_{kij} \cdot \frac{\partial U_4}{\partial X_k} \tag{7}$$

$$D_i = e_{ikl} \cdot \frac{1}{2}\left(\frac{\partial U_k}{\partial X_l} + \frac{\partial U_l}{\partial X_k}\right) - \varepsilon_{ik} \cdot \frac{\partial U_4}{\partial X_k} \tag{8}$$

Replacing (7) and (8) in (5) and (6), we obtain the piezoelectric tensorial equations:

$$C_{ijkl}\frac{\partial^2 u_k}{\partial \cdot X_i \partial \cdot X_l} + e_{lij}\frac{\partial^2 \phi}{\partial \cdot X_k \partial \cdot X_i} = \rho\frac{\partial^2 u_j}{\partial \cdot t^2} \tag{9}$$

$$e_{ikl}\frac{\partial^2 u_k}{\partial \cdot X_i \partial \cdot X_l} - \varepsilon_{ik}\frac{\partial^2 \phi}{\partial \cdot X_k \partial \cdot X_i} = 0 \tag{10}$$

3. THE FORM OF SOLUTION

Consider the following form of the surface wave (partial wave):

$$U_i = u_i \exp(jk \cdot \alpha_i Y) \exp -j\left[\omega \cdot t - \beta(1 + j\gamma)X\right] \tag{11}$$

where u_i ($i = 1, 2, 3$) are the displacement amplitudes, u_i ($i = 4$) is the amplitude of the electric potential, k is the constant of propagation, the α_i are the penetration coefficients of the wave inside the piezoelectric substrate (Figure 1), γ is the coefficient of longitudinal attenuation and ω is the angular pulsation.

Equations (9) and (10) can be written in a matrix form as:

$$[A][U] = [0] \tag{12}$$

with $[U] = [u_1,\, u_2,\, u_3,\, u_4]^T$, $[A]$ is a matrix (4×4)

The determinant of the matrix $[A]$ must be zero to ensure a non trivial solution, it can be written as:

$$\sum_{i=0}^{8} \beta_i \cdot \alpha^i = 0 \tag{13}$$

where B_i depends on the piezoelectric material features (C_{ijkl}, ε_{ik}, e_{lij}) and of the acoustic velocity V_S.

The determinant of the matrix $[A]$ must be zero, we have eight complex roots ($i = 1 \ldots 8$):

$$\alpha_i = a_i + jb_i \tag{14}$$

where a_i: is the real part and b_i: is the imaginary part, with $a_m = a_{m+1}$ and $b_m = -b_{m+1}$ where $m = 1, 3, 5, 7$.

In the surface mode (or Rayleigh wave) the α_i ($i = 1 \ldots 8$) are conjugated by pairs and only the complex roots with negative imaginary part are taken into consideration (for convergence reasons).

Let us first neglect the longitudinal attenuation ($\gamma = 0$) and insert (14) in (11) to obtain:

$$U_i = u_i \exp -jb_i\beta \cdot Y \exp -j\left[\omega \cdot t - \beta(X + a_i Y)\right] \tag{15}$$

If we go inside the crystal (Y tends to $-\infty$), the wave U_i tends to zero. This corresponds to surface acoustic waves (S.A.W) (Figure 1). In the opposite case (Y tends to $+\infty$), U_i tends $+\infty$ (without physical signification).

4. LEAKY WAVES

The variation of the acoustic velocity V_S allows us to obtain $b_i = 0$ (imaginary part) and (15) becomes:

$$U_i = u_i \exp -j\left[\omega \cdot t - \beta \cdot (X + a_i Y)\right] \tag{16}$$

The wave nature leaky waves (**L.W**) depend on the sign of the real part of $\alpha_i(a_i)$.

If a_i is positive, we have leaky waves (radiation out of the crystal (Figures 1 and 2)).

5. GENERAL FORM OF THE ACOUSTIC WAVE SOLUTION

The general form of U_i ($i = 1, 2, 3, 4$) is expressed by:

$$U_i(k, Y) = \sum_{n=1}^{4} C_n \cdot A_i^{(n)} \cdot \exp^{-j[\omega t - k \cdot \alpha_n \cdot Y]} \tag{17}$$

where $A_i^{(n)}$ are the components of eigen vectors associated to coefficients α_n [3]. C_n are the constants determined by the traction-free boundary conditions [3]. The product $C_n \cdot A_i^{(n)}$ depends on C_{ijkl}, ε_{ik}, e_{lij} and of the acoustic velocity V_S.

For $i = 4$, we have a potential electric given by:

$$U_4(k, Y) = \sum_{n=1}^{4} C_n \cdot A_4^{(n)} \cdot \exp^{-j[\omega t - k \cdot \alpha_n \cdot Y]} \tag{18}$$

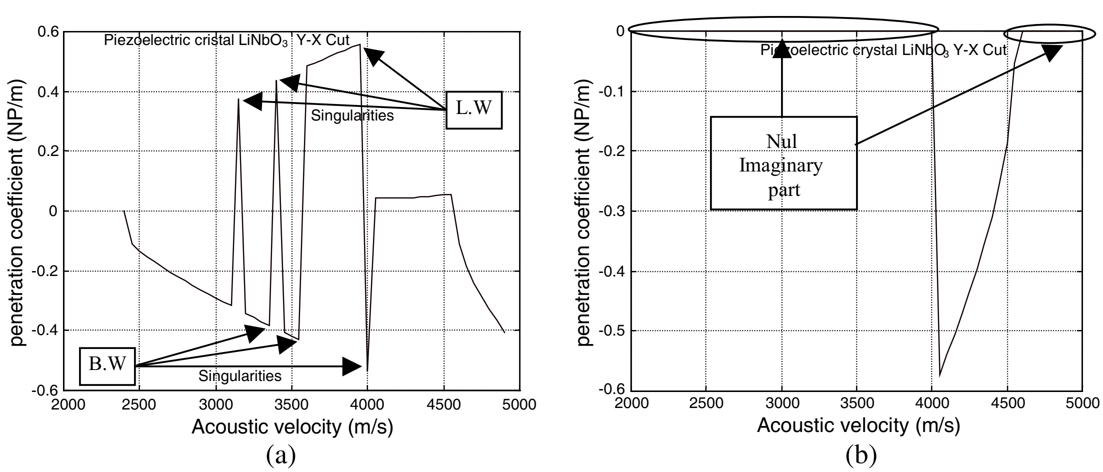


Figure 2: Penetration coefficient α_4. (a) Real part. (b) Imaginary part.

6. WAVELET TRANSFORM OF SIGNAL

The signal chosen for analysis by wavelet transform is given by Equation (18). It represents an electric potential "U_4" coupled with an elastic wave of components "U_1, U_2, U_3". This wave with the frequency "f" propagates along the X direction and guided on the free surface of the piezoelectric (Figure 1).

The space-scale type wavelet transform of U_4 (in the neighbourhood of the surface: $Y \approx 0$) is given by this convolution product [8]:

$$T_{U_4}(Y \approx 0, X, a) = U_4(Y \approx 0, X) \otimes \frac{1}{\sqrt{a}} \psi^* \left(\frac{X}{a} \right) \tag{19}$$

where U_4 $(Y \approx 0, X)$ is a signal in the neighbourhood of the material surface $(Y \approx 0)$. $\psi * (X)$ is the complex conjugate of the wavelet (Mexican-hat): $\psi(X) = d^2/dX^2(e^{-X^2/2})$.

The frequency-scale type wavelet transform of U_4 becomes a simple product:

$$T_{U_4}(Y \approx 0, f, a) = \sqrt{a} \cdot U_4(Y \approx 0, f) \cdot \psi^* (a \cdot f) \tag{20}$$

where U_4 $(Y \approx 0, f)$ is the Fourier transform of a signal in the neighbourhood of the material surface. $\psi * (f)$ is the Fourier transform of $\psi * (X)$: $\psi * (f) = (2 \cdot \pi)^{1/2} \cdot e -^{(4 \cdot \pi^2 \cdot f^2/2)} \cdot (2 \cdot \pi \cdot f)^2$.

Replacing the expression of $\psi * (a \cdot f)$ in Equation (20), the wavelet transform of U_4 in this case becomes:

$$T_{U_4}(Y \approx 0, f, a) = |U_4(k, 0)| \cdot \underbrace{\sqrt{2 \cdot a \cdot \pi} \cdot e^{-\frac{4 \cdot \pi^2 a^2 \cdot f^2}{2}} \cdot (2\pi f \cdot a)^2}_{\psi*(a \cdot f)} \tag{21}$$

with $|U_4(k, 0)| = \sum_{n=1}^{4} C_n \cdot A_4^{(n)}$, $k = 2 \cdot \pi \cdot f/V_S$.

7. RESULTS AND DISCUSSION

The detection of leaky waves appears at the level of the penetration coefficients when the acoustic velocity Vs change its value. This change results in an annulation of the imaginary parts of the penetration coefficients. We note that when the imaginary part becomes null, it appears at the level of signal wavelet transform an ***abrupt variations*** called ***singularities***. These singularities are not always observable. The use of the scale of this transform permits to visualise them. Once these singularities are detected, the sign of the real part can inform us about the leaky wave ($a_i > 0$).

The analysis of the signal by wavelet transform (Equation (21)) clearly shows these singularities at the level of the contour of three-dimensional figure (the above view). This figure englobes the Wavelet Transform of "U_4", the Acoustic Velocity "V_S" and the Scale "a" with the frequency "f" as parameter (Figures 3, 4 and 5). At the frequency $f = 1$ GHz, it is impossible to detect the singularities for a scale superior to 10^{-8}, In this case we can't detect the singularities (Figure 3). For a good detection, it is necessary to reduce the scale from 10^{-8} to 10^{-10} (Figures 4 and 5), these singularities appear more and more clearly.

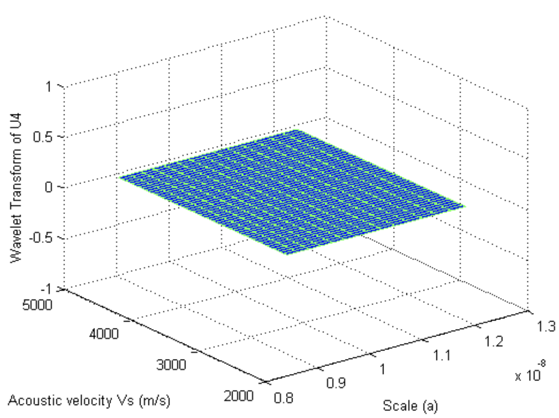


Figure 3: The wavelet transform of the U_4 scale order (10^{-8}).

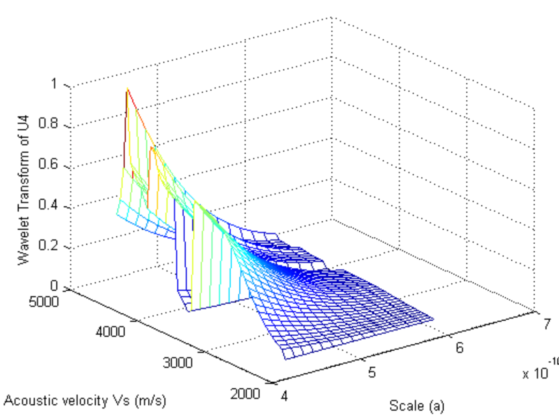


Figure 4: The wavelet transform of U_4 scale order (10^{-10}).

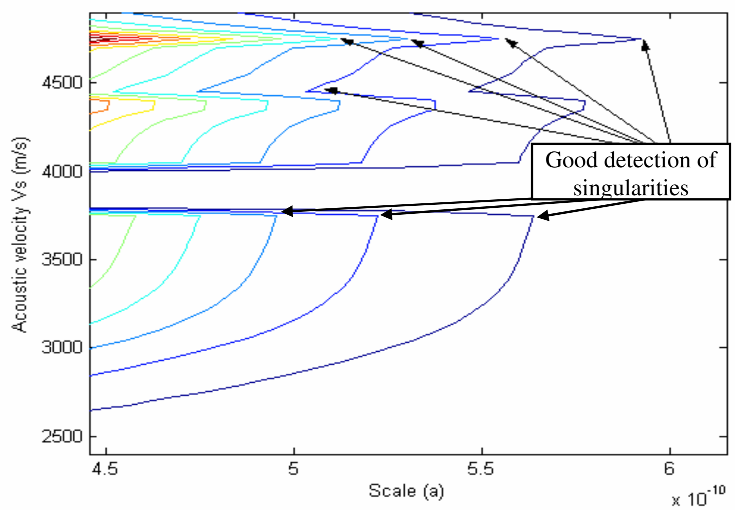


Figure 5: The above view of Figure 4 (Good detection of singularities).

8. CONCLUSION

In this work, we have developed a model for studying the behaviour of electroacoustic waves at the level of singularities which appear for some velocities giving the leaky waves. This model that insures the detection of singularities using the wavelet transform. This latter has the property of being locally maximal around the points where the signal is singular or singular. In this paper, we were interested in the influence of the scale "a" of this transform. The decrease of the scale allowed us to detect the singularities whatever the frequency so that we can obtain all the details about the propagation of acoustic waves in particular on the leaky waves. This information can be useful for many applications such us the antenna device.

REFERENCES

1. Greeb, C. A. J., et al., "Electrical coupling properties of acoustics and electric surface waves," *Physical Report*, Vol. 1C, No. 5, 235–268, 1971.
2. Lakin, K. M., "Perturbation theory for Electromagnetic coupling to Elastic surface waves on piezoelectric substrates," *J. Appl. Phys.*, Vol. 42, No. 3, 899–906, 1971.
3. Milson, R. F., N. H. C. Reilly, and M. Redwood, "Analysis of generation and detection of surface and bulk acoustic waves by interdigital transducers," *IEEE Trans. Sonics and Ultrason.*, Vol. 24, No. 3, 147–166, 1977.
4. Jhunjhunwalla, A. and J. F. Vetelino, "Spectrum of acoustic waves emanating from an IDT on a piezoelectric half space," *Proceedings of IEEE Ultrasonics Symposium*, 945–952, 1979.
5. Yashiro, K. and N. Goto, "Analysis of generation of acoustic waves on the surface of a semi-ifinite piezoelectric solid," *IEEE Trans. Sonics and Ultrason.*, Vol. 25, No. 3, 146–153, 1978.
6. Lightil, J., *Waves in Fluid*, University Press, Cambridge, 1978.
7. Mallat, S. and W. L. Wan, "Singularity detection and processing with wavelets," *IEEE Trans. on Information Theor.*, Vol. 38, 617–643, 1992.
8. Arneodo, A., F. Argoul, E. Bacry, J. Elezgaray, and J. F. Muzy, *Ondelettes, Multifractales et Turbulences*, Diderot (ed.), Art et Sciences, Paris, France, 1995.
9. Fusero, Y., S. Ballandras, J. Desbois, J. M. Hodé, and P. Ventura, "SSBW to PSAW conversion in SAW devices using heavy mechanical loading," *IEEE Trans. on Ultrason., Ferroelectrics, and Frequency Control*, Vol. 49, No. 6, 805–814, 2002.
10. Martin, F., "Propagation characteristics of harmonic surface skimming bulk waves on ST," *Electron Lett.*, Vol. 38, 941–942, August 2002.

Electromagnetic Study of Planar Periodic Structures Using a Multi-scale Approach

S. Mili and T. Aguili
Laboratory of the Communication Systems, National Engineering School of Tunis, Tunisia

Abstract— A Multi-Scale Approach combined to the generalized equivalent circuit modeling is used in this paper to alleviate the electromagnetic study of planar periodic structures. In fact, when the elementary cell contains fine details, the study of the whole large structure is very time consuming and leads to badly scaled matrices due to the critical dimensions ratio. To circumvent such difficulties, the multi-scale (MS) method dissociates fine details from the whole structure; the elementary cell is studied separately and then replaced by its surface impedance operator. The main advantage of the MS method is the significant reduction of the problem's high aspect ratio since fine details are studied separately of the larger structure. Moreover, the manipulated matrices are well conditioned leading then to a better numerical convergence. When the structure complexity increases, the MS method requires less processing time and memory storage than the moment method (MoM). Values obtained with the MS method are in agreement with those obtained with the MoM when the elementary cell is well described.

1. INTRODUCTION

The electromagnetic behavior of planar structures with periodically arranged elements was widely investigated during this century. Such periodic structures play a key role in many antennas systems since they can offer multi-band and/or broadband characteristics [1] which can be tuned by integrating PIN Diodes into the elementary cell [2, 3]. Classical methods (MoM, FEM, and TLM) found their limits when applied to complex structures (i.e., containing many details) since they require very long solution time and important memory storage. Moreover, if these structures present critical dimensions ratio, the manipulated matrices become badly scaled and the numerical convergence is hard to establish. An alternative method adapted to the electromagnetic study of complex structures is presented in this paper. It is called the MS-GEC approach and consists of a Multi-Scale (MS) approach combined to the Generalized Equivalent Circuit (GEC) Modeling.

2. DESCRIPTION OF THE MS-GEC APPROACH

The MS-GEC method can be used for the electromagnetic study of any complex structure. In this paper, it will be applied particularly to compute the input impedance of some periodic structures. Instead of treating the whole structure at once, the main idea of the MS-GEC method is to distinguish the elementary cell which will be studied separately in order to compute its surface impedance matrix. Next, the latter matrix is converted to an impedance operator. The last step consists of replacing the unit cell within the initial structure by its surface impedance operator. When the structure contains PIN diodes, these latter can be easily integrated in the MS-GEC approach thanks to their equivalent impedance model [2].

Since fine details are located in the unit cell, to study it separately solves the problem of badly scaled matrices because there is no critical dimensions ratio.

3. THE MS-GEC APPROACH APPLIED TO PERIODIC STRUCTURES

The MS-GEC method is applied to the two periodic structures (p: periodicity number) depicted in Fig. 1. These structures are located in the cross section of a parallel plates EMEM waveguide: two perfect electric walls to the top and the bottom, lateral walls are magnetic.

3.1. First Step of the MS-GEC Method: Study of the Unit Cell

When applied to periodic structures, the first step of the MS-GEC approach is to study the unit cell separately. The aim is to compute its surface impedance matrix. For that, the elementary cell is enclosed by convenient boundary conditions which can be chosen between the following: (a) Perfect Electric boundaries, (b) Perfect Magnetic boundaries, (c) periodic boundaries, or (c) a combination of these latter boundary conditions. In this work, we choose to enclose the unit cell by EMEM boundaries: two Perfect Electric boundaries to the top and the bottom, lateral boundaries are Perfect Magnetic. Next, N modal sources chosen among the modes of the local

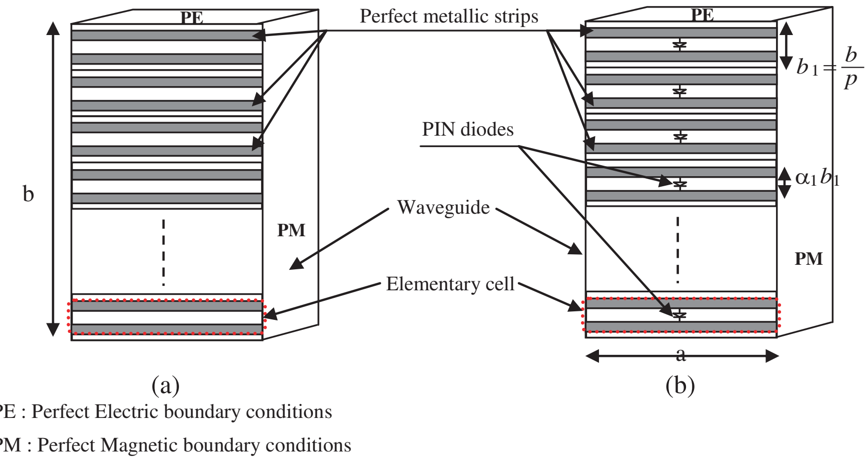


Figure 1: Periodic structures: (a) structure without PIN diodes, (b) structure with PIN diodes.

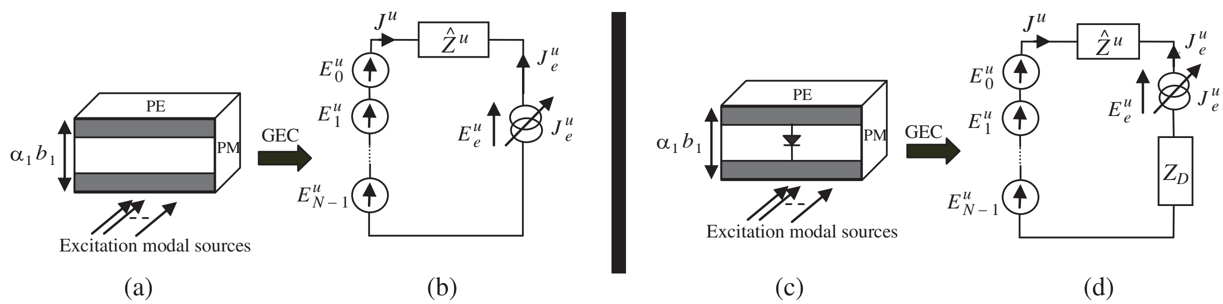


Figure 2: (a) The unit cell of the structure Fig. 1(a), (b) its generalized equivalent circuit, (c) the unit cell of the structure Fig. 1(b), and (d) its generalized equivalent circuit.

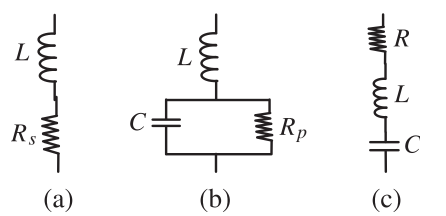


Figure 3: The PIN Diode: (a) forward bias equivalent circuit, (b) reverse bias equivalent circuit, (c) reverse bias equivalent RLC circuit [2].

modal basis will artificially excite the unit cell in order to compute its surface impedance matrix. These modal sources are called also active modes. The GEC method [5, 6] on which relies the MS-GEC approach consists of representing the Maxwell equations and the continuity conditions with a generalized equivalent circuit model.

Lets consider the unit pattern of the structures in Fig. 1. Their generalized equivalent circuits are shown Fig. 2.

Lets (f_n^u) be the local modal basis of the EMEM waveguide enclosing the unit pattern. The excitation modal sources E_i^u, $i \in [0, N-1]$ are expressed as follows: $E_i^u = V_i^u f_i^u$ where the (f_i^u) are the active modes of the waveguide. The impedance operator $\hat{Z}^u$ is expressed as a function of higher-order modes $|f_m^u\rangle$ and their modes' impedances z_m^u [4–6]. The unknown of the problem J_e^u describes the electromagnetic state on the discontinuity interface. It is presented by a current virtual source expressed as a series of known test functions g_p^u weighted by unknown coefficients. The virtual source is defined on the metallic strips and on the PIN diode domain (if it exists), it is null elsewhere.

$$\begin{cases} \hat{Z}^u = \sum\limits_{\substack{m \\ m \neq actif}} |f_m^u\rangle \, z_m^u \, \langle f_m^u| \\ J_e^u = \sum\limits_p x_p^u \, g_p^u \end{cases} \tag{1}$$

If PIN diodes are incorporated in the structure, they can be integrated within the GEC thanks to their equivalent impedance Z_D [2]. In fact, a PIN diode can be modeled using one of the following equivalent circuit models presented in Fig. 3.

In this paper, the values used for forward bias are $R = 5\,\Omega$ and $L = 0.4\,\text{nH}$. For reverse bias, a capacitance $C = 0.27\,\text{pF}$ is added. Let w be the width of the PIN diode and d its height. The equivalent impedance of the PIN diode is expressed in (2).

$$ZD = \begin{cases} \frac{w}{d}\left(R + jL\omega\right): & \text{forward bias} \\ \frac{w}{d}\left(R + jL\omega - \frac{j}{C\omega}\right): & \text{reverse bias} \end{cases} \tag{2}$$

Let consider the unit cell of Fig. 2(a). Based on its corresponding equivalent circuit model depicted Fig. 2(b), the generalized Ohm and Kirchhoff laws give the equations system (3).

$$\begin{cases} J^u = -J_e^u \\ E_e^u = E_0^u + E_1^u + \ldots + E_{N-1}^u + \hat{Z}^u J_e^u \end{cases} \tag{3}$$

A formal relation between sources (real and virtual) and their duals is expressed in (4).

$$\begin{pmatrix} J^u \\ \\ \\ E_e^u \end{pmatrix} = \begin{bmatrix} 0 & -1 & -1 & \ldots & -1 & \\ & & & & & \\ & & & & & \\ 1 & 1 & 1 & \ldots & 1 & \hat{Z}^u \end{bmatrix} \begin{pmatrix} E_0^u \\ E_1^u \\ \vdots \\ E_{N-1}^u \\ J_e^u \end{pmatrix} \tag{4}$$

Next, we apply the Galerkin method to Equation (4). Consequently, the impedance matrix $[Z_u]$ of the unit cell is given by (5).

$$[Z_u] = \frac{1}{2}\left([A]\ [Z]^{-1}\ [A]^T\right)^{-1} \quad \text{where } A(i,p) = \langle f_i^u \mid g_p^u \rangle, \quad Z(p,q) = \left[\left\langle g_p^u \,\middle|\, \hat{Z}^u g_q^u \right\rangle\right] \tag{5}$$

The same development can be applied to the unit cell with PIN diode of Fig. 2(c) and its GEC depicted Fig. 2(d). In this case, the impedance matrix $[Z_u]$ is given by (6).

$$[Z_u] = \frac{1}{2}\left([A]\,[Z]^{-1}\,[A]^T\right)^{-1} \text{ where } A(i,p) = \langle f_i^u \mid g_p^u \rangle, \quad Z(p,q) = \left[\left\langle g_p^u \,\middle|\, \left(\hat{Z}^u + ZD\right) g_q^u \right\rangle\right] \tag{6}$$

3.2. Second Step of the MS-GEC Method: Study of the Equivalent Total Structure

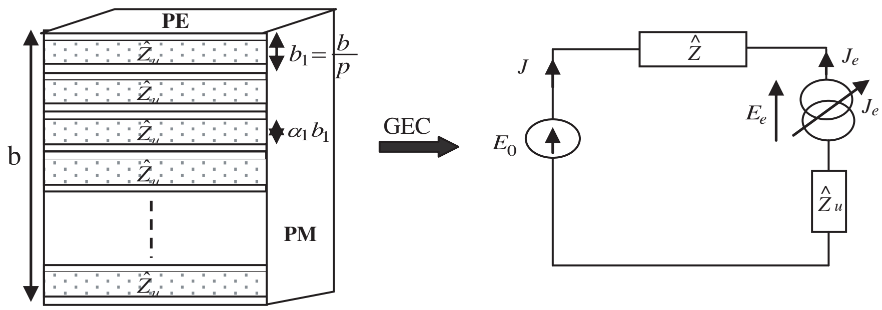
In the first step, the unit cell was studied separately in order to compute its surface impedance matrix. The main idea of the MS-GEC method is to convert the latter matrix into an impedance operator as expressed in (7) and then replace the unit cell with its impedance operator within the initial structure.

$$\hat{Z}_u = \sum_{i=1}^{N}\sum_{j=1}^{N} \left|f_{i-1}^u\right\rangle Z_u(i,j) \left\langle f_{j-1}^u\right| \quad \text{where } \left(f_i^u\right) \text{ are the excitation sources of the unit cell} \tag{7}$$

The Equivalent structure obtained by replacing the unit cell by its impedance operator is depicted Fig. 4(a). The corresponding GEC is shown in Fig. 4(b).

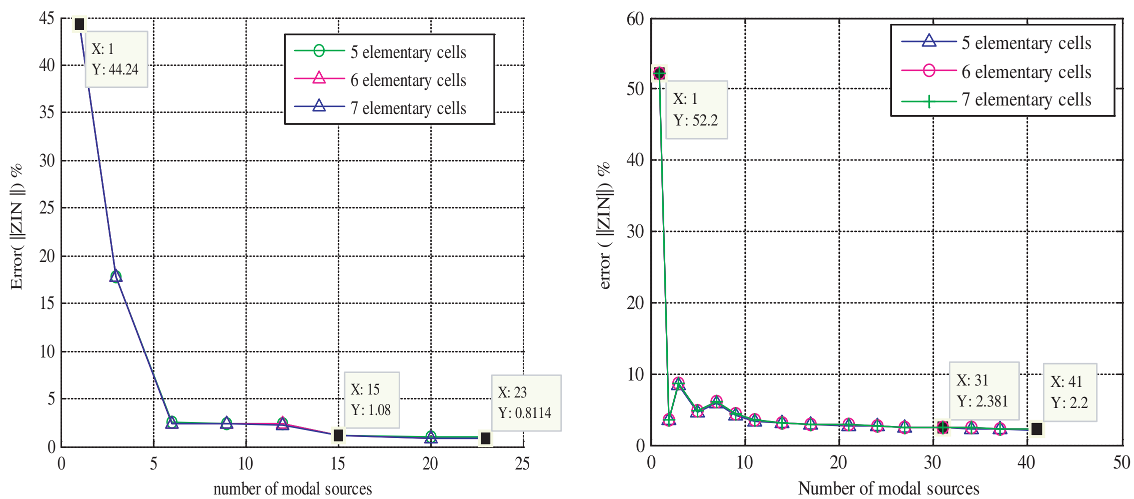
A similar development as done for the unit cell leads to the expression of the input impedance as stated in (8).

$$Z_{IN\_MS-GEC} = \frac{1}{2}\frac{1}{[A][Z]^{-1}[A]^T} \text{ where } A(1,p) = \langle f_0 \mid g_p \rangle, \quad Z(p,q) = \left[\left\langle g_p \,\middle|\, \left(\hat{Z} + \hat{Z}_u\right) g_q \right\rangle\right] \tag{8}$$



(a) (b)

Figure 4: (a) The equivalent structure with impedance operator, (b) the corresponding GEC.



(a) (b)

Figure 5: Variation of the relative error on the input impedance with the number of active modes: (a) structure of Fig. 1(a), (b) structure of Fig. 1(b).

4. NUMERICAL RESULTS

The input impedance computed with the MS-GEC method applied to the two proposed structures is compared to those given by the Moment method. Let $\xi = 100\frac{||Z_{IN\_MoM}||-||Z_{IN\_MS-GEC}||}{||Z_{IN\_MoM}||}$ be the relative error between the MoM and the MS-GEC methods. The relative error variation with the number of active modes has been investigated. Fig. 5(a) shows that the relative error decreases to be less than 1% when the number of active modes is sufficiently enough to sharply describe the unit cell of the structure Fig. 1(a). In this case, we can say that the impedance operator of the unit cell is equivalent to the unit cell itself.

When the considered structure contains PIN diodes (Fig. 1(b)), the relative error is less than 2.3% when a sufficient number of active modes are used. We notice that when the structures' complexity increases, more active modes are required.

5. CONCLUSIONS

In this paper, a Multi-Scale approach combined to the generalized equivalent circuit has been applied to compute the input impedance of some periodic structures with various complexities. We have shown that when the unit cell is well described, then the MS-GEC and the MoM method are in agreement. A better description of the unit cell is ensured when this latter is excited by a sufficient number of active modes. In fact, increasing the number of active modes to excite the unit cell enables a better description of the coupling between the unit cell and the higher level.

REFERENCES

1. Vardaxoglou, J. C., *Frequency Selective Surfaces, Analysis and Design*, John Wiley and Sons, 1997.

2. Kiani, G. I., K. P. Esselle, A. R. Weily, and K. L. Ford, "Active frequency selective surface using PIN diodes," *IEEE Antennas and Propagation Society International Symposium*, 4525–4528, Honolulu, HI, June 9–17, 2007.
3. Tennant, A. and B. Chambers, "Adaptive radar absorbing structure with PIN diode controlled active frequency selective surface," *Smart Materials and Structures*, Vol. 13, 122–125, 2004.
4. Aubert, H. and H. Baudrand, *L'électromagnétisme par les Schémas Équivalents*," Cépaduès ed., 2003.
5. Aguili, T., "Modélisation des composants S.H.F planaires par la méthode des circuits équivalents généralisés," Thesis, National Engineering school of Tunis ENIT, May 2000.
6. BenSalah, T., C. L. Aguili, and T. Aguili, "Renormalization group application to multi-port model for studying fractal-shaped structures' diffraction," *PIERS Proceedings*, 1629–1633, Beijing, China, March 23–27, 2009.

Study of Edge Effect of 4340 Steel Specimen Heated by Induction Process Using Axi-symmetric Simulation

N. Barka, A. Chebak, and J. Brousseau
Département de mathématiques, d'informatique et de génie
Université du Québec à Rimouski, Rimouski, Québec G5L 3A1, Canada

Abstract— Induction heating is a case hardening process used to improve the performance of machine components by generating a hard martensitic microstructure and high compressive residual stresses at the surface. A typical industrial application of induction hardening consists in heating specific regions of a part above the austenitizing temperature. The hardness profiles have to be controlled to improve wear resistance and enhance contact and bending fatigue life. This paper presents a study of electromagnetic edge effect on cylindrical steel specimen heated by induction heating process using high frequency generator. This research is performed by 2D finite element simulation and it conducts to develop a uniform hardness profile. The simulation is based on coupling of electromagnetic field and heat transfer. The axi-symmetric model is developed using Comsol software based on an adequate formulation and taking into account the material properties and process parameters such as the initial inductor current density (A/m^2), the frequency (kHz) and the heating time (s). The obtained induced currents and temperatures distributions in the part are analyzed versus the inductor dimensions. Experimental results are also given for the hardness profile. The originality of this paper is that it introduces a new method to eliminate completely the electromagnetic edge effect in the high frequency heating case and therefore to have a uniform hardness profile. The proposed method consists principally to place two thin cylindrical discs (flux concentrators) such a way the part is taken in sandwich, taking care to leave a small gap that can be optimized by simulation. The obtained results demonstrate that the final hardness profile is quasi-uniform as the temperature and induced currents are almost identical across the part.

1. INTRODUCTION

The induction hardening process principle is based on the application of electric AC power into the coil which is coupled with the part to be treated. The resultant current in the inductor generates a variable electromagnetic field that induces eddy currents at the surface of the metallic component placed nearby. This induced heating process results from a unique coupling of distinct physical phenomena: electromagnetic, thermal and metallurgical. Figure 1 illustrates the overall process of induction hardening. For a given heating time, the case depth of the part depends on the frequency and the current intensity in the coil. The cooling rate affects directly the final hardness profile and the compressive residual stresses [1, 2]. A typical industrial application of induction hardening consists in heating specific regions of a part above the austenitizing temperature (Ac_3) to allow martensite to be formed after quenching. The hardness profile has to be controlled to improve wear resistance and enhance contact fatigue life. Induction hardening is also an environmental friendly technology since it does not use any plating, nor produce carbon monoxide [1–3].

The application of high frequency power (HF) is quite interesting than medium frequency power (MF) because the hardness profile can be better controlled during heating since the induced currents are concentrated at small layer close to the surface and the hardness profile is less sensitive to machine power [3, 4]. The mechanical components such as the cams, bearing surface, friction wheels could be treated by HF in order to have a thin hard layer and compressive residual stresses beneficent for contact fatigue. However, it is very difficult to realize uniform hardness profile using HF power even if the part has a simple geometry. In fact, the electromagnetic edge effect is present in this case because the magnetic field is concentrated on the edges of the heated part and it influences the eddy currents distribution. Consequently, the temperature and hence the case depth are more important in the edges than the middle plane of the part. It is then necessary to analyze the edge effect, to quantify the sensitivity of the hardness profile to coil thickness and to find specific methods to reduce this effect. In the literature, there is no reference that has studied the problem of edge effect in induction heating process and its effects on the hardness profile. Also, the reduction of the edge effect has never before been thoroughly covered and documented.

This paper presents a study of electromagnetic edge effect on cylindrical steel specimen heated by induction heating process. This research is based on an axi-symmetric model developed using Comsol multi-physics software. The developed model considers electromagnetic and thermal

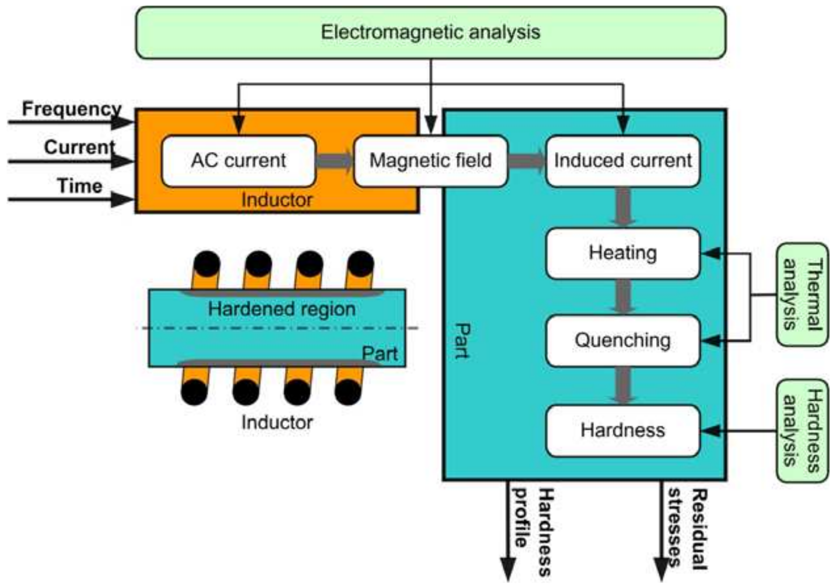


Figure 1: Schematic representation of the induction hardening process.

properties of the material. The machine parameters, used in this work, are principally the initial inductor current density (A/m^2), the generator frequency (kHz) and the heating time (s). The resultant induced currents and temperatures distributions in the part are analyzed versus the inductor dimensions. An experimental validation is also realized for the hardness profile. The paper presents also an original method able to eliminate completely the electromagnetic edge effect of the part and ensure a uniform hardness profile. The proposed method consists mainly to add two discs acting as flux concentrators and placed very close to the part edges. The simulation permits then to optimize the gap between each disc and the part. The obtained results demonstrate that the final hardness profile is uniform as the temperature and induced currents are almost identical across the part. The developed method could greatly help mechanical designer to optimize the heating process and enhance the hardness quality of produced part.

2. FINITE ELEMENT METHOD FORMULATION

The global system of electromagnetic equations used to derive the magnetic field distribution is based on the Maxwell's equations that are combined with the constitutive relations that introduce the material electromagnetic properties. Considering the magnetic potential vector ($\mathbf{A}$) formulation and neglecting the hysteresis and magnetic saturation, the general equation governing the electromagnetic behavior can be expressed as [3, 5].

$$\frac{1}{\mu(T)}\nabla^2\mathbf{A} = -j\omega\sigma(T)\mathbf{A} + \mathbf{J}_0 \tag{1}$$

The resolution of the electromagnetic problem is used to calculate the thermal energy generated during induction heating. The heat ($\mathbf{Q_{Ind}}$) is generated in the surface layer and transferred into the part core by conduction mode. The following equation shows the heat can be expressed in function with the magnetic potential vector [3, 5]

$$\mathbf{Q_{Ind}} = \frac{\left\|(\nabla^2\mathbf{A})^2\right\|}{\mu(T)\cdot\sigma(T)} \tag{2}$$

The thermal analysis is coupled with the electromagnetic problem. The mode of heat transfer by conduction is most important during induction heating [2, 3]. The heat transfer mechanism can be described by the following equation of Fourier-Kirchhoff. In addition, a small fraction of this energy is lost by convection and radiation modes. The convection is approximated to the conduction in the environment and the radiation is neglected during heating due to the short heating time [6].

$$k(T)\nabla^2 T = \gamma C_P(T)\frac{\partial T}{\partial z} + \dot{\mathbf{Q}}_{\mathbf{Ind}} \tag{3}$$

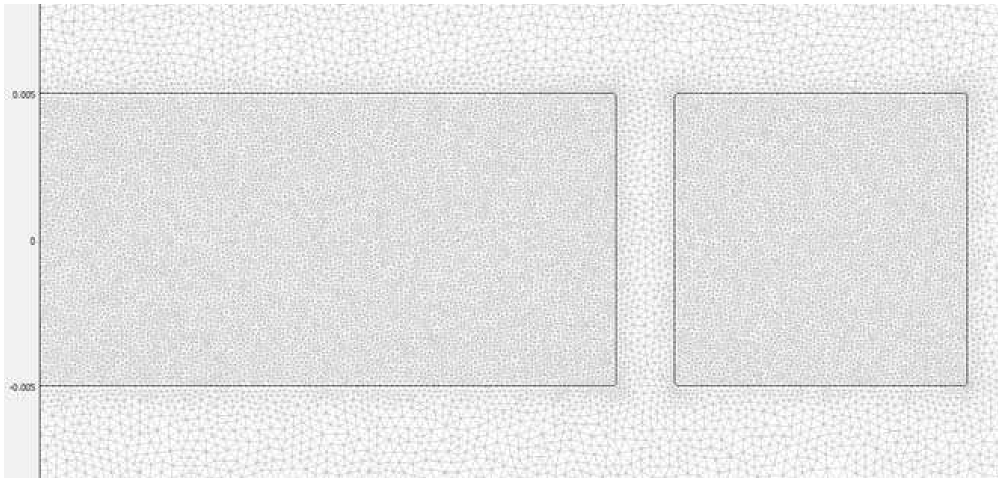


Figure 2: Finite element method mesh.

The simulation efforts are done using Comsol software (axi-symmetrical model). The simulation parameters considered for the modeling are the initial currents density in the coil ($\mathrm{A \cdot m^{-2}}$), the heating time (s), and the frequency (kHz). The material properties behavior versus temperature is pondered in this study since New Finite Element Analysis (FEA) software can combine the required physical phenomena into a global model. The temperature dependence of the electrical conductivity, the relative magnetic permeability, the specific heat, the thermal conductivity and the coefficient of thermal expansion of the material are considered in this study. These physical properties are only known at thermodynamic equilibrium and transient phenomena taking place during fast heating or cooling are not taken into consideration. The selected reference data are used for thermodynamic equilibrium, which cannot exist in induction heating due to a high heating rate [7]. A disc specimen (ϕ 60 mm of diameter and 10 mm of thickness) made of 4340 low-alloy steel has been used for the simulation. A square section coil has been used. The material is considered as homogeneous and isotropic. The FE analysis takes into account the following boundary limits: (1) the ambient temperature has been set to 293°K, (2) the main components are surrounded by a local dielectric environment that is magnetically isolated along with vacuum permittivity and permeability. The convection is supposed equivalent to conduction in the air at interface due to the very short time of treatment. The loss of heat by radiation is neglected because of a very short time of heating. An initial current density (J_0) is applied at the inductor and it is varied to assess the temperature profile. The preliminary mesh used for the FEA simulation is illustrated in Figure 2. The mesh is dense inside the part and the coil because of the high induced currents and temperature gradients between the surface and the core. A convergence study was conducted by refining step by step the mesh size in function of the temperature and current density in the part and in the coil to converge toward to the optimized mesh model. The final mesh consists of 405304 elements and the optimal model needs to 1014969 degrees of freedom to be resolved.

3. SIMULATION RESULTS

During this study, the frequency is adjusted at 200 kHz and the heating time is fixed at 0.50 second. This time is generally very short since the heat treatment must be superficial and the part core should not be affected by the final temperature distribution. Moreover, the recent developments in terms of power electronics allow having very powerful induction machine that can transform part within 0.1 second. The initial current density (J_0) in the coil is varied within a range permitting to have a desired surface temperature and consequently the appropriate hardness profile. The coil thickness is varied three times from 8 mm to 12 mm and the (J_0) is adjusted to obtain 1000°C Kat the surface. This temperature is considered as the reference that permits to compare the three cases. Figures 3, 4 and 5 show the distributions of the total current density and temperature after a heating time of 0.5 s in the HF case (200 kHz). In each case, the (J_0) in the coil are adjusted at $46 \times 10^9\,\mathrm{A \cdot m^{-2}}$, $41.45 \times 10^9\,\mathrm{A \cdot m^{-2}}$ and $38.7 \times 10^9\,\mathrm{A \cdot m^{-2}}$, respectively. When the coil thickness is the same than the part (10 mm), the induced currents are concentrated at thin layer around the edges of both the inductor and the part. Since the temperatures are higher at the part edges than at the middle plane, the regions close the edges will be first transformed in hard martensite.

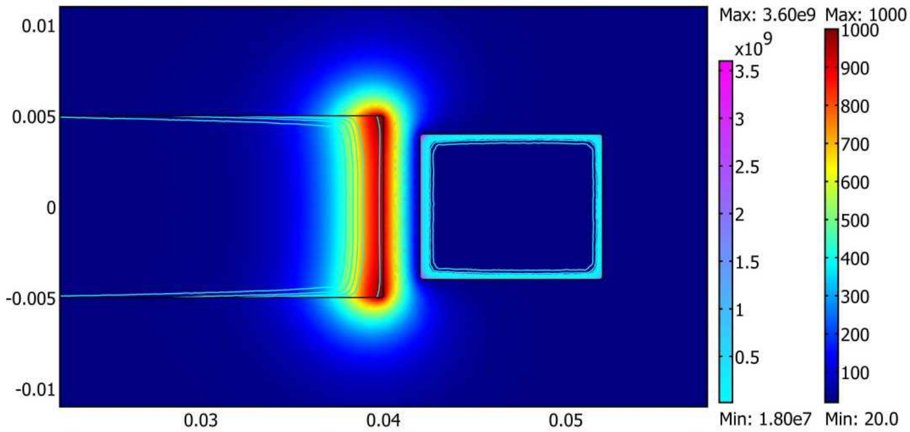


Figure 3: Distributions of induced currents ($A \cdot m^{-2}$) and temperature (°C) — coil thickness of 8 mm.

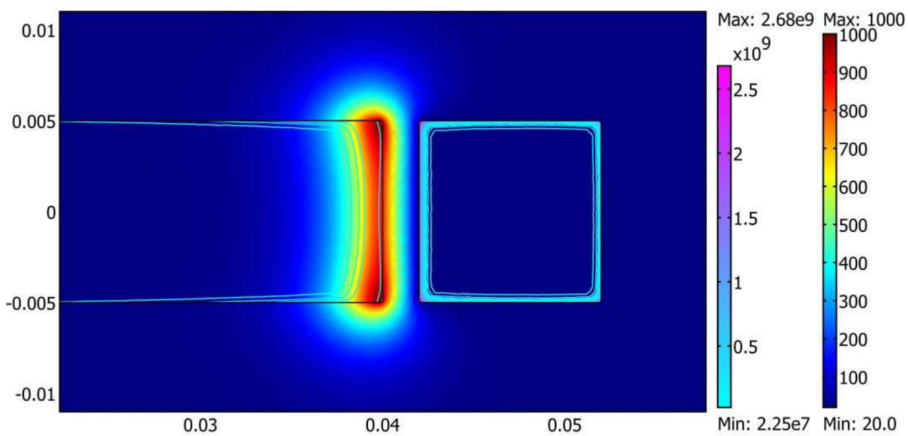


Figure 4: Distributions of induced currents ($A \cdot m^{-2}$) and temperature (°C) — coil thickness of 10 mm.

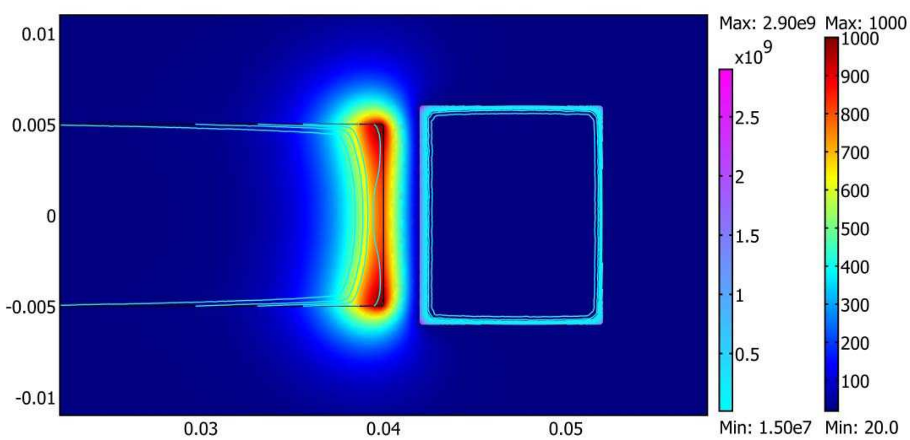


Figure 5: Distributions of induced currents ($A \cdot m^{-2}$) and temperature (°C) — coil thickness of 12 mm.

Increasing the thickness of the inductor grows up the edge effect. However, when the thickness decreases, the temperature profile becomes more uniform. In this case, the entire thickness of the part is transformed into hard martensite. These results show that the uniformity of the temperature distributions is a direct consequence of both the induced currents and frequency [8, 9].

The simulation results demonstrate that the edge effect is very sensitive to the coil thickness.

Figure 6 presents the difference in temperature (ΔT) between the edges (T_E) and the middle plan (T_M) of the part during the heating process. The ΔT becomes two times greater when the thickness changes from 8 mm to 10 mm at the end of heating process (at 0.5 s). The difference is very small at heating start and becomes more important with time. The results demonstrate also that ΔT is very small when the thickness coil is 8 mm at the heating start and doesn't exceed 75°C at the heating end. However, when the coil thickness becomes the same than the part, ΔT increases greatly with the heating time and reaches 150°C; while this difference is about 200°C when the coil thickness is 12 mm. It is also interesting to note that the temperature evolution is non linear in function with the heating time since the material properties have a major influence on the electromagnetic and thermal behavior.

The hardness profile can be obtained from the austenized region through the assumption stipulating that all region heated above 850°C will transformed to hard martensite upon cooling. Figure 7 presents the hardness profile is the three cases cited previously. The simulation results reveal that the hardness profile is quasi-uniform when the coil thickness is 8 mm. The increasing of coil thickness by only 2 mm changes completely the profile hardness that becomes deeper at the edges comparing with the middle plane. Finally, when the thickness is 12 mm, only a small region around the edges is hardened.

To validate the developed numerical models, it is necessary to perform experimental tests at HF with modulating the machine power. Figure 8 presents the validation tests results for two cases: low and medium machine power. It can be seen that a good concordance exist between the simulation and the experiment results. While a small power amount is applied, only the edges are transformed to hard martensite. This case corresponds to the simulation results presented in

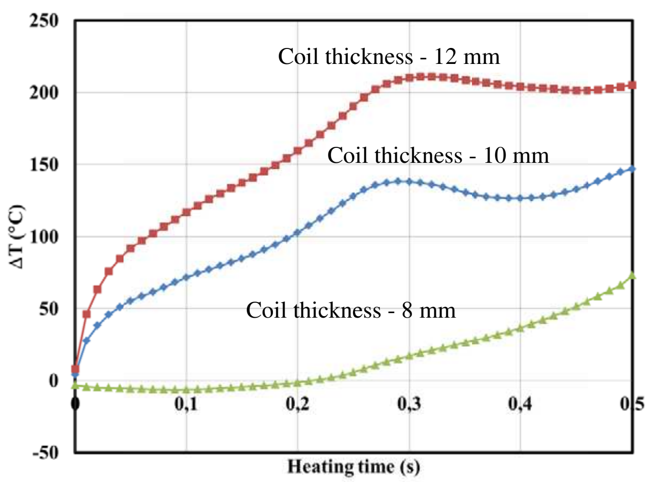


Figure 6: ΔT versus coil thickness.

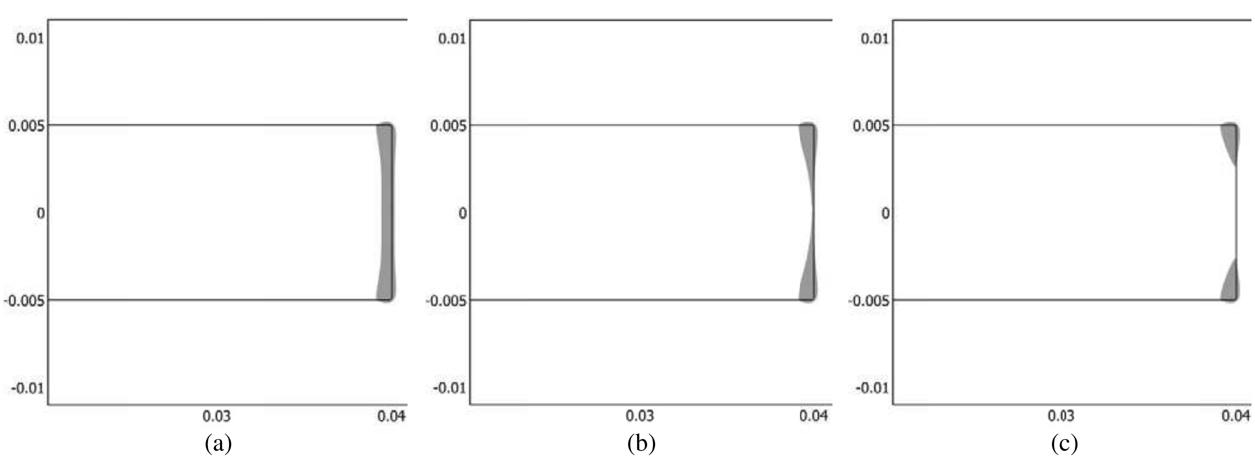


Figure 7: Hardness profiles obtained by simulation. (a) 8 mm, (b) 10 mm, and (c) 12 mm.

(a) (b)

Figure 8: Hardness profile obtained by experiments with 10 mm thickness coil, (a) low machine power and (b) medium machine power.

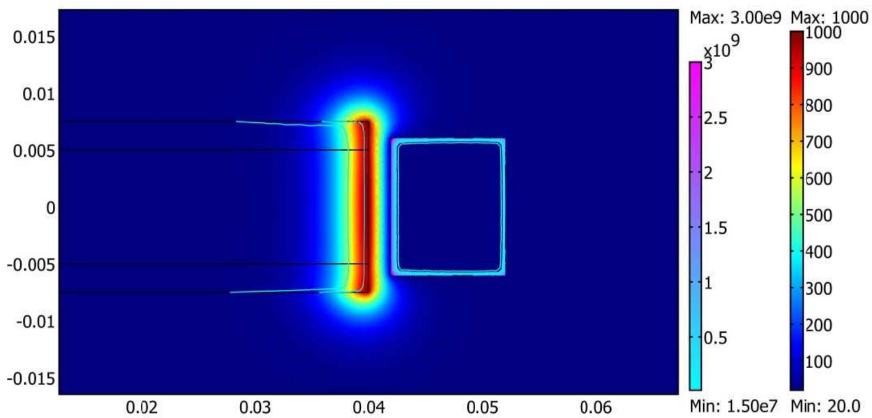


Figure 9: Distributions of induced currents ($A \cdot m^{-2}$) and temperature (°C).

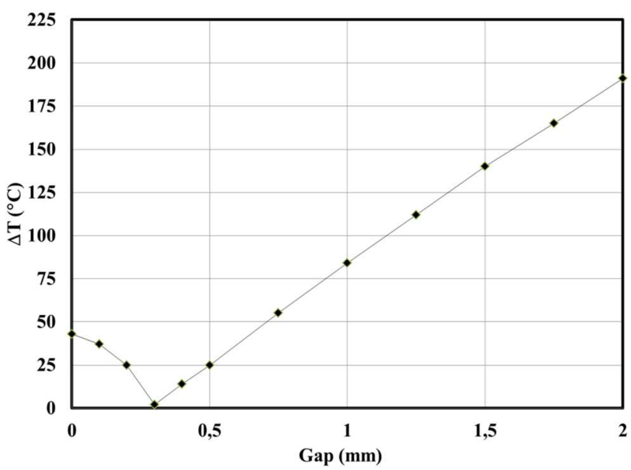


Figure 10: ΔT versus gap between part and discs.

Figure 7(b) (coil thickness of 8 mm) and obtained with a small initial current value. At medium machine power, the middle plane region starts to transform and that correspond to medium of J_0. One can conclude that the hardness profile remains still deeper at the edges than the middle plane.

4. REDUCTION OF EDGE EFFECT

The original solution to reduce the edge effect consists to add two thin discs (washers) on each edge of the part These two discs are acting as flux concentrators. Figure 9 presents the simulation results when the coil thickness is 12 mm and the discs thickness is 5 mm. The two washers are in permanent contact with the heated part (there is no gap between the part and washers). The obtained results are very interesting and show at first sight that is possible to eliminate completely the edge effect.

The gap between the part and the discs used as flux concentrators is varied from 0 to 2 mm and the simulation is performed to analyze its effect on the hardness profile. Figure 10 presents ΔT versus the gap between the part and washers. It is important to note that the difference in temperature is minimal (4°C) when the gap is 0.3 mm. This value represents the optimal gap permitting to have a uniform hardness profile. Moreover, the complete contact between the part and the discs could disturb the heating process since the residual stresses distribution at the edges due to the permanent contact.

5. CONCLUSIONS

The paper has presented an original and comprehensive approach permitting to reduce the edge effect on the hardness profile using FE simulation. First, an axis-symmetric model has been developed using Comsol and used to study of the electromagnetic and thermal effects. Second, the sensitivity of hardness profile to the coil thickness has been analyzed by a comparison of three cases representing a thin coil, a medium thickness coil and thick coil. The simulation model has been also validated by a comparative analysis with experiments results. Finally, an optimization study has been performed to eliminate the edge effect using additional discs (flux concentrators) and to find the optimal gap between the part and discs that permits to obtain a perfect and uniform hardness profile. The obtained results show that the simulation represents a powerful tool to better understand how the current are distributed in the part and how they are affected by the material properties behavior. They are useful also to understand the effect of the coil thickness and how to optimize the edge effect. Consequently, they help certainly induction-heating engineers to choose the good parameters for the induction machine and achieve the desired hardness profile. The developed method could be used to optimize the surface treatment of spur and helical gears.

REFERENCES

1. Haimbaugh, R. E., "Practical induction heat treating," *ASM International*, Materials Park, OH, 2001.
2. Zinn, S., "Elements of induction heating: design, control, and applications," *ASM International*, Metals Park, OH, 1988.
3. Rudnev, V., D. Loveless, R. Cook, and M. Black, *Handbook of Induction Heating*, Marcell Dekker Inc., New York, 2003.
4. Bark, N., P. Bocher, J. Brousseau, M. Galopin, and S. Sundararajan,"Modeling and sensitivity study of the induction hardening process," *Advanced Materials Research*, Vol. 15–17, 525–530, 2007.
5. Favennec, Y., V. Labbe, and F. Bay, "Induction heating processes optimization a general optimal control approach," *Journal of Computational Physics*, Vol. 187, No. 1, 68–94, 2003.
6. Yuan, J., J. Kang, Y. Rong, and R. D. Sisson, Jr., "FEM modeling of induction hardening process in steel," Worcester Polytechnic Institute, MA, 2003.
7. U.S. Defense Departement, *Metallic Materials and Elements for Aerospace Vehicle Structures*, MIL-HDBK-5H, Military Handbook, 1998.
8. Barka, N., P. Bocher, J. Brousseau, M. Galopin, and S. Sundararajan, "Sensitivity study of induction hardening machine parameters," *3rd International Symposium on Aerospace Materials and Manufacturing Processes*, 781–790, Montreal, Canada, 2006.
9. Bark, N., A. J. Khelalfa, P. Bocher, and J. Brousseau, "Analysis of the effect of frequency using numerical modeling of induction hardening process," *1st International Conference on Sustainable Manufacturing*, Montreal, Canada, October 2007.

Optimization of Hardness Profile of Bearing Seating Heated by Induction Process Using Axisymmetric Simulation

N. Barka, A. Chebak, and J. Brousseau
Département de mathématiques, d'informatique et de génie
Université du Québec à Rimouski, Rimouski, Québec, G5L 3A1, Canada

Abstract— The mechanical components are always expected to be hardened in order to ameliorate their behavior in service. Various surface heat treatment processes are commonly used, such as thermo-chemical processes (carbonizing, nitruding) and induction heating process. In particular, the induction heating is very promising process since it exhibits several industrial advantages as heating is done into superficial layer and during a short time not exceeding 1 second. In addition, the process can be easily implemented in automated manufacturing cells and it makes use of any greenhouse gas emissions. However, the process is not completely under control since the electromagnetic effects influence greatly the final temperature profile in the part. Consequently, it is difficult to produce uniform hardness profiles which are design requirements. This paper presents an analysis of temperature and hardness profiles of a bearing shoulder made of 4340 steel and heated by induction process. The work is essentially done with Comsol software and implies the coupling between Maxwell equations and heat transfer equations. The obtained results are very beneficial and permit to better control the hardness profile by proposing a method which consists to introduce some useful modifications of the part and the coil. In this method, two flux concentrators are added to the part and the coil geometry is optimized in order to have a uniform hardness profile. In summary, the study is carried out in few steps. First, an axis-symmetrical model representing the bearing shoulder is built using the adequate formulation and taking into account the material properties. This model allows to determinate the temperature and induced currents density in the part in function of the machine parameters. Second, the effect of added washers on the temperature profile is studied. Finally, the optimized solution leading to a uniform case depth is presented and discussed.

1. INTRODUCTION

The bearings are always used to allow constrained relative motion between shaft and housing and then this technique give good friction efforts resistance. During their operations, these components exhibit axial or radial cyclic loading depending of the excitations nature. A hard surface with compressive residual stresses could improve wear resistance and enhance contact fatigue life. Generally, a shoulder is machined with specific tolerance to mount bearing adequately on the rotating shaft [1, 2].

Induction heating process is usually used to treat bearing surface using the induced currents. The surface hardening is very beneficial since it gives a combination of hard surface layer and toughness in the part core. Even if it is a very simple geometry, the hardness profile obtained experimentally is not uniform and it is difficult to have a sufficient case depth in the inner corner of the shoulder (Fig. 1). The experimental efforts done on the induction machine using medium (MF) or high frequency (HF) are demonstrated that HF power is very suitable to perform the final hardness profile than MF power [3, 4]. This practical method could greatly help mechanical designer to optimize the heating process and enhance the hardness quality of produced part. However, some modifications on the induction system (part, coil) must be done in order to have a uniform hardness profile. The literature has demonstrated the complete absence of method conducting to optimize the hardness profile [5, 6].

The present paper presents an analysis of the induction heating results of a bearing shoulder made from 4340 steel and proposes a new method that allows having a uniform hardness profile. The work is essentially done with Comsol software and implies the coupling between Maxwell equations and heat transfer equations. The obtained results show that the hardness profile can be controlled by making some modifications of the part and the coil. In fact, flux concentrators are added to the part and the coil geometry is optimized in order to have a uniform hardness profile. This study is carried out in few steps. First, an axis-symmetrical model representing the bearing shoulder was built using the adequate formulation and taking into account the material properties. This model allows to determinate the temperature and induced currents density in the part in function of the machine parameters. Second, the effect of added flux concentrators on the temperature profile is studied. Finally, the optimized solution lead to a uniform case depth is presented and discussed.

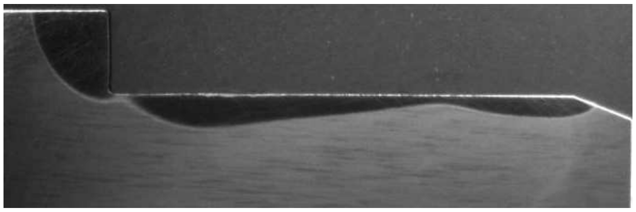


Figure 1: Hardness profile of bearing seating obtained by experiments [ETS development tests].

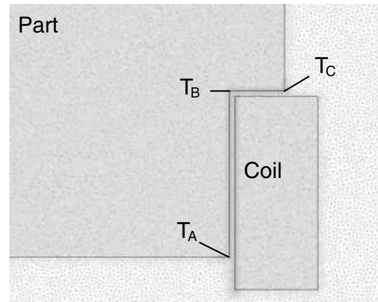


Figure 2: Finite element method mesh.

2. FINITE ELEMENT METHOD FORMULATION

The simulation efforts are done using Comsol software (axis-symmetrical model). This simulation is based on the coupling between the transient FE analysis of the electromagnetic field and the thermal problem. The simulation parameters considered for the modeling are the initial currents density in the coil ($A \cdot m^{-2}$), the heating time (s), and the frequency (kHz). The material properties behavior versus temperature is pondered in this study since New Finite Element Analysis (FEA) software can combine the required physical phenomena into a global model. The temperature dependence of the electrical conductivity, the relative magnetic permeability, the specific heat, the thermal conductivity and the coefficient of thermal expansion of the material are considered in this study. These physical properties are only known at thermodynamic equilibrium and transient phenomena taking place during fast heating or cooling are not taken into consideration. The selected reference data are for thermodynamic equilibrium, which cannot exist in induction heating due to a high heating rate [7]. A shoulder specimen ($\phi 40$ mm minor diameter, $\phi 50$ mm major diameter and 8 mm thickness) made of 4340 low-alloy steel has been simulated. A rectangular section inductor with a section of 100 mm^2 has been used as illustrated in Fig. 2. The material is regarded as homogeneous and isotropic.

The FEM analysis takes into account the following boundary limits: (1) the ambient temperature has been set to 293∘K, (2) the main components are surrounded by a local dielectric environment that is magnetically isolated along with vacuum permittivity and permeability. The convection is supposed equivalent to conduction in the air at interface due to the very short time of treatment. The loss of heat by radiation is neglected because of a very short time of heating. An initial current density (J_0) is applied at the inductor and it is varied to assess the temperature profile. The preliminary mesh used for the FEA simulation is illustrated in Fig. 2. The temperatures T_A, T_B and T_C are measured in three locations as indicated in this figure. The mesh is dense inside the part and the coil because of the high induced currents and temperature gradients between the surface and the core. A convergence study was conducted by refining step by step the mesh size in function of the temperature and current density in the part and in the coil and that conducts toward an optimized mesh model.

3. SIMULATION RESULTS

Only high frequency is considered in this study (200 kHz). The heating time is fixed at 0.50 s. The heating time is very small since the heat treatment must be superficial and the part core should not be affected by the final temperature distribution. Moreover, the new induction machine is very powerful and can transform part within 0.1 sec. The J_0 in the inductor is adjusted to $2.15 \times 10^{10}\,A \cdot m^{-2}$ in order to have a maximal temperature of about 1200°C at end of heating and that permits to have a case depth of about 1 mm. Then, the simulated case depth is deducted at the edges and at the middle plan of the part for comparison and extract the effect of J_0 and time. Fig. 3 shows the distributions of the total current density and temperature at the heating end. At first sight, the induced currents are more concentrated at the edges than the corner. Consequently, the temperatures are higher at the part edges and the transformation will be in these zones. The results show that the corner is not transformed as can be compared to experimental tests (Fig. 1). The temperature distributions are a direct consequence of the induced currents and the frequency.

The simulation efforts reveal that the corner of the shoulder doesn't reach the austenitization

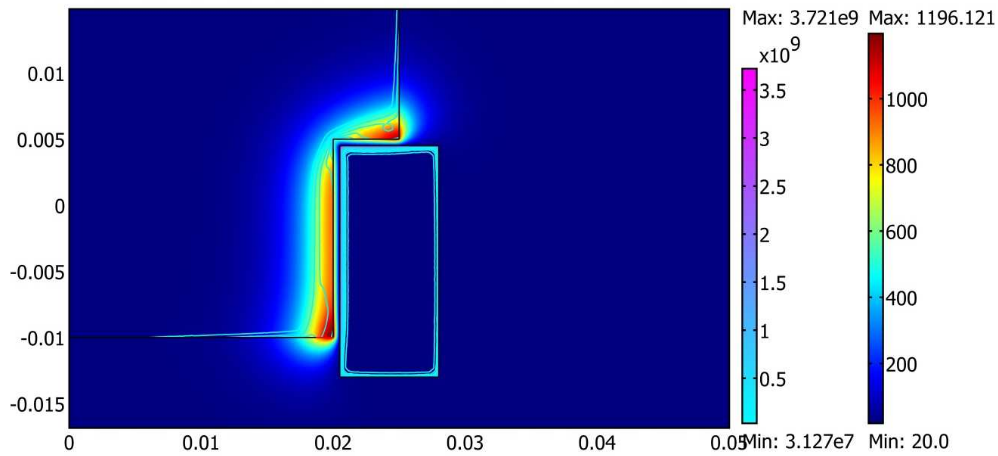


Figure 3: Distributions of induced currents ($A \cdot m^{-2}$) and temperature (°C).

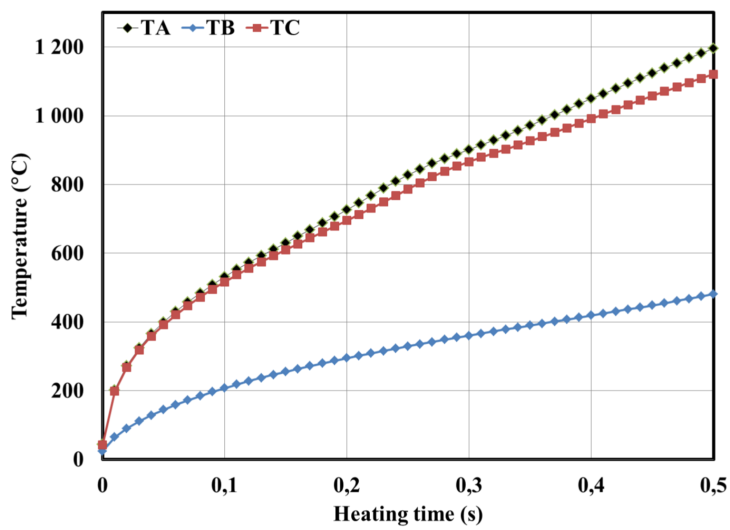


Figure 4: Temperatures T_A, T_B and T_C versus heating time.

temperature (Ac_3) (Figs. 3 and 4). Consequently, the corner is not transformed to martensite because it is very difficult to concentrate induced currents in this region. It is then important to find a practical method to enhance the hardness profile without changing the experimental setup. These results can be compared to hardness profile obtained by experiments showed in Fig. 1.

4. OPTIMIZITATION STUDY

In order to concentrate more induced currents in the corner region, it is interesting to add two flux concentrators made of 4340 steel with 2.5 mm of thickness and placed nearby the part and the coil (Fig. 5). The discs are mounted by keeping the same gap between the part and the coil (0.5 mm). The obtained results demonstrate that the hardness profile is quasi-uniform. However, the corner zone is hardly transformed into martensite and the flux concentrator's solution appears to be interesting but not sufficient (Figs. 5 and 6). The temperatures T_A, T_B and T_C versus the heating time demonstrate that the edge regions are transformed to martensite while the corner region does not reach the transformation temperature. However, the temperatures T_A and T_B have less difference than the initial configuration (without flux concentrators).

The final solution that allows having a uniform hardness profile consist not only to add the two flux concentrator but also to modify the coil geometry as shown in Fig. 7. The coil geometry can be optimized by cutting the surface close to the part and ensure to keep a particular form close to the part corner. The results show that the induced currents are concentrated at the entire exposed surface of the bearing seating. The temperature profile is uniform and the hardness profile

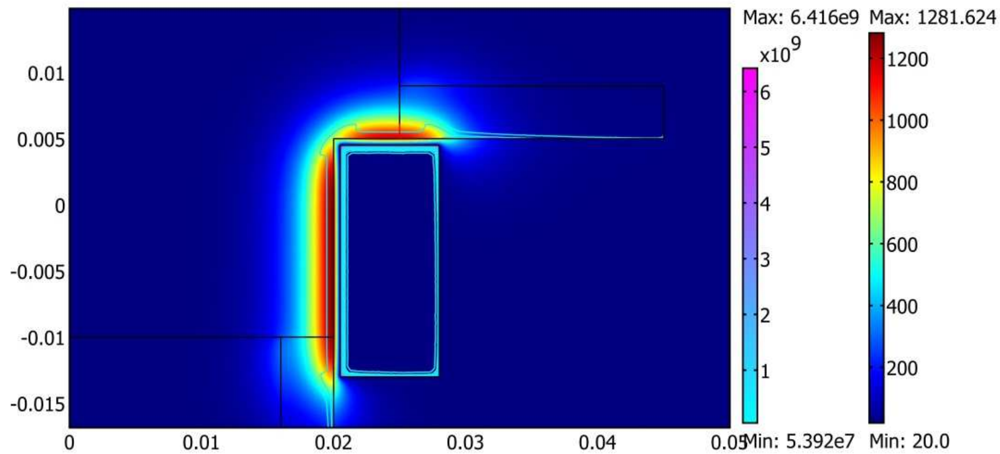


Figure 5: Distributions of induced currents ($A \cdot m^{-2}$) and temperature (°C) (for a part with flux concentrators).

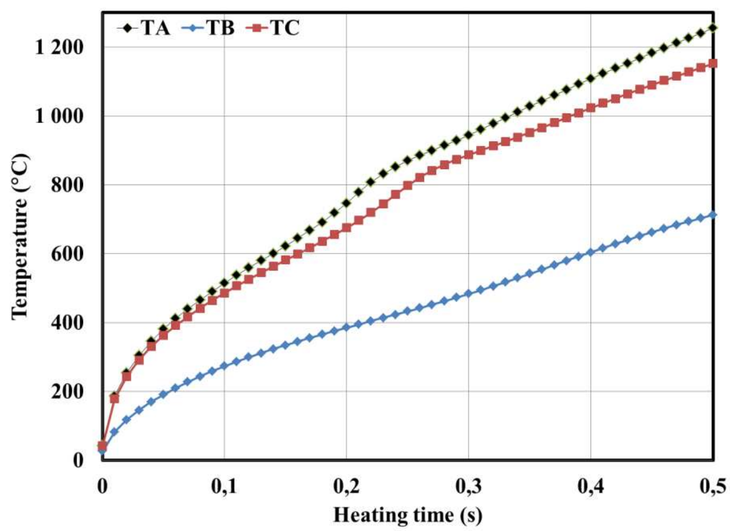


Figure 6: Temperatures T_A, T_B and T_C versus heating time (for a part with flux concentrators).

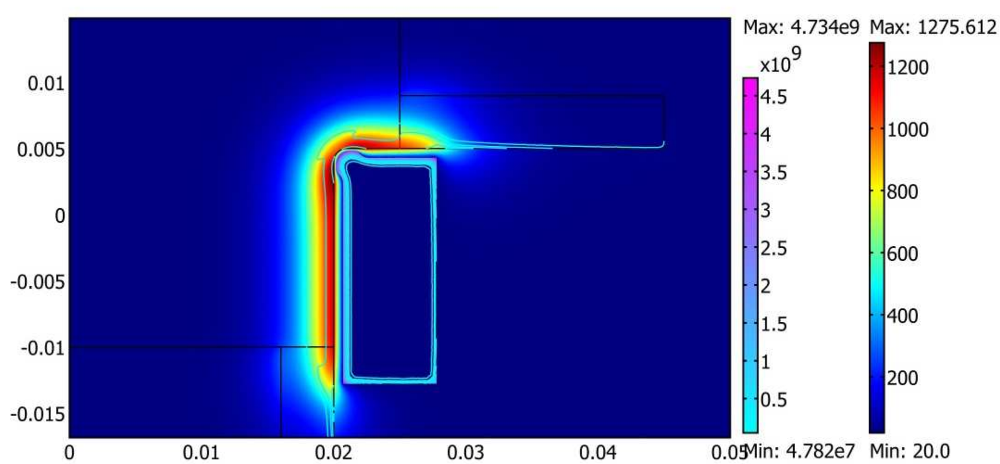


Figure 7: Distributions of induced currents ($A \cdot m^{-2}$) and temperature (°C) (for a part with flux concentrators and modified coil geometry).

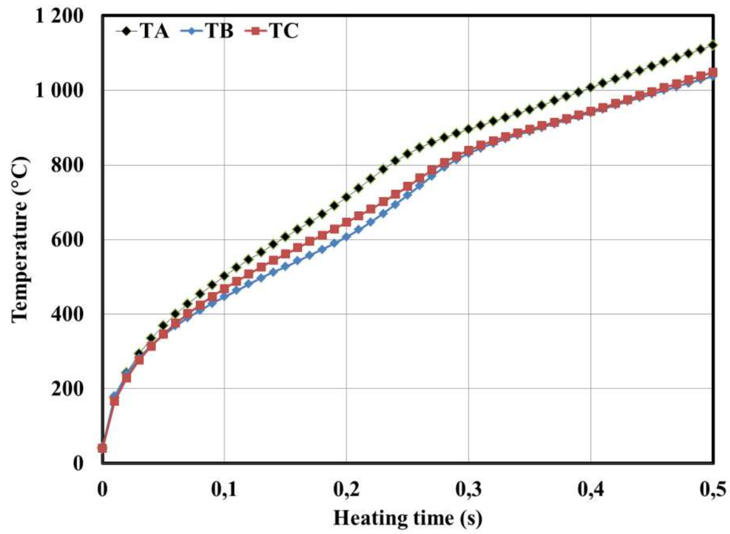


Figure 8: Temperatures T_A, T_B and T_C versus heating time (for a part with flux concentrators and modified coil geometry).

becomes then uniform even in the corner region. These results are confirmed by the temperatures curves presented in Fig. 8. The three temperatures are very close and they confirm that the flux concentrators use and the coil modification represent a good solution that ameliorate greatly the hardness profile.

5. CONCLUSIONS

The paper has presented an original approach permitting to enhance the hardness profile of bearing seating using the electromagnetic field and thermal simulation. The axis-symmetric model developed has taken into account the the machine parameters and various configurations of coil and part. The obtained results have been analyzed and that conducted to emphasize the electromagnetic and thermal effects. The distributions of induced current and temperature are exposed and compared in order to converge step by step to optimal solution using the flux concentrator and the modified coil. The simulation represents a powerful tool to better understand how the current are distributed in the part and how they affected the heat energy generated during the heating process. This way could permit to save development time and reduce wastage due the trial and error method.

REFERENCES

1. Haimbaugh, R. E., "Practical induction heat treating," *ASM International*, Materials Park, OH, 2001.
2. Zinn, S., "Elements of induction heating: Design, control, and applications," *ASM International*, Metals Park, OH, 1988.
3. Rudnev, V., D. Loveless, R. Cook, and M. Black, *Handbook of Induction Heating*, Marcell Dekker Inc., New York, 2003.
4. Barka, N., P. Bocher, J. Brousseau, M. Galopin, and S. Sundararajan, "Modeling and sensitivity study of the induction hardening process," *Advanced Materials Research*, Vol. 15–1, 525–5307, 2007.
5. Favennec, Y., V. Labbe, and F. Bay, "Induction heating processes optimization a general optimal control approach," *Journal of Computational Physics*, Vol. 187, No. 1, 68–94, 2003.
6. Yuan, J., J. Kang, Y. Rong, and R. D. Sisson, Jr., "FEM modeling of induction hardening process in steel," Worcester Polytechnic Institute, MA, 2003.
7. U.S. Defense Departement, "Metallic materials and elements for aerospace vehicle structures," *Military Handbook*, MIL-HDBK-5H, 1998.

A Time Domain Hybrid Approach to Study Buildings Connected by Cables

N. Muot[1], **E. Bachelier**[2], **X. Ferrieres**[2], **and C. Girard**[1]

[1]AxesSim, Strasbourg, France
[2]ONERA, Toulouse, France

Abstract— In term of application for large and complex systems, many hybridizations and multidomain approaches have been studied in different ways and have shown their capabilities to deal with electromagnetic problems. However, as far as we are interested in applications like lightning electromagnetic induced effects on buildings, interconnected by powerlines or communication networks, the usual multidomain methods needs to be improved. For those problems, the challenge is to reduce the computational problem size (for the area) between buildings. The problem is to take into account the interactions between soil, buildings and cables, without a 3D meshing of the global scene. The paper presents a multidomain strategy based on a hybrid approach in the time domain: by coupling 3D methods (FDTD, FVTD ...) with a transmision line method, first step of the general problem.

The idea is to define a 3D problem around each building and a 1D problem along the transmission line. Then, the hybridization strategy consists in making exchanges of some quantities at each interface between the two problems. In our way of coupling, the considered interfaces are reduced to a dipole on which we propose to exchange data of both domain by the introduction of a local Thevenin equivalent circuit. Through the impedance, the local condition of continuity at the virtual interface of the line is ensured. On the other hand, the generator send the signal relative to the other computational domain.

Considering the applied problem we are interested in, the next step to be investigated is the effect of the soil conductivity on the induced effects, between buildings, grounding and networks. The paper will also present the actual fields of investigation for these specific aspects.

1. INTRODUCTION

Except for a few special cases, the general analytic solution of the time dependant Maxwell's equations is unknown. Among the many numerical methods that have been proposed to solve it, the Finite Difference Time Domain method (FDTD) is one of the most commonly used. Indeed, it has been demonstrated (showed) to be powerful, versatile and has the benefits of being used by the electromagnetism community since a long time now. However, some class of problems are still difficult to solve, and in particular the computation of electromagnetic induced effects of lightning on large and complex systems like buildings connected by cables.

The complexity of this problem is due to the addition of several specific points: the modeling of the lighting strike itself, the computation of the currents guided along overhead and buried cables, and the effect of the soil. This paper is focused on buildings interconnected structure, and in particular on the induced effects between buildings. It presents a numerical process to build a multi-method strategies that could be performed for this kind of electromagnetic problem.

By the fact that lightning problem is a Low Frequency (LF) problem (high frequency $f_h \approx 10\,\mathrm{kHz}$), it induces specific issues when numerical methods are used. Indeed, this particularity has an influence on the stepsize used for the spatial discretization, with regards to the wavelength. Usually, the spatial discretization is given by the wave length ($\lambda_p = C_0/f_h \approx 30\,\mathrm{km}$) but in the lightning context, the geometrical criterion (order of magnitude of object is meter) imposes that the spatial step (and consequently the time step) is quite lower. Added to the fact that the lightning signals are long in time, the use of methods like FDTD to solve lightning problems needs a huge computational time and memory.

It is possible to reduce very large problem by using a hybrid approach. This kind of approaches can use for example the Huygens-Fresnel principle [1], the Gaussian quadrature or reciprocity theorems. A hybrid approach is very interested in the way one can split the whole problem in several simpler problems. Each sub-problem is then solved by the most adapted numerical method. However, hybridization by Huygens-Fresnel approach or the ones listed above are well defined for unconnected problems, but limited to those applications. The hybridization in a context of interconnected objects is still a difficult point and is an actual topic of interest. As an example we can notice their application in finite differences method (MR-FDTD) [2], or the use of MR with hybridization in a multiresolution [3].

For all these reasons, we have chosen to build a convenient hybrid approach to solve lightning electromagnetic induced effects on buildings interconnected by powerlines or communication networks. For those problems, the challenge is to reduce the computational problem size for the area between the buildings and use the most adapted numerical method to solve sub-parts. The problem is to take into account the interactions between soil, buildings and cables, without a 3D meshing of the global scene. This paper presents a multidomain strategy based on a hybrid approach in time domain, by coupling 3D methods (FDTD, FVTD ...) with a transmision line method, first step of the strategy for solving the general and complex problem.

A brief overview of interfaces used to take into account the interactions between cables and 3D mesh is presented in the first section. Then, in Section 2, the numerical discretization for some methods are proposed. The last section presents the results obtained by the hybrid approach and a comparison with global numerical results (using one main method).

2. ANALYSIS

In this section we present the way to reduce the initial problem, that is to solve two separate parts of a single transmission line by using two different numerical methods. Thus, to ensure the computation of a global solutions, we need to define an interface on which will be applied specific boundary conditions. For this purpose, we will focus first on the possible strategies for hybridizing two time domain methods for solving a transmission line problem. We will see here the strategie to do that.

Some simple approach have already proposed [4] to hybrid transmission line with FDTD method on Holland wire [5, 6] using as an antenna. This approach which consists in sharing the unknowns at the interfaces. This approach is attractive but is not very accurate in the transmission line case.

The power waves formalism is helpful to describe the energy exchanges between two transmission lines. This formalism, introduced by Kurokawa in [7], defines two waves: the incident wave and the reflected wave. In our case, the incident wave is the one transmitted to the other part of the line, the reflected wave for other way.

Usually those waves are defined as follows, where Z is the wave impedance:

$$W^i = \frac{V + ZI}{2\sqrt{\Re[Z]}} \tag{1}$$

$$W^s = \frac{V - Z^* I}{2\sqrt{\Re[Z]}} \tag{2}$$

It is not necessarily easy to properly define these waves in numerical methods and it is neither quite simple to implement. This is why we have tried to improve this approach, by implementing this formalism in the form of a Thevenin equivalent model. We can therefore simply hybridize a code if it can add discrete components on a wire as a voltage generator and a resistor. The following figure illustrates how to use equivalent Thevenin models to make hybridization. Note this hybridation is made in one way only. We will justify this point in hypothesis depict in next step.

The computation of the elements of the model is made by the denormalized power wave equation, where we impose the wave impedance Z is the characteristic impedance of the transmission line, define as following.

$$Z_c = \sqrt{\frac{R + jL\omega}{G + jC\omega}} \tag{3}$$

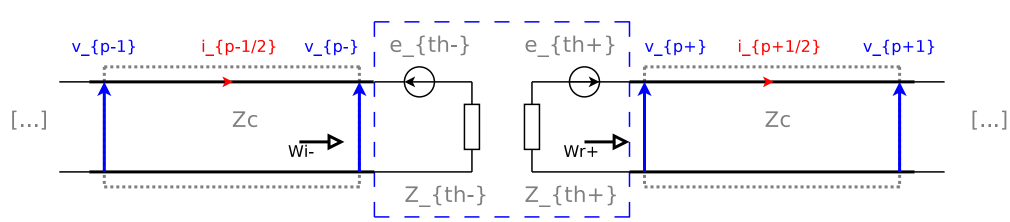


Figure 1: Equivalent models developed by hybridization of Thevenin equivalent model.

Thus, imposes constraints on the impedances $Z = Z_c$ we limit the study to the case without reflected wave. We have several use cases, either the transmission line is adapted, or reflected wave has not come back yet (clear time).

Impose $Z = Z_c$ restrict denormalized power wave equations are:

$$w^i = v + Z_c i \tag{4}$$

$$w^r = v - Z_c i \tag{5}$$

Finally, if v is voltage across the generator and i is the current flowing through it, then, on the side of the Thevenin generator, the equation is:

$$v_p = e_{th} - z_{th} i_p \tag{6}$$

Applying those equations to point $p-$ we get the incident and reflected waves:

$$w_p^i = e_{th} + (Z_c - Z_{th}) \cdot i_p = \frac{Z_{th} - Z_c}{Z_{th}} v_p + \frac{Z_c}{Z_{th}} e_{th} \tag{7}$$

$$w_p^r = e_{th} - (Z_0 + Z_{th}) \cdot i_p = \frac{Z_{th} + Z_c}{Z_{th}} v_p - \frac{Z_c}{Z_{th}} e_{th} \tag{8}$$

We obtain, equations hybridization provide by scattering parameters of a matched networks. Those equations are equivalent to impose the continuity of current and voltage at the point of hybridization, moreover we have no reflected wave in the right part of the transmission line ($w_{p+}^i = 0$).

$$w_{p+}^r = w_{p-}^i \tag{9}$$

$$w_{p-}^r = w_{p+}^i = 0 \tag{10}$$

So we have all the information needed to carry out the hybridization. The Equation (10) ($w_{p-}^r = 0$) is a matched boundary that is equivalent to $e_{th-} = 0$ and $z_{th-} = Z_c$ into Thevenin equivalent model. The Equation (9) in turn, is reduced to $e_{th+} = 2v_{p-}$ for $z_{th+} = Z_c$. The equation system of our hybridization is:

$$e_{th+} = 2v_{p+} = -2i_{p+}/z_{th+} \tag{11}$$

$$e_{th-} = 0 \tag{12}$$

$$Z_{th-} = Z_{th+} = Z_c \tag{13}$$

3. NUMERICAL APPLICATION

In this section we will focus particularly on the methods that can be used for the discretization of the hybridization equations in different situations. We will see in particular the FDTD method in 1D (line code) and in 3D.

3.1. Numerical Application to FDTD 1D

FDTD code for a transmission line of the dipole is done directly by a boundary condition of Thevenin type. However, this model requires knowledge of the source voltage at half time step, property imposed by the FDTD scheme. Unfortunately, the transmission line code gives us the voltage at a full time step. Like it is shown below, we must do a half discretization, to evaluate this voltage.

$$-\partial il = Gv + C\partial vt \tag{14}$$

$$\Rightarrow \quad V_{p-}^{n+1/2}\left(\frac{G\Delta t + 2C}{\Delta t} + \frac{2}{Z_{th}\Delta l}\right) = V_{p-}^{n}\frac{2C}{\Delta t} + \left(I_{p-1/2}^{n+1/2} + \frac{E_{th-}^{n+1/2}}{Z_{th}}\right)\frac{2}{\Delta l}$$

3.2. Numerical Application to FDTD 3D

In contrast, for 3D FDTD codes, we have 'just' to mesh the geometry of the transmission line. The dipole is modeled by a wire-like Holland on which we will dispose lumped elements (generator voltage, resistance). One side of this dipole is connected to the conductor and the other side is connected to the reference of the transmission line.

To compute the Thevenin equivalent voltage at the extremity we just have to measure the current on the wire modeling hybridization dipole (Equation (11)).

A similar approach can be used in any other numerical method like FVTD.

4. NUMERICAL VALIDATION

In this section, we present the results of a numerical simulation on a problem studied in order to validate the Thevenin Equivalent Formalism for hybridization between FDTD and a transmission line code by comparisons with simulation using a single FDTD code on the whole domain.

4.1. Description of the Case Studied, Derived from Lightning Problems

We describe here the problem we are interested in for the test case. Some elements of the modeling still need to be improved to lead to a realistic problem, but we can find in the chosen test case the same main difficulties as in a real lightning problem on connected buildings. The interest here is also to limit the complexity of the results interpretation, and to concentrate the analysis of the results on the validation of the Thevenin equivalent formalism. The following figure depict the geometry of the scene. We find in this case, two metal boxes, connected at ground by a resistance equal to the impedance of the transmission line. These two boxes are connected together by a wire of radius 0.01m. The illumination is performed by a current generator like a lightning channel.

For this case the generator is a current generator of normalized Gaussian derivative type. The magnitude is provide by the equation $i(t) = -\alpha\sqrt{2}(t-\tau)e^{1/2-\alpha^2(t-\tau)^2}$ where the parameters are choosen to get maximum magnitude at the characteristic frequency ($f_0 = c_0/\lambda$) and insure zero start value.

The following Fig. 3(a) shows the current magnitude on link point between right box and transmission line. The red one is provided when we solve the problem come from the full FDTD method. The blue one come from the hybrid method.

We can do two remarks. First, we have a short time compression between the two magnitudes. This discrepancy is probably due to the difference in dispersion between the two digital patterns. Second, the current on hybrid approach taller that for the single FDTD approach. This point is

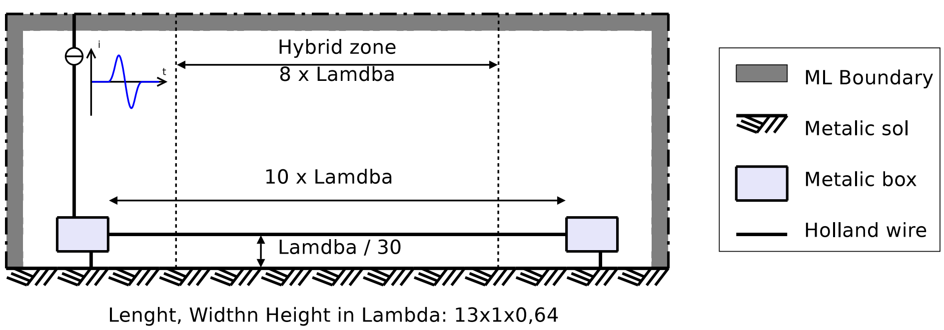


Figure 2: Geometry of the case study.

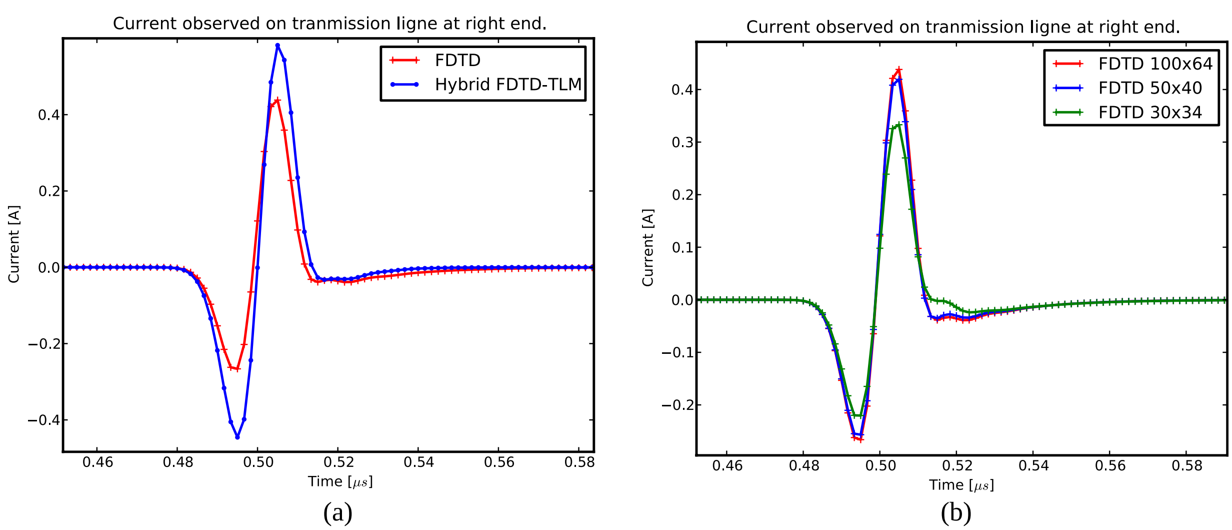


Figure 3: Current observed on transmission line at right end. (a) Comparison between the FDTD and Hybrid approach. (b) Bring out boundary effect on line loss for tree cross section.

	run time [min]	memory [Mo]
FDTD	≈ 51	≈ 387
Hybrid	≈ 12	≈ 125

fallout of absorption along the transmission line of the transverse wave into ML boundary. To ensure this point if we reduce cross section of the computational domain and we have more and more of line loss (Fig. 3(b)).

The last result is the run time and memory consumed during the computation. The following table gives this information.

So we gain in using the hybrid approach. This gain stems directly from the fact that we use a more compact for model a large part of the problem. Moreover, the time performance is increased by implied parallelization between the different methods.

5. CONCLUSION

In this paper, we have applied a Hybried approch to resolve problem where a large part of the problem can be reduced into a transmission line. The Hybridization are provided by an equivalent Thevenin model is easy to depict on wire by adding a voltage generator and a impedance derived by hybridyzation equations.

Finally, this paper shows that this approach is really useful for reducing the complexity of a computation of the electromagnetic induced effects of lightnings on connected buildings. However, this approach is penalized by the fact that it is unidirectional. Thus, the work must be made in this way.

REFERENCES

1. Huygens, C., *Trait de la lumire*, Gauthier-Villars et Cie Editeurs, Paris, 1920.
2. Johnson, J. M. and Y. Rahmat-Samii, "Multiple region FDTD (MR/FDTD) and its application to microwave analysis and modeling," *IEEE MTT-S International Microwave Symposium Digest*, Vol. 3, 1475–1478, 1996.
3. Pascaud, R., R. Gillard, R. Loison, J. Wiart, and M. F. Wong, "A new multiresolution FDTD approach based on the hybridization of MR-FDTD and DG-FDTD methods," *IEEE/MTT-S International Microwave Symposium*, 883–886, June 3–8, 2007.
4. Maloney, J. G., K. L. Shlager, and G. S. Smith, "A simple FDTD model for transient excitation of antennas by transmission lines," *IEEE Transactions on Antennas and Propagation*, Vol. 42, No. 2, 289–292, Feb. 1994.
5. Holland, R. and L. Simpson, "Finite-difference analysis of emp coupling to thin struts and wires," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 23, No. 2, 88–97, May 1981.
6. Collino, F. and F. Millot, "Fils et méthodes déléments finis pour les équations de Maxwell," Le modèle de Holland revisité, Research Report RR-3472, INRIA, 1998.
7. Kurokawa,K., "Power waves and the scattering matrix," *IEEE Microwave Theory and Techniques*, Vol. 13, No. 2, 194–202, Mar. 1965.

The Effect of Metamaterial Patterning to Improve the Septum GTEM Chamber Performance

H. X. Araújo and L. C. Kretly
University of Campinas, Brazil

Abstract— In this work, a metamaterial cell patterning is applied to the *septum* of a GTEM chamber in order to obtain a better coupling between the EM field and the DUT. The basic cells consist of a ring resonator — RR (Ring Resonator) or CLL (Capacitively Loaded Loop) in SRR (Split Ring resonator) or CSRR (Complementary Split Ring Resonator) configuration topologies. They are oriented depending on the wave direction from the device to the measuring port or from the exciter side — APEX to the device under test. Actually, resonant type metamaterial transmission lines are very similar to CLL lines. However, the first one exhibit a transmission zero at a finite frequency (to the left of the left handed band) due to the coupling between the host line and the loading resonators (magnetic coupling in SRR-loaded lines and electric coupling in CSRR-loaded lines). The simulated results show the direction property of the patterned *septum* in comparison with the useful one. This new and promising technique appears to improve the GTEM chamber performance.

1. INTRODUCTION

In many cases, integrated circuits and electronic boards are the mainly causes of interference in others electronics equipments [1]. The EMC — Electromagnetic Compatibility of an electrical device is its capability to operate safely in an electromagnetic environment without interfering and be not susceptible by other interference sources [2]. In this context, the EMC work is on increase demand.

Different approaches support the pre-compliance tests (EMC/EMI/EMS) setups which have consolidated standards and regulations. These setups include TEM/GTEM cells, Magnetic Loop, Magnetic probe, Workbench Faraday Cage, OATS — Open Area Test Site and others. Based on the frequency range, from 500 MHz–18 GHz (this work), the GTEM — Gigahertz Transverse Electromagnetic chamber has advantages over the others test setups.

A GTEM chamber, used for testing EMI and EMS of electronic devices and systems has a thin metallic baffle, called *septum* to conduct the radiation over the device under test. Similar to a rectangular coaxial transmission line with outer conductors closed and joined together, the TEM cell has tapered ends acting as transition to adapt to the standard 50 Ω coaxial connectors. The central conductor of the cell provides the electromagnetic propagation in TEM mode. Therefore electromagnetic field is uniform only in a certain region of the chamber normally in the middle, place where the EUT — equipment under test, should be located. Figure 1 shows a GTEM cell.

In this work, a metamaterial patterning is applied to the *septum* (internal conductor) of a GTEM chamber in order to obtain a better tuning in a desire frequency. The basic cells consist of a ring resonator — RR (Ring Resonator) or CLL (Capacitively Loaded Loop) in SRR (Split Ring resonator) or CSRR (Complementary Split Ring Resonator) configuration topologies. They are oriented depending on the wave direction from the device to the measuring port or from the exciter side — APEX to the device under test. Actually, resonant type metamaterial transmission lines are very similar to CLL lines.

2. GTEM CHAMBER

The origin of the Gigahertz Transverse Electromagnetic chamber (GTEM) was based on the transverse electromagnetic chamber (TEM) which is basically a planar expanded transmission line operating in the TEM mode to simulate a free space planar wave. This expanded wave guide with a flat and wide center conductor and coated by electromagnetic waves absorbers, is commonly used in analyses of products that are physically small and compacts, especially electronic components.

At low frequencies only the TEM mode propagates on the chamber. However, with the increase of the operational frequency, TE and TM modes can be excited inside the chamber. The maximum frequency is calculated from the first lower resonant of the higher modes, which depends on the size and shape of the chamber. The main advantage of the TEM cell is the small size and low cost,

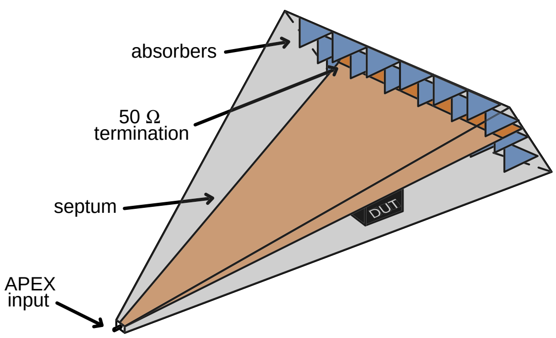


Figure 1: GTEM chamber — internal arrangement.

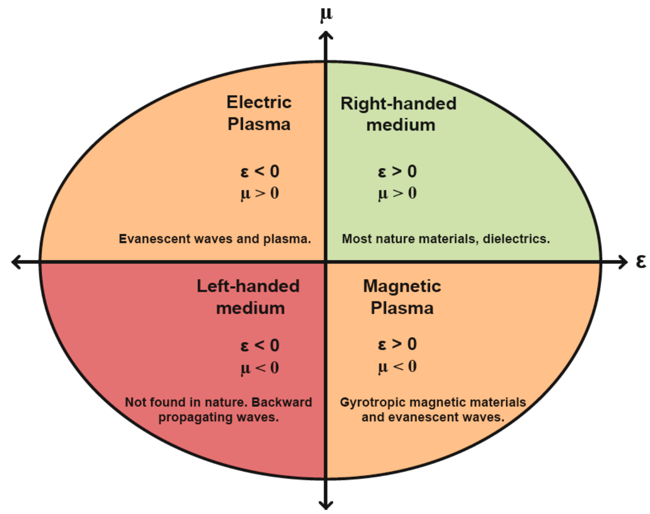


Figure 2: Material classification based on the macroscopic parameters — adapted from [8].

and further it is not necessary any additional external shielding. However, the TEM chamber is restricted in frequency, according to IEC 1000-4-3 [3], the upper frequency limit is below 1 GHz.

The need for EMC measurements on upper frequencies, Hansen et al. [4] proposed a new concept of TEM cells. The idea behind the new chamber was in avoiding the tapered corners which are the main reason for the TEM cell frequency limitations. So, it becomes a new chamber called GTEM (Gigahertz Transverse Electromagnetic) in which is possible to do measurements from some Hertz through 18 GHz.

The restriction on upper frequency limit is eliminated by tapering the transmission line continuously outward from the feed point to a termination system. The tapered point and the absorbers at the end of the chamber allow the chamber to operate in high frequency.

This cell consist basically in a tapered section of a rectangular transmission line, RF absorbers at the end of the cell, and a tapered *Septum* which is a metallic plate constructed by brass, cooper or aluminum. The larger extremity ends with a 50 Ω resistive load, and has RF absorbers, while the narrow extremity has a 50 Ω coaxial adapter.

The RF absorbers are basically polyurethane foams with low density in pyramidal shape, filled by a high losses material. The RF power is absorbed in a range starting from a frequency that depends mainly on the length of the pyramids. So, the lowest frequency attenuated by the absorbers, is defined by the following expression:

$$f_{\min} = \frac{c}{2h} \tag{1}$$

where c is the speed of light and h is the height of the absorbers.

A crucial challenge on the design of the GTEM chamber is to establish the correct size of the Apex, which takes about 10% of the overall length of the cell, and works in the transition from the 50 Ω coaxial cable (input power) to the body of the chamber through the septum. To improve the performance of the septum and the same time assure the APEX efficiency, the metamaterial technology, working as artificial magnetic and electric conductor is employed on the GTEM structure. In the next section, is shown the basic metamaterial concepts.

3. METAMATERIAL APPLICATION

The concept of artificial material, which defines the metamaterial technology, was synthesized by Rodger M. Walser, in 1999 on his work about macroscopic composites with synthetic and periodic cellular architecture. However, the first attempt to explore the concept of "artificial" materials was in 1898 with Jagadis Chunder Bose by his experiment about twisted structures [5]. Later, in 1914, the work about artificial chiral media was done by Lindman [6]. After these, in the past 20 years, the interest on metamaterial technology had strong increased, with researches on superlens and telecommunication environment, including transmission lines and antennas applications [7].

Actually, metamaterial is a macroscopic composite of periodic or non-periodic structure, whose function is due to both the cellular architecture and the chemical composition [8]. Therefore, the behavior of a material, in the presence of an electric field, is determined by the macroscopic parameters, permittivity ε and permeability μ. In Figure 2 is shown the material classification based on its macroscopic parameters.

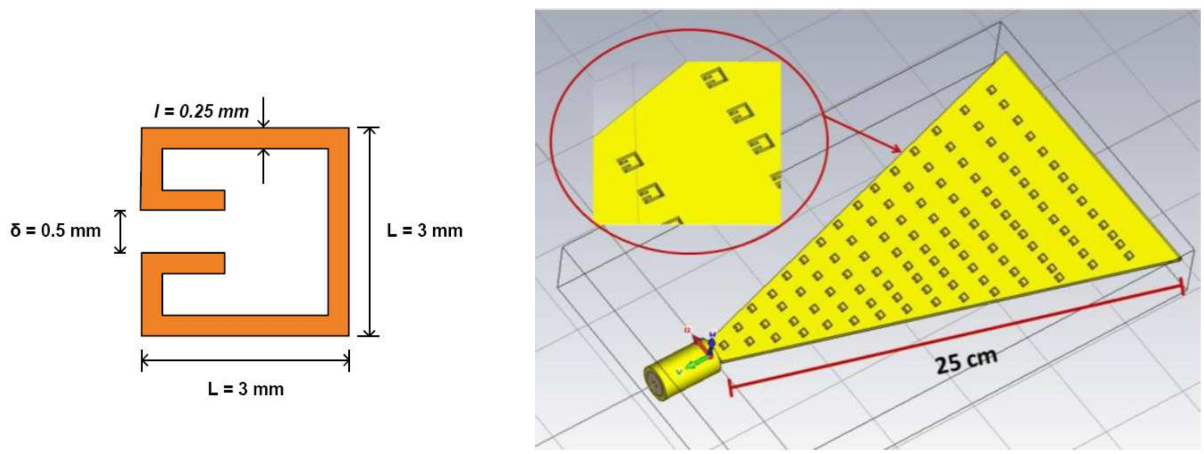
In this work, the authors used metamaterial cells, acting as an artificial magnetic conductor (AMC). This strategy improves even the electric or the magnetic field. The artificial magnetic conductor is obtained when the plane wave is incident upon the capacitive gaps, and act as an artificial electric conductor (AEC) when the plane wave is incident from the opposite direction. This structure was proposed by Erentok et al. in [9].

The main idea is to use the metamaterial patterning to improve the septum performance in a specific frequency range. Basically, the metamaterial patterning was printed on the septum in order to act as artificial magnetic conductor. The designed structure is shown in Figure 3.

The behavior of the metamaterial pattern stamped on the septum has been investigated. From the dimensions of the metamaterial cells ($3\,\text{mm} \times 3\,\text{mm}$), it was expected a best response in 10 GHz. As a rule of thumb, the cells are arranged in such way that the period (distance between cells) is around three cells dimensions. The dimensions of the metamaterial cells are given by:

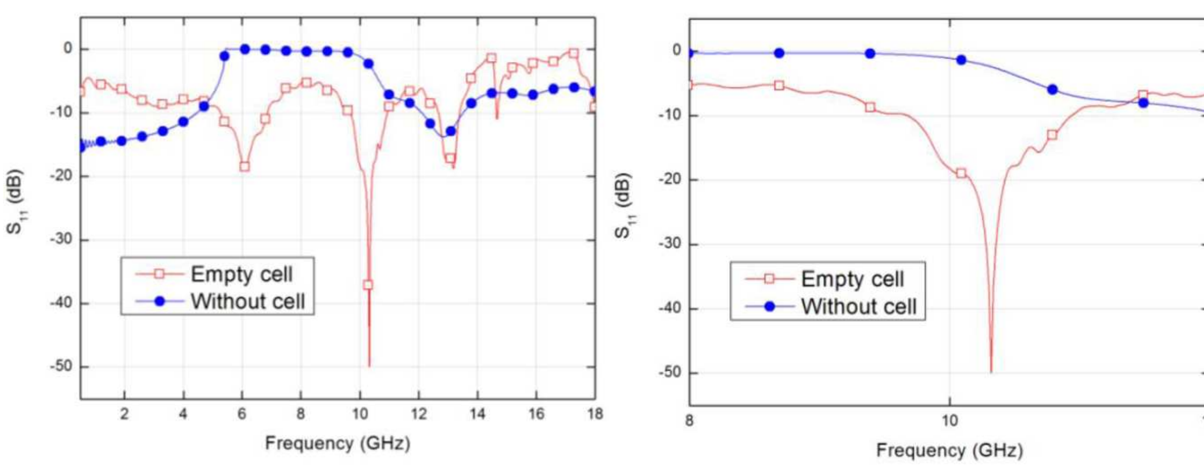
$$L = \frac{\lambda_0}{10}, \quad \delta = \frac{L}{6}, \quad l = \frac{\delta}{2} \tag{2}$$

A more realistic relationship between the frequency and the dimensions of the cell can be obtained by numerical simulations. Based on the boundary conditions, Equation (2) is satisfactory.



(a) (b)

Figure 3: (a) RR or CLL metamaterial cell; (b) Metamaterial cell stamped on the *septum*.



(a) (b)

Figure 4: S_{11}-parameter comparison between the proposed technique and the useful structure. (a) Full range; (b) Limited range.

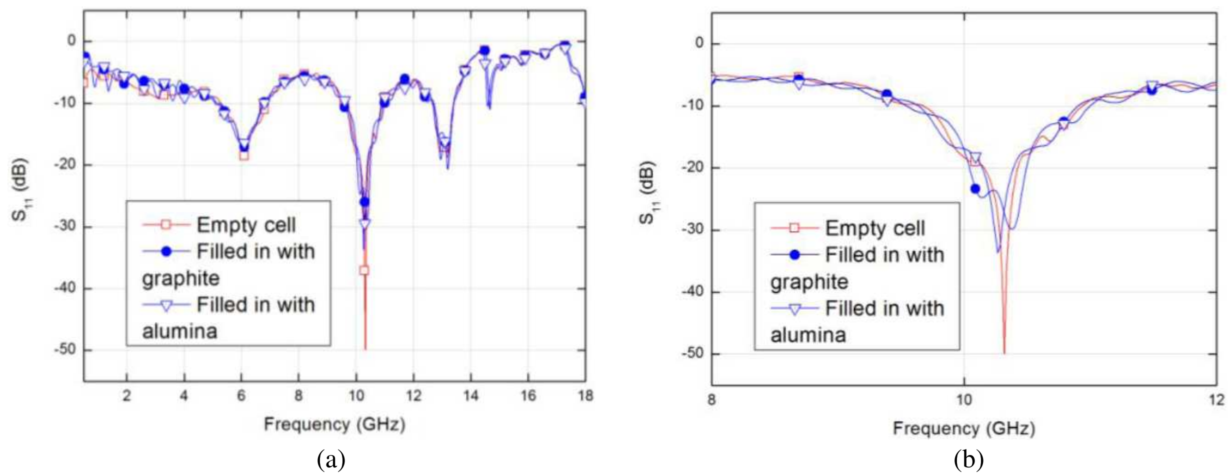


Figure 5: S_{11}-parameter for the filled in and empty cells. (a) Full range; (b) Limited range.

Figure 4 shows the comparison between the proposed structure and the useful one.

As expected, from Figure 4 is easy to observe the effect of the cells on the *septum* in 10 GHz. Other scenarios were designed in other to obtain a better performance. To investigate and proving the influence of the cells on the RF and microwave propagation, additional simulation were done to check out the real effect of the cells filled in with different materials of various resistivity. Basically three options to fill the metamaterial cells were analyzed — filled in with graphite, filled in with alumina and empty cells.

As can be shown in Figure 5 as the cell is filled in with different materials the behavior of S_{11} parameter is modified but the solution follows the same tuning format near the designed frequency and converges to the empty (vacuum) solution. This demonstrates unambiguously the real influence of the cell on the signal propagation through the modified *septum*.

4. CONCLUSIONS

In this work guide lines for the design of periodic structures on the septum of GTEM were proposed.

Several approximations were done based on electromagnetic metamaterial cells and some rule of thumb design equations were proposed. The use of this methodology showed a good correlation between the approximately solution and the simulation results.

This new and promising technique can be taken into account to strengthen the electric field in a given frequency improving the GTEM chamber performance. In this work, the designed cells allowed a compatible response in 10 GHz. An investigation of the best material to fill in the cell was also done. Three types of cells were analyzed and based on the results, the empty configuration proved to be the most appropriate.

REFERENCES

1. Dhia, S. B., M. Ramdani, and E. Sicard, "Techniques for low emission and susceptibility," *Electromagnetic Compatibility of Integrated Circuits*, Springer Science, 2006.
2. Montrose, M. I. and E. M. Nakauchi, *Testing for EMC Compliance*, Wiley Interscience, New York, 2004.
3. Electromagnetic Compatibility (EMC) — Part 4: Testing and Measurement Techniques — Section 3: Radiated, Radio-frequency, Electromagnetic Field Immunity Test, International Standard CEI/IEC 1000-4-3, Geneva, Feb. 1995.
4. Konigstein, D. and D. Hansen, "A new family of TEM-cells with enlarged bandwidth and optimized working volume," *Proceedings of the 7th International Zurich Symposium on Electromagnetic Compatibility*, 127–130, Zurich, Mar. 1987.
5. Bose, J. C., "On the rotation of plane of polarization of electric waves by a twisted structure," *Proc. Roy. Soc.*, Vol. 63, 146–152, 1898.
6. Lindell, I. V., A. H. Sihvola, and J. Kurkijarvi, "Karl F. Lindman: The last Hertzian, and a Harbinger of electromagnetic chirality," *IEEE Antennas Propag. Mag.*, Vol. 34, No. 3, 24–30, 1992.

7. Cui, T. J., "Electromagnetic metamaterials: Recent advances on the theory, experiments, and applications," *IEEE Antennas Propag. Mag.*, 2008.
8. Cui, T. J., D. V. Smith, and R. Liu, *Metamaterials: Theory, Design and Applications*, Springer, New York, 2010.
9. Erentok, A., P. Luljak, and R. W. Ziolkowski (editors), "Antenna performance near a volumetric metamaterial realization of an artificial magnetic conductor," *IEEE Trans. Antennas Propagat.*, Vol. 53, 160–172, 2005.

The Integration of the Multihoming Concept in Ad Hoc MANET Mobile Networks

A. Messai, S. Sadouni, S. Aris, H. Mokhtari, and M. Benslama
Electromagnetism and Telecommunication Laboratory, University of Constantine, Algeria

Abstract— In the field of telecommunications, mobility represents new practical ways of communication that are known to be vaguely flexible in terms of use; for users expect to enjoy constant and good quality service. We will make an attempt; all along this file; to introduce the concept of Multihoming to mobile networks in order to increase their strength and resistance toward the effects of mobility, along with the optimization of network resources. For this purpose, we opted for the BGP protocol, the one which takes into charge the introduction of the Multihoming concept to Ad Hoc MANET networks; all this is achieved by its cohabitation with other Ad Hoc routing protocols.

At first, we will proceed by AODV for the reactive protocol category, afterwards we will move to the OLSR for the proactive ones. This; in turn, will allow us to determine the routing protocol category that fits the most the Multihoming integration as a new concept; in addition to the crucial mechanisms that ought to be taken into consideration.

1. INTRODUCTION

THE most prominent and remarkably embarrassing mobile network disadvantages are known to be the following: the constant instability (or change) of its topology, along with the varying characteristics and performances of the link relating two nodes. Consequently, mobile networks fall into two main categories:

Networks provided with fixed preexisting infrastructure, Networks without infrastructure

Regarding the first category, the centralized infrastructure takes in charge the control and the organization of all the communication processes that take place between various elements of the network. It is also important for the achievement of some extremely crucial tasks such as: security, and the control of intensity, and even the intercellular automatic transfer by having recourse to the help of *Handover* [1] mechanism. Practically speaking, for such an architecture the radio resources are extremely scarce; to optimize them; a mobile station would get solely connected at a one given moment in time and throughout only one interface. This leads us to move to the second category of Ad Hoc MANET networks. Regarding the latter, all the units change their position freely and with no centralized administration making it necessary to automatically organize. Every node may contribute in the routing process, and in order to again transmit the packets from one node whose destination is out of reach.

Ad Hoc MANET protocols use varying routing techniques in order to correctly transport the packets. However, no limitation has been imposed in terms of the number of the links relating two nodes. As a result, it will be possible to integrate the Multihoming concept without having to face any contradiction with the IEEE 802.11 standard norms [2].

The main concern of this article is to reveal the extent to which the integration of Multihoming with a new topology strengthens the network as it was previously intended, all this is to overcome the effects of mobility. Besides, it is to be recognized that mechanisms play a major part in improving the functioning that results from the constraints related to this type of networks: such as the optimization of the charge load and the network resources. In order to do so, the second section of this paper will be mainly devoted to tackle the concept of Multihoming, and its integration to mobile networks; focusing all along this part on the functioning of the BGP protocol. Regarding the third section, it will be devoted to both: the introduction of AD Hoc MANET networks of the IEEE 802.11 norm, and to the information routing. However, we will be particularly interested in studying the reactive protocols introduced by AODV, and the proactive ones illustrated by OLSR.

To display the statistics revealing the advantages expected from the integration of the Multihoming to mobile networks, we should simulate an Ad Hoc network on which we will ought to achieve multiple scenarios: some with the with the reactive protocols, and others to the proactive ones all included in the fourth section. Finally, the fifth section will allow us to make a conclusion, and to drag the attention to the perspectives and expectations of this modest research.

2. MULTIHOMING AND BGP ROUTING PROTOCOLS

The concept of Multihoming was previously used among the wired networks technology. It gives a node the opportunity to simultaneously use more than one network interface; moving this way from one interface to another without the interruption of the current data transmission [3]. When it comes to internet networks, such configuration allows the user to have access to a permanent and simultaneous Internet connection with various ISPs "Internet Service Provider" in order to improve its QoS, assure the availability of a support network, and to share the degree of load that provides in turn Higher-in-speed internet connection.

Finally, as another advantage we may list the achievement of "Smart Routing". This may be realized by choosing different BGP routings for prefixed data in relation QoS, suggested by various ISP on the different ways.

2.1. Multihoming with Mobile Networks

A mobile network is said to be multi-homed when it has many Internet fixing points throughout different MRs; or when an MR has several extern interfaces, otherwise, several addresses on its extern interface [4].

The expected perspectives are the same as they are for a fixed network, but mobility is the element which gives more advantages to this configuration. In fact, such configuration allows us to make up for the break-downs, to share the intensity and the load, to add up some preferences, or to simply guarantee a better access to internet connection; along with the use of different technologies [5].

2.2. The BGP Protocol

Border Gateway Protocol is considered as being the core of internet networks. Its main objective is to exchange data on the availability of networks by maintaining an up-to-date list of autonomous "AS" [6] from which certain As or loops that are less interesting might be excluded. Connections between different neighbors communicate with one another via a TCP session, by exchanging dynamically routes announcements [RFC1771].

The BGP protocol is based on *Bellman-ford* algorithm [7], it supports unclassified routing. It is also provided with a path vector protocol [9]. As it is the case for any protocol, BGP maintains the routing tables, transmits the routes updates to their neighbors and bases its routing decisions on a metric system.

The main BGP disadvantage is the fact that it is not provided with a global vision that covers the whole routing topology. Thus, it solely sends the routes announcements to its neighbors. Still, it is not provided with a balancing system that handles the load born by different links which may result in a probable congestion in certain ones. The BGP might again be affected by the rapid and sudden oscillation of the routes, that may cause an over-load; or may disturb the routing stability [8].

3. AD HOC MANET NETWORKS

The concept of Ad Hoc mobile networks has been developed on the basis of extending the notions of mobility to all the components of the environment. It may contain mobile platforms called Nodes that can change their position freely and without constraints. Consequently, it constitutes an autonomous system capable of organizing itself without a previously defined infrastructure; through its nodes; which may directly communicate with one another when they are in reach of a radio. This can be by using a certain number of routing protocols [RFC 2501].

Different from wireless networks, an Ad Hoc MANET mobile network is featured with a set of particular characteristics among which we may list the following:

- **A dynamic topology**
- **Multi-hop**

This will cause a number of constraints: a limited bandwidth, an increased rate of errors; a varying capacity of links, a low-speed Internet connection, a limited duration of use ... etc.

3.1. Routing among Ad Hoc MANET Networks

After having considered all of the previously stated constraints, being actually functional in an Ad Hoc mobile environment requires from the routing protocol to take into account three different stages [10]:

Dissemination of the Routing Data: In order to be vaguely acquainted with the elements constituting the network topology.
Selection of the Route: Among the some obtained in accordance with some criteria.
Routes Maintenance: Take into account the occurring changes of the topology, and up-to-date the routes that have just been cut.

There are three main categories of routing protocols which belong to the Ad Hoc MANET type:

3.2. Proactive Routing Protocols [11]

This category of protocols attempts to maintain the best existing ways to reach all of the possible destinations. The routes are saved even when not used. The most efficient and successful of this category of protocols is OLSR (Optimized Link State Routing): the one which is based on the *State Link* algorithm. It uses a unique form for all its messages.

HELLO as a mechanism detecting the presence of neighbors **TC (Topology Control).** It helps the other nodes to build up their topology and routing tables.

In order to decrease the number of pointless control messages during a flooding (or inundation) in the network, OLSR uses the concept of Multipoint relays MPR; in which every node uses its minimal sub-part **(MPR-SET)** among its symmetric neighbors to one hop [12].

3.3. Reactive Routing Protocols

These ones, shortly after the demand, create and maintain the routes accordingly with the needs. The most widely-known of these protocols is the AODV type (Ad Hoc On Demand Distance Vector Routig). It uses a broadcast mechanism in order to find out the valid routes. It defines two types of operations [12, 13]:
The Routes Discovery: It uses two types of messages:

- RREQ: (Route Request)
- RREP: (Route Reply)

The Routes Maintenance: In order to fix (or repair) a route that is broken, it uses the message:

- RERR Route Error.

When this whole process is taking place, this type of protocol predicts a mechanism to locally repair a route, in order to limit the transmission of the routing messages [14].

4. AN AD HOC MANET NETWORK SIMULATION INTEGRATING THE MULTIHOMING CONCEPT

As a step toward a full consideration and study of the impact and effects of the integration of the Multihoming concept to a mobile network, we have used a software of simulation of OPNET networks. (Optimized Networks Engineering Tools), to follow the functioning of an Ad Hoc MANET network. They are of an average size expanding over a surface of 16 square KM, and constituted of 9 mobile stations of the type ***MANET_station_adv*** with the communication interfaces WLAN 1, 2, 5, 5 11 Mbps. Plus three wireless routers of the type ***Wlan_ethernet_router***, with an IEEE 802.11 interface, configured to support the BGP and constitute an AS.

The BGP has as a function the integration of the Multihoming concept; thus it maintains an up-to-date list of Ass. However, it is not provided with a global vision of the whole topology. Then, AODV and OLSR intervene so as to discover and to predict the changes that may occur in the topology. Our study is supposed to cover the realization of two scenarios for every protocol:

- An Ad Hoc network with simple mobility integrating Multihoming.
- An Ad Hoc network with developed mobility integrating Multihoming.

4.1. Collected Statistics

This involves the collection of global statistics including the total number of network nodes: AODV/ OLSR ROUTING TRAFIC SENT. Regarding the individual statistics: Source node/ destination node: IP TRAFFIC SENT TRAFIC SENT/ IP TRAFFIC RECEIVED. Finally, three mobile router nodes: IP TRAFFIC RECEIVED.

The transmission of the data packets starts up right after the 100th second by the source node, while the simulation lasts 500 seconds. The unit of measurement is the packet per second.

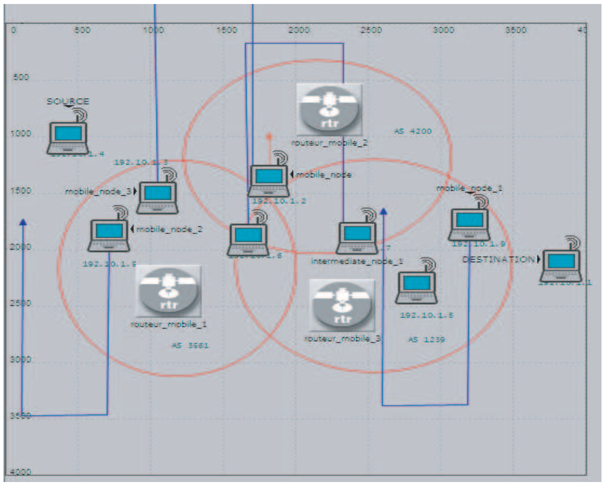


Figure 1: Simulated topology of Ad Hoc network.

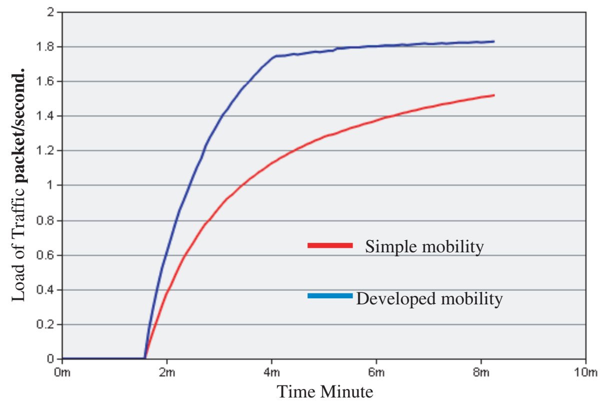


Figure 2: The load traffic received by the 3rd router.

4.2. Discussions of Results

4.2.1. Part One Reactive Routing Protocols AODV

Comparing and contrasting the between the previously collected statistics on both scenarios, we have noticed the absence of the traffic right from the beginning of the simulation and on till the 100th second (protocol on demand). Then, it is to be noticed that the measures reveal that both scenarios are similar for the routing traffic AODV with a slight increase for the developed mobility. The latter might be due to the transmission of the extra messages of the route discovery and maintenance. This, in turn, is caused by the rapid and important changes occurring in the topology.

However, for both scenarios, we have noticed a full similarity for the traffic sent by the source; and that received by the destination. This; in turn; proves that despite the effects of the developed mobility, the traffic sent and received is always the same.

Coping with the changes of topology, and maintaining continuous communications involves the increase of the routing traffic load: this can be noticed in Figure 2. Which introduces a contrast between the traffic load received by the 3rd router (next to the node destination).

In the beginning of the packets transmission, there is a noticeable increase of the traffic received by the Router 3 attaining the value of 1.82 packets/Second. Comparing the scenario that involves a developed mobility with the one involving a simple one; the former exceeds the latter with a value of 1.51 packets/second.

4.2.2. Part Two OLSR Proactive Routing Protocols

This part will be devoted to the study of the impact of integrating Multihoming with the same topology, but this once the OLSR proactive protocol will be used for the data routing.

The first remark to be noted is the automatic and simultaneous transmission of a big number of OLSR routing packets ever since the beginning of the simulation (proactive protocol). However, the comparison between the measures taken from the OLSR routing traffic shows an important difference of the load born among the two scenarios. This difference is due to the changing topology involving the recalculation of the MPR_SET and the routes to all of the destinations. Besides, TC and HELLO messages overload the network as an effect to Multihoming.

As far as the traffic sent by the source is concerned, the measures taken from both scenarios are totally identical. Whereas for the destination, the traffic received in the case of developed mobility is more noticeable than that of the simple one. This is due to the OLSR routing traffic.

In order to define the impact of mobility on such topology, it will be necessary to spot the light of our study on the individual statistics of the router 3 as a sample in the Figure 3.

A brief look at the two curves shows the difference in load. For the simple mobility, the traffic is relatively little; while for the developed one, it is nearly double the traffic of the former. All in all, it is more crucial to note that despite the important load born by the AS of the router 3, the traffic of the received packets among the destination node is not the same for both scenarios.

The Multihoming allows a given node the opportunity of preserving multiple active connections, meanwhile, it permits the automatic switch from one deficient link to another. As for the OLSR

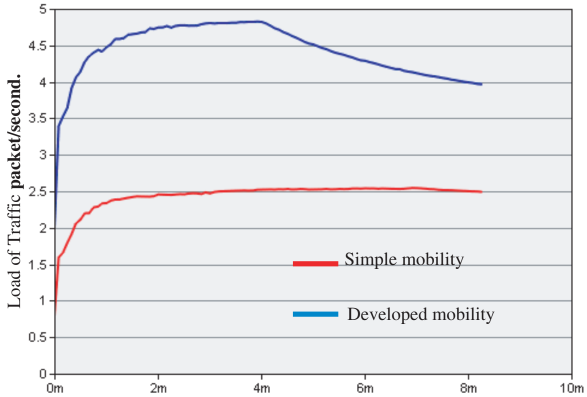


Figure 3: The traffic load received by the router 3.

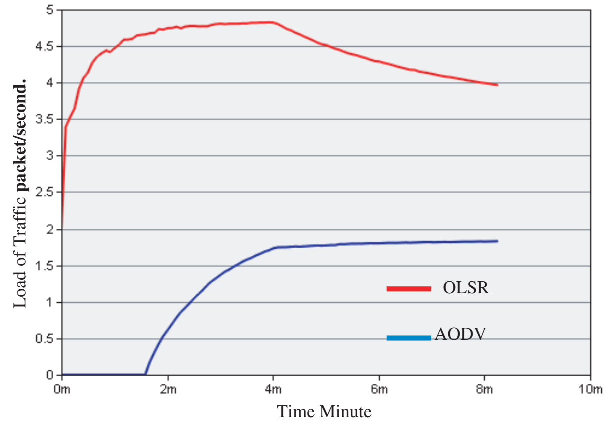


Figure 4: The traffic received by the 3rd router for both protocols.

proactive protocol, the routes are to be saved by a continuous exchange of updating messages of the routes, even when they are not used. On the other hand, AODV uses the local repairing mechanism on the level of the node detecting the breakdown. It also limits the transmission of research and response messages of the routes in order to avoid the problem of congestion.

To define the type of routing protocol with which the integration of Multihoming fits the most, a comparison of the load received by the 3rd router and the two routing protocols was an option to reach this end.

OLSR entices an remarkable increase of the traffic load, still it might be affected by the effects of mobility. Whereas for half the AODV routing traffic the same traffic received by the destination node is preserved independently of the extension of the mobility.

5. CONCLUSION

In this humble research, we have provided evidence proving the fact that the integration of the concept of Multihoming to an Ad Hoc MANET network strengthens it, and gives more resistance to the effects of mobility. However, this can be more useful in terms of optimization of the load and network resources with the type of reactive routing protocols than with the proactive ones' type.

However, in this article, we have been offered the opportunity to unfold some of our humble suggestions that may serve in the optimization of the integration of the Multihoming concept to mobile networks. From the former we may list:

The integration of a local maintenance mechanism of the routes in every node, comparable with that of AODV protocol. Use a balancing load among Ass set with routers of the type: *Inter-Autonomous System.* Decrease the time allocated for the reply to the demands of the route by creating automatic neighboring tables for every AS. Secure the network with the help of an identification procedure on ASs. Still, achieving a fully mobile network requires the absolute deletion of the ASs from our topology, and the attempt to adopt a new approach such as the SCTP routing protocol.

REFERENCES

1. Colin, E. and F. Berna, "Routage et re-routage dans les réseaux mobiles," Universite Montpellier 2, IUP Génie Mathématiques et Informatique, 2001.
2. *IEEE Standard for Information Technology — Telecommunications and Information Exchange between Systems, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications,* 2003.
3. Fabrizi, S. and B. Wertlen, "Roaming in the mobile internet," *Science Direct,* 2008.
4. Caballero, X. F., "Étude d'IPsec projet d'une API_IPsec pour la mobilité et le multihoming," 2008.
5. Ernst, T., "Le support des réseaux mobiles dans IPv6," 2001.
6. Fabrizi, S. and B. Wertlen, "Roaming in the mobile internet," *Science Direct Telecommunications Policy,* 2008.

7. De Launois, C., B. Quoitin, and O. Bonaventure, "Leveraging network performance with IPv6 multihoming and multiple provider-dependent aggregatable prefixes," *Science Direct, Computer Networks*, Vol. 50, 2006.
8. Ioannidis, I., B. Carbunar, and N.-R. Cristina, "High throughput routing in hybrid cellular and Ad-Hoc networks," Purdue University, West Lafayette, 2007.
9. Dhar, S., "MANET: Applications, issues and challenges for the future," *International Journal of Business Data Communications and Networkin*, June 2005.
10. Atmani, M., F. Ameza, and N. Aassam, "Les technologies sans fil: Le routage dans les réseaux Ad Hoc(OLSR et AODV)," Licence Académique en Informatique, 2007.
11. Atir, A., "Etude de L'Attaque du Trou de Ver avec le Routage Proactif dans les Réseaux Ad Hoc," Université de Bejaia, Algérie, 2006.
12. Timóteo, R., A. Adnane, C. Bidan, and L. Mé, "Analyse de la confiance implicite requise dans le routage Ad Hoc OLSR," Campus de Rennes, 2008.
13. Marti, S., T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile Ad Hoc networks," ACM MOBICOM, Boston MA, USA, 2000.
14. Perkins, C., E. Belding-Royer, and S. Das, "Ad Hocon-demand distance vector (AODV) routing," RFC 3561, IETF, July 2003.

Electromagnetic Compatibility of CMOS Circuits along the Lifetime

R. Fernández-García[1], **J. M. Ruiz**[2], **I. Gil**[1], **and M. Morata**[2]

[1]Electronic Engineering, UPC Barcelona Tech. 08222 Colom 1, Terrassa, Spain
[2]Escuela Universitaria Salesiana de Sarriá, Barcelona 08017, Spain

Abstract— The continuous scaling of CMOS circuits has set the MOSFET transistor in the nanoelectronic era. In this context, the functionality and complexity of integrated circuits (ICs) are growing up. However, the operation voltage has been continuously reduced. The higher complexity of ICs has allowed including electronic systems in a lot of safety critical applications (i.e., automotive, aeronautics and/or medical applications). Therefore, the functionality of these electronic equipments must be assured and the risk of electromagnetic interference (EMI) must be reduced during their lifetime. Nowadays, circuits' robustness to electromagnetic interference is checked in a burn-in component, without taking into account the impact of the natural devices' aging. However, shrunk dimensions imply the appearance of several wear out mechanisms, which can limit the functionality of the circuits and modify their electromagnetic performance. Therefore, the time dependence of electromagnetic behaviour, which is known as Electromagnetic Robustness or Electromagnetic Reliability (EMR), should be evaluated. The switching noise is probably one of the main EMC emission problems in CMOS circuits. It is know that wear out mechanism affects the switching behaviour of CMOS circuits. Therefore, some effects on EMC performance of these circuits should be expected. In this work, the switching noise behaviour or CMOS circuits under one of the most important reliability problems are analysed by means of electrical simulation. In order to do that, a characterization of wear out mechanism on single MOSFET is presented and modelled. The results show a reduction on the frequency switching noise emission in circuits subjected to wear out, due to the reduction of the drain current of MOSFET.

1. INTRODUCTION

The continuous scaling of CMOS circuits has set the MOSFET transistor in the nanoelectronic era. In this context, the functionality and complexity of integrated circuits (IC) are growing up. However, the operation voltage has been continuously reduced. The higher complexity of IC has allowed including electronic system in a lot of safety critical applications (i.e., automotive, aeronautics and/or medical application). Therefore, the functionality of these electronics equipment must be assured and the risk of electromagnetic interferences must be reduced during their lifetime.

Nowadays, circuits' robustness to electromagnetic interferences (EMI) is checked in a burn-in component [1], without taking into account the impact of the natural devices' aging. However, shrunk dimensions cause the appearance of several time dependent failure mechanisms, which can limit the functionality of the circuits and modify their electromagnetic behaviour [2]. Therefore, the time dependence of electromagnetic behaviour, which is known as Electromagnetic Robustness or Electromagnetic Reliability (EMR), should be evaluated. In this paper, for the first time, the impact of one of the main time dependent reliability problem, the negative bias temperature instability (NBTI) on the switching noise emission has been analysed.

It is mainly accepted that NBTI is ascribed to the formation of Si/SiO_2 interface states and the oxide positive charge [3]. Regarding to this problem, the dominating work has been concentrated on discrete transistor parameter shift, rather than on circuit performance. At device level, it has been demonstrated that NBTI effect on pFET is manifested as a gradually increment of the threshold voltage (V_{th}). Therefore, the drive current is reduced [4]. The NBTI under static and dynamic conditions has also been reported. These investigations show that the voltage threshold shift (ΔV_{th}) under dynamic stress is almost half of DC case, due to the recovery properties of NBTI [5]. At circuit level, the NBTI effects have not been deeply investigated. However, some works have pointed out that NBTI provoke a signal noise margin (SNM) reduction on SRAM cell [2]. According to the authors' knowledge there is not any work focused on the effects of NBTI in the EMC behaviour. In this paper, the impact of NBTI on the overall behaviour and the switching noise emissions of CMOS ring oscillator have been evaluated.

This paper is structured as follows: in Section 2, the impact of NBTI on single pFET is obtained experimentally and the voltage threshold shift is experimentally quantified. In Section 3, the impact of NBTI on oscillation frequency and switching noise of a 41 stages ring oscillator has been evaluated through simulation, using the data measured on Section 2. In Section 4, the conclusions are summarized.

2. MEASURING THE IMPACT OF NBTI

This investigation has started by measuring the voltage threshold shift (ΔV_{th}) due to NBTI on a single pFET. In order to do that, a stress-measure-stress sequence has been followed in the experiment, with exponentially increasing periods of time. During the measurement phase, the ID-VG pFET characteristic has been measured. During the NBTI stress phase, a constant voltage of 0 V has been applied to the pFET gate and the rest of the terminals have been biased at 2 V. In order to accelerate the NBTI effects the experiments have been done at 125°C. In Fig. 1, the ID-VG pFET behaviours at different stress time are plotted, As it is shown, a voltage threshold shift is produced in the transistor due to NBTI, the ΔV_{th} time dependence shows a power law dependence, depicted in Fig. 2, which matches with the theoretical prediction [2].

3. EMC BEHAVIOUR UNDER NBTI

The switching noise is probably one of the main EMC emission problems in CMOS circuits. It is well known that switching noise is a common impedance coupling [6]. In order to reduce it, several strategies can be followed, such as reducing the common impedance and/or including intra or extra-chip decoupling capacitance. In the previous paragraph it has been shown that NBTI affects the ID-VG behaviour of the pFET. Therefore, some effect on switching noise behaviour should be expected. In order to analyse the impact of NBTI on switching noise, a CMOS ring oscillator has been simulated with commercial *Agilent ADS* software. The BSIM4 model extracted from PTM website [7] has been used and the voltage threshold shift of the pFET has been shifted following the time dependence response obtained experimentally in Fig. 2. In Fig. 3, the schematic of the internal oscillator under investigation is shown. The oscillator is based on 40 inverters plus 2 input NAND gate (the NAND gate is included as enable). The aspect ratios for the CMOS inverters are 1.0 μm/90 nm and 3 μm/90 nm for nFET and pFET, respectively.

In Fig. 4(a), the output voltage behaviour of CMOS ring oscillator has been plotted without

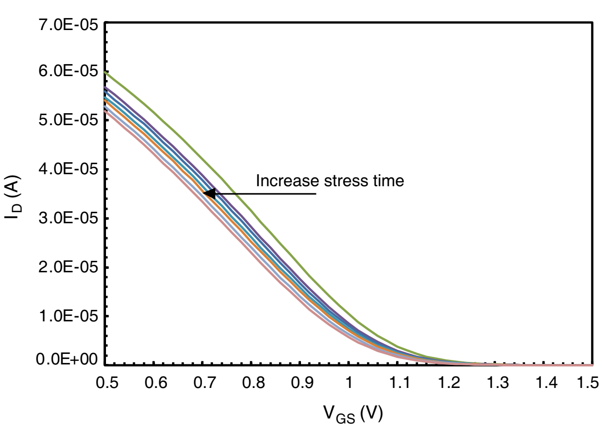


Figure 1: I_D-V_G behavior of pFET for different stress time.

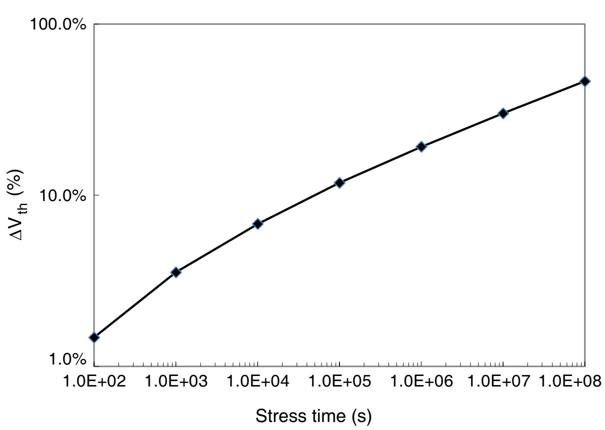


Figure 2: ΔVth of single pFET against NBTI stress time.

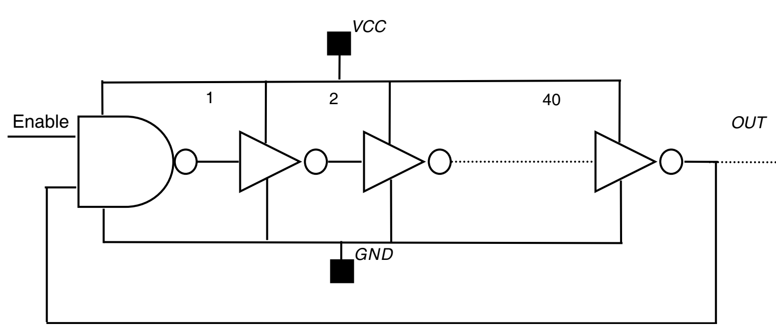


Figure 3: Ring oscillator under evaluation.

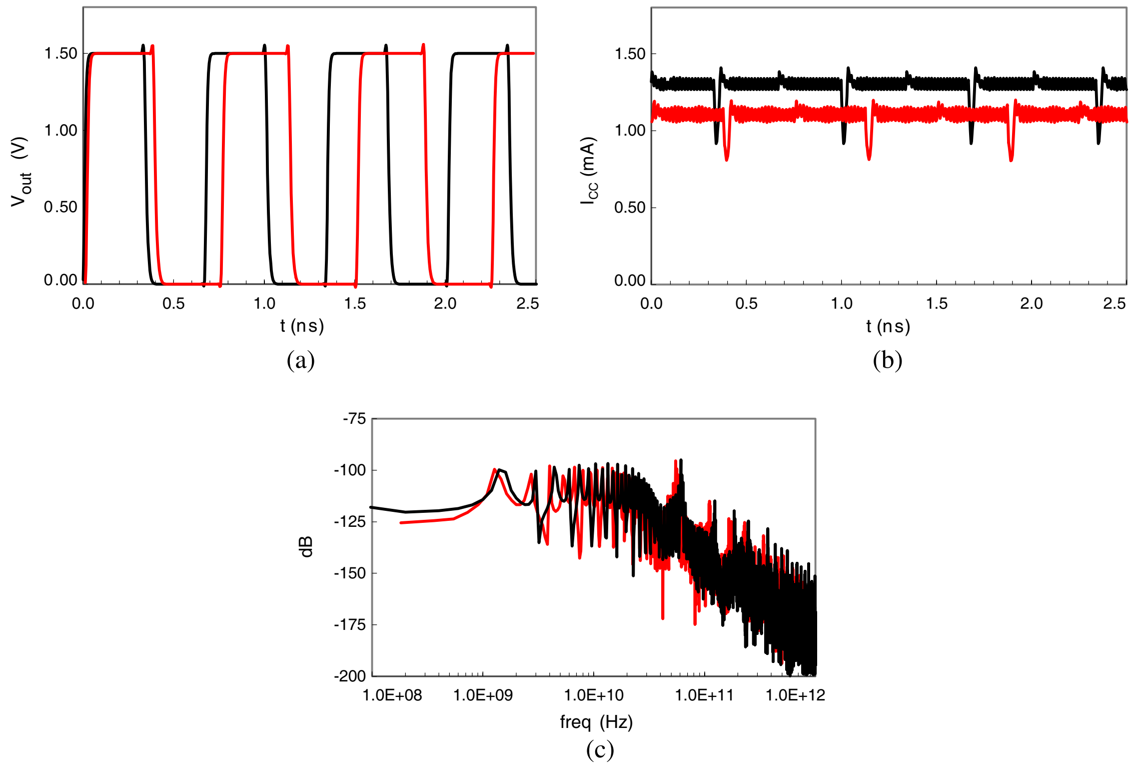


Figure 4: Ring oscillator behaviour before (black) after stress (red). (a) Voltage output. (b) Current consumption. (c) Spectrum of current consumption.

stress (black line) and after 10^8 seconds, almost 3 years NBTI stress (red line). In both cases, the ring oscillator is working. However, a reduction of oscillation frequency from 1.5 GHz to 1.3 GHz after 3 years is observed, which is expected since the NBTI provokes the reduction of the pFET drain current (Fig. 1). In order to have an idea about the switching noise behaviour the current consumption it is shown for similar stress time (Fig. 4(b)). As it is observed, the current consumption is reduced after 3 years, due to the reduction of oscillation frequency. Moreover, some sharp transitions on ICC are observed during the output transition which provokes a shift on the supply voltage due to the common impedance coupling. In Fig. 4(c), the noise spectrum on ICC has been plotted, before and after 10^8 s stress. Before stress (black line) a high power density about 1.5 GHz (and his respective harmonics) is observed due to the sharp point on ICC during the output transitions, and a high density power is also observed about 61 GHz, with correspond with the transitions of each CMOS inverter. After stress, the same behaviour is observed with a frequency oscillation shift; 1.3 GHz instead of 1.5 GHz and 53 GHz instead of 61 GHz, due to the frequency oscillation reduction, as expected.

4. CONCLUSION

In this paper, the impact of aging on the electromagnetic behaviour of CMOS ring oscillator has been evaluated. Specifically, the NBTI effect on switching noise has been evaluated. After 3 years NBTI stress, the ring oscillator is still working, with a reduction about 13% of the oscillation frequency. Therefore, a current consumption reduction is observed. Regarding to the electromagnetic emission, the energy is mainly in 61 GHz before stress and 53 GHz after stress; which corresponds with the transition of each CMOS inverter.

ACKNOWLEDGMENT

This work has been supported by the Spain-MICINN under Project TEC2009-09994 and AGAUR 2009 SGR 1425.

REFERENCES

1. Ben Dhia, S., A. C. Ndoye, A. Boyer, L. Guillot, and B. Vrignon, "IC emission spectrum drifts after burn-in cycles," *Asia-Pacific Symposium on Electromagnetic Compatibility and*

19th International Zurich Symposium on Electromagnetic Compatibility, APEMC 2008, 255–258, May 19–23, 2008.

2. Kang, K., H. Kufluoglu, K. Roy, and M. Alam, "Impact of negative-bias temperature instability in nanoscale SRAM array: Modeling and analysis," *IEEE Transactions on Computer-aided Design of Integrated Circuits and Systems*, Vol. 26, No. 10, 1770–1781, Oct. 2007.
3. Alam, M. A. and S. Mahapatra, "A comprehensive model of PMOS NBTI degradation," *Microelectronics Reliability*, Vol. 45, No. 1, 71–81, Jan. 2005.
4. Stathis, J. H. and S. Zafar, "The negative bias temperature instability in MOS devices: A review, microelectronics and reliability," Vol. 46, No. 2–4, 270–286, Feb.–Apr. 2006.
5. Fernandez, R., B. Kaczer, A. Nackaerts, S. Demuynck, R. Rodriguez, M. Nafria, and G. Groeseneken, "AC NBTI studied in the 1 Hz ý 2 GHz range on dedicated on-chip CMOS circuits," *IEDM Tech. Digest*, 1, 2006.
6. Ben Dhia, S., M. Ramdani, and E. Sicard, *Electromagnetic Compatibility of Integrated Circuits*, Springer, 2005, ISBN 0-387-26600-3.
7. http://www.eas.asu.edu/∼ptm.

Reconfigurable RF-MEMS Metamaterials Filters

I. Gil[1], **M. Morata**[2], **R. Fernández-García**[1], **X. Rottenberg**[3], **and W. De Raedt**[3]

[1]Department of Electronic Engineering, UPC Barcelona Tech., Colom 1, Terrassa 08222, Spain
[2]Escuela Universitaria Salesiana de Sarriá, Barcelona 08017, Spain
[3]IMEC, Leuven B-3001, Belgium

Abstract— In this work, the design procedure, modelling and implementation of reconfigurable filters based on RF microelectromechanical systems metamaterials is presented. Specifically, tunable stop-band and pass-band frequency responses are obtained by combining RF-MEMS with metamaterials based in complementary split rings resonators. These particles, allow for the design of negative effective permittivity transmission lines, providing forbidden propagation frequency bands. Moreover, CSRRs properly combined with metal vias in transmission lines, generate a simultaneously $\varepsilon < 0$ and $\mu < 0$ effective media which involves an allowed frequency band. These two phenomenons have been used in order to implement stop-band and pass-band filters. Since CSRRs present a LC-tank behaviour and are electrically coupled to the host line, the tunability is achieved by means of the RF-MEMS, which modify the electrical characteristics of the CSRRs and the electric coupling. A full electrical model for the description of the proposed structures is presented. The circuit model take into account the electrical characteristics of the RF-MEMS, CSRRs and transmission lines as well as the involved electromagnetic coupling and are used for accurate prediction of switchable filters response.

1. INTRODUCTION

RF microelectromechanical systems (RF-MEMS) tunable filters have been developed in recent years for military and commercial applications such as multiband communication systems or wide-band transceivers due to their low-power consumption, high quality factor and good linearity [1]. RF-MEMS are typically implemented by means of several lithographic steps which perform RF-MEMS in conventional distributed transmission lines. Alternatively, the combination of metamaterial structures together with transmission lines (i.e., artificial lines consisting of a host line loaded with reactive elements) have been revealed as good candidates in order to improve the performance of conventional distributed passive devices, and specifically those corresponding to microwave filters. In order to deepen the knowledge of designers concerning both RF-MEMS tunable filters and tunable metamaterial transmission lines in PCB, several works have been published including performance and accurate electrical models [2–5]. In this paper, the design procedure, modelling and implementation of reconfigurable filters based on RF-MEMS metamaterials is presented. Specifically, tunable stop-band and pass-band frequency responses are obtained by combining RF-MEMS with complementary split rings resonators (CSRRs) [6]. CSRRs, which are the dual image of split rings resonators (SRRs) [7], allow for the design of negative effective permittivity transmission lines, providing forbidden propagation frequency bands due to the effective behavior of $\varepsilon < 0$ and $\mu > 0$ in a certain frequency band. Moreover, CSRRs properly combined with metal vias in transmission lines, generate a simultaneously $\varepsilon < 0$ and $\mu < 0$ effective media which involves an allowed frequency band. These two phenomenons have been used in order to implement stop-band and pass-band filters.

2. STOP-BAND CSRR/RF-MEMS FILTERS

Figure 1 shows the layout of the proposed stop-band filter prototype. It consists of a $50\,\Omega$ coplanar waveguide structure (CPW) loaded with 4-stages of rectangular shaped CSRRs etched in the metal strip and RF-MEMS bridges over them. The actual RF-MEMS are based on electrostatic parallel-plate varactors which are constituted of a movable electrode mechanically anchored on the substrate and suspended above a second fixed electrode. In fact, RF-MEMS are implemented in the CSRRs by using an electrostatic floating bridge anchored on the substrate in holes of the CPW ground planes. Under DC-bias, the device tends to close, adjusting the capacitance defined by the two plates [8]. Due to the intrinsic instability of this actuation scheme, these devices are used in switch-mode. Only up- and down-states (i.e., small and large capacitance) are functional. A stripped-down RF-MEMS technology using only 3 lithographic steps [9] has been used to define the structures depicted in Fig. 1. First, a $1\,\mu\text{m}$ thick Al layer is sputter-deposited and patterned

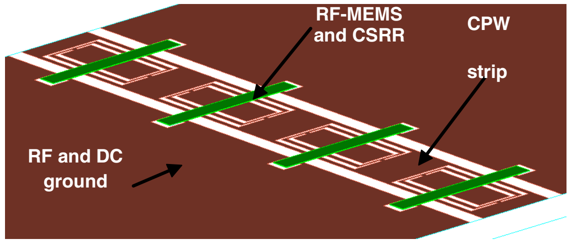


Figure 1: Layout of the 4-stages CSRR/RF-MEMS switchable notch filter.

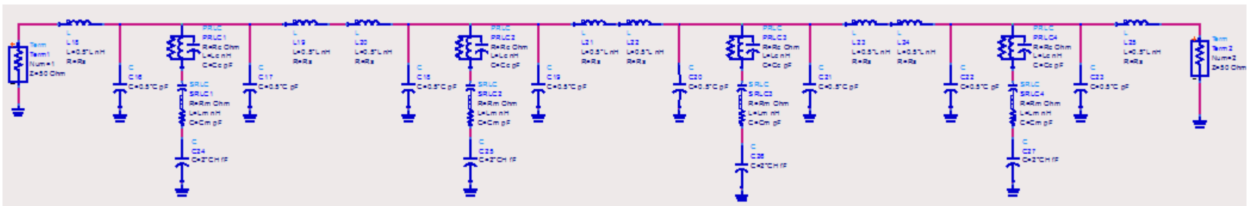


Figure 2: Equivalent lumped circuit model of the filter depicted in Fig. 1.

on a 650 µm thick AF45 glass substrate ($\varepsilon_r = 5.9$) to define mainly the CPW structures. Then, a 3 µm thick sacrificial photoresist layer is spun and patterned to define the anchoring regions of the MEMS devices before a second Al layer is deposited and patterned in the same way as the first one. Therefore, the MEMS beams are defined. Finally, the sacrificial photoresist is ashed to release the devices. Due to the variable capacitance of the RF-MEMS and its coupling to the host transmission line, the effective capacitance of the CSRRs can be modified, and hence, their intrinsic resonance frequency. A full electrical model for the description of the proposed structures is presented in Fig. 2. The circuit models take into account the electrical characteristics of the RF-MEMS, CSRRs and transmission lines as well as the involved electromagnetic coupling and are used for accurate prediction of switchable filters response. The lienRF-MEMS bridges are modelled by means of a lumped RLC series circuit, with a variable capacitance, C_M, (having an up-state and a down-state value), L_M is the bridge inductance and the resistor R_M involves the microelectromechanical system losses. The anchoring capacitance of the CPW holes is modelled by C_H. The CPW line is described by means of the per-section inductance, L, and capacitance, C. The etched CSRRs are modelled by means of a parallel RLC tank, L_C and C_C being the reactive elements which constitute the intrinsic resonance frequency of the resonators and RC takes into account the eventual losses associated with the resonator. CSRRs are directly connected to the host line and electrically coupled to the capacitance of the RF-MEMS. Therefore, the intrinsic resonance frequency of the subwavelength resonators, f_o, (1) is directly affected by the RF-MEMS actuation. As a result of this behaviour, a reconfigurable stop-band frequency response is achieved, which corresponds to the condition which nulls the impedance of the shunt equivalent branch, Z_S, constituted by the impedances of the RF-MEMS, CSRRs, anchoring capacitance and host line capacitance, according to (2):

$$f_o = \frac{1}{2\pi\sqrt{L_C C_C}}, \tag{1}$$

$$Z_S = 0 \Rightarrow \frac{j\omega L_C R_C}{R_C - \omega^2 R_C L_C C_C + j\omega L_C} + \frac{R_C - \omega^2 L_M\left[2C_H C_M/(2C_H + C_M)\right] + j\omega R_M\left[2C_H C_M/(2C_H + C_M)\right]}{j\omega\left[2C_H C_M/(2C_H + C_M)\right]} = 0. \tag{2}$$

The presented model has been electrically simulated by using the commercial *Agilent ADS* and validated by comparison with electromagnetic simulation (i.e., layout simulation with *Agilent Momentum*) and experimental test (a HP8510C vector network analyzer and G-S-G test probes having coaxial connectors type 2.4 mm, have been used). Fig. 3 shows the comparison between the measured data, electromagnetic simulations and the proposed lumped model insertion losses (S_{21}). Both, the up and down state have been electromagnetically simulated by considering a parallel plate to the host line. Layout simulation has been carried out by considering planar plate heights of 0.5 μm (down-state) and 2 μm (up-state), respectively and no losses have been assumed. As expected, the structure exhibits stop-band behaviour with reconfigurable capability. The simulated frequency range corresponds to 39–48.8 GHz (Q-band), which implies to a switching range of roughly 20%. As can be seen, good matching level between the involved notch frequencies concerning the model prediction is achieved. However, the band edges, related with the quality factor, as well as the rejection level are not correctly reproduced due to the intrinsic lossless simplicity. Electromagnetic simulated rejection levels are $IL < -50\,\mathrm{dB}$ whereas model predicts higher notch rejection levels. Regarding experimental results, the RF-MEMS have been subjected to actuation voltage, caused by an external polarization, from 17 V (down-state) to 0 V (up-state). A significant agreement with respect to the full equivalent circuit model including losses is achieved, not only concerning the tuning range (experimental values are 39–48.1 GHz, whereas model correspond to 38.9–47.5 GHz), but also in terms of rejection level (experimental $ILD = -43\,\mathrm{dB}$; $ILU = -40.9\,\mathrm{dB}$ and model $ILD = -50.4\,\mathrm{dB}$; $ILU = -40.2\,\mathrm{dB}$) and quality factor (experimental $QD = 30.9$; QU is not determined due to the limitation up to 50 GHz in the setup, and model $QD = 29.7$; $QU = 29.5$).

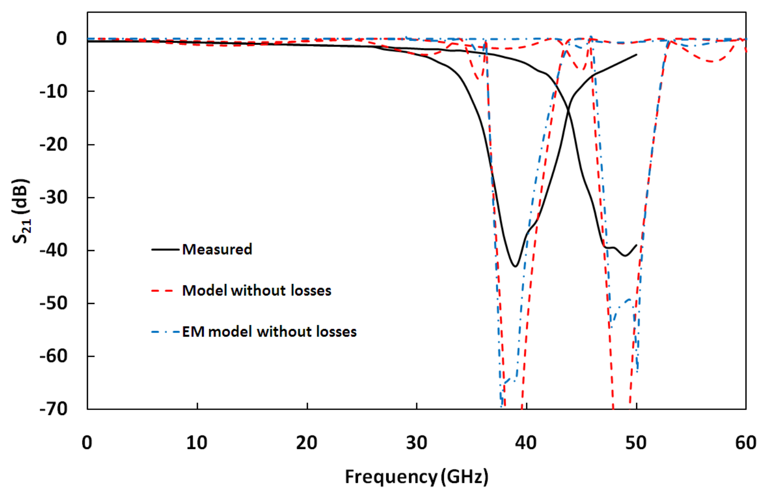


Figure 3: Frequency response transmission (insertion losses) corresponding to the up and down-state of the stop-band CSRR/RF-MEMS filter. Experimental data, electromagnetic simulations and electrical model are shown.

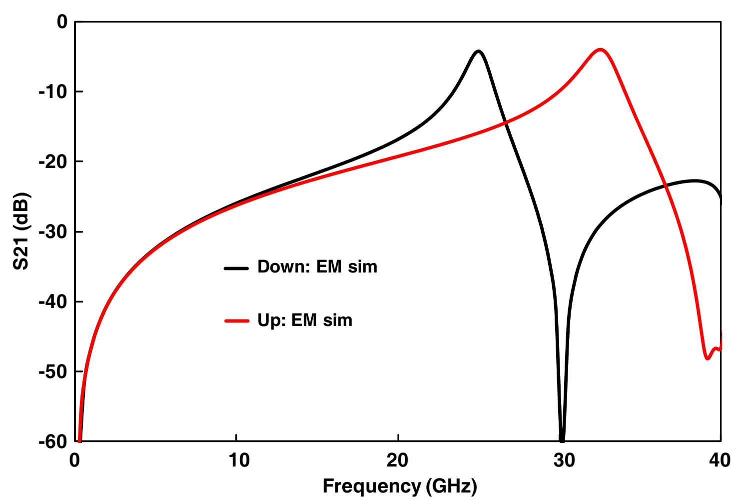


Figure 4: Frequency response transmission (insertion losses) corresponding to the up and down-state of the pass-band CSRR/RF-MEMS filter. Electromagnetic simulations are shown.

3. PASS-BAND CSRR/RF-MEMS FILTERS

In order to obtain a band-pass filter response, a similar unit cell topology than in Fig. 1 can be reused by adding two metal wires (per stage) connecting the central CPW strip and the ground planes in order to achieve an effective band pass frequency response. In fact, due to the equivalent $\mu < 0$ behaviour of the metal wires combined with the $\varepsilon < 0$ of the CSRRs, an allowed band in a certain frequency band is obtained. Fig. 4 shows the electromagnetic simulation of a $50\,\Omega$ CPW structure loaded with 1-stages of rectangular shaped CSRR etched in the metal strip and RF-MEMS bridges over them including the extra metal wires connecting host line and ground. As expected, a pass-band CSRR/RF-MEMS filter is obtained. The frequency range is different with regard to Fig. 3, since other subwavelength resonators' dimensions have been considered. This structure is now pending of fabrication and work is in progress in order to evaluate its experimental performance.

4. CONCLUSIONS

Several tunable stop-band and pass-band filter, based on the combination of complementary split ring resonators and RF-MEMS switchable capacitors have been presented and characterized as proof-of concept. The measured frequency responses of the device for different actuation voltages are good and significant tuning range has been achieved for stop-band responses. Results are promising for pass-band structures and work is currently in progress.

ACKNOWLEDGMENT

This work has been supported by the Spain-MICINN under Project TEC2009-09994 and AGAUR 2009 SGR 1425.

REFERENCES

1. Rebeiz, G. M., *RF MEMS: Theory, Design and Technology*, John Wiley & Sons, 2003.
2. Entesari, K. and G. M. Rebeiz, "A 12–18 GHz three-pole RF MEMS tunable filter," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 53, 2566–2571, August 2005.
3. Entesari, K., K. Obeidat, A. R. Brown, and G. M. Rebeiz, "A 25–75 MHz RF MEMS tunable filters," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 55, 2399–2405, November 2007.
4. Bonache, J., I. Gil, J. García-García, and F. Martín, "Novel microstrip band pass filters based on complementary split rings resonators," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 54, 265–271, January 2006.
5. García-García, J., F. Martín, E. Amat, F. Falcone, J. Bonache, I. Gil, T. Lopetegi, M. A. G. Laso, A. Marcotegui, M. Sorolla, and R. Marqués, "Microwave filters with improved stop band based on sub-wavelength resonators," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 53, 1997–2006, June 2005.
6. Falcone, F., T. Lopetegi, J. D. Baena, R. Marqués, F. Martín, and M. Sorolla, "Effective negative-ε stopband microstrip lines based on complementary split ring resonators," *IEEE Microwave Wireless Compon. Lett.*, Vol. 14, 280–282, June 2004.
7. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 47, 2075–2084, November 1999.
8. Varadian, V. K., K. J. Vinoy, and K. A. Jose, *RF MEMS and Their Applications*, John Wiley & Sons, 2003.
9. Rottenberg, X., S. Brebels, P. Ekkels, P. Czarnecki, P. Nolmans, R. P. Mertens, B. Nauwelaers, R. Puers, I. De Wolf, W. De Raedt, and H. A. C. Tilmans, "Electrostatic fringing-field actuator (EFFA): Application towards a low-complexity thin-film RF-MEMS technology," *J. Micromech. Microeng.*, Vol. 17, S204–S210, 2007.

Optimization of Coherence Multiplexed Coding for High Density Signal Processing

S. Elwardi[1,2], **M. Zghal**[1], **and B.-E. Benkelfat**[2]

[1]Cirta'Com Laboratory, Engineering School of Communication of Tunis (Sup'Com)
University of Carthage, Ghazala Technopark, Ariana 2083, Tunisia
[2]Institut Télécom, Télécom SudParis SAMOVAR UMR INT-CNRS 5157
9 rue Charles Fourier 91011 Evry Cedex, France

Abstract— In this paper, we study the effects of the divergence of a Gaussian beam in coherence multiplexed coding technique. We use this optical system for high density signal parallel arithmetic operations. We illustrate the variation of the measurement coherence length as function of the angular divergence for different wavelength λ, beam-waist W_0, and z_1, the distance travelled by the first beam of the interferometer. We show that the divergence of a Gaussian beam introduces an additional crosstalk between light fields.

1. INTRODUCTION

Real time optical arithmetic operation of signals has found a wide range of applications in both temporal and spatial domain. The most commonly used method for this fundamental arithmetic uses interferometric configurations [1], a noninterferometric method [2] and also the basic joint transform correlator architecture [3].

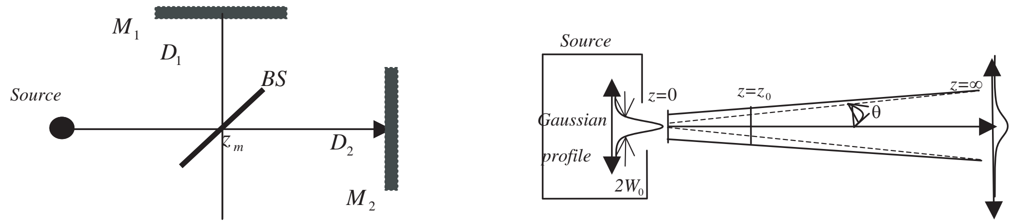
In our research work, in a novel and original application of coherence multiplexed coding technique, we propose a high density signal processing [4]. Coherence multiplexed technique is an interesting alternative method among the existing different techniques for optical communication and signal processing. It has been used for sensing of high intensity and wideband electric fields [5], for transmitting simultaneously a several signals through a single light beam [6] and also for processing data such as faster matrix-vector, matrix-matrix products [7, 8]. Let's recall briefly that the N-multiplexed coherence system is composed by a cascade of N encoding modules (EM) and a decoding module (DM). Each EM_i is formed by a birefringent electro-optic modulator (EOM_i) and a birefringent slab Q_i set between two polarizers (P) and introducing a static optical path difference (OPD_i), $i = 1, \ldots, N$. The main condition to have coherence modulation is the choice of the static OPD, which must be larger than the coherence length of the source (Lc) [6].

In this paper, in the context of increased need of optimization and performance in optical coherence multiplexed technique, and in order to reduce noise and crosstalk, we study the effect of diffraction in coherence multiplexed systems. In the second section, by using a simple Gaussian model, we report the measurement of the coherence length of the source as function of the angular divergence θ for difference value of the distance traversed by the first beam z_1, wavelength λ and beam-waist size W_0. Finally, we find out the value of the additional crosstalk introduced in the coherence multiplexed system as function of the divergence and suggest a solution based on the optical properties of the birefringent elements to reduce the crosstalk.

2. DIFFRACTION EFFECT ON COHERENCE LENGTH MEASUREMENT

To have coherence modulation of light, it is mandatory to introduce a static OPD greater than the coherence length of the source. In this context, we have performed a numerical study of the measurement of coherence length $L_{c,meas}$ of the source, which is usually measured, in laboratory, through the Michelson interferometer. The schematic diagram of this interferometer is illustrated in Fig. 1(a). As we can see, that the incident light is splitted into two equal parts via the beam splitter (BS). They traverse, respectively, the distance $z_1 = z_m + 2D_1$ and $z_2 = z_m + 2D_2$. As a result, the important parameter that characterizes the interference pattern, at the output of the Michelson interferometer, is the fringe visibility V. If we consider, a cylindrical coordinates system (ρ, z), After some approximations the fringe visibility will take the form [9, 10],

$$V \approx \frac{2}{\frac{W_1}{W_2} + \frac{W_2}{W_1}} \tag{1}$$



(a) (b)

Figure 1: (a) Michelson interferometer, BS: Beam Splitter, M_1 and M_2: Mirror, D_1 and D_2 lengths of the arms. (b) Gaussian beam diffraction.

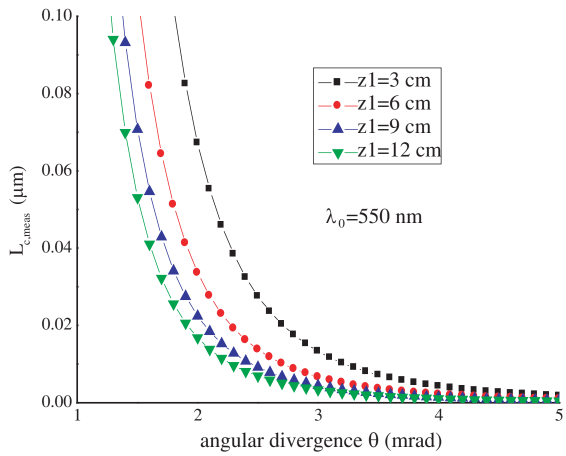


Figure 2: Variation of the $L_{c,meas}$ as a function of θ for $z_1 = 3, 6, 9$ and 12 cm.

where $W_i = W(z_i)$, $i = 1$ or 2, is the Gaussian beam radius and it expressed as [9],

$$W_i^2(z) = W_0^2 \left[1 + \left(\frac{z_i}{z_0} \right)^2 \right] \tag{2}$$

where $z_0 = \pi W_0^2/\lambda$ is the Rayleigh range of the input beam, W_0 is its beam-waist size. If the incident beam diverges with an angular divergence $\theta = \lambda/\pi W_0$, the expression of the Rayleigh range of the input beam will become $z_0 = W_0/\theta$, as shown in Fig. 1(b). The FWHM width corresponding to visibility degradation to 0.5 is given by $\delta\nu = cln2/\pi\Delta z_{0.5}$, where $\Delta z_{0.5} = z_2 - z_1$ is the OPD defined for $V = 0.5$. After some algebra calculation, the expression of $\Delta z_{0.5}$ is expressed as,

$$\Delta z_{0.5} \approx \sqrt{\left(4\sqrt{3} - 6\right) z_0^2 + z_1^2} - z_1 \tag{3}$$

The expression of the measurement coherence length $L_{c,meas}$ is given by $L_{c,meas}(\theta) = \pi\Delta z_{0.5}(\theta)$ $/2\ln(2)$. According to (3), we obtain,

$$L_{c,meas}(\theta) \approx \frac{\pi}{2\ln(2)} \left(\sqrt{\frac{(4\sqrt{3} - 6)\, W_0^2}{\theta^2} + z_1^2} - z_1 \right) \tag{4}$$

The variation of the measurement coherence length as function of the angular divergence θ is illustrated in Fig. 2 for $z_1 = 3, 6, 9$, and 12 cm. We observe that the $L_{c,meas}$ is more affected for a $z_1 = 6$ cm. In Fig. 3(a) and Fig. 3(b), similar results are obtained for $z_1 = 6$ cm, respectively, for $W_0 = 1, 10, 20$ and 30 μm and for different wavelength $\lambda = 550, 630, 825$ and 1300 nm.

3. CROSSTALK

Coherence modulation of light allows an original multiplexing technique of several signals through a single light beam. For simplicity, we limit the discussion to only two encoding module. It

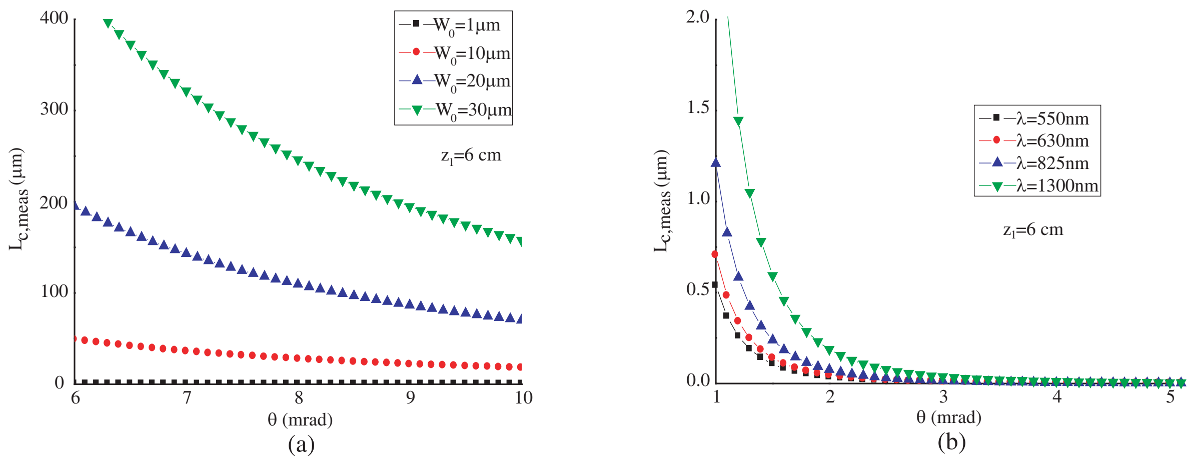


Figure 3: Variation of the $L_{c,meas}$ as a function of θ, (a) for $W_0 = 1, 10, 20$ and 30 µm, (b) for $\lambda = 550, 630, 825$ and 1300 nm.

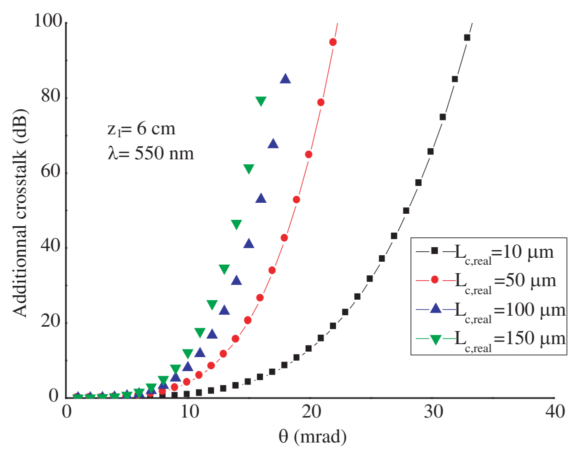


Figure 4: Variation the additional Crosstalk as a function of θ for $L_{c,real} = 10, 50, 100$ and 150 µm.

consists of two cascaded encoding module (EM) or coherence modulators (CM) powered by a CW superluminescent diode (SLD), one decoding module (DM), and a photodiode. To reduce crosstalk between channels the optical path difference (OPD_1) introduced by the first EM, OPD_1 must be equal to $3 * OPD_2$ [5]. Furthermore, it was showed that the optical crosstalk between channels is closely dependent on the spectrum envelope of the source used. It has been shown that the best performance in terms of crosstalk is obtained using a source with a Gaussian emission spectrum. By using the reference [11] the expression of the crosstlak level at the first detector is expressed as,

$$CT_{1,L_{c,real}} \approx 10 \log_{10} \left\{ \exp \left(-\pi * OPD_1 / L_{c,real} \right) \right\} \tag{5}$$

where $L_{c,real}$ is the real and the correct value of the coherence length. If we take into account that the Gaussian beam diverges, we obtain that the crosstalk can be expressed as,

$$CT_{1,L_{c,meas}} \approx 10 \log_{10} \left\{ \exp \left(-\pi * OPD_1 / L_{c,meas} \right) \right\} \tag{6}$$

So the additional crosstalk is given by the difference between $CT_{1,L_{c,real}} - CT_{1,L_{c,meas}}$, so it can be expressed as,

$$CT_\theta \approx 10 \log_{10} \left\{ \exp \left(L_{c,real} / L_{c,meas} \right) \right\} \tag{7}$$

Or also,

$$CT_\theta \approx 10 \log_{10} \left\{ \exp \left(L_{c,real} / \frac{\pi}{2 \ln(2)} \left(\sqrt{\frac{\left(4\sqrt{3} - 6\right) W_0^2}{\theta^2} + z_1^2} - z_1 \right) \right) \right\} \tag{8}$$

This additional crosstalk might have an impact on the performance of the system. In Fig. 4, we plot the variation of the additional crosstalk as function of θ for $L_{c,real} = 10, 50, 100$ and $150\,\mu\text{m}$. For example, when we use a white light $\lambda = 550\,\text{nm}$ and with a coherence length $L_{c,real} = 50\,\mu\text{m}$ to illuminate a 2-D coherence coding and we want to have a maximum of crosstalk equal to $-40\,\text{dB}$, if the Gaussian beam diverge with a 15 mrad, we would have an additional crosstalk $= 20\,\text{dB}$. As a result, the given crosstalk will be equal to $-20\,\text{dB}$. We observe that it is important to use a light with a very smaller coherence length to reduce the effect of the divergence in the system. We can compensate this additional crosstalk by using a birefringent slab in the encoding module which introduces an additional delay. The exact value of the additional delay is deduced from Fig. 4.

4. CONCLUSIONS

We have carried out a numerical study of the effects of the diffraction of a Gaussian beam on optical crosstalk level in coherence multiplexed coding technique. We have demonstrated that the transverse effects can lead to significant variations in the optical crosstalk. So, the multiplexed configuration will suffer from an important additional crosstalk level.

REFERENCES

1. Ebersole, J. F. and J. C. Wyant, "Real-time optical subtraction of photographic imagery for difference detection," *Appl. Opt.*, Vol. 15, 871, 1976.
2. Shoop, B. L., B. Pezeshki, J. W. Goodman, and J. S. Harris, Jr., "Noninterferometric optical subtraction using reflection-electroabsorption modulators," *Opt. Lett.*, Vol. 17, 58, 1992.
3. Pati, G. S., G. Unnikrishnan, and K. Singh, "Multichannel image addition and subtraction using joint-transform correlator architecture," *Opt. Com.*, Vol. 150, 33, 1998.
4. Elwardi, S., M. Zghal, and B.-E. Benkelfat, "Coherence modulation of light for mathematical operations," *IEEE Int. Nat. Conf. on Elec. Circ. and Sys., ICECS'08*, 878–881, 2008.
5. Rodríguez-Asomoza, J., R. Rosas-Romero, S. E. Lobato-Larios, R. Alejos-Palomares, J. A. Navarro-Martínez, and C. Gutiérrez-Martínez, "Electric signal sensor using an electrooptic coherence modulator," *IEEE Trans. on Ins. & Meas.*, Vol. 54, 347–351, 2005.
6. Goedgebuer, J.-P., H. Porte, and A. Hamel, "Electrooptic modulation of multilongitudinal mode laser diodes: Demonstration at 850 nm with simultaneous data transmission by coherence multiplexing," *IEEE J. of Quant. Elec.*, Vol. 23, No. 7, 1135–1144, 1987.
7. Giust, R., J.-P. Goedgebuer, and N. Butterlin, "A digital processor formed by cascaded spatial light modulators," *Opt. Comm.*, Vol. 181, 279–285, 2000.
8. Goedgebuer, J.-P., N. Butterlin, and H. Porte, "Electro-optic systolic processor with a pipeline architecture," *Opt. Lett.*, Vol. 20, 1719–1721, 1995.
9. Aït-Ameur, K. and F. Sanchez, "Transverse effects as source of error in coherence length measurements," *Opt. Com.*, Vol. 233, 39–43, 2004.
10. Al-Saeed T. A. and D. A. Khalil, "Diffraction effects in optical microelectromechanical system Michelson interferometers," *Applied Optics*, Vol. 49, No. 20, 3960–3966, 2010.
11. Geodgebuer, J.-P., A. Hamel, H. Porte, and N. Butterlin, "Analysis of optical crosstalk in coherence multiplexed systems employing a short coherence laser diode with arbitrary power spectrum," *IEEE J. of Quan. Elec.*, Vol. 26, No. 7, 1217–1226, 1990.

Highly Birefringent Photonic Crystal Fiber for Coherent Infrared Supercontinuum Generation

A. Ben Salem[1], **R. Cherif**[1], **M. Zghal**[1], **and J. Burger**[2]

[1]Cirta'Com Laboratory, Engineering School of Communication of Tunis (Sup'Com)
University of Carthage, Ghazala Technopark, Ariana 2083, Tunisia
[2]National Metrology Institute of South Africa (NMISA), Pretoria 0001, South Africa

Abstract— We report on numerical and experimental studies of supercontinuum (SC) generation in a silica-based nonlinear polarization-maintaining photonic crystal fiber (PM-PCF). The optical properties of the fundamental mode of the PM-PCF were numerically determined in terms of birefringence, chromatic dispersion and nonlinear coefficient. The fiber has an air hole diameter of 0.6 μm and a pitch of 1.25 μm exhibiting a nonlinear coefficient of $90\,\mathrm{km}^{-1}\mathrm{W}^{-1}$ and a high value of birefringence which we found to be equal to $4 \cdot 10^{-4}$ at $\lambda = 820\,\mathrm{nm}$. We also determined the positions of the slow and fast axes of the PM-PCF according to the laboratory frame, using two different techniques based on near and far fields. The SC was generated by coupling ultra-short pulses delivered by a mode-locked femtosecond Ti: Sapphire laser centered at 820 nm with a repetition rate of 202 MHz at the average power of 20 mW and having duration of 56 fs. The generation of over 800 nm bandwidth of visible and infrared light on both axes (slow and fast) was observed in the 2-m-long silica-based nonlinear PM-PCF. Finally, the coherence and noise structure of the supercontinua generated in the PM-PCF have been discussed.

1. INTRODUCTION

Supercontinuum (SC) generation has been observed and investigated since 2000 in photonic crystal fibers (PCF) [1]. In fact, many studies have experimentally demonstrated the generation of broadband supercontinuum in nonlinear photonic crystal fibers that broadly agree with the numerical simulations which widely used the generalized scalar nonlinear Schrödinger equation (NLSE). The NLSE describes all the nonlinear effects and takes various inputs namely the pulse parameters, the high order dispersion coefficients of the fiber and the temporal Raman response [2]. However, the majority of these studies have demonstrated unpolarized SC spectra which are considered to be an important issue for many applications that require polarized SC spectra from birefringent fibers [3]. Thus, the modeling of the nonlinear propagation in birefringent fibers should take into account the coupling between polarizations so that a full vectorial approach has to be applied on the NLSE instead of the scalar one [3]. By this way, it is easy to see the effect of the input polarization orientation on the SC generation and the polarization properties of the output spectra.

In this paper, we characterize the optical properties of the PCF namely the birefringence, the chromatic dispersion and the nonlinear coefficient. We propose two experimental techniques for the determination of the far and near fields in order to recognize the position of the fast and slow axes. Based on the previous study, we launch the light separately through the two axes of the PCF and we generate the correspondent supercontinua. We discuss the interplay of the nonlinear effects and we analyze how the SC was constructed in the polarization maintaining PCF. Finally, we numerically study the effect of the input quantum noise on the coherence properties of the SC generated in the highly birefringent PCF.

2. OPTICAL CHARACTERIZATION OF THE PM-PCF

A full-vectorial finite element method (FEM), which assures high solution accuracy, was used to determine the propagation properties of the PCF. The silica photonic crystal fiber presents an internal stress in the first corona and particularly in two specific holes alongside the core enlarged during the fabrication process. The silica PCF has an average hole diameter $d_1 = 0.6\,\mu\mathrm{m}$ and a distance Λ between the hole centers about 1.25 μm. The two enlarged holes have an average diameter $d_2 = 0.66\,\mu\mathrm{m}$ which makes the core area elliptical so that the two axes of the PCF match the axes of the mode propagation distribution field. Figure 1 depicts the cross section of the PCF and the field distribution of the two polarization fundamental modes.

The analysis of the optical fundamental mode shows that the fast axis of the fiber is the axis crossing the two large holes consequently, the slow axis is the perpendicular one. In fact, we simulated the propagation of the two polarization fundamental modes and we determined the effective

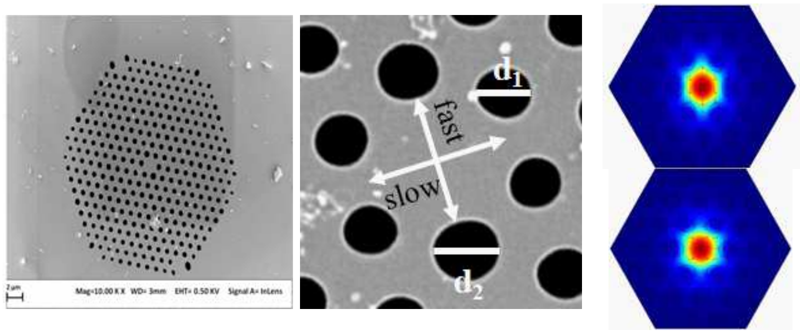


Figure 1: (a) SEM photos of the silica PCF. (b) Simulated field distribution of the two polarization modes.

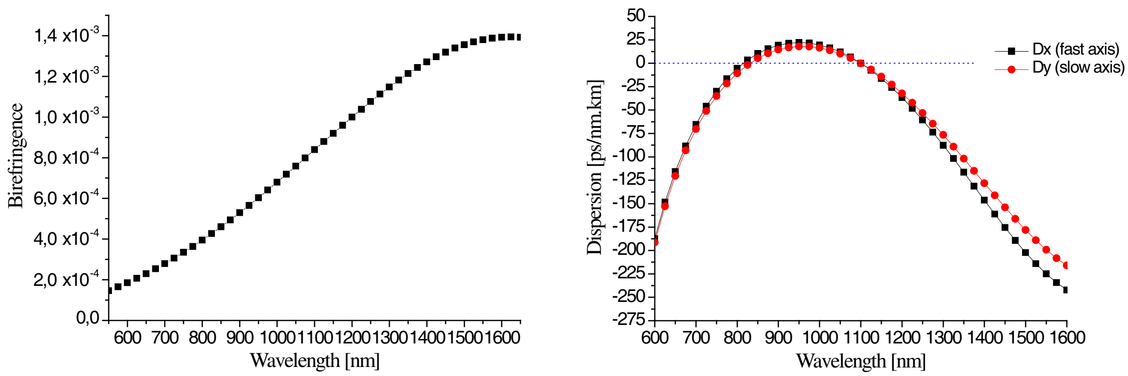


Figure 2: (a) Birefringence of the silica PCF. (b) Chromatic dispersions of the PM-PCF axes as a function of the wavelength.

indices of both axes of the fiber. We recall that the birefringence is defined as the difference between the two effective indices of the x- and y-polarized modes and it is given by: $B = |n_{effy} - n_{effx}|$. Figure 2(a) shows the variation of the birefringence as a function of the wavelength.

We found that the silica PCF is a highly birefringent fiber with birefringence ranging from $2 \cdot 10^{-4}$ at $\lambda = 600$ nm to $1.4 \cdot 10^{-3}$ at $\lambda = 1600$ nm. The strong birefringence is due to the absence of rotational symmetry of $\pi/3$ and the existence of geometric imperfections related to the two large holes that creates asymmetries and breaks the degeneracy of the propagation modes [4]. This perturbation, namely the two opposite holes in the first ring having a diameter greater than the average one ($d_2 > d_1$), is responsible for increasing the birefringence and obtaining a PM (Polarization Maintaining) PCF. In fact, it is a technique for making such a highly birefringent PCF. At $\lambda = 820$ nm, we calculated a birefringence equal to $4 \cdot 10^{-4}$. We note that the manufacturer mentioned a measured birefringence $> 3 \cdot 10^{-4}$ at $\lambda = 780$ nm while our modeling gives a birefringence of $3.7 \cdot 10^{-4}$ at the same wavelength.

To determine the chromatic dispersion, the effective indices (n_{effx} and n_{effy}) of the fundamental modes of the PCF are computed as a function of wavelength. Figure 2(b) gives the spectral evolution of the chromatic dispersion of both fast and slow axes of the PM-PCF.

We clearly see that our PM-PCF exhibits two zero dispersion wavelengths (λ_{ZDW}). The short $\lambda_{ZDW,f}$ is equal to 815 nm for the HE_{11x} mode and 831 nm for the HE_{11y} mode whereas the long zero dispersion wavelength is around 1100 nm for both modes.

3. DETERMINATION OF THE AXES ORIENTATION OF THE FIBER

Experimentally, using a CCD camera, the determination of the fast and slow axes can be recognized by viewing the profile of the electric field into the structure. The experimental determination of the axis of the PCF is based on the collection of the electric field distribution and the analysis of the elliptical profile from which we could deduce the direction of the minor axis. The minor axis corresponds to the fast axis and the major axis corresponds to the slow axis of the PM-PCF. We

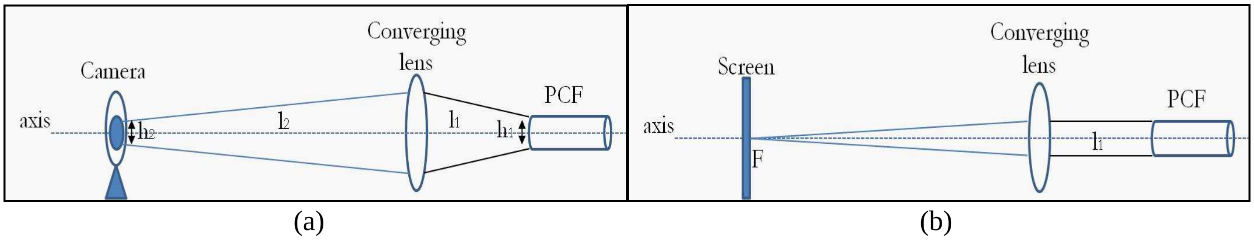


Figure 3: Experimental setups to collect the (a) near and (b) far fields.

performed the study on both near and far fields to ensure the axes recognition. Figure 3 shows the experimental setups.

The CCD camera was arranged to be able to directly collect the near field (as shown in Figure 3(a)). We choose the parameters l_1 and l_2 so that the image collected by the camera does not exceed the size of its screen h_2 which is equal to 4.8 mm and taking into account the magnification factor of the lens $M = -h_2/h_1 = -l_2/l_1$ for $h_1 \approx 2$ microns. The image collected by CCD camera is imported and analyzed to remove the added white noise in order to determine the accurate orientation of the elliptical distribution of the electric field intensity. Thus, the position of the major and minor axes will be determined. We verified our calculation by computing the second order moments along lines traversing the mass centers of the images in order to compensate for inaccuracies in imaging due to aberrations and diffraction limitations, as well as digital imaging noise.

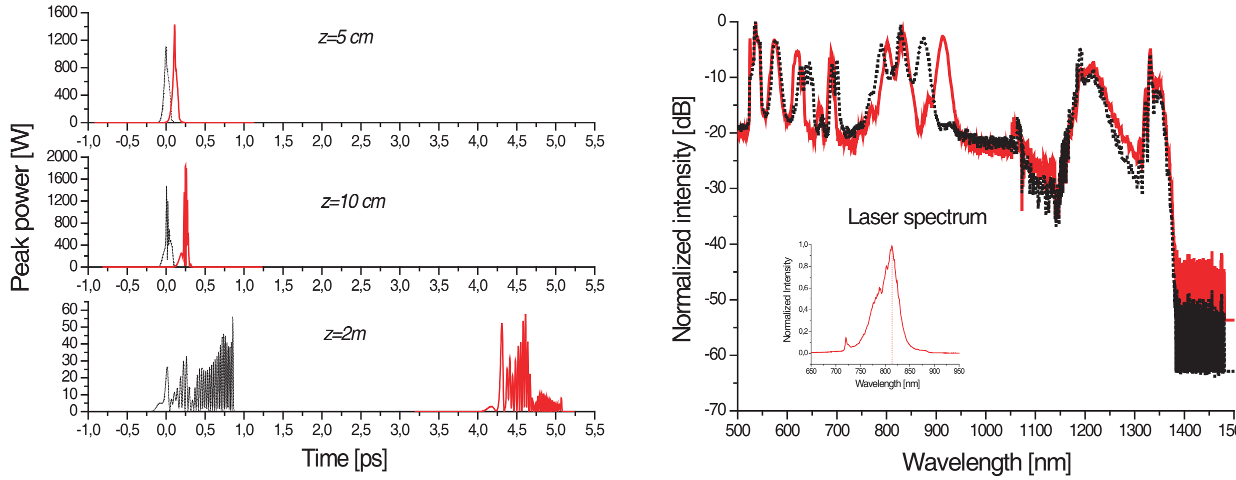
To determine the far field, we put a projection plane in the focal point of the lens. This configuration is presented in Figure 3(b). It allows the collection of an image presenting the Fourier transform of the near-field. In fact, the most remarkable and useful properties of the used converging lens is its inherent ability to perform two-dimensional Fourier transforms. By applying the Fourier transforms on the elliptical distribution, the two axes will be changed so that the major axis will be the minor axis and vice versa. The computation of the second order moment confirms the elliptical distribution orientation. From this mode analysis, we numerically evaluated the effective mode area of the fundamental mode to be equal to $A_{eff} = 2.12\,\mu\text{m}^2$ and by using $n_2 = 2.5 \cdot 10^{-20}\,\text{m}^2 \cdot \text{W}^{-1}$, we deduced a nonlinear coefficient $\gamma = 2\pi n_2/(\lambda A_{eff})$ of $90\,\text{km}^{-1} \cdot \text{W}^{-1}$ at $\lambda = 820\,\text{nm}$.

4. SUPERCONTINUUM GENERATION AND COHERENCE PROPERTIES

After determining the optical properties (birefringence, chromatic dispersion and nonlinear coefficient) of the PM-PCF, we will be able to simulate the nonlinear process and the pulse propagation along the fiber. In this section, we will apply the scalar NLSE to solve the nonlinear propagation on both axes of the fiber (slow and fast). The NLSE which takes into account the contribution of the linear and nonlinear effects describes the temporal and longitudinal dependence of the pulse envelope $A(z, t)$ [5].

We noticed that two pulses launched separately around an average power of 20 mW on each axis were decoupled in time after about 10-cm propagating distance (see Figure 4(a)). Thus, the scalar approach for the resolution of the nonlinear Schrödinger equation can be applied for the 2-m PM-PCF length. We considered a Gaussian input pulse shape $A(0, T) = \sqrt{P_0}\exp(-T^2/2T_0^2)$ centered at 820 nm with a fixed intensity full width at half-maximum of 56 fs duration measured with an autocorrelator and delivered at an average power $P = 20\,\text{mW}$ with a repetition rate of 202 MHz. Figure 4(b) shows the experimental supercontinua generated in both fast and slow axes for an average power $P = 20\,\text{mW}$. As we can see from the laser spectrum, the laser bandwidth centered at the wavelength $\lambda = 820\,\text{nm}$ is set in the both anomalous and normal dispersion regimes. Thus, the over 800 nm bandwidth generated in the visible and the infrared region for both fiber axes are mainly ruled by the effects of self phase modulation, the Raman effects as a responsible for solitonic fissions and the dispersive wave radiation [6].

After studying the SC generation, we are also in the process of appraising coherence and noise as these are critical in applications like frequency measurement and coherent communications. We tried to evaluate the coherence degradation caused by severe fluctuations in the spectral phase.



(a) (b)

Figure 4: (a) Temporal evolution of the SC as a function of the distance. (b) Supercontinua generated in 2-m fiber length at an average power of 20 mW. Solid red (dashed black) curve, along the fast (slow) axis.

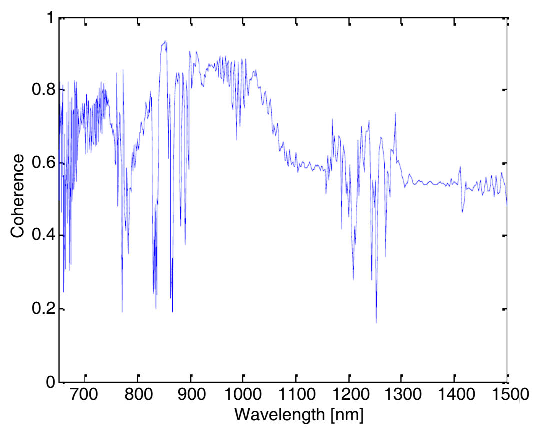


Figure 5: Degree of coherence calculated from an ensemble average of 20 SC pairs.

The first order coherence at each wavelength in the generated SC is given by [7]:

$$|g_{12}(\lambda, t_1 - t_2)| = \frac{|\langle E_1^*(\lambda, t_1) E_2(\lambda, t_2)\rangle|}{\sqrt{\left\langle |E_1(\lambda, t_1)|^2 \right\rangle \left\langle |E_2(\lambda, t_2)|^2 \right\rangle}}$$

where E_1 and E_2 are two electric fields of two immediately successive SC spectra. The maximum value of 1 corresponds to no significant coherence degradation. In order to focus on the wavelength dependence of the coherence, the calculation of g_{12} is considered at $t_1 - t_2 = 0$. The input pulse shot noise is modeled by adding one photon per mode with a random phase noise [7]. From an average ensemble of 20 independent simulations (Figure 5), we clearly see that the spectrum is highly coherent over the entire supercontinuum bandwidth.

5. CONCLUSION

We performed a numerical and experimental study of SC generation in a polarization PM-PCF. We applied two techniques in order to determine the fast and the slow axes of the fiber. We found more than 800 nm bandwidth generated at an average power of 20 mW on both axes of the PM-PCF at $\lambda = 820$ nm. The modeling of the noise structure with independent random phases shows a high degree of coherence over the generated bandwidth.

REFERENCES

1. Ranka, J. K., R. S. Windeler, and A. J. Stentz, "Visible continuum generation in air silica microstructure optical fibers with anomalous dispersion at 800 nm," *Opt. Lett.*, Vol. 25, 25–27, 2000.
2. Agrawal, P., *Nonlinear Fiber Optics*, 3rd Edition, Academic Press, San Diego, 2001.
3. Zhu, Z. and T. G. Brown, "Polarization properties of supercontinuum spectra generated in birefringent photonic crystal fibers," *J. Opt. Soc. Am. B*, Vol. 21, 249–257, 2004.
4. Zghal, M. and R. Cherif, "Impact of small geometrical imperfections on chromatic dispersion and birefringence in photonic crystal fibers," *Opt. Eng.*, Vol. 46, 128002:1–7, 2007.
5. Cherif, R., M. Zghal, I. Nikolov, and M. Danailov, "High energy femtosecond supercontinuum light generation in large mode area photonic crystal fiber," *Optics Communications*, Vol. 283, 4378–4382, 2010.
6. Cherif, R., M. Zghal, L. Tartara, and V. Degiorgio, "Supercontinuum generation by higher-order mode excitation in a photonic crystal fiber," *Opt. Express*, Vol. 16, 2147–2152, 2008.
7. Dudley, J. M. and S. Coen, "Coherence properties of supercontinuum spectra generated in photonic crystal fiber and tapered optical fiber," *Opt. Lett.*, Vol. 27, 1180–1182, 2002.

On the Role of Maxwell Fields in the Resonant Transfer of Energy

A. Salam
Department of Chemistry, Wake Forest University
Winston-Salem, NC 27109-7486, USA

Abstract— Response theory within the framework of molecular quantum electrodynamics is used to calculate the rate of resonant energy transfer between a pair of chromophores in vacuum as well as in a chiral dielectric medium that can undergo dispersion and absorption. Discriminatory transfer rates occurring between chiral molecules are also evaluated. The viewpoint adopted is one in which an unexcited acceptor moiety responds to the driving fields of an excited source donor molecule. The fields are expanded in a power series solution from the Heisenberg equations of motion for the fermion and boson creation and annihilation operators. The role of Maxwell fields in the migration of excitation energy is examined.

1. INTRODUCTION

The theory of molecular quantum electrodynamics (MQED) continues to have considerable success when applied to problems in theoretical chemistry, and atomic, molecular and optical physics. Broadly speaking, these cover a vast array of spectroscopic techniques and intermolecular interactions, both of which involve the coupling of electrons and photons. Particular examples of one or more electromagnetic fields interacting with matter include single- and multi-photon absorption and emission processes, coherent and incoherent elastic and inelastic scattering of light, four-, five- and six-wave mixing, Kerr effect and numerous other linear and nonlinear quantum optical phenomena. Meanwhile fundamental inter-particle couplings comprise the resonant transfer of excitation energy as well as electrostatic, induction and dispersion forces. Foundations of MQED and specific applications may be found in a number of well written and comprehensive textbooks and monographs [1–8], each of which contains extensive bibliographies of the primary literature.

The molecular version of QED is a non-relativistic approximation to covariant QED [9], and is applicable to bound electrons moving at speeds significantly less than that of light, and with energies below mc^2. Its characteristic feature, like that of QED itself, is that the radiation field along with matter is quantized, the kinematics varying according to whether a non-covariant or a relativistic formulation is employed. Hence the photon emerges naturally as a consequence of quantizing the electromagnetic field, with signals properly retarded due to their propagation at the speed of light. Compared to its semi-classical antecedent, in which only electrons are quantized, QED and MQED offer a rigorous description of the interaction of radiation with matter and between matter itself.

As the name QED implies, its basis lies in classical mechanics (Newtonian or relativistic) and electromagnetism, to which are applied the laws of quantum mechanics. Hence Maxwell's equations lie at the heart of the theory, with the free radiation field readily quantized and expressed as a sum of simple harmonic oscillators. Along with the quantum mechanical description of charged particles, the quantum electrodynamical Hamiltonian operator for the total system comprised of matter, radiation field and mutual interaction may be straightforwardly constructed via the familiar canonical quantization procedure and applied to specific problems. The role of the electromagnetic field doesn't end here as is largely the case in the usual formulation of MQED in terms of charged particles interacting with the radiation field. An alternative, often more powerful description is to choose a fully field theoretic picture of coupled electron and photon wavefields, with observables computed from knowledge of the Maxwell field operators in the neighbourhood of a charged source [8, 10, 11]. One obvious application is to intermolecular interactions [12], in which the response of one test body to the MQED fields of the other entity may be computed. Such an approach is adopted in the present work, where resonant migration of electronic excitation energy between two molecules, each of which may also be chiral — leading to discriminatory transfer, is studied. The influence of a chiral dielectric medium on the rate is also evaluated.

2. SECOND QUANTIZED RADIATION-MOLECULE HAMILTONIAN

In the field theoretic approach, the fully second quantized radiation-molecule Hamiltonian operator forms a convenient starting point for the calculation of a number of properties, including resonant energy transfer using response theory. It takes the form [8, 10, 11]

$$H = \sum_n b_n^\dagger b_n E_n + \sum_{\vec{k},\,\lambda} \left[a^{\dagger(\lambda)}(\vec{k}) a^{(\lambda)}(\vec{k}) + \frac{1}{2} \right] \hbar ck - \varepsilon_0^{-1} \sum_{m,\,n} b_m^\dagger b_n \vec{\mu}^{mn} \cdot \vec{d}^{\perp}(\vec{R},\, t), \tag{1}$$

where the self-energy term has been ignored, and b_n and $b_n^\dagger$ are implicitly time-dependent fermion annihilation and creation operators for the state $|n\rangle$ of the electron, of energy E_n. Analogously, $a^{(\lambda)}(\vec{k})$ and $a^{\dagger(\lambda)}(\vec{k})$ are time-dependent boson lowering and raising operators, which respectively decrease and increase by one the number of photons of mode $(\vec{k},\, \lambda)$, with $\vec{k}$ the propagation wavevector, and λ the polarization index. The first two terms of (1) therefore represent the total electron energy and the energy of the radiation field, respectively. The last term of (1) describes the coupling of radiation with matter within the electric dipole approximation of the multipolar framework of MQED [1–3, 5, 8, 11], with $\vec{\mu}^{mn}$ the transition electric dipole moment matrix element between states $|m\rangle$ and $|n\rangle$. It goes without saying that the field theoretic method is entirely equivalent to an occupation number representation in many-body theory. Appearing in Eq. (1) is the time-dependent transverse electric displacement field, whose mode expansion is given by

$$\vec{d}^{\perp}(\vec{r},\, t) = i \sum_{\vec{k},\,\lambda} \left(\frac{\hbar ck\varepsilon_0}{2V} \right)^{1/2} \left[\vec{e}^{(\lambda)}(\vec{k}) a^{(\lambda)}(\vec{k},\, t) e^{i\vec{k}\cdot\vec{r}} - \bar{\vec{e}}^{(\lambda)}(\vec{k}) a^{\dagger(\lambda)}(\vec{k},\, t) e^{-i\vec{k}\cdot\vec{r}} \right], \tag{2}$$

where $e^{(\lambda)}(\vec{k})$ is a unit electric polarization vector and V is the box quantization volume. A similar expansion may be written for the microscopic magnetic field operator $\vec{b}(\vec{r},\, t)$ in terms of the unit magnetic polarization vector $\vec{b}^{(\lambda)}(\vec{k}) = \hat{k} \times e^{(\lambda)}(\vec{k})$.

On employing the fundamental commutator between $a^{(\lambda)}(\vec{k})$ and $a^{\dagger(\lambda)}(\vec{k})$, and the anti-commutator relating b_n and $b_n^\dagger$, and the Heisenberg operator equation of motion, coupled integro-differential equations are obtained for the photon and electron creation and destruction operators. These are typically solved by iteration, generating a power series solution in the electric dipole moment operator. Substituting into the Maxwell field operators yields an expansion of the electric displacement and magnetic field operators in powers of $\vec{\mu}$. Illustrating for the $\vec{d}$-field, the first two terms are the vacuum field, which is independent of the source moment, and the electric displacement field linear in $\vec{\mu}$. Their ith Cartesian components are

$$d_i^{(0)}(\vec{r},\, t) = i \sum_{\vec{k},\,\lambda} \left(\frac{\hbar ck\varepsilon_0}{2V} \right)^{1/2} \left[e_i^{(\lambda)}(\vec{k}) \alpha(0) e^{i\vec{k}\cdot\vec{r} - i\omega t} - H.C. \right], \tag{3}$$

and

$$d_i^{(1)}(\vec{\mu};\, \vec{r},\, t) = \begin{cases} \dfrac{1}{4\pi} \displaystyle\sum_{m,\,n} \beta_m^\dagger(0) \beta_n(0) \mu_j^{mn} (-\nabla^2 \delta_{ij} + \nabla_i \nabla_j) \dfrac{e^{ik} nm^{(r-ct)}}{r}, & r < ct \\ 0, & r > ct \end{cases}, \tag{4}$$

for a dipole situated at the origin, with $a(t) = \alpha(t) e^{-i\omega t}$, and $b_n(t) = \beta_n(t) e^{-i\omega t}$, and similarly for the Hermitian-conjugates. Note the first-order field is strictly causal. Similar expressions may be derived for the magnetic field correct to first-order in $\vec{\mu}$, as well as higher-order terms for both fields, in addition to higher multipole dependent contributions such as those involving magnetic dipole and electric quadrupole terms [10, 13–15]. This simply means accounting for higher multipole terms in the interaction Hamiltonian at the outset, as in

$$H_{int} = -\varepsilon_0^{-1} \sum_{m,\,n} b_m^\dagger b_n \left[\mu_i^{mn} + Q_{ij}^{mn} \nabla_j + \ldots \right] d_i^{\perp}(\vec{R},\, t) - \sum_{m,\,n} b_m^\dagger b_n \left[m_i^{mn} + \ldots \right] b_i(\vec{R},\, t), \tag{5}$$

where $\vec{m}$ is the magnetic dipole moment operator and Q_{ij} is the electric quadrupole tensor.

3. ELECTRIC DIPOLE-DIPOLE RESONANCE ENERGY TRANSFER

Having now obtained the Maxwell fields correct up to first-order in the vicinity of an electric dipole source species, it is straightforward to evaluate the matrix element for electric dipole-dipole resonant migration of energy between an excited donor moiety, D, in state $|p\rangle$, located at $\vec{R}_D$, and an initially unexcited acceptor entity, A, positioned at $\vec{R}_A$ which becomes excited on transfer to state $|p\rangle$ with D decaying to the ground state $|0\rangle$. Body A is viewed as a test dipole molecule responding at $\vec{R}_A$, to the electric dipole source driving field of donor D which is undergoing a downward transition from $|p\rangle$ to $|0\rangle$ with energy $E_{p0} = E_p - E_0 = \hbar\omega_{p0}$, where ω_{p0} is the circular frequency. Taking the expectation value of the linear field (4) between states $|p\rangle$ and $|0\rangle$ and inserting into the coupling energy

$$-\varepsilon_0^{-1}\mu_i^{p0}(A)e^{-i\omega_{0p}t}d_j^{(1)}(\vec{\mu};\, \vec{R}_A,\, t), \tag{6}$$

yields the matrix element [12, 15]

$$M = \mu_i^{0p}(D)\mu_j^{p0}(A)V_{ij}(\omega_{p0},\, \vec{R}), \tag{7}$$

where the retarded resonance interaction tensor is

$$V_{ij}(\omega,\, \vec{R}) = -\frac{1}{4\pi\varepsilon_0}(-\nabla^2\delta_{ij} + \nabla_i\nabla_j)\frac{e^{i\omega R/c}}{R}, \tag{8}$$

and $\vec{R} = \vec{R}_A - \vec{R}_D$. Substituting (7) into the Fermi Golden Rule and orientationally averaging over dipole moments results in the isotropic transfer rate

$$\Gamma = \frac{4\pi\rho}{9\hbar}\frac{1}{(4\pi\varepsilon_0 R^3)^2}|\vec{\mu}^{0p}(D)|^2|\vec{\mu}^{p0}(A)|^2\left[k_{p0}^4R^4 + k_{p0}^2R^2 + 3\right], \tag{9}$$

where ρ is the density of final states and $\omega_{p0} = ck_{p0}$. Easily obtainable are the near-zone R^{-6} radiationless transfer rate due to Förster [16], and the inverse square long-range radiative transfer rate. MQED therefore provides a unified theory of resonant energy transfer, incorporating both mechanisms at the extremes of inter-chromophore separation distance, and clearly explicates the role of Maxwell fields in the process.

4. INFLUENCE OF A CHIRAL DIELECTRIC MEDIUM ON THE TRANSFER RATE

An interesting extension to the transfer rate calculated in the previous Section is to consider the effect of an homogeneous, dispersive chiral dielectric medium that is also absorptive, on the migration of energy. Such a medium is characterized by the electric permittivity, $\varepsilon(\omega)$, magnetic permeability, $\mu(\omega)$, and chirality admittance, $\beta(\omega)$ — each of which is complex, with the last quantity describing the magneto-electric response generated by the optically active molecules of the medium. The electromagnetic fields may be calculated from the time-harmonic Maxwell equations and the Drude-Born-Fedorov equations. Hence the $0p$-th matrix element of the electric displacement field in the medium is

$$\left\langle 0|d_j^{(1)}(\vec{\mu};\, \vec{r},\, t)|p\right\rangle = \frac{-1}{4\pi}\sum_{L,R}|\gamma^{L/R}(\omega_{0p})|^2|\varepsilon(\omega_{0p})|^{-3/2}|\mu(\omega_{0p})|^{-1/2}\frac{\mu_i^{p0}e^{i\omega p0t}}{\omega_{0p}\gamma^{L/R}(\omega_{0p})}V_{ij}^{med}(\omega_{0p},\, \vec{r}), \tag{10}$$

where the retarded coupling tensor in a chiral medium is

$$V_{ij}^{med}(\omega,\, \vec{r}) = (-\nabla^2\delta_{ij} + \nabla_i\nabla_j)\frac{e^{i\gamma^{L/R}(\omega)r}}{r}, \tag{11}$$

$\gamma^{L/R}(\omega) = k(\omega)/1 \mp k(\omega)\beta(\omega)$, is the wavevector length, $k^2(\omega)/\omega^2 = \varepsilon(\omega)\mu(\omega)$, and the sum in (10) is taken over the two circularly polarized modes. The energy transfer matrix element is obtained from (7), giving the rate [17]

$$\Gamma^{med} = \sum_{L,R}\frac{1}{72\pi\hbar|\varepsilon(\omega_{p0})|^3|\mu(\omega_{p0})|}|\vec{\mu}^{0p}(D)|^2|\vec{\mu}^{p0}(A)|^2\mathrm{Re}\left[\frac{\overline{\gamma}^{L/R}(\omega_{p0})}{\omega_{p0}}V_{ij}^{med}(\omega_{p0},\, \vec{R})\right]^2. \tag{12}$$

Recent work has extended this result to include the effect of a chiral medium in modifying the pure magnetic dipole and discriminatory contributions to the transfer rate [18]. Assuming no absorption due to the environment, the influence of the medium is to enhance the rate in the near-zone by a few percent relative to the rate in an achiral medium.

5. DISCRIMINATORY ENERGY TRANSFER

Resonant exchange of energy between two chiral molecules is discriminatory, depending on the handedness of the respective enantiomer. Response theory may be applied to calculate the transfer rate in a manner similar to that outlined in Section 3 for the electric dipole-dipole case. To leading order magnetic dipole coupling must be accounted for, and the previously derived [13–15] magnetic dipole dependent electric displacement and magnetic fields need to be employed, as well as the electric dipole dependent first-order magnetic field. The resulting rate is

$$\Gamma^{disc} = \frac{\rho}{18\pi\hbar\varepsilon_0^2 c^2 R^6}\left[\vec{\mu}^{0p}(D)\cdot\vec{m}^{p0}(D)\right]\left[\vec{\mu}^{0p}(A)\cdot\vec{m}^{p0}(A)\right]\left[2k_{p0}^4 R^4 + 2k_{p0}^2 R^2 + 3\right], \tag{13}$$

which changes sign when one optical isomer is replaced by its antipode.

6. CONCLUSION

MQED successfully treats long-range resonant energy transfer between molecules, both in free space and in the presence of a chiral dielectric medium that is absorptive and dispersive. Use of response theory allows the role of source dependent Maxwell field operators to clearly come to the fore when interpreting migration of electronic energy.

REFERENCES

1. Power, E. A., *Introductory Quantum Electrodynamics*, Longmans, London, 1964.
2. Healy, W. P., *Non-Relativistic Quantum Electrodynamics*, Academic Press, London, 1982.
3. Craig, D. P. and T. Thirunamachandran, *Molecular Quantum Electrodynamics*, Dover, New York, 1998.
4. Compagno, G., R. Passante, and F. Persico, *Atom-Field Interactions and Dressed Atoms*, Cambridge University Press, Cambridge, 1995.
5. Cohen-Tannoudji, C., J. Dupont-Roc, and G. Grynberg, *Photons & Atoms*, Wiley, New York, 1989.
6. Milonni, P. W., *The Quantum Vacuum*, Academic Press, San Diego, 1994.
7. Andrews, D. L. and P. Allcock, *Optical Harmonics in Molecular Systems*, Wiley-VCH, 2002.
8. Salam, A., *Molecular Quantum Electrodynamics*, Wiley, New Jersey, 2010.
9. Schwinger, J. S., *Selected Papers on Quantum Electrodynamics*, Dover, New York, 1958.
10. Power, E. A. and T. Thirunamachandran, *Phys. Rev. A*, Vol. 28, 2663, 1983.
11. Salam, A., *Int. Rev. Phys. Chem.*, Vol. 27, 405, 2008.
12. Power, E. A. and T. Thirunamachandran, *Phys. Rev. A*, Vol. 28, 2671, 1983.
13. Thirunamachandran, T., *Phys. Scr.*, Vol. T21, 123, 1988.
14. Salam, A. and T. Thirunamachandran, *Phys. Rev. A*, Vol. 50, 4755, 1994.
15. Salam, A., *J. Chem. Phys.*, Vol. 122, 044113, 2005.
16. Förster, T., *Ann. der Physik*, Vol. 6, 55, 1948.
17. Rodriguez, J. J. and A. Salam, *Chem. Phys. Lett.*, Vol. 498, 67, 2010.
18. Rodriguez, J. J. and A. Salam, *J. Phys. Chem. B*, in Press.

Raman Response of a Highly Nonlinear As_2Se_3-based Chalcogenide Photonic Crystal Fiber

A. Ben Salem, R. Cherif, and M. Zghal
Cirta'Com Laboratory, Engineering School of Communication of Tunis (Sup'Com)
University of Carthage, Ghazala Technopark, Ariana 2083, Tunisia

Abstract— We characterize the nonlinear propagation of ultra-short femtosecond pulses in a highly nonlinear As_2Se_3-based chalcogenide photonic crystal fiber (PCF). We propose an accurate fit of the Raman response function of the As_2Se_3 chalcogenide glass in order to evaluate the effect of the stimulated Raman scattering on the supercontinuum (SC) generation in the PCF. We have investigated the interplay of the nonlinear effects leading to the SC generation as a function of the input power and the fiber length. The SC spanning from 1.9 μm to 4 μm was generated in 1-cm fiber length with 200 fs pulses at an input average power of 500 μW. The generation of the SC has mainly been attributed to the Raman effect which is responsible for the solitonic fission and the dispersive wave radiation.

1. INTRODUCTION

Supercontinuum (SC) generation is a spectral broadening obtained by propagating intense short pulses in a nonlinear medium such as small core photonic crystal fiber (PCF). Silica fibers are characterized by a low nonlinear refractive index n_2 around $2.6 \cdot 10^{-20}\,\mathrm{m^2W^{-1}}$ and most of the investigations employing silica PCF require kilowatt peak powers to generate octave spanning SC from the visible to the longest wavelength which is not greater than 2 μm. For this purpose, SC generation above 2 μm wavelength requires fibers with longer infrared transmission window. Highly nonlinear glasses such as chalcogenide have been introduced to offer numerous applications in the infrared region. It has been shown that these materials exhibit many attractive optical properties and are characterized by a high refractive index, large nonlinearities and excellent transmission at the infrared wavelengths [1, 2]. Depending on their composition based on chalcogen elements, (Sulphur, Selenium, and Tellurium), the nonlinear coefficient n_2 has been measured between two to three orders of magnitude higher than for silica glass [3]. In this context, recent works have investigated the nonlinear propagation in the chalcogenide glasses and tried to determine the stimulated Raman scattering (SRS) which is known to be the major effect impacting ultra-short pulses and inducing broadband spectrum. For instance, Xiong et al. [4] have determined the temporal Raman response of the As_2S_3 material. Shaw et al. [5] have reported experimental work that demonstrates the SC generation from 2.1 μm to 3.2 μm in a 1 m section As_2Se_3-based chalcogenide PCF with one ring of air holes in a hexagonal structure.

In this work, we study the Raman response of the As_2Se_3 glass by presenting an accurate fit of the temporal Raman response based on the experimental data recently published by Hu et al. [6]. We also characterize the nonlinear propagation of ultra-short femtosecond pulses in a sub-cm length highly nonlinear As_2Se_3-based chalcogenide photonic crystal fiber and evaluate the effect of the stimulated Raman scattering on the supercontinuum generation in this PCF. The reminder of this paper is organized as follow. In Section 2, we present an optical characterization of a highly nonlinear As_2Se_3 chalcogenide PCF namely the effective mode area and the chromatic dispersion. Section 3 presents the numerical modeling of the temporal Raman response in order to solve the nonlinear propagation in the chalcogenide PCF. Therefore, SC spectra are generated in the mid infrared region in sub-centimeter length PCF. In Section 4, we discuss the interplay of the nonlinear effects and determine how the SC is constructed in the highly nonlinear As_2Se_3-based chalcogenide photonic crystal fiber.

2. CHALCOGENIDE PCF CHARACTERIZATION

In order to solve the nonlinear propagation, we determine the optical properties of an As_2Se_3-based chalcogenide PCF, having an average air hole diameter d of 1.26 μm and a distance Λ between the hole centers of 1.77 μm. We calculate the effective indices of both fundamental polarization modes using the finite element method (FEM) which is highly suited to the analysis of our periodic structure [7]. The chromatic dispersion, the effective area and the nonlinear coefficient are then deduced from the determination of the effective indices. The effective mode area of a guided mode for a PCF is given by:

$$A_{eff} = \frac{\left(\int\int_{-\infty}^{\infty} \left|\vec{E}(x,y)\right|^2 dxdy\right)^2}{\int\int_{-\infty}^{\infty} \left|\vec{E}(x,y)\right|^4 dxdy} \tag{1}$$

The integrals involving field distribution, obtained from mode analysis, are numerically evaluated to calculate the above effective mode area of the fundamental mode. Consequently, the obtained small effective areas enhance the effective fiber nonlinearity and the nonlinear coefficient γ which is given by:

$$\gamma = \frac{2\pi n_2}{\lambda A_{eff}} \tag{2}$$

We found the nonlinear coefficient of the highly nonlinear As_2Se_3-based chalcogenide photonic crystal fiber equal to $\gamma = 17.2\,\mathrm{m}^{-1}\mathrm{W}^{-1}$ at $\lambda = 2.8\,\mu\mathrm{m}$ by using $n_2 = 2.3 \cdot 10^{-17}\,\mathrm{m}^2\mathrm{W}^{-1}$ [3]. We also focus on the second-order chromatic dispersion, also referred to as group velocity dispersion (GVD). Fig. 1 gives the spectral evolution of the group velocity dispersion parameter $\beta_2 = -\frac{\lambda^2 D_c}{2\pi c}$ and the effective mode area of the chalcogenide PCF, where c is the velocity of light in vacuum and λ is the wavelength.

We note that the zero dispersion wavelength (λ_{ZDW}) of the PCF was found to be equal to $\lambda_{ZDW} = 2.67\,\mu\mathrm{m}$ and we calculated $\beta_2 = -4.058 \cdot 10^{-2}\,\mathrm{ps}^2/\mathrm{m}$ at $\lambda = 2.8\,\mu\mathrm{m}$.

3. RAMAN RESPONSE MODELING OF THE AS_2SE_3 CHALCOGENIDE GLASS

The simulation of the nonlinear propagation is based on the resolution of the generalized nonlinear Schrödinger equation (GNLSE):

$$\frac{\partial U}{\partial z} = -\frac{\alpha}{2}U - \sum_{m\geq 2}\frac{j^{m-1}\beta_m}{m!}\frac{\partial^m U}{\partial t^m} + j\left(\gamma_0 + j\gamma_1\frac{\partial}{\partial t} - \frac{\gamma_2}{2}\frac{\partial^2}{\partial t^2}\right)\left(1 + \frac{j}{\omega_0}\frac{\partial}{\partial t}\right)\times\left(U(z,t)\int_{-\infty}^{+\infty} R(t')\left|U(z,t-t')\right|^2 dt'\right) \tag{3}$$

where $U(z,t)$ is the slowly varying envelope, α is the attenuation coefficient and β_m the higher order dispersion coefficients of the propagation constant β. The total response function $R(t)$ including the instantaneous electronic and the vibrational Raman contributions is given by:

$$R(t) = (1 - f_r)\delta(t) + f_r h_R(t) \tag{4}$$

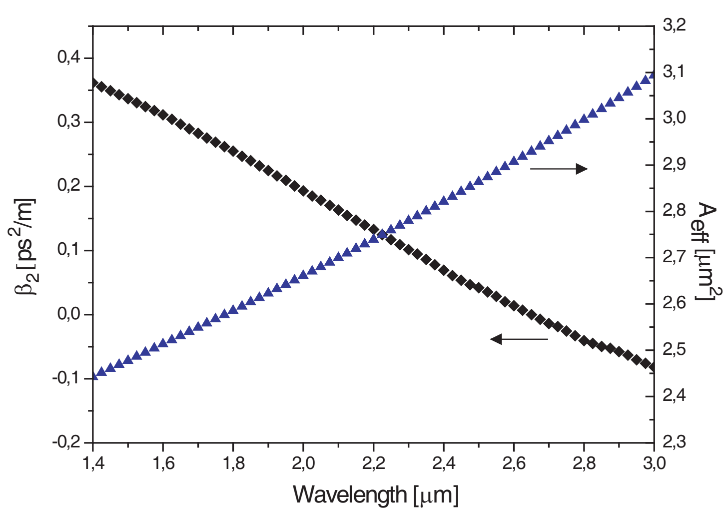


Figure 1: Group velocity dispersion and effective mode area of the As_2Se_3 PCF as a function of the wavelength.

Let us recall that the optical properties of the chalcogenide glasses are strongly affected by the photoinduced effects such as photodarkening and photoinduced anisotropy [8]. Thus, the response times and the third order susceptibility $\chi^{(3)}$ will change. It has been shown that, as a function of the irradiation time and when increasing the input power intensity, the Raman gain spectrum fluctuates and depicts oscillatory changes which disappear after reaching a final saturation state [9]. In our calculations, the temporal Raman response $h_R(t)$ was extracted from a measured Raman gain spectrum of an As_2Se_3 fiber [6] for which we consider that the input laser powers are sufficiently high to produce a stable Raman spectrum with a Stokes component centered around $230\,\mathrm{cm}^{-1}$. The temporal Raman response $h_R(t)$ is related to the Raman gain spectrum by the expression [10]:

$$g_R(f) = \frac{2\omega_p}{c} n_2 f_r Im[H_R(f)] \tag{5}$$

where ω_p is the pump frequency and $\mathrm{Im}[H_R(f)]$ is the imaginary part of the Fourier transform of $h_R(t)$. In fact, by taking the inverse Fourier transformation of the Raman gain and fitting it with a lorentzian profile, we can simply deduce the delayed Raman response $h_R(t)$ which is expressed through the Green's function of the damped harmonic oscillator [11]:

$$h_R(t) = \frac{\tau_1^2 + \tau_2^2}{\tau_1 \tau_2^2} \exp\left(-\frac{t}{\tau_2}\right) \sin\left(\frac{t}{\tau_1}\right) \tag{6}$$

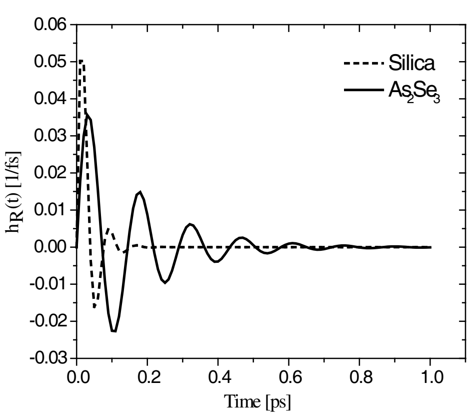
where the calculated parameters $\tau_1 = 23\,\mathrm{fs}$ and $\tau_2 = 164.5\,\mathrm{fs}$ correspond respectively to the inverse of the phonon oscillation frequency and the bandwidth of the Raman gain spectrum. They have been chosen to provide a good fit with the lorentzian spectral profile (see Fig. 2(a)). The fraction f_r can be calculated from the *Kramers-Kroning* relation [10] and has been found to be equal to 0.148.

$$f_r = \frac{\lambda}{2\pi^2 n_2} \int_0^{\infty} \frac{g_R(f)}{f} df \tag{7}$$

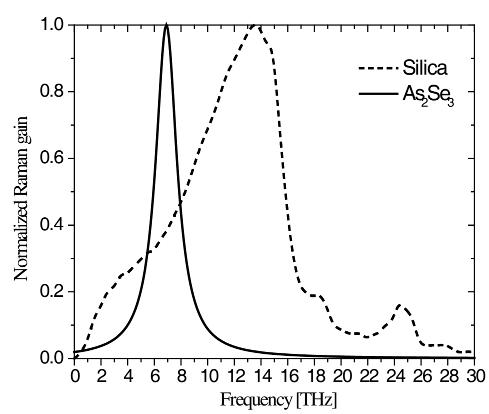
Figure 2 shows the temporal variation of the Raman response function $h_R(t)$ for the silica [11] and the As_2Se_3 chalcogenide materials and their corresponding Raman gains [6, 10]. We can observe from Fig. 2(a) that the temporal Raman response of the As_2Se_3 is longer than the response time of the silica material. Thus, the Raman response has to be considered for pulses having duration around one picosecond in opposition to the silica fibers where the time response can be neglected for pulses longer than 300 fs. The Raman gain spectra in Fig. 2(b) show that the Raman frequency shift for the silica material ($\approx 13.2\,\mathrm{THz}$) is larger than the chacogenide one ($\approx 6.9\,\mathrm{THz}$). Good agreement with the temporal Raman response published in Refs. [6, 12] is found.

4. SUPERCONTINUUM GENERATION

Equation (3) has been solved numerically using the symmetrized split-step Fourier method [13]. In our study, we consider the injection of the power laser pulses with secant hyperbolic field profile



(a)



(b)

Figure 2: (a) Temporal Raman response functions calculated for both As_2Se_3 chalcogenide (solid curve) and silica materials (dashed curve) and (b) the corresponding normalized Raman gains.

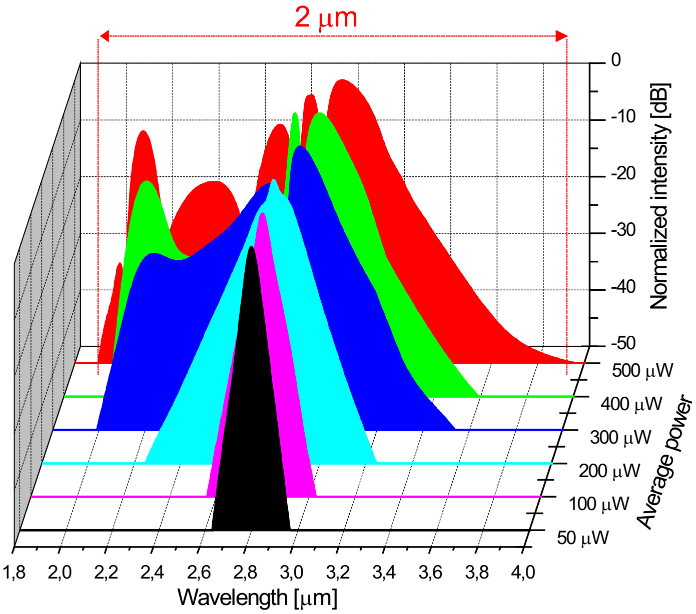


Figure 3: Spectra generated as a function of the average power in the As_2Se_3 PCF.

emerging from a femtosecond laser $U(0,T) = \sqrt{P_0}\sec h(T/T_0)$ having an average power $P = 500\,\mu\mathrm{W}$, a pulse width of 200 fs and a repetition rate of 82 MHz [14]. In our calculations we set the pump wavelength equal to 2.8 μm, close to λ_{ZDW}, in the anomalous dispersion regime in order to get broadband spectrum generation. In order to study the construction of the SC, we determined the SC generated by increasing the average power in 1-cm length of PCF, which corresponds to the fiber length giving a maximum spectral broadening. In fact, we simulated the nonlinear propagation for different PCF lengths starting from 4 mm length and found that the optimal spectral broadening is obtained for 1-cm PCF length. Fig. 3 depicts the output spectra generated for different average powers at a pump wavelength equal to 2.8 μm. We notice that the total generated bandwidth increases as the average power increases.

For low input power, a symmetric broadening is observed mainly due to the effect of self phase modulation (as shown for $P = 50\,\mu\mathrm{W}$, $P = 100\,\mu\mathrm{W}$ and $P = 200\,\mu\mathrm{W}$). When increasing the input power at $P = 300\,\mu\mathrm{W}$, asymmetry induced by the SRS effect is depicted generating the first order soliton. After that, the first soliton formed is subsequently self-frequency shifted to longer wavelengths as the pumping power is increased, giving rise to the broadband spectrum in the mid-IR region. By increasing the pump power, we show that the SC can reach more than 2 microns bandwidth extending from 1.9 μm to 4 μm. The obtained spectrum with 1-cm PCF length and with low average power reveals interesting features of the generated stable broadband covering mid-IR wavelengths.

5. CONCLUSION

We presented a numerical model of the temporal Raman response of the As_2Se_3 glass and demonstrated the SC generation in few millimeters highly nonlinear As_2Se_3-based chalcogenide PCF. We found that the stimulated Raman scattering is responsible for the extra-broadening of supercontinuum spectra. We showed that over two octaves-spanning SC from 1.9 μm to 4 μm can be easily generated by using only 1-cm fiber length and low average powers. The intrinsic properties of the As_2Se_3 glass and the microstructure provide enhanced nonlinearities that can offer numerous applications in the infrared field including ultrafast all optical switching, Raman fiber lasers and high resolution spectroscopy.

REFERENCES

1. Savage, J. A., "Optical properties of chalcogenide glasses," *J. Non-Cryst. Solids*, Vol. 47, 101–116, 1982.
2. Petkov, K. and P. J. S. Ewen, "Photoinduced changes in the linear and non-linear optical properties of chalcogenide glasses," *J. Non-Cryst. Solids*, Vol. 249, 150–159, 1999.

3. Harbold, J. M., F. O. Ilday, F. W. Wise, J. S. Sanghera, V. Q. Nguyen, L. B. Shaw, and I. D. Aggarwal, "Highly nonlinear As-S-Se glasses for all-optical switching," *Opt. Lett.*, Vol. 27, 119–121, 2002.
4. Xiong, C., E. Magi, F. Luan, A. Tuniz, S. Dekker, J. S. Sanghera, L. B. Shaw, I. D. Aggarwal and B. J. Eggleton, "Characterization of picosecond pulse nonlinear propagation in chalcogenide As_2S_3 fiber," *Appl. Opt.*, Vol. 48, 5467–5474, 2009.
5. Shaw, L. B., V. Q. Nguyen, J. S. Sanghera, I. D. Aggarwal, P. A. Thielen, and F. H. Kung, "IR supercontinuum generation in As-Se photonic crystal fiber," *Proc. Advanced Solid State Photonics*, paper TuC5, Vienna, Austria, 2005.
6. Hu, J., C. R. Menyuk, L. B. Shaw, J. S. Sanghera, and I. D. Aggarwal, "Maximizing the bandwidth of supercontinuum generation in As_2Se_3 chalcogenide fibers," *Opt. Express*, Vol. 18, 6722–6739, 2010.
7. Cherif, R., A. B. Salem, M. Zghal, P. Besnard, T. Chartier, L. Brilland, and J. Troles, "Highly nonlinear As_2Se_3-based chalcogenide photonic crystal fiber for midinfrared supercontinuum generation," *Optical Engineering*, Vol. 49, 095002-1–095002-6, 2010.
8. Abdulhalim, I., R. Beserman, and R. Weil, "Photodarkening, structural instabilities and crystallization of glassy As_2Se_3 induced by laser irradiation," *Phys. Rev. B*, Vol. 40, 12476–12486, 1989.
9. Abdulhalim, I. and R. Beserman, "Raman scattering study of laser induced structural transformations in glassy As_2Se_3," *Sol. Stat. Commun.*, Vol. 64, 951–955, 1987.
10. Kärtner, F. X., D. J. Dougherty, H. A. Haus, and E. P. Ippen, "Raman noise and soliton squeezing," *J. Opt. Soc. Am. B*, Vol. 11, 1267–1276, 1994.
11. Blow, K. J. and D. Wood, "Theoretical description of transient stimulated Raman scattering in optical fibers," *IEEE J. Quantum Electron.*, Vol. 25, 2665–2673, 1989.
12. Ung, B. and M. Skorobogatiy, "Chalcogenide microporous fibers for linear and nonlinear applications in the mid-infrared," *Opt. Express*, Vol. 18, 8647–8659, 2010.
13. Cherif, R., M. Zghal, L. Tartara, and V. Degiorgio, "Supercontinuum generation by higher-order mode excitation in a photonic crystal fiber," *Opt. Express*, Vol. 16, 2147–2152, 2008.
14. Lohner, A., P. Kruek, and W. W. Riihle, "Generation of 200 femtoseeond pulses tunable between 2.5 and 5.5 μm," *Appl. Phys. B: Photophys.Laser Chem.*, Vol. 59, 211–213, 1994.

Generation and Detection of Terahertz Radiation by Field Effect Transistors

M. I. Dyakonov
CNRS, Université Montpellier II, Place E. Bataillon, Montpellier 34095, France

Abstract— This is an overview of the main physical ideas for application of field effect transistors for generation and detection of THz radiation. Resonant frequencies of the two-dimensional plasma oscillations in FETs increase with the reduction of the channel dimensions and reach the THz range for sub-micron gate lengths. When the mobility is high enough, the dynamics of a short channel FET is dominated by plasma waves. This may result, on the one hand, in a spontaneous generation of plasma waves by a dc current and on the other hand, in a resonant response to the incoming radiation. In the opposite case, when plasma oscillations are overdamped, the FET can operate as an efficient broadband THz detector.

1. INTRODUCTION

The channel of a field effect transistor (FET) can act as a resonator for plasma waves with a typical wave velocity of 10^8 cm/s. The plasma frequency of this resonator depends on its dimensions and for gate lengths of a micron and sub-micron size can reach the Terahertz (THz) range. The interest in the THz applications of FETs was initiated at the beginning of '90s by the theoretical work of Dyakonov and Shur [1] who predicted that a steady current flow in an asymmetric FET channel can lead to instability against spontaneous generation of plasma waves. This will, in turn, produce the emission of electromagnetic radiation at the plasma wave frequency. Later, it was shown [2] that the nonlinear properties of the 2D plasma in the transistor channel can be used for detection and mixing of THz radiation. The resonant case of high electron mobility, when plasma oscillation modes are excited in the channel, and the non-resonant case of low mobility, where plasma oscillations are overdamped were analysed.

Both THz emission [3–6] and detection, resonant [7–9] and non-resonant [10, 11], were observed experimentally at cryogenic, as well as at room temperatures, clearly demonstrating effects related to the excitation of plasma waves. At the moment, the most promising application appears to be the broadband THz detection and imaging in the overdamped regime, where plasma waves are non-existent. However, THz emission and resonant detection by excitation of plasma waves are also quite interesting phenomena that deserve further exploration.

2. PLASMA WAVES IN A FET

Plasma waves are oscillations of the electron density. Generally, they can be obtained from the continuity equation:

$$\frac{\partial \rho}{\partial t} + \mathrm{div}\mathbf{j} = 0, \tag{1}$$

where ρ is the charge density and $\mathbf{j}$ is the current density, related to the local electric field $\mathbf{E}$ by Ohm's law $\mathbf{j} = \sigma \mathbf{E}$ where σ is the conductivity. These equations must be complemented by the relation between the electric field and the charge density. In three dimensions, this relation obviously is $\mathrm{div}\mathbf{E} = 4\pi\rho/\varepsilon$, where ε is the background dielectric constant (we use Gaussian units everywhere). In two and one dimensional structures, while this equation obviously remains true, it does not help because now the field $\mathbf{E}$ is *not* the total electric field, but rather its component that can drive the current, e.g., for a two-dimensional electrons it is the part of electric field that lies in the 2D plane.

It should be taken into account that plasma waves exist in the high-frequency limit $\omega\tau > 1$, where ω is the frequency, and τ is the momentum relaxation time, which is also the damping time for plasma waves. Accordingly, if damping is completely ignored, we should use the high frequency limit for the complex conductivity: $\sigma(\omega) = ine^2/m\omega$, where n is the electron concentration, e and m are the electron charge and effective mass respectively. This formula can be derived by writing the Drude equation for the mean electron velocity as $\partial \mathbf{v}/\partial t = e\mathbf{E}/m$ and neglecting the "friction" term $\square\mathbf{v}/\tau$. Equivalently, one can use the following equation for the current density $\mathbf{j} = en\mathbf{v}$:

$$\frac{\partial \mathbf{j}}{\partial t} = \frac{ne^2}{m}\mathbf{E}. \tag{2}$$

Combining Eqs. (1) and (2), we obtain

$$\frac{\partial^2 \rho}{\partial t^2} + \frac{ne^2}{m}\mathrm{div}\mathbf{E} = 0. \tag{3}$$

The relation between the charge density and the electric field in the channel is readily obtained from the plane capacitor formula:

$$\rho = en = CU, \tag{4}$$

where C is the gate-to-channel capacitance per unit area, and U is the so called gate voltage swing: $U = V_g - V_{th}$, V_g is the gate voltage, and V_{th} is the threshold voltage, at which the channel becomes completely depleted. From Eq. (4) we obtain: $\mathbf{E} = -\nabla U = -\frac{1}{C}\nabla\rho$.

It is important to understand that Eq. (4) holds not only when U is a constant, but also when the scale of the spatial variation of U is large compared to the gate-to-channel separation (the graduate channel approximation). Eq. (3) now gives a linear dispersion relation for plasma waves with velocity s:

$$\omega(k) = sk, \quad s = \sqrt{\frac{ne^2}{mC}} = \sqrt{\frac{eU_0}{m}} \tag{5}$$

where U_0 the dc part of the gate voltage swing related to the electron concentration n by Eq. (4).

The above considerations are based on Eq. (4), which is valid when the wavelength is much greater than the gate-to-channel separation d, or $kd \ll 1$.

3. INSTABILITY OF THE STEADY STATE WITH A DC CURRENT IN FET

This instability was predicted in Ref. [1]. The conditions for instability are:

A) The plasma wave damping is small, $\omega\tau > 1$, where ω is the plasma oscillation frequency on the order of s/L, L is the channel length, and τ is the momentum relaxation time defining the electron mobility. B) The boundary conditions at the source and the drain are asymmetric. An extreme case of such asymmetry, considered in Ref. [1], consists in the open circuit condition at the source and the short circuit condition at the drain. C) The steady state electron drift velocity v must exceed a threshold value depending on the damping time τ. For $\omega\tau \gg 1$, the threshold value of the drift velocity is *much smaller* than the plasma wave velocity s.

The physical origin of this instability is related to the difference in velocities of plasma waves propagating upstream $(s - v)$ and downstream $(s + v)$. Because of this difference, the reflection coefficients at the boundaries may be *greater than 1*. It can be shown that for the boundary conditions mentioned above the net amplification due to plasma wave reflections during a round trip is equal to $(s + v)/(s - v)$. The time t_0 needed to make a round trip is obviously $t_0 = L/(s+v) + L/(s-v)$. For $t \gg t_0$ the number of round trips can be estimated as t/t_0, and the total increase of the plasma wave amplitude during time t is $[(s + v)/(s - v)]^{t/t_0}$. We can now rewrite this expression as $\exp(\gamma t)$, where the instability increment γ is given by the formula:

$$\gamma = \frac{s}{2L}\left(1 - \frac{v^2}{s^2}\right)\ln\left(\frac{s+v}{s-v}\right). \tag{6}$$

This is the result obtained in Ref. [1] by the standard method of studying what happens to small perturbations of the steady state with a given drift velocity v. For low drift velocities, $v \ll s$, Eq. (4) reduces to $\gamma = v/L$. In the absence of damping, the steady state is always, however if damping is taken into account, the instability occurs when $\gamma > 1/\tau$ and this condition defines the threshold value of the drift velocity. If τ is small enough the steady state is stable.

The instability of the current-carrying steady state results in generation of plasma waves at the resonator modes. Essentially, the device operates like a laser, with an interesting difference: contrary to what happens in a laser, the gain is due to amplification *during reflections* from the "mirrors", while the losses occur during the propagation of the plasma wave between the mirrors.

The experimental results on THz emission from FETs in Refs. [3–6] and other work cannot be directly compared with the one-dimensional theory [1]. In the standard experimental situation, the width W of the gate is much larger than the gate length L, typically $W/L \sim 100$. Under such conditions the one-dimensional model, where the plasma density and velocity depend on the coordinate x only, is not appropriate, since obviously oblique plasma waves with a non-zero component of the wave vector in the y direction can propagate. In such geometry, the gated region is not a resonator, but rather a waveguide with a continuous spectrum of plasma waves.

In Ref. [12], the analysis of stability was extended to the more realistic case when $W \gg L$, and it was shown that, somewhat unexpectedly, in such a geometry an additional new mode of instability dominates, which is localized near the gate boundaries. Moreover, a similar instability should exist near a single boundary of current-carrying two-dimensional plasma.

Certainly, the linear theory cannot predict the outcome of this instability. However, since the spectrum of plasma waves is continuous, it seems likely that the instability will result in a turbulent motion of the electron fluid near the boundary of the gated region. The spectrum of the plasma oscillations should be broad, as it is observed in experiments. This is similar to what one can see in a river, when the water flows with sufficient velocity across an abrupt step in the waterbed: waves with wave vectors perpendicular to the flow are excited, while the wave vectors in the direction of the flow are purely imaginary, which accounts for the localization of the turbulent region near the step. It would be interesting to verify these predictions in specially designed experiments.

4. DETECTION OF THZ RADIATION BY FET

The idea of using a FET for detection of THz radiation was put forward in Ref. [2]. The possibility of the detection is due to nonlinear properties of the transistor, which lead to the rectification of an ac current induced by the incoming radiation. As a result, a photoresponse appears in the form of dc voltage between source and drain which is proportional to the radiation power (photovoltaic effect). Obviously, some asymmetry between the source and drain is needed to induce such a voltage. There may be various reasons of such an asymmetry. One of them is the difference in the source and drain boundary conditions due to some external (parasitic) capacitances. Another one is the asymmetry in feeding the incoming radiation, which can be achieved either by using a special antenna, or by an asymmetric design of the source and drain contact pads. Thus the radiation may predominantly create an ac voltage between the source and the gate (or between the drain and the gate) pair of contacts. Finally, the asymmetry can naturally arise if a dc current is passed between source and drain, creating a depletion of the electron density on the drain side of the channel.

In most of the experiments carried out so far, the THz radiation was applied to the transistor channel, together with contact pads and bonding wires. In such a case, it is obviously difficult to define how exactly the radiation is coupled to the transistor. Theoretically, we will consider the case of an extreme asymmetry, where the incoming radiation creates an ac voltage with amplitude U_a only between the source and the gate, see Fig. 1. We will also assume that there is no dc current between the source and drain. Generally, the FET may be described by an equivalent circuit presented in Fig. 1. The obvious elements are the distributed gate-to-channel capacitance and the channel resistance, which depends on the gate voltage through the electron concentration in the channel, according to Eq. (4).

As mentioned above, this equation is valid locally, so long as the scale of the spatial variation of $U(x)$ is larger than the gate-to-channel separation d (the gradual channel approximation). Under

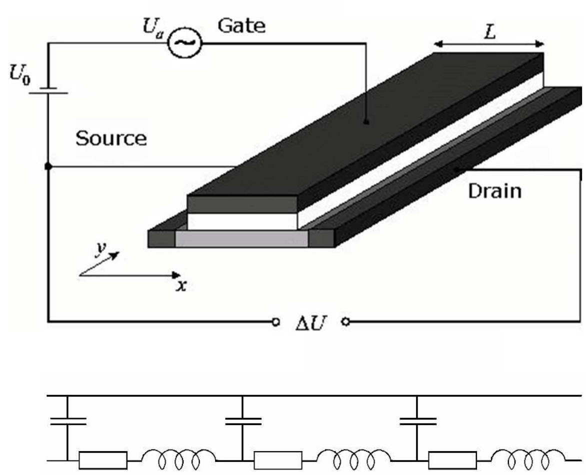


Figure 1: Schematics of a FET as a THz detector (above) and the equivalent circuit (below).

static conditions and in the absence of the drain current, $U = U_0 = V_g - V_{th}$, where U_0 is the static voltage swing. The inductances in Fig. 1 represent the so-called *kinetic* inductances, which are due to the electron inertia and are proportional to m, the electron effective mass. Depending on the frequency ω, one can distinguish two regimes of operation, and each of them can be further divided into two sub-regimes depending on the gate length L.

1. High frequency regime occurs when $\omega\tau > 1$, where τ is the electron momentum relaxation time, determining the conductivity in the channel $\sigma = ne^2\tau/m$. In this case, the kinetic inductances in Fig. 2 are of primordial importance, and the plasma waves analogous to the waves in an RLC transmission line, will be excited. The plasma waves have a velocity $s = (eU_0/m)^{1/2}$ and a damping time τ Thus their propagation distance is $s\tau$.

1*a*. *Short gate, $L < s\tau$.* The plasma wave reaches the drain side of the channel, gets reflected, and forms a standing wave with enhanced amplitude, so that the channel serves as a high-quality resonator for plasma oscillations. The fundamental mode has the frequency $\sim s/L$, with a numerical coefficient depending on the boundary conditions.

1*b*. *Long gate, $L \gg s\tau$.* The plasma waves excited at the source will decay before reaching the drain, so that the ac current will exist only in a small part of the channel adjacent to the source.

2. Low frequency regime, $\omega\tau \ll 1$. Now, the plasma waves cannot exist because of over-damping. At these low frequencies, the inductance in Fig. 1 become simply short-circuits which leads to an RC line. Its properties further depend the gate length, the relevant parameter being $\omega\tau_{RC}$, where τ_{RC} is the RC time constant of the whole transistor. Since the total channel resistance is $L\rho/W$, and the total capacitance is CWL (where W is the gate width and $\rho = 1/\sigma$ is the channel resistivity), one finds $\tau_{RC} = L^2\rho C$.

2*a*. *Short gate, $L < (\rho C\omega)^{1/2}$.* This means that $\omega\tau_{RC} < 1$, so that the ac current goes through the gate-to-channel capacitance practically uniformly on the whole length of the gate. This is the so-called "resistive mixer" regime [20–22]. For the THz frequencies this regime can apply only for transistors with extremely short gates smaller than 70 nm at 1 THz in silicon).

2*b*. *Long gate, $L \gg (\rho C\omega)^{1/2}$.* Now $\omega\tau_{RC} \gg 1$, and the induced ac current will leak to the gate at a small distance l from the source, such that the resistance $R(l)$ and the capacitance $C(l)$ of this piece of the transistor channel satisfy the condition $\omega\tau_{RC}(l) = 1$, where $\tau_{RC}(l) = R(l)C(l) = l^2\rho C$. This condition gives the value of the "leakage length" l on the order of $(\rho C\omega)^{1/2}$ (which can also be rewritten as $s(\tau/\omega)^{1/2}$). If $l \ll L$, then neither ac voltage, nor ac current will exist in the channel at distances beyond l from the source.

Thus, the characteristic length where the ac current exists is $s\tau$ for $\omega\tau > 1$, and $s(\tau/\omega)^{1/2}$ for $\omega\tau < 1$ [2]. If the conditions of the case 1*a* are satisfied, the photoresponse will be resonant, corresponding to the excitation of discrete plasma oscillation modes in the channel. Otherwise, the FET will operate as a broad-band detector.

For a long gate, there is no qualitative difference between the low-frequency regime ($\omega\tau \ll 1$), when plasma waves do not exist (the case 2*b*) and the high frequency regime ($\omega\tau \gg 1$), where plasma oscillations are excited (the case 1*b*). However, and their excitation of plasma waves in the case 1*b* has been clearly confirmed by recent detection experiments in magnetic field [13]. Plasma waves cannot propagate below the cyclotron frequency. Therefore, in experiments with a fixed radiation frequency the photoresponse is strongly reduced when the magnetic field goes through the cyclotron resonance [14]. This is probably the most spectacular manifestation of the importance of plasma waves in THz detection by FETs.

The maximal photovoltage is achieved at $U_0 \approx 0$, where the relative ac modulation of the electron concentration in the channel is the strongest (note that Eq. (4) is not valid in the near vicinity of $U_0 = 0$). A theoretical study of the photoresponse in this region is presented in Ref. [15].

REFERENCES

1. Dyakonov, M. I. and M. S. Shur, *Phys. Rev. Lett.*, Vol. 71, 2465, 1993.
2. Dyakonov, M. I. and M. S. Shur, *IEEE Transactions on Electron Devices*, Vol. 43, 380, 1996.
3. Knap, W., J. Łusakowski, T. Parenty, S. Bollaert, A. Cappy, and M. S. Shur, *Appl. Phys. Lett.*, Vol. 84, 2331, 2004.
4. Dyakonova, N., F. Teppe, J. Łusakowski, W. Knap, M. Levinshtein, A. P. Dmitriev, M. S. Shur, S. Bollaert, and A. Cappy, *J. Appl. Phys.*, Vol. 97, 114313, 2005.
5. Dyakonova, N., A. El Fatimy, J. Łusakowski, W. Knap, M. I. Dyakonov, M. A. Poisson, E. Morvan, S. Bollaert, A. Shchepetov, Y. Roelens, C. Gaquiere, D. Theron, and A. Cappy, *Appl. Phys. Lett.*, Vol. 88, 141906, 2006.

6. El Fatimy, A., N. Dyakonova, Y. Meziani, T. Otsuji, W. Knap, S. Vandenbrouk, K. Madjour, D. Théron, C. Gaquiere, M. A. Poisson, S. Delage, P. Prystawko, and C. Skierbiszewski, *J. Appl. Phys.*, Vol. 107, 024504, 2010.
7. Lu, J.-Q., M. S. Shur, J. L. Hesler, L. Sun, and R. Weikle, *IEEE Electron Device Lett.*, Vol. 19, 373, 1998.
8. Lu, J.-Q. and M. S. Shur, *Appl. Phys. Lett.*, Vol. 78, 2587, 2001.
9. Knap, W., Y. Deng, S. Rumyantsev, and M. S. Shur, *Appl. Phys. Lett.*, Vol. 81, 4637, 2002.
10. El Fatimy, A., F. Teppe, N. Dyakonova, W. Knap, D. Seliuta, G. Valusis, A. Shchepetov, Y. Roelens, S. Bollaert, A. Cappy, and S. Rumyantsev, *Appl. Phys. Lett.*, Vol. 89, 131926, 2006.
11. El Fatimy, A., N. Dyakonova, F. Teppe, W. Knap, N. Pala, R. Gaska, Q. Fareed, X. Hu, D. B. Veksler, S. Rumyantsev, M. S. Shur, D. Seliuta, G. Valusis, S. Bollaert, A. Shchepetov, Y. Roelens, C. Gaquiere, D. Theron, and A. Cappy, *Electron. Lett.*, Vol. 42, 1342, 2006.
12. Dyakonov, M. I., *Semiconductors*, Vol. 42, 984, 2008.
13. Boubanga-Tombet, S., M. Sakowicz, D. Coquillat, F. Teppe, W. Knap, M. I. Dyakonov, K. Karpierz, J. Łusakowski, and M. Grynberg, *Appl. Phys. Lett.*, Vol. 95, 072106, 2009.
14. Lifshits, M. B. and M. I. Dyakonov, *Phys. Rev. B*, Vol. 80, 121304(R), 2009.
15. Knap, W., V. Kachorovskii, Y. Deng, S. Rumyantsev, J.-Q. Lv, R. Gaska, M. S. Shur, G. Simin, X. Hu, M. Asif Khan, C. A. Saylor, and L. C. Brunel, *J. Appl. Phys.*, Vol. 91, 9346, 2002.

Terahertz Emission, Detection and Modulation Using Two-dimensional Plasmons in High-electron-mobility Transistors Featured by a Dual-grating-gate Structure

Taiichi Otsuji
Research Institute of Electrical Communication, Tohoku University, Sendai, Japan

Abstract— This paper reviews recent advances in emission, detection and modulation of terahertz (THz) radiation using two-dimensional plasmons in high-electron-mobility transistors (HEMMTs) featured by a dual-grating-gate (DGG) structure. The dual grating gates can alternately modulate the 2D electron densities to periodically distribute the plasmonic cavities along the channel, acting as a broadband antenna. The sample was fabricated with standard GaAs- or InP-based heterostructure material systems, succeeding in emission and detection of broadband (0.5–6.5 THz) radiation at room temperature with a maximum available THz output power and detection responsivity of $\sim \mu$W and $\sim 10^{-10}$ W/Hz$^{0.5}$ from a $30 \times 75\,\mu\text{m}^2$ single die, respectively, with an excellent power conversion efficiency of 10^{-4}. The device also can work for intensity modulation of incident THz radiation. We also demonstrate that the DGG-HEMT can work for a fast broadband THz intensity modulator.

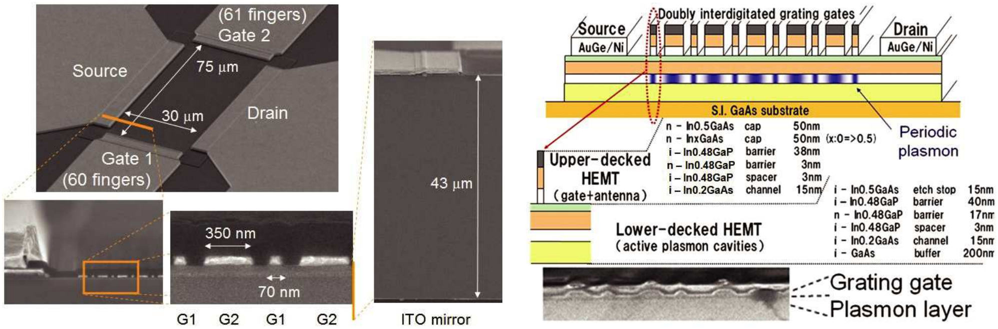
1. INTRODUCTION

In the research of modern terahertz (THz) electronics, development of compact, tunable and coherent sources and detectors operating at THz frequencies is one of the hottest issues [1]. Two-dimensional plasmons (2DPs) in submicron transistors have attracted much attention due to their nature of promoting emission/detection of electromagnetic radiation in the THz range [2, 3]. Therefore different devices/structures of micron and submicron sizes supporting low-dimensional plasmons were intensively studied as possible candidates for solid-state far-infrared (FIR)/THz sources [4–9] and detectors [10–16]. Mechanisms of plasma wave excitation/emission can be divided (by convention) into two types: (i) incoherent and (ii) coherent type. The first is related to thermal excitation of broadband nonresonant plasmons by hot electrons [4–6]. The second is related to the plasma wave instability mechanisms like Dyakonov–Shur Doppler-shift model [2] and/or Ryzhii-Satou-Shur transit-time model [17, 18], where coherent plasmons can be excited either by hot electrons or by optical phonon emission under near ballistic electron motion [19]. It is known that hydrodynamic nonlinearities in 2DPs in the HEMT channel can be used for detection of THz radiation [3]. Resonant THz detection [12–14] is related to the excitation of high quality plasmons in 2D electron channel, whereas a broadband non-resonant THz detection [11] is related to over damped plasmons in depleted regions of the channel. The metal grating coupler is a conventional tool that can ensure strong coupling between the plasmons in 2D electron channel and THz radiation [4–6]. THz imaging utilizing 2DP detectors has been recently demonstrated [20–22], attracting great interest on safety/security applications and even more on future wireless communications. When the frequency of THz incoming radiation is resonant (off-resonant) to the 2DP, it will be well absorbed in (transmitted through) the HEMT channel. Thus one can control the transmissivity of THz radiation by the gate bias, leading to an intensity modulation function. Numerical simulations demonstrate a fast and broadband THz intensity modulation with a practically acceptable extinction ratio [23].

This paper reviews recent advances in emission, detection and modulation of terahertz (THz) radiation using two-dimensional plasmons in high-electron-mobility transistors (HEMMTs) featured by a dual-grating-gate (DGG) structure [23–31].

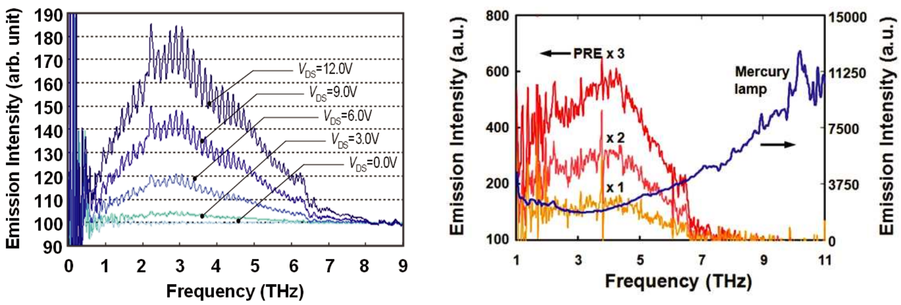
2. PLASMON-RESONANT THz EMITTER

The device was fabricated with InGaP/InGaAs/GaAs material systems in a double-deck HEMT with semiconducting 2D electron gas (2DEG) DGG [24–31]. The schematic device cross section and its SEM image are shown in Fig. 1. The device structure is based on a HEMT and incorporates (i) interdigitated DGGs (G1 and G2) that periodically localize the 2D plasmon in stripes on the order of 100 nm with a micron-to-submicron interval and (ii) a vertical cavity structure in between the top grating plane and a THz mirror at the backside. The structure (i) works as a THz antenna and (ii) works as an amplifier. When the DC drain-to-source bias VDS is applied,



(a) (b)

Figure 1: SEM images of a fabricated (a) metal grating-gate and (b) semiconducting grating-gate plasmon-resonant emitters (after Ref. [28]).



(a) (b)

Figure 2: (a) FTIR measured emission spectra for a DGG-HEMT $L_{g1}/L_{g2} = 150\,\text{nm}/1850\,\text{nm}$ at room temperature (after Ref. [26].)). (b) DGG-HEMT vs. HP-Hg lamp. Double- and havingtriple-sample chip DGG-HEMT operation almost doubles and triples the emission power, respectively (after Ref. [30]).

2D electrons are accelerated to produce a constant drain-to-source current IDS. Due to such a distributed plasmonic cavity systems in periodic 2D electron-density modulation, the DC current flow may excite the plasma waves in each plasmonic cavity. Asymmetric cavity boundaries make plasma-wave reflections as well as abrupt change in the density and the drift velocity of electrons, which may cause the current-driven plasmon instability [2, 9, 17, 18] leading to excitation of coherent resonant plasmons. Thermally excited hot electrons also may excite incoherent plasmons [6]. The DGG act also as a THz antenna that converts non-radiative longitudinal plasmon modes to radiative transverse electromagnetic modes [4–6, 9, 24]. The 2D plasmon layer is formed with a lower-deck quantum well at the heterointerface between a 15-nm thick InGaAs channel layer and a 60-nm thick, InGaP carrier-supplying layer. The upper-deck InGaAs channel, serving as the dual grating-gate electrodes, periodically etched to form the uncapped region where the 2DEG concentration becomes lower than the capped region without any external gate bias. The intrinsic device area has geometry of $30\,\mu\text{m} \times 75\,\mu\text{m}$. The DGG consists of 150-nm lines and 1850-nm lines aligned alternately with a spacing of 100 nm. The number of fingers is 37 (38) for the 1500-nm (1850-nm) grating. The substrate thickness was 260 μm, corresponding to the fundamental vertical cavity resonance of 80 GHz and the odd harmonics with a 160-GHz spacing.

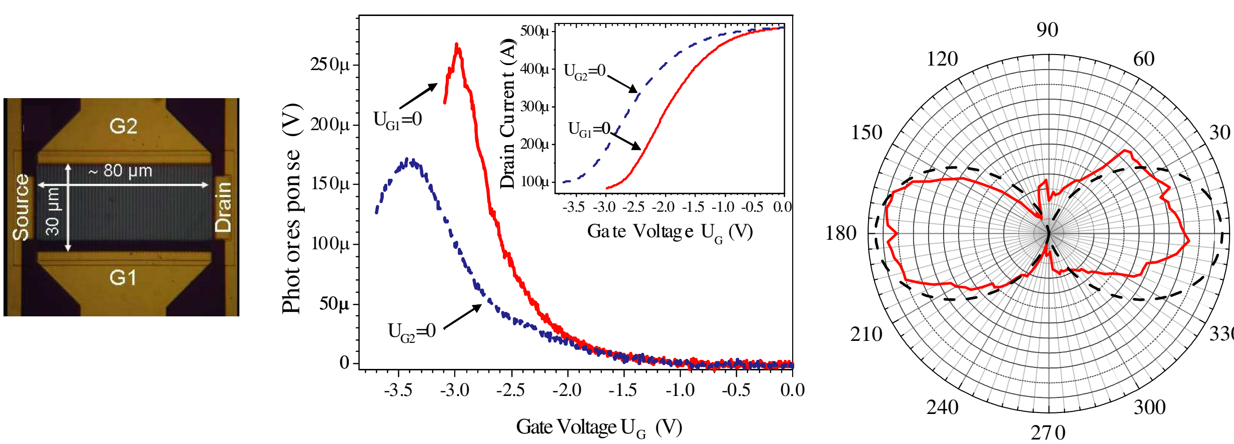
The field emission properties of the fabricated samples were measured by a Fourier transform infrared spectrometer (FTIR) at room temperature. The detector was a 4.2 K-cooled Si bolometer. Typical measured emission spectra for a sample having grating finger sizes (L_{g1}, L_{g2}) of $L_{g1}/L_{g2} = 150\,\text{nm}/1850\,\text{nm}$ for different drain-bias conditions are shown in Fig. 2(a). One can see relatively broad spectra starting from 0.5 THz with maxima around 3.0 THz. The emission dies off abruptly around 6.5 THz, which is thought to be due to the Reststrahlen band of optical phonon modes of the InGaAs channel [27]. The emission intensity is not linear function of the applied VDS bias, thus the drain current, but close to a quadratic function with a threshold property [29]. The maximum emission power at $V_{DS} = 12.0\,\text{V}$ is estimated to be on the order of 1 μW at 300 K. Taking account of the monitored power consumption of the order of 100 mW, the energy conversion efficiency (from DC to THz) is on the order of 10^{-5}. Idealistic coherent plasmon modes originated from Dyakonov-Shur and/or transit-time-driven instabilities [2, 17, 18] should result in sharp emission peaks on the spectra. The realistic operating condition of the device electrically biased at room temperature, however, produces additional spectral broadening effects. Thermally excited incoherent plasmons [27] and dispersion of the plasmon-resonant frequency depending on the drain bias potential [28, 29] contribute to broaden the emission spectra. The fine spectral profile exhibits a Fabry-Perot longitudinal-mode with 160-GHz spaces.

The emission spectral profile of the fabricated device was compared to that of a standard water-cooled high-pressure mercury (HP-Hg) lamp used as a THz light source in FTIR systems [29, 30]. Fig. 2(b) plots the typical results [30]. It is clearly seen that the main lobe of the emission spectra of the fabricated device stays around 1 to 6 THz, which is far apart from and lower than that for the HP-Hg lamp. The emission spectrum of the HP-Hg lamp traces the black-body radiation due to its nature of thermally heated emission so that the emission at lower THz region is substantially weakened. The DGG-HEMT exhibits a superior (two orders higher) energy conversion efficiency of 10^{-4} to 10^{-5} although the output power is by one order lower than the HP-Hg lamp. We confirmed dual- and triple-chip operation of the DGG-HEMTs successfully doubles and triples the emission intensity as shown in Fig. 2(b). The results encourage us to propose the source that can exceed the output power of the HP-Hg lamp by integrating the DGG-HEMTs in an array of 10×10.

3. PLASMON-RESONANT THz DETECTOR

The device structure is equivalent to the emitter based on InGaP/InGaAs/GaAs single-deck DGG-HEMT (see Fig. 3(a)) [43]. The DGGs having $L_{G1}/L_{G2} = 100/300$, 100/300, 100/1300, 100/1300 in nm are prepared [31]. The spacing between the DGG fingers is 100 nm for all four structures. Each DGG structure was irradiated at normal incidence by THz beam at frequency 0.24 THz. We measured a photovoltaic DC response between the source and drain contacts (the source contact was grounded and no drain bias DC voltage was applied) at room temperature. The electron momentum relaxation time in 2D electron channel at $T = 300\,\mathrm{K}$ is $\tau_m = \mu m^*/e = 0.17\,\mathrm{ps}$, which yields the quality factor smaller than unity, $\omega\tau = 0.25$, for the incident radiation frequency $\omega/2\pi = 0.24\,\mathrm{THz}$. Therefore, the non-resonant detection should be expected in our experiments at room temperature. Terahertz radiation was generated by a Backward Wave Oscillator (BWO) with maximum output power of 5 mW.The radiation was linearly polarizedand chopped at 170 Hz. The radiation-induced photo voltage ΔU was measured as a function of the gate voltage U_{G1} or/and U_{G2} as 170 Hz ac voltage component of the source-to-drain voltage using the standard lock-in technique.

Figure 3(b) demonstrates the photo response measured as a function of the gate voltage U_{G1} (for $U_{G2} = 0$) and that as a function of the gate voltage U_{G2} (for $U_{G1} = 0$) for the structure with



(a) (b) (c)

Figure 3: (a) Top view of the DGG-HEMT. The external THz radiation is incident normally from the top. (b) The photo response at 0.24 THz in the structure with $L_{G2} = 300\,\mathrm{nm}$ as a function of the gate voltage U_{G1} (for $U_{G2} = 0$) and as a function of the gate voltage U_{G2} (for $U_{G1=0}$). The inset shows the transfer characteristics of the DGG-HEMT. (c) Photo response at 0.24 THz as a function of the azimuthal angle φ between the electric field vector of the incident THz wave and the source-to-drain direction (solid curve). It follows the $\cos 2(\varphi)$ dependence (dashed curve) with the strongest photo response occurred for $\varphi = 0$ (after Ref. [31]).

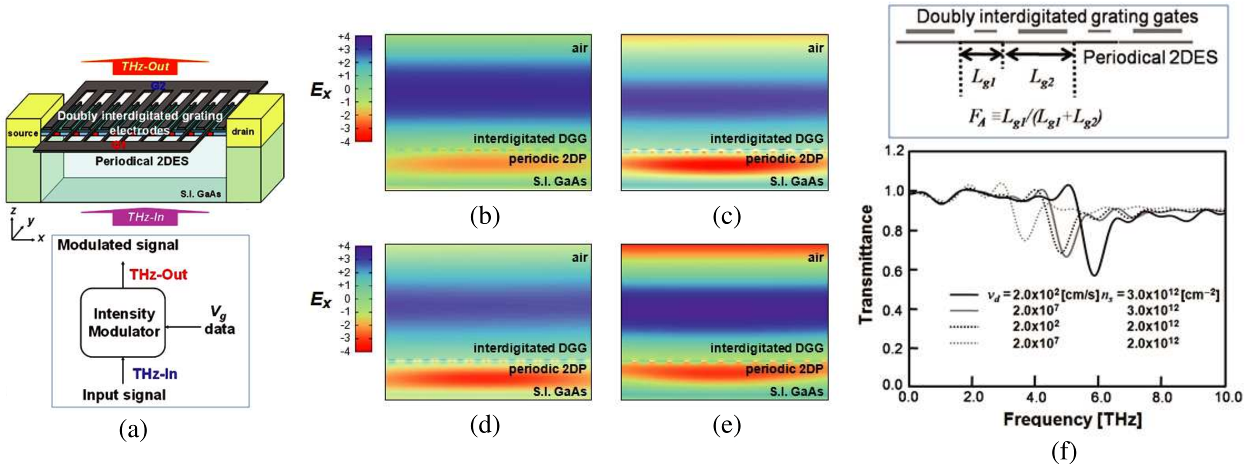


Figure 4: (a) Schematic of intensity modulator using a DGG-HEMT. (b)–(e) 2D electric field intensity distributions (x component) for different electron drift velocity v_d conditions under a constant electron density n_s of $3 \times 10^{12}\,\mathrm{cm}^{-2}$. (b) $v_d = 2 \times 10^2\,\mathrm{cm/s}$, off-resonant at 4.8 THz, (c) $v_d = 2 \times 10^2\,\mathrm{cm/s}$, resonant at 5.7 THz, (d) $v_d = 2 \times 10^7\,\mathrm{cm/s}$ resonant at 4.8 THz, (e) $v_d = 2 \times 10^7\,\mathrm{cm/s}$ off-resonant at 5.7 THz. (f) Transmittance vs. radiation frequency for various vd/ns conditions. Area factor defined by the equation in the inset of (f) is fixed at 0.2 for all the results (after Ref. [23]).

$L_{G2} = 300\,\mathrm{nm}$ when the THz electric field was directed across the grating-gate fingers (i.e., in the source-to-drain direction) [16, 31]. The photo response grows when one of the grating gates, either $G1$ or $G2$, is biased to the threshold voltage (seethe inset in Fig. 3(b)). Fig. 3(c) shows the dependence of the photo response on the azimuthal angle φ between the THz electric field and the source-to-drain direction for the structure with $L_{G2} = 1300\,\mathrm{nm}$ [31]. A two-lobe-shape angle dependence that relatively well follows the $\cos^2\varphi$ law, was obtained with the maximal photo response for $\varphi = 0$. This demonstrates that the coupling between THz radiation and 2D electron channel is ensured by the DGG structure.

4. PLASMON-RESONANT THz INTENSITY MODULATION

One can consider a different operation scheme of the DGG-HEMT for use in an intensity modulator where the device is subject to THz EM wave irradiation from one side and manages its transmissivity to the other side by controlling the gate/drain biases as shown in Fig. 4(a) [23]. Since the thickness of the DGG structure is much smaller than the THz wavelength, such structure can be modeled as a 2D resonant layer whose transmissivity, reflectivity and absorbance of THz radiation coupled with the n-th 2DP is formulated in Ref. [32]. If the frequency of incident THz EM waves matches to the 2DP resonant frequency, the incoming THz waves are well absorbed to excite the 2DPs. In such resonant conditions, the absorbance as well as the reflectivity becomes maximal whereas the transmissivity becomes minimal. If not, the THz waves are transmitted through the system. The 2DP resonant frequency is tunable by means of electrical control of the device parameters like the sheet electron densities and/or the electron drift velocity in 2DP cavities. This means the transmissivity could be alternately controllable for a narrow-band incoming THz wave whose bandwidth is well limited within the 2DP resonant bandwidth. Therefore the DGG-HEMT device can serve as an intensity modulator for THz waves.

We numerically analyzed the excitation of 2DPs by the THz radiation and resultant field distribution profiles as functions of the densities and the drift velocity of electrons in 2DPs. Typical results are shown in Figs. 4(b)–(f) [23]. It is revealed that the transmittance is deeply dominated by the coupling/excitation of the 2DPs and that the condition is controlled by the 2D plasmon density and the electron drift velocity. The effective area responsible for the electromagnetic coupling of the 2DPs and THz waves is defined by the area factor $F_A = L_{g1}/(L_{g1} + L_{g2})$. All the results shown in Fig. 4 are for the case of $F_A = 0.2$. With increasing F_A value up to 0.7 the extinction ratio increases from ~ 0.25 to 0.8, which will be applicable to the data coding (on/off keying) for the THz wireless communications.

5. CONCLUSION

Recent advances in emission, detection and intensity modulation of THz radiation using 2DPs in HEMTs featured by a DGG structure were reviewed. These three distinctive functionalities will make the 2DP DGG-HEMTs to be possible transmitter/receiver front-end key devices for THz wireless communication systems.

ACKNOWLEDGMENT

The authors thank Profs. E. Sano, W. Knap, V. V. Popov, V. Ryzhii, Drs. F. Teppe, D. Coquillat, V. Ryzhii, A. Satou, Y. Tsuda, T. Nishimura, T. Komori for their contributions. They also thank Prof. M. Dyakonov and M. S. Shur for valuable discussion. This work was supported in part by the JST-ANR WITH and the JSPS-SAKURA programs.

REFERENCES

1. Tonouchi, M., *Nature Photon*, Vol. 1, 97–105, 2007.
2. Dyakonov, M. and M. Shur, *Phys. Rev. Lett.*, Vol. 71, 2465–2468, 1993.
3. Dyakonov, M. and M. Shur, *IEEE Trans. Electron Devices*, Vol. 43, 1640–1645, 1996.
4. Tsui, D. C., E. Gornik, and R. A. Logan, *Solid State Comm.*, Vol. 35, 875–877, 1980.
5. Wilkinson, R. J., C. D. Ager, T. Duffield, H. P. Hughes, D. G. Hasko, H. Armed, J. E. F. Frost, D. C. Peacock, D. A. Ritchie, A. C. Jones, C. R. Whitehouse, and N. Apsley, *J. Appl. Phys.*, Vol. 71, 6049–6061, 1992.
6. Hirakawa, K., K. Yamanaka, M. Graysonm, and D. C. Tsui, *Appl. Phys. Lett.*, Vol. 67, 2326, 1995.
7. Knap, W., J. Lusakowski, T. Parenty, S. Bollaert, A. Cappy, V. V. Popov, and M. S. Shur, *Appl. Phys. Lett.*, Vol. 84, 2331, 2004.
8. Dyakonova, N., F. Teppe, J. Lusakowski, W. Knap, M. Levinshtein, A. P. Dmitriev, M. S. Shur, S. Bollaert, and A. Cappy, *J. Appl. Phys.*, Vol. 97, 114313, 2005.
9. Mikhailov, S. A., *Phys. Rev. B*, Vol. 58, 1517–1532, 1998.
10. Allen, Jr., S. J., D. S. Tsui, and R. A. Logan, *Phys. Rev. Lett.*, Vol. 38, 980–983, 1977.
11. Knap, W., V. Kachorovskii, Y. Deng, S. Rumyantsev, J.-Q. Lu, R. Gaska, M. S. Shur, G. Simin, X. Hu, M. Asif Khan, C. A. Saylor, and L. C. Brunel, *J. Appl. Phys.*, Vol. 91, 9346–9353, 2002.
12. Knap, W., Y. Deng, S. Rumyantsev, J.-Q. Lu, M. S. Shur, C. A. Saylor, and L. C. Brunel, *Appl. Phys. Lett.*, Vol. 80, 3433–3435, 2002.
13. Otsuji, T., M. Hanabe, and O. Ogawara, *Appl. Phys. Lett.*, Vol. 85, 2119, 2004.
14. El Fatimy, A., F. Teppe, N. Dyakonova, W. Knap, D. Seliuta, G. Valusis, A. Shchepetov, Y. Roelens, S. Bollaert, A. Cappy, and S. Rumyantsev, *Appl. Phys. Lett.*, Vol. 89, 131926, 2006.
15. Knap, W., F. Teppe, N. Dyakonova, D. Coquillat, and J. Sakowski, *J. Phys.: Condens. Matter*, Vol. 20, 384205, 2008.
16. Knap, W., M. Dyakonov, D. Coquillat, F. Teppe, N. Dyakonova, J. Lusakowski, K. Karpierz, M. Sakowicz, G. Valusis, D. Seliuta, I. Kasalynas, A. El Fatimy, Y. M. Meziani, and T. Otsuji, *J. Infrared Milli. Terahz. Waves*, Vol. 30, 1319–1337, 2009.
17. Ryzhii, V., A. Satou, and M. Shur, I*EICE Trans. Electron.*, Vol. E89-C, 1012–1019, 2006.
18. Ryzhii, V., A. Satou, M. Ryzhii, T. Otsuji, and M. S. Shur, *J. Phys.: Condens. Matter*, Vol. 20, 384207, 2008.
19. Ryzhij, V., N. A. Bannov, and V. A. Fedirko, *Sov. Phys. Semicond.*, Vol. 18, 481–491, 1984.
20. Lisauskas, A., W. von Spiegel, S. Boubanga-Tombet, A. El Fatimy, D. Coquillat, F. Teppe, N. Dyakonova, W. Knap, and H. G. Roskos, *Electron. Lett.*, Vol. 44, 408–409, 2008.
21. Lisauskas, A., U. Pfeiffer, E. Öjefors, P. H. Bolìvar, D. Glaab, and H. G. Roskos, *J. Appl. Phys.*, Vol. 105, 114511, 2009.
22. Öjefors, E., U. R. Pfeiffer, A. Lisauskas, and H. G. Roskos, *IEEE J. Solid-State Circuits*, Vol. 44, 1968–1976, 2009.
23. Nishimura, T., N. Magome, and T. Otsuji, *Jpn. J. Appl. Phys.*, Vol. 49, 054301-1–7, 2010.
24. Otsuji, T., M. Hanabe, T. Nishimura, and E. Sano, *Opt. Express*, Vol. 14, 4815–4825, 2006.
25. Otsuji, T., Y. M. Meziani, M. Hanabe, T. Ishibashi, T. Uno, and E. Sano, *Appl. Phys. Lett.*, Vol. 89, 263502, 2006.

26. Otsuji, T., Y. M. Meziani, M. Hanabe, T. Nishimura, and E. Sano, *Solid State Electron.*, Vol. 51, 1319–1327, 2007.
27. Meziani, Y. M., H. Handa, W. Knap. T. Otsuji, E. Sano, V. V. Popov, G. M. Tsymbalov, D. Coquillat, and F. Teppe, *Appl. Phys. Lett.*, Vol. 92, 201108, 2008.
28. Otsuji, T., Y. M. Meziani, T. Nishimura, T. Suemitsu, W. Knap, E. Sano, T. Asano, and V. V. Popov, *J. Phys.: Condens. Matter*, Vol. 20, 384206, 2008.
29. Tsuda, Y., T. Komori, A. El Fatimy, T. Suemitsu, and T. Otsuji, *J. Opt. Soc. Am. B*, Vol. 26, A52–A57, 2009.
30. Otsuji, T., *Proc. SPIE*, Vol. 7671, 767102-1–12, 2010.
31. Coquillat, D., S. Nadar, F. Teppe, N. Dyakonova, S. Boubanga-Tombet, W. Knap, T. Nishimura, T. Otsuji, Y. M. Meziani, G. M. Tsymbalov, and V. V. Popov, *Opt. Express*, Vol. 18, 6024–6032, 2010.
32. Popov, V. V., O. V. Polischuk, T. V. Teperik, X. G. Peralta, S. J. Allen, N. J. M. Horing, and M. C. Wanke, *J. Appl. Phys.*, Vol. 94, 3556–3562, 2003.

Spin Related Effect in Terahertz Photovoltaic Response of Si-MOSFETs

H. Videlier[1], **N. Dyakonova**[1], **F. Teppe**[1], **C. Consejo**[1], **W. Knap**[1], **J. Lusakowski**[2], **D. Tomaszewski**[3], **J. Marczewski**[3], **and P. Grabiec**[3]

[1]GES-UMR5650, Université Montpellier 2 and CNRS, Montpellier 34095, France
[2]Institute of Experimental Physics, Warsaw University, Warsaw, Poland
[3]Institute of Electron Technology, Al. Lotnikow 32/46, Warsaw 02-668, Poland

Abstract— We report on investigations of photovoltaic response of Si-MOSFETs subjected to Terahertz radiation in high magnetic fields. The MOSFETs develop a dc drain-to-source voltage that shows singularities in magnetic fields corresponding to paramagnetic resonance conditions.

These singularities are investigated as a function of incident frequency, temperature and two-dimensional carrier density. We tentatively attribute these resonances to spin transitions of the electrons bounded to Si dopants and discuss the possible physical mechanism of the photovoltaic signal generation.

1. INTRODUCTION

Photoresponse to sub-Terahertz and Terahertz (THz) radiations in Silicon Metal Oxide Semiconductor Field Effect Transistors (Si-MOSFETs) has been studied previously [1–3]. The obtained values of responsivity and Noise Equivalent Power have shown the potential of Si-MOSFETs as sensitive detectors at room temperature [1]. The photovoltaic response dependences on the gate length and the gate bias were in good agreement with the Dyakonov-Shur plasma wave detection theory [4]. The detection signal results from the rectification of high frequency currents induced by the incident radiation in the transistor channel. The rectification takes place due to a nonlinear response of the gated two dimensional electron gas and a source-drain asymmetry. The main source of nonlinearity is the superposition of two radiation-induced effects: i) the modulation of the carrier density in the channel by variations of the gate potential and ii) drift velocity modulation due to variations of the drain potential. Recently, monolithically integrated THz focal-plane arrays including antennas and amplifiers on a single silicon die has been designed to work at room temperature [5].

In this work, we have studied the effect of high magnetic field on the Si-MOSFETs photovoltaic response subjected to sub-Terahertz radiation.

2. RESULTS

The measurements were carried out on silicon-on-insulator (SOI) MOSFETs. Structures used in our study had gate length and width respectively, $Lg = 10\,\mu\text{m}$ and $Wg = 5\,\mu\text{m}$. The transistors were bonded to a commercially available multi pin support with a gold wire. The transistor were placed at the centre of a superconducting coil in the Faraday configuration and cooled by a helium exchange gas. Sub-Terahertz radiation in the range 185–300 GHz was brought from a Backward Wave Oscillator source or a Gunn diode to the transistor through a polished pipe. No special coupling antennas were used and the radiation was coupled to the device through metallization pads. The photovoltaic signal, ΔU, was measured between source and drain contacts using a standard lock-in technique. The temperature range was from 4 K to 300 K.

The photoresponse depends on the gate bias in a typical way showing a maximum close to the threshold voltage [1]. With temperature decrease the photoresponse maxima shifts to higher gate bias following the shift of the transistor threshold. Also the amplitude of the photoresponse grows by approximately one order of magnitude with lowering temperature from 300 K down to 10 K.

The magnetic field affected the photoresponse in two manners (Fig. 1). First, we observed a smooth decreasing of the signal with magnetic field. Second, we observed well pronounced structures that shift to higher magnetic fields with increasing light frequency.

The smooth decreasing behavior can be explained by plasma waves damping that takes place at magnetic fields B higher than the cyclotron resonance one $Bc = \omega cm/e$, where ωc is the cyclotron frequency, m is the effective mass. For silicon at 300 GHz, $Bc \approx 2\,\text{T}$. One can see in Fig. 1 that the signal effectively begins to decrease around 2 T in a good agreement with above estimation. More

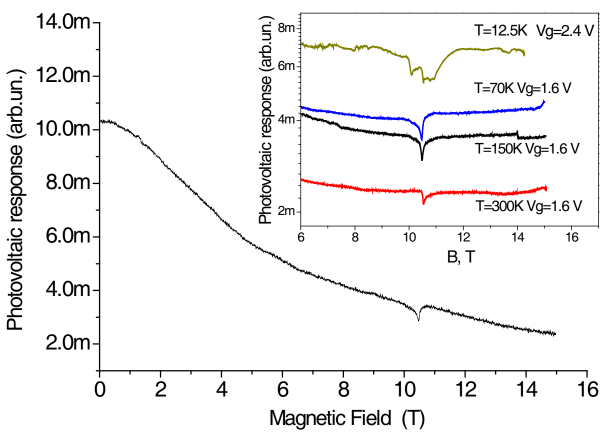


Figure 1: Photovoltaic signal at 10 K as a function of applied magnetic field at incident frequency 300 GHz. Gate voltage is 1.6 V. Inset: temperature dependence of photovoltaic response in the vicinity of singularity.

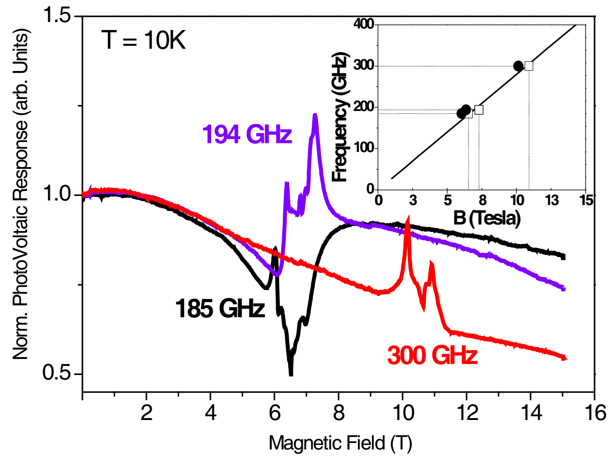


Figure 2: Photovoltaic signal at 10 K as a function of applied magnetic field at three different incident frequencies: 185 GHz, 194 GHz, 300 GHz. Gate voltage is 2.4 V. Inset shows the position of the 2 peaks (full circles and empty squares). Straight line is a calculation with $g = 2$.

detailed description of the influence of magnetic field on the THz detection by field effect transistor can be found in Ref. [6].

The inset of Fig. 1 shows also the singularity and its evolution with temperature. One can see that at low temperature (12.5 K) the singularity has a complex structure, composed of several lines. With temperature increase side lines disappear, the structure as a whole narrows and its relative amplitude decreases (note the logarithmic scale of the photoresponse axe).

In Fig. 2, photovoltaic signal is presented at 10 K as a function of applied magnetic field for three different incident frequencies. A constant bias of 2.4 V was applied on the gate, corresponding to its threshold voltage value and hence to the transistor's maximum of sensitivity. For the sake of clarity the curves are normalized to the signal value at zero field. A very well pronounced structure is visible on each curve with two main peaks shifting to higher fields by increasing incident light frequency. The position of these peaks is plotted in the inset of Fig. 2 where the straight line is a calculation of the frequency dependence on the magnetic field with the following equation $f = g\mu_B B/h$ and $g = 2$ corresponding to the g factor of typical Silicon dopants. One can see that the experimental points follow the calculated dependence indicating a possible link between the observed structures in the photovoltaic response and paramagnetic resonances (spin resonance transitions).

The exact physical mechanism leading to photovoltaic response related to spin resonance transitions is not clear up to now. In addition, from experimental point of view we have some unclear points, for example, the signal can change its shape dramatically with temperature, light frequency or between sample cooling cycles (compare Figs. 1 and 2). Further experimental and theoretical investigations are needed.

Looking for theoretical explanation of the observed phenomenon, here we can only speculate that the signal is related to the change of the free carrier mobility. The change of the mobility can take place at the spin resonance via one of the spin dependent scattering mechanisms. This change of the mobility can lead to the photovoltaic response in a similar way as the change in the carrier density. The effect of the photovoltaic signal amplification due to the modulation of the mobility was already observed in the Shubnikov de Hass related experiments in III-V HEMTS [6–8].

However, a full quantitative interpretation of our results requires more complete experimental and theoretical developments.

In conclusion we observed a new phenomenon in the Terahertz photovoltaic response of Silicon MOSFETs at low temperature and high magnetic field. The magneto-photovoltaic response exhibits a structure moving linearly with the incident radiation frequency. Preliminary studies show possible link between the observed structure in the photovoltaic response and a spin related phenomenon.

ACKNOWLEDGMENT

The authors from the University of Montpellier acknowledge support of the region of Languedoc-Roussillon "Terahertz Platform." This work was also supported by GDR-E "Terahertz Sources and Detectors," and by ANR program TeraGaN. We thank M. I. Dyakonov for valuable discussions. The authors from Warsaw acknowledge Polish Ministery of Higher Education (Project No. OR00015509).

REFERENCES

1. Tauk, R., F. Teppe, S. Boubanga, D. Coquillat, W. Knap, Y. M. Meziani, C. Gallon, F. Boeuf, T. Skotnicki, C. Fenouillet-Beranger, D. K. Maude, S. Rumyantsev, and M. S. Shur, "Plasma wave detection of terahertz radiation by silicon field effects transistors: Responsivity and noise equivalent power," *Appl. Phys. Lett.*, Vol. 89, 253511, 2006.
2. Knap, W., F. Teppe, Y. Meziani, N. Dyakonova, J. Lusakowski, F. Boeuf, T. Skotnicki, D. Maude, S. Rumyantsev, and M. S. Shur, "Plasma wave detection of sub-terahertz and terahertz radiation by silicon field-effect transistors," *Appl. Phys. Lett.*, Vol. 85, 675–677, 2004.
3. Stillman, W., M. S. Shur, D. Veksler, S. Rumyantsev, and F. Guarin, "Device loading effects on nonresonant detection of t, erahertz radiation by silicon MOSFETs," *Electronics Letters*, Vol. 43, No. 7, 2007.
4. Dyakonov, M. and M. Shur, "Detection, mixing, and frequency multiplication of terahertz radiation by two-dimensional electronic fluid," *IEEE Transactions on Electron Devices*, Vol. 43, No. 3, 380–387, 1996.
5. Ojefors, E., U. R. Pfeiffer, A. Lisauskas, and H. G. Roskos, "A 0.65 THz focal-plane array in a quarter-micron CMOS process technology," *IEEE Journal of Solid-State Circuits*, Vol. 44, No. 7, 1968–1976, 2009.
6. Lifshits, M. B. and M. I. Dyakonov, "Photovoltaic effect in a gated two-dimensional electron gas in magnetic field," *Phys. Rev. B*, Vol. 80, 121304(R), 2009.
7. Boubanga-Tombet, S., M. Sakowicz, D. Coquillat, F. Teppe, N. Dyakonova, W. Knap, K. Karpierz, and J. Lusakowski, "Terahertz detection by field effect transistor in high magnetic fields: Influence of shubnikov-de haas and cyclotron resonance effects," *Appl. Phys. Lett.*, Vol. 95, 072106–072108, 2009.
8. Klimenko, O. A., Y. A. Mityagin, H. Videlier, F. Teppe, N. V. Dyakonova, C. Consejo, S. Bollaert, V. N. Murzin, and W. Knap, "Terahertz response of InGaAs field effect transistors in quantizing magnetic fields," *Appl. Phys. Lett.*, May 2010.

Design of 2 x 2 U-shape MIMO slot antennas with EBG material for mobile handset applications

Z. Z. Abidin[1,2], **Y. Ma**[1], **R. A. Abd-Alhameed**[1], **K. N. Ramli**[1,2], **D. Zhou**[3], **M. S. Bin-Melha**[1], **J. M. Noras**[1], **and R. Halliwell**[1]

[1]Mobile and Satellite Communications Research Centre, University of Bradford
Bradford, West Yorkshire, BD7 2DF, UK
[2]Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia
Parit Raja, Batu Pahat, Johor, Malaysia
[3]Surrey Space Centre, University of Surrey, Guildford, GU2 7XH, UK

Abstract— A compact dual U-shaped slot PIFA antenna with Electromagnetic Bandgap (EBG) material on a relatively low dielectric constant substrate is presented. Periodic structures have found to reduce mutual coupling and decrease the separation of antenna and ground plane. A design with EGB material suitable for a small terminal mobile handset operating at 2.4 GHz was studied. Simulated and measured scattering parameters are compared for U-shaped slot PIFA antenna with and without EBG structures. An evaluation of MIMO antennas is presented, with analysis of the mutual coupling, correlation coefficient, total active reflection coefficient (TARC), channel capacity and capacity loss. The proposed antenna meets the requirements for practical application within a mobile handset.

1. INTRODUCTION

The potential for MIMO antenna systems to improve reliability and enhance channel capacity in wireless mobile communications has generated great interest [1]. A major consideration in MIMO antenna design is to reduce correlation between the multiple elements, and in particular the mutual coupling, electromagnetic interactions that exist between multiple elements, is significant, because at the receiver end this could largely determine the performance of the system. Lower mutual coupling can result in higher antenna efficiencies and lower correlation coefficients. The effect of mutual coupling on capacity of MIMO wireless channels is studied in [2]. Spatial diversity is strongly affected by mutual coupling and correlation. For minimal coupling, it has been shown in [3], than the separation between multiple antenna elements should be at least 0.5λ. A 2×2 MIMO meander planar inverted-F antenna (PIFA) at 2.6 GHz is reported in [4] to obtain a mutual coupling of -15 dB, with a separation between two antenna elements of 0.23λ. Authors in [5], introduced a U-shaped slot patch antenna operating at 2.6 GHz for mobile handset applications, with isolation of -20 dB achieved by pattern diversity and a capacity loss of 0.2 bits/s/Hz.

EBG structures have the ability to act like perfect magnetic conductors (PMC), so that the distance between the antenna and ground plane can be smaller than $\lambda/4$. This can be compared to perfect electric conductors (PEC) where a distance of $\lambda/4$ is essential so that the reflected wave interferes constructively with the emitted one [6]. When the height of the antennas to the ground plane is reduced, mutual coupling can be expected to be decreased [7, 8]

This paper presents a compact dual U-shaped slot PIFA antenna with EBG material on a relatively low dielectric constant substrate, operating at 2.4 GHz and suitable for for compact mobile handsets S-parameters for U-shaped slot PIFA antennas with and without EBG materials are compared. In addition, the MIMO antennas are analysed in terms of their mutual coupling, correlation coefficient, total active reflection coefficient (TARC), channel capacity and capacity loss.

2. MIMO

2.1. Basic Theoretical Concepts

2.1.1. Total Active Reflection Coefficient (TARC)

TARC is defined as the ratio of the square root of total reflected power divided by the square root of total incident power. The TARC for a 2×2 antenna array can be directly calculated from the scattering matrix elements as follows [5]

$$TARC = \sqrt{(|(s_{11} + s_{12}e^{j\theta})^2 + |(s_{21} + s_{22}e^{j\theta})^2)/\sqrt{2}}, \tag{1}$$

where θ represents the phase from 0 to 2π.

2.1.2. Correlation Coefficient

Previous work shows that the correlation coefficient, ρ, of a 2×2 antenna system can also be determined using S-parameters [9]

$$\rho = \frac{S_{11}^* S_{12} + S_{21}^* S_{22}^2}{\left(1 - S_{11}^2 - S_{21}^2\right)\left(1 - S_{22}|^2 - S_{12}^2\right)}. \tag{2}$$

2.1.3. Channel Capacity

Based on the channel transfer matrix, H, the Shannon capacity, C for the MIMO system channel is [1, 10]

$$C = \log_2 \left(det \left(1 + \frac{SNR}{M} H H^{\dagger} \right) \right) \tag{3}$$

where $H^{\dagger}$ is the Hermitian of the matrix H, M is is the number of receivers and SNR is the estimated channel signal-to-noise ratio.

2.1.4. Capacity Loss

In case of high SNR, the capacity loss is given by [11]

$$C(loss) = -\log_2 det \left(\Psi^R\right) \tag{4}$$

where Ψ^R is the receiving antenna correlation matrix.

3. DESIGN OF 2 × 2 U-SHAPED SLOT PIFA ANTENNA INTEGRATED WITH EBG MATERIAL

The basic geometrical configuration and dimensions of the PIFA with the U-shape slot structure is shown in Figure 1. The design frequency in this study is 2.4 GHz, and the antenna assembly is mounted on a $0.36\lambda \times 0.68\lambda$ ground plane. The antennas are constructed from 0.5 mm thick plate,

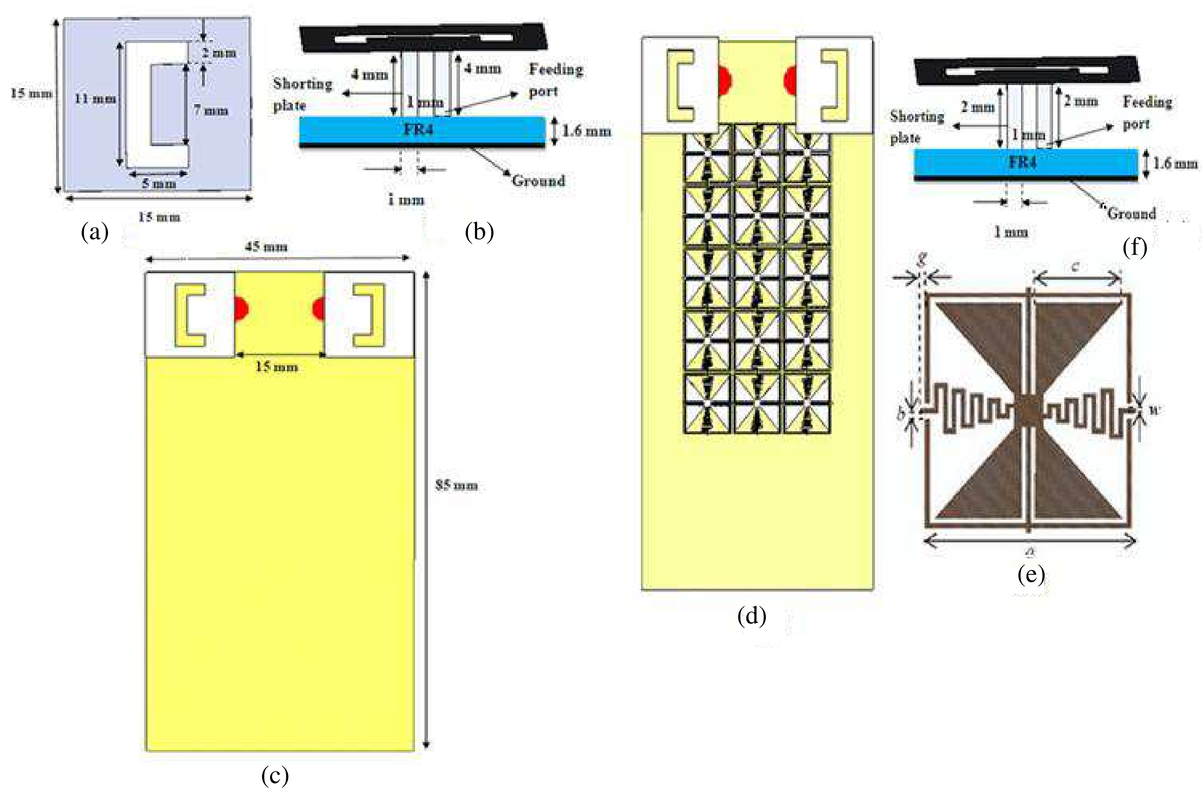


Figure 1: Geometries and dimensions: U-shape antenna (a) top view, (b) side view, (c) 2-D schematic; U-shape antenna with EBG (d) 2-D schematic, (e) EBG unit cell, (f) side view.

Table 1: Measured results for correlation coefficient, TARC and capacity loss at 2.4 GHz.

	PIFAs	PIFAs with EBG
Correlation Coefficient (dB)	−43.99	−49.53
TARC (dB)	5.65	−7.17
Capacity Loss (bits/s/Hz)	0.45	0.43

with a maximum area of $0.12\lambda \times 0.12\lambda$. The antenna is shorted to the ground plane by a metallic strip and fed by a standard 50 Ω SMA connector. The antennas are mounted on FR4 substrate with relative permittivity of 4.5, and loss tangent of 0.002 at 2.4 GHz. The substrate thickness is 1.6 mm, and the distance between two antenna elements is 0.24λ.

Surface waves play a dominant role in the mutual coupling between the antenna array elements. Since the EBG structure has the ability to suppress surface waves, an EBG structure [12] is implemented with the antennas as shown in Figure 1. The EBG unit cell is shown in detail in Figure 1(e), with the dimensions of $a = 9.7$ mm, $b = 0.2$ mm, $c = 4.0$ mm, $g = 0.2$ mm, and $w = 0.2$ mm. In [7], the reduction in height between the PIFAs and the EBG material is shown to mitigate the effects of mutual coupling, and thus improves the antenna efficiency. Therefore, the height of the antennas in this study has been reduced to 2 mm, with other dimensions of the antenna unchanged

4. SIMULATED AND MEASURED PERFORMANCE

Figures 2 and 3 show the simulated and measured s-parameters output for PIFAs without and with EBG material respectively. Measurements show that the predictions for return loss and mutual coupling for both PIFAs are quite accurate with impedance bandwidth approximately

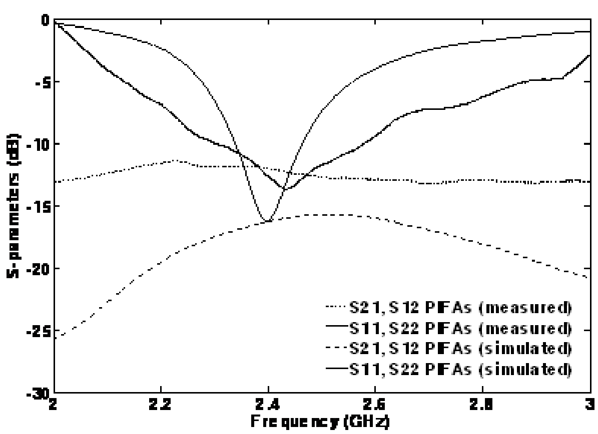


Figure 2: Comparative plot of s-parameters output for simulated and measured results for PIFAs without EBG.

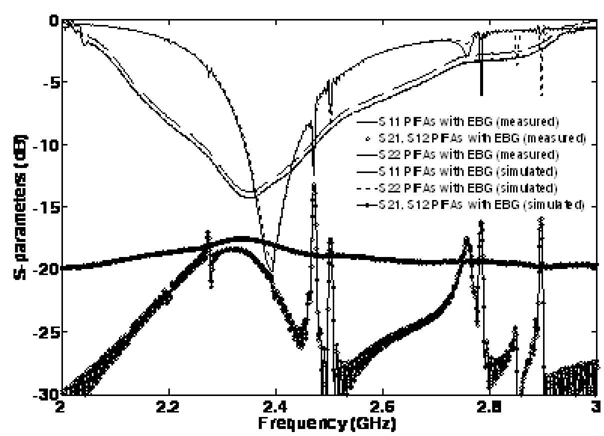


Figure 3: Comparative plot of s-parameters output for simulated and measured results for PIFAs with EBG.

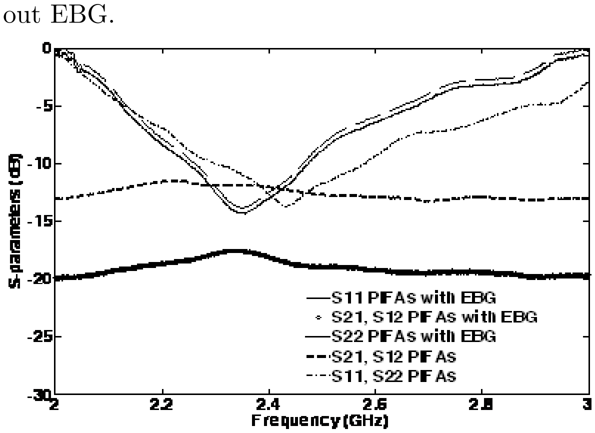


Figure 4: Measured s-parameters of PIFAs with and without EBG.

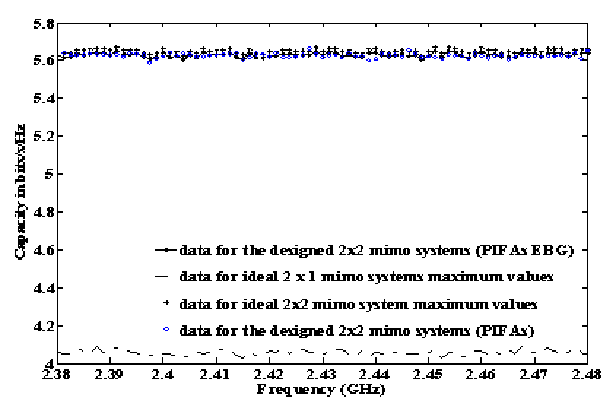


Figure 5: Measured capacity of PIFAs with and without EBG.

11.2% for both designs. Apparent discrepancies are believed to be due to minor inconsistencies in prototyping. It can be seen that an improved isolation of -6 dB is achievable against the PIFAs without EBG as shown in Figure 4. Next, to show the performance of the MIMO system, the correlation coefficient, TARC, capacity loss and capacity are presented in Table 1. This shows that the PIFAs with EBG material have lower loss of capacity and better performance for TARC and correlation coefficient compared to PIFAs without EBG. The channel capacity of PIFAs with and without EBG is 5.62 bits/s/Hz and 5.64 bits/s/Hz at 2.4 GHz, respectively, as depicted in Figure 5.

5. CONCLUSION

Isolation improvement using an EBG structure on 2×2 U-shape slot patch traditional PIFA antennas for 2.4 GHz WLAN operation in mobile application has been verified The proposed antenna can be effectively implemented on a thinner profile FR4 substrate with a low cost and is thus particularly suitable for compact mobile handsets The use of EBG material plays an important role in reducing the mutual coupling between dual antenna elements at 2.4 GHz. The proposed antenna design was mounted on a reasonable ground size of $0.36\lambda \times 0.68\lambda$ in which 3×5 EBG unit cells are used, has achieved a good isolation of 6 dB at 2.4 GHz. Further, the correlation coefficient, TARC, capacity and capacity loss have been analysed for PIFA with and without EBG. It has been shown that the proposed antenna has met the requirements for practical application in mobile handsets.

REFERENCES

1. Foschini, G. J. and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, Vol. 6, No. 3, 311–335, 1998.
2. Thaysen, J. and K. B. Jakobsen, "Wireless system: Design considerations for low antenna correlation and mutual coupling reduction in multi antenna terminals," *Euro. Trans. Telecomms.,* 319–326, 2007.
3. Stutzman, W. L. and G. A. Thiele, *Antenna Theory and Design*, 2nd Edition, Wiley, New York, 1998.
4. Qiong, W., P. Dirk, Z. Hui, and W. Klaus, "Diversity performance of an optimized meander PIFA array for MIMO handsets," *Proceedings of the Forth European Conference on Antennas and Propagation (EuCAP)*, 1–5, 2010.
5. Chae, S. H., S.-K. Oh, and S.-O. Park, "Analysis of mutual coupling, correlations, and TARC in WiBro MIMO array antenna," *IEEE Antennas and Wireless Propagation Letters*, Vol. 6, 122–125, 2007.
6. Sievenpiper, D., "High impedance electromagnetic surfaces," Ph.D. Dissertation, Department of Electrical Engineering, University of California, Los Angeles, 1999.
7. Abidin, Z. Z., R. A. Abd-Alhameed, N. J. McEwan, and M. B. Child, "Analysis of the effect of EBG on the mutual coupling for two-PIFA assembly," *Loughborough Antennas & Propagation Conference*, 473–476, 2010.
8. Zhao, Y., Y. Hao, and C. G Parini, "Radiation properties of PIFA on electromagnetic bandgap substrates," *Microwave and Optical Technology Letters*, Vol. 44, 21–24, 2005.
9. Blanch, S., J. Romeu, and I. Corbella, "Exact representation of antenna system diversity performance from input parameter description", *Electron. Lett.*, Vol. 39, 705–707, May 2003.
10. Moustakas, A. L., H. U. Baranger, L. Balents, A. M. Sengupta, and S. H. Simon, "Communication through a diffusive medium: Coherence and capacity," *Science*, Vol. 287, 287–290, 2000.
11. Shin, H. and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole", *IEEE Trans. Inform. Theory*, Vol. 49, 2636–2647, 2003.
12. Abidin, Z. Z., R. A. Abd-Alhameed, N. J. McEwan, S. M. R. Jones, K. N. Ramli, and A. G. Alhaddad, "Design and analysis of UC-EBG on mutual coupling reduction," *Loughborough Antennas & Propagation Conference,* 693–696, 2009.

Silhouette Coverage Analysis for Multi-modal Video Surveillance

S. Verstockt[1,2], C. Poppe[1], P. De Potter[1], C. Hollemeersch[1], S. Van Hoecke[2], P. Lambert[1], and R. Van de Walle[1]

[1]ELIS Department Multimedia Lab, Ghent University, IBBT, Belgium
[2]ELIT Lab, University College West Flanders, Ghent University Association, Belgium

Abstract— In order to improve the accuracy in video-based object detection, the proposed multi-modal video surveillance system takes advantage of the different kinds of information represented by visual, thermal and/or depth imaging sensors. The multi-modal object detector of the system can be split up in two consecutive parts: the registration and the coverage analysis.

The multi-modal image registration is performed using a three step silhouette-mapping algorithm which detects the rotation, scale and translation between moving objects in the visual, (thermal) infrared and/or depth images. First, moving object silhouettes are extracted to separate the calibration objects, i.e., the foreground, from the static background. Key components are dynamic background subtraction, foreground enhancement and automatic thresholding. Then, 1D contour vectors are generated from the resulting multi-modal silhouettes using silhouette boundary extraction, cartesian to polar transform and radial vector analysis. Next, to retrieve the rotation angle and the scale factor between the multi-sensor image, these contours are mapped on each other using circular cross correlation and contour scaling. Finally, the translation between the images is calculated using maximization of binary correlation.

The silhouette coverage analysis also starts with moving object silhouette extraction. Then, it uses the registration information, i.e., rotation angle, scale factor and translation vector, to map the thermal, depth and visual silhouette images on each other. Finally, the coverage of the resulting multi-modal silhouette map is computed and is analyzed over time to reduce false alarms and to improve object detection.

Prior experiments on real-world multi-sensor video sequences indicate that automated multi-modal video surveillance is promising. This paper shows that merging information from multi-modal video further increases the detection results.

1. INTRODUCTION

The growing demand for security has given raise to the increased use of video surveillance systems in recent years. Surveillance cameras are rapidly appearing in all sort of places and a huge number of visual object detection algorithms, which automatically process these camera images, have been proposed in literature. However, due to the variability of shape, motion, colors, and patterns of moving objects, and also due to the dynamic character of the background, many of these visual object detectors are still vulnerable to false and missed detections. To avoid the disadvantages of using visual sensors alone, we believe the use of other types of imagery, e.g., thermal infrared (IR) and Time-of-Flight (ToF) depth images, can be of added value. The combination of this types of imagery yields information about the scene that is rich in color, motion, depth and/or thermal detail. Once such information is registered, i.e., aligned with each other, it can be used to improve detection performance and activity analysis in the scene. Since each type of sensor has its own type of detection limitations, misdetections in one sensor can be corrected by the other sensors. As such, the combination of multi-sensor information is considered to be a win-win.

In order to combine multi-modal images, it is required that the corresponding objects in the scene are aligned, or registered. The goal of registration is to establish geometric correspondence between the images so that they may be transformed, compared, and analyzed in a common reference frame. Usual features used for multi-sensor registration are edges, corners, and contours [1]. Since contours representing the region boundaries are preserved in most cases, object silhouettes form the most reliable correspondence between objects in color, thermal and/or depth image pairs [2]. For this reason, they are also used in our multi-modal video surveillance system.

The remainder of this paper is organized as follows. Section 2 gives a global description of the silhouette-based registration of multi-modal images, which is based on moving object silhouette extraction, contour vector generation, contour mapping and binary correlation. As an example, the registration of visual and long-wave infrared (LWIR) images is shown. Subsequently, Section 3 discusses the silhouette coverage analysis, i.e., the multi-modal merging of the detection results from the visual, thermal and/or depth image sensors. By two use cases, i.e., a shadow removal and a smoke detection experiment, we show how the coverage analysis of multi-modal images can be used to obtain better object detection results than either sensor alone. Next, in Section 4, we provide details of the experimental setup. Finally, Section 5 ends this paper with the conclusions.

2. SILHOUETTE-BASED REGISTRATION OF MULTI-MODAL IMAGES

The multi-modal image registration (Fig. 1) starts with a moving object silhouette extraction [2] to separate the calibration objects, i.e., the moving foreground, from the static background. Key components are the dynamic background subtraction, automatic thresholding and (iterative) morphological filtering. The dynamic background subtraction [3] extracts the moving foreground (FG) out of the visual and thermal video frames using a visual background estimation, which is updated dynamically. By subtracting the frames with everything in the scene that remains constant over time, i.e., the background, only the moving part of those images remains. After this background subtraction, the resulting foreground images are thresholded automatically using automatic gamma correction, (adaptive) k-means clustering and morphological filtering with growing structuring elements, which grow iteratively until the resulting silhouette is suitable for multi-modal silhouette matching. The combination of all these steps achieves favorable results, as is shown by the visual and the LWIR silhouette extraction in our experiments (Fig. 4). Similar results can be expected for ToF depth silhouette extraction.

After the silhouettes are extracted, registration of both images is performed using a three step registration algorithm. Like in [4], the registration algorithm assumes that the geometric transformation between the multi-sensor images is a rigid transformation, which can be decomposed into a 2D rotation, scaling and translation. To estimate each of these three geometric parameters, the contours and the correlation of the visual and thermal silhouettes are analyzed. First, the rotation is computed using silhouette contour extraction and circular cross correlation [5], which analyzes the translation of the 1-D contour centroid distance (CCD) of both silhouettes. As such, the 2D silhouette matching problem is converted to a one-dimensional signal matching problem. After rotating, the scale factor between both views is estimated by analyzing the ratio of the thermal

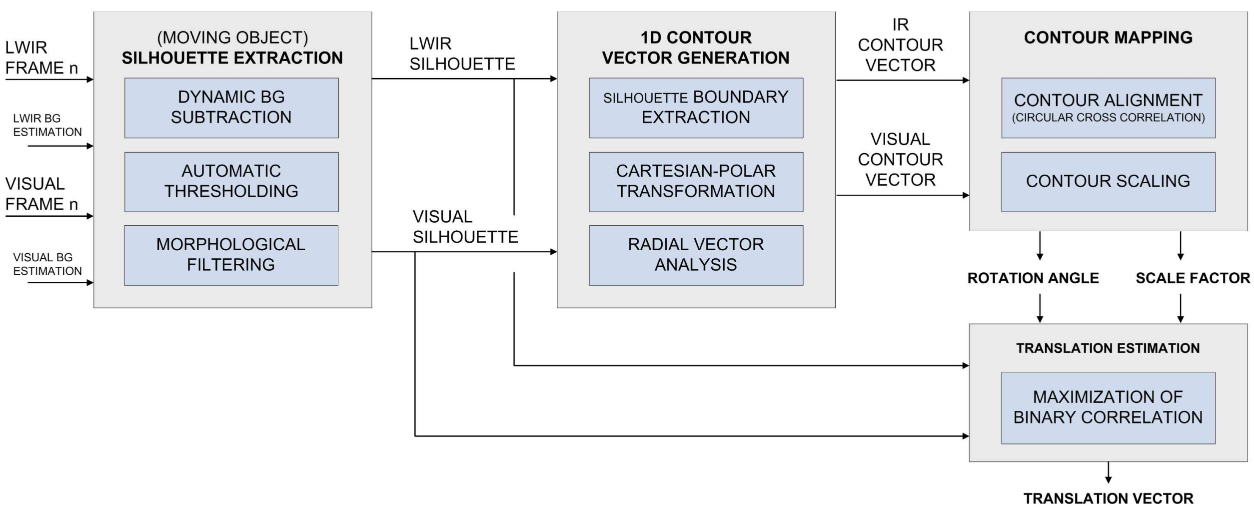


Figure 1: Silhouette-based image registration of thermal and visual images.

Figure 2: Experimental results of LWIR-visual registration.

and visual aligned CCDs. Since the thermal-visual CCD ratios are not constant and show some kind of disorder, the median ratio is chosen as an adequate scale factor. Finally, the translation vector is estimated using the binary correlation technique proposed by Chen et al. [2], which is based on template matching in the frequency domain. As the registration result in (Fig. 2) show, the proposed registration algorithm is able to coarsely map visual and thermal object silhouettes.

3. SILHOUETTE COVERAGE ANALYSIS

The silhouette coverage analysis (Fig. 3) also starts with the moving object silhouette extraction, which was already discussed in the previous section. Then, it uses the registration information, i.e., rotation angle, scale factor and translation vector, to map the thermal and visual silhouette images on each other. As soon as this mapping is finished, the combined LWIR-visual silhouette map is analyzed over time using a temporal coverage analysis algorithm. Depending the video surveillance application for which the multi-modal analysis is used, this silhouette coverag analysis (SCA) can be performed in different ways. In the following subsections, two exemplary use cases of how the SCA can be used are given. In the first use case, the SCA is used for shadow removal in visual images. In the second use case, the SCA is used as a first warning method for smoke detection.

3.1. Use Case 1: Shadow Removal

Shadows are a main drawback for all visual surveillance applications and affect the accuracy of the system performance. Since shadows do not occur in thermal or ToF depth images, both types of imagery can be used to discard them in visual images. This is also shown by the first experiment in (Fig. 4(a)). In this experiment, the multi-modal SCA is used to count the number of people in a room. Due to their shadows, the visual silhouettes of both persons overlap in the visual images. Without the LWIR-visual SCA, a visual people counter could miscount the number of people as 1. By using the LWIR-visual SCA we can correct this mistake. As can be seen in (Fig. 4(a)), the registered visual and thermal silhouettes do not overlap in the shadow regions, i.e., the gray

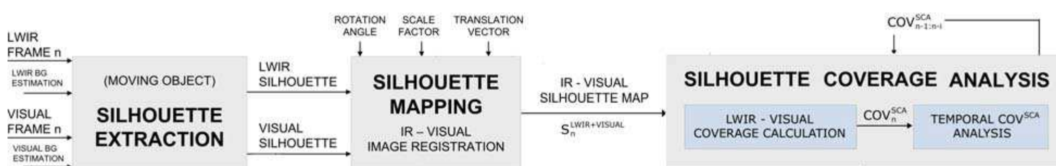
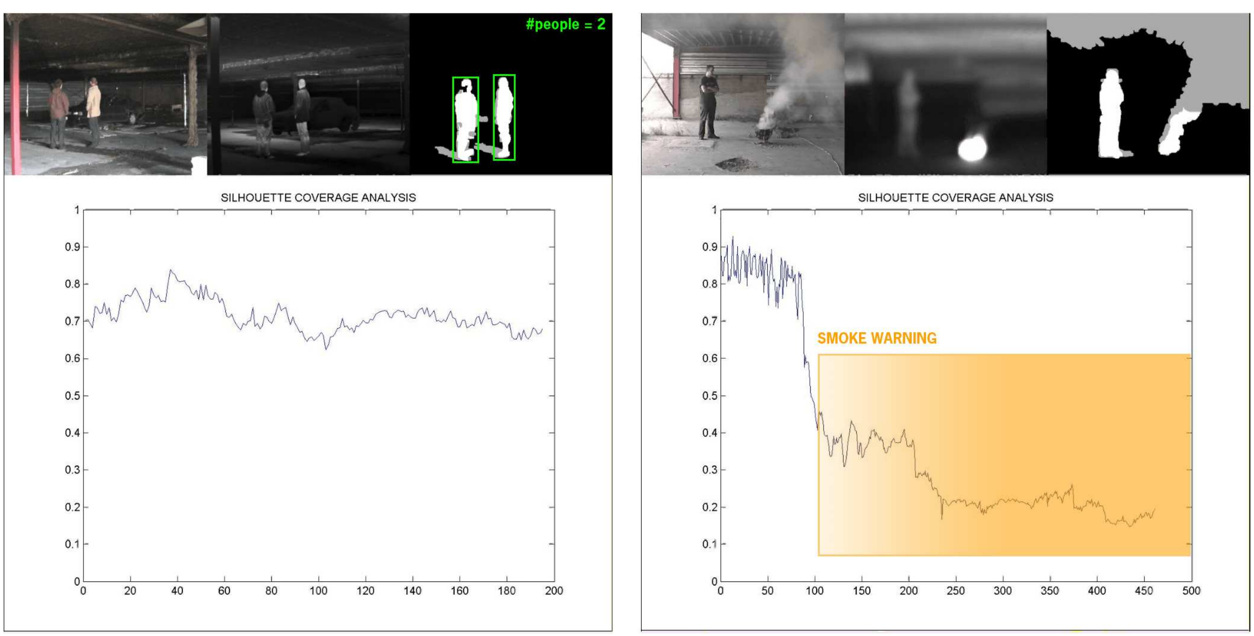


Figure 3: Silhouette coverage analysis.



(a) (b)

Figure 4: Experimental results of silhouette coverage analysis for (a) shadow removal and (b) smoke detection.

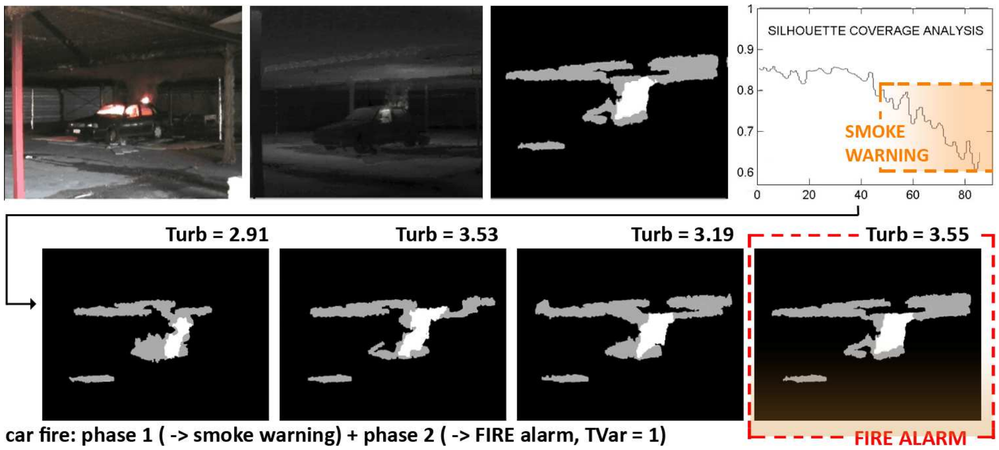


Figure 5: "Car park fire [8]" test results of SCA-based smoke detection.

regions. As such, by only counting the regions which occur in both thermal and visual images, i.e., the white regions, and by analyzing if this regions are stable over time, the SCA results in a more robust and efficient people counter. Similar results are expected with visual-ToF depth SCA analysis. The bounding boxes, shown in the figure, were created by calculating the smallest enclosing rectangle (whose sides are parallel to the x and y axes) around the common, i.e., white, visual-LWIR regions.

3.2. Use Case 2: Smoke Detection

Although smoke is almost transparent in LWIR images, we can make use of its absence to detect it. Since ordinary moving objects, such as people, cars, etc., produce similar silhouettes in background-subtracted visual and thermal IR images, the coverage between these images is quasi constant. This can also be seen in the coverage graph of experiment 1. The coverage for the moving people stays quasi constant over all the frames. Smoke, contrarely, will only be detected in the visual images, and as such the coverage will start to decrease (Fig. 4(b)). This decrease can be detected using a sequence/scene independent technique based on slope analysis of the linear fit, i.e., trend line, over the most recent silhouette coverage values. If the slope of this trend line is negative and decreases continuously, smoke warning is given. Due to its dynamic character, the visual silhouettes of a smoke region will also show a high degree of turbulence [6]. By focusing on both the visible-invisble character of smoke and its visual disorder, a multi-sensor detector can detect smoke very accurately.

Compared to the results of any individual detector in [7], the 2-phase multi-sensor smoke detector is able to detect the smoke more accurate, i.e., with less missdetections and false alarms. This is also illustrated by the test results of a car park fire [8] in (Fig. 5). Due to the low-cost of the silhouette coverage analysis and the visual disorder analysis, which is only performed if smoke warning is given, the algorithm is also less compuational expensive as many of the individual detectors.

4. EXPERIMENTAL SETUP

The multi-modal sequences were acquired by a Xenics Gobi-384 LWIR camera and a CANON MD110 camera, which works in the 8–14 µm spectral range and the visible spectrum respectively. The Gobi thermal imager has a resolution of 384×288 pixels, and a frame rate of 28–30 fps. The CANON its resolution is 576×720 and its framerate is 25 fps. In order to cope with the different frame rates and resolutions, and also whith the differences in the the field of view of the cameras, the multi-modal frames are spatio-temporal registered using temporal frame alignment and the silhouette-based registration proposed in this paper.

5. CONCLUSIONS

Multi-modal video surveillance takes advantage of the different kinds of information represented by thermal, visual and/or depth images in order to accurately detect moving objects. By fusing the different modalities and using the strengths of each medium, object detection can be done more accurate and with less false detections, as is shown by two use cases in this paper. Merging information from multiple types of image sensors has, as such, proven to be a win-win.

To detect the presence of objects, the detector analyzes the silhouette coverage of moving objects in multi-modal registered images. In order to register the multi-sensor images, the proposed algorithm analyses the contours and the correlation of visual and thermal FG silhouettes. First, the rotation is computed using silhouette contour extraction and circular cross correlation. Next, contour scaling is used to estimate the thermal-visual scale factor. Finally, the translation vector is estimated by maximization of binary correlation.

The geometric parameters found during this registration phase are further used by the detector to coarsely map the silhouette images and coverage between them is calculated. Depending the video surveillance application for which the multi-modal analysis is used, this coverage can then be further used to improve the detection results, as is shown by the people counter and the smoke detection experiment.

Future work will mainly focus on the improvement of the registration results. Currently, only the binary silhouettes of the calibration objects are used to do the registration. We expect that a first improvement can be made by also incorporating their gray-scale information, especially in the translation estimation. As the contour mapping is based on the boundary correspondences it is not expected that grayscale information will lead to better results in the rotation and scale estimation. However, further testing is necessary to confirm this. Also the use other types of class clustering classifiers will be further investigated.

ACKNOWLEDGMENT

The research activities as described in this paper were funded by Ghent University, the Interdisciplinary Institute for Broadband Technology (IBBT), University College West Flanders, Warrington Fire Ghent, the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT), the Fund for Scientific Research-Flanders, the Belgian Federal Science Policy Office (BFSPO), and the European Union.

REFERENCES

1. Zitova, B. and J. Flusser,, "Image registration methods: A survey," *Image and Vision Computingn*, Vol. 21, 977–1000, 2003.
2. Chen, H.-M., S. Lee, R. M. Rao, M.-A. Slamani, and P. K. Varshney, "Imaging for concealed weapon detection," *IEEE Signal Processing Magazine*, 52–61, March 2005.
3. Toreyin, B. U., Y. Dedeoglu, U. Gdkbay, and A. E. Cetin, "Computer vision based method for real-time fire and flame detection," *Pattern Recognition Letters*, Vol. 27, 49–58, 2006.
4. Han, J. and B. Bhanu, "Fusion of color and infrared video for moving human detection," *Pattern Recognition*, Vol. 40, 1771–1784, 2007.
5. Hamici, Z., "Real-time pattern recognition using circular cross-correlation: A robot vision system," *International journal of Robotics and Automation*, Vol. 21, pp 174–183, 2006.
6. Verstockt, S., A. Vanoosthuyse, S. van Hoecke, P. Lambert, and R. van de Walle, "Multi-sensor fire detection by fusing visual and non-visual flame features," *International Conference on Image and Signal Processing (ICISP)*, 333–341, June 2010.
7. Verstockt, S., B. Merci, B. Sette, P. Lambert, and R. van de Walle, "State of the art in vision-based fire and smoke detection," *14th International Conference on Automatic Fire Detection*, Vol. 2, 285–292, September 2009.
8. Merci, B., "Fire safety and explosion safety in car parks," http://www.carparkfiresafety.be/.

Structural Entropy Based Localization Study of Wavelet Transformed AFM Images for Detecting Background Patterns

Sz. Nagy[1], **A. Fehér**[1], **and L. M. Molnár**[2]

[1]Department of Telecommunications, Széchenyi István University, H-9026 Győr, Egyetem tér 1, Hungary
[2]Department of Electronics Technology, Budapest University of Technology and Economics
H-1111 Budapest, Goldmann Gy. tér 3, Hungary

Abstract— By defining the structural entropy the von Neumann entropy of a charge distribution on a finite grid is divided into two parts. The first one, the extension entropy, is simply the logarithm of the occupation number (i.e., the number of the average grid sites occupied by the charge distribution), while the second part is the structural entropy itself. On a structural entropy versus participation ratio map the different types of localizations follow different, well characterized curves, and every distribution is represented by a vector on the map.

By a structural entropy-filling factor map of any charge distributions on a finite grid (e.g., finite representation of an electron density, or a grayscale atomic force microscope (AFM) image) superstructures of different scale topologies with different decay types can be traced as well. However it is rather hard to distinguish elements of an additive superstructure, especially if the numerical parameters of the different scale patterns are necessary. In the AFM practice the background patterns are sometimes hard to compensate, and by simple structural entropy based calculations it is almost impossible to separate the superstructure of the atomic scale and the image-scale pattern. The reason is the following. Superstructures manifest on the structural entropy map as sum of the sub-structures vector, thus since none of the structures are known, only the sum of their vectors, the sub-vectors are not unique.

Multiresolution or wavelet analysis (MRA) uses a system of basis functions with various time and frequency (space and spatial frequency) parameters for expanding functions. This system makes the time-frequency localization possible. Using some selected MRA resolution levels of the AFM image and carrying out the structural entropy based localization study on each of these levels will determine the decay type of the image at the length scales corresponding to the selected frequencies. This approach is promising for determining the large-scale patterns on AFM pictures.

1. INTRODUCTION

As the manufacturing processes tend to the nanometer domain, atomic force microscope (AFM) is almost an everyday tool in the hand of the producers. AFM usually has built-in image processing facilities for correcting background slope or bendings, which work usually very well, but in some cases, the background pattern is not detected by them. In a previous article [3], we have studied the possibility of applying a structural entropy [1, 2] based topology analysis to detect the superstructure of such images. Structural entropy — filling factor plots are used primarily for detecting the localization type of probabilistic density distributions on a grid [4, 5], but it can be applied for topological characterization of nano-structures as well [6–10]. In case of multiplicative superstructures, when the image is a result of two (or more) structures at different length-scales, the superposition manifests at the structural entropy versus filling factor (log) map as the sum of the vectors corresponding to the separate structures. Since detecting additive superstructures by that method is not a straightforward task, and a lot of previous calculation is needed, we decided to see superstructures from another point of view. Since in the AFM practice the disturbing background patterns are usually much larger scaled than the image, separating the length-scales and detecting the structure of the separated, different length-scale images can be the key for background pattern elimination. In order to separate length-scales, multiresolution analysis (MRA) or wavelet analysis is a very effective tool. Wavelets are extensively used in image processing from fingerprint databases to the JPEG format's advanced version [11–13].

2. MULTIRESOLUTION ANALYSIS

Multiresolution analysis [14] of the Hilbert space $\mathrm{L}^2(\mathbb{R})$ (i.e., space of the square integrable functions) is a sequence of closed subspaces $\{V_m, m \in \mathbb{Z}\}$ that meets the properties

- $\forall\, m \in \mathbb{Z} \quad V_m \subset V_{m+1}$, i.e., the sequent subspaces are embedded in each other, V_m contains all the subspaces V_n with $n < m$.

- $\exists\, a > 1\ v(r) \in V_m \iff v(ar) \in V_{m+1}$, i.e., shrinking the functions of any subspace V_m by a constant a, which is the same for all m, leads into the next, finer subspace.

- $\exists\, b > 0\ v(r) \in V_0 \iff v(r-b) \in V_0$, i.e., translation of a function of the zero level subspace at a uniform grid of spacing b does not lead out of the subspace. (This statement together with the previous one gives similar conditions for any subspace V_m, with the grid spacing $2^{-m}b$.)

- $\bigcap\limits_{m=-\infty}^{\infty} V_m = \{0\}$ and $\bigcup\limits_{m=-\infty}^{\infty} V_m$ is dense in $\mathrm{L}^2(\mathbb{R})$.

- there exists a *scaling function* $s_0 \in V_0$ with a non-zero integral, such that the functions $\{s_0(r - b\ell) |\, \ell \in \mathbb{Z}\}$ form an orthonormal basis of V_0. (It is not necessary to demand orthogonality.)

The orthonormal basis of subspace V_m is the set of functions: $\{s_{m,\ell} |\, \ell \in \mathbb{Z}\}$, where $s_{m,\ell}(r) = a^{m/2} s_0(a^m r - b\ell)$. For the dilation and translation constants, the values $a = 2$ and $b = 1$ are commonly used. The key of the applicability of MRA is the fact, that the basis functions of a rougher resolution level m are expandable by the scaling functions of the finer levels $n > m$. For $n = m + 1$, this formula is called *refinement equation*

$$s_{m,\ell} = \sum_{k=0}^{N_s} p_k s_{m+1\, k+2\ell} \tag{1}$$

with p_k being the scaling function filter coefficients and $\sum_{k=1}^{N_s} p_k = 2$.

A function $f \in \mathrm{L}^2(\mathbb{R})$ can be projected onto any V_m subspace of resolution level m, thus an mth level approximation of the function is

$$f^{[m]}(r) = \sum_{\ell} c_{m\ell} s_{m\ell}(r), \tag{2}$$

where $c_{m\ell} = \langle \tilde{s}_{m\ell} | f \rangle$ are the inner products of the function and the basis function. The dual $\tilde{s}_{m\ell}(r)$ of the scaling function can be calculated, and have similar scaling properties to the original $s_{m\ell}$, with different filter coefficients. (The filter coefficients of this dual scaling function are usually called analysis low-pass filter coefficient, whereas those of Eq. (1) are the synthesis low-pass filter coefficients). Self-dual scaling functions also exist, and we are going to use such $\tilde{s}_{m\ell} = s_{m\ell}$ basis set. The 2 dimensions scaling functions functions are mostly direct products of 1D functions s_{mk}.

In order to introduce wavelets, one has to consider the *detail space* W_m, which is the orthogonal complement of the rougher resolution subspace V_m in the finer V_{m+1}

$$V_{m+1} = V_m \oplus W_m. \tag{3}$$

The Hilbert space is thus the infinite system of subspaces W_m. In these subspaces W_m a basis set can be defined similarly to those of the subspaces V_m, i.e., as $\{w_{m\ell}(r) | \ell \in \mathbb{Z}\}$ with the basis functions $w_{m\ell}(r) = 2^{m/2} w_0(2^{-m} r - \ell)$ generated from the *mother wavelet* w_0. Wavelets can also be generated from the coefficients p_k and the scaling functions using the following equation

$$w_{m,\ell} = \sum_{k=1-N_s}^{1} q_k s_{m+1\, 2\ell-k} \tag{4}$$

with $q_k = (-1)^k p^*_{-k+1}$ being the wavelet filter coefficients. (These coefficients are also called synthesis high-pass filter coefficients.) Wavelets in more dimensions can also be the direct products of 1D wavelets and scaling functions of various resolution levels.

The projection of a function $f \in \mathrm{L}^2(\mathbb{R})$ onto the resolution level m_1 can be given either in the basis $\{s_{m_1 k} | k \in \mathbb{Z}\}$ or in $\{s_{m_0 k}, w_{mk} | k \in \mathbb{Z}, m = m_0, m_0 + 1, \ldots, m_1 - 1\}$ as

$$f^{[m_1]}(r) = \sum_k c_{m_1 k} s_{m_1 k}(r) \tag{5}$$

$$f^{[m_1]}(r) = \sum_k c_{m_0 k} s_{m_0 k}(r) + \sum_{m=m_0}^{m_1-1} \sum_k d_{mk} w_{mk}(r). \tag{6}$$

Here $d_{m\ell} = \langle \tilde{w}_{m\ell} | f \rangle$ are also the inner product with the dual wavelet (that has the high-pass analysis filter coefficients in its refinement equation). The two approximations mean that we can have the function represented by purely its scaling function coefficients $\{c_{m_1 k} | k \in \mathbb{Z}\}$ or by a rough level scaling function coefficient set $\{c_{m_0 k} | k \in \mathbb{Z}\}$ and a series of wavelet coefficients $\{d_{mk} | m = m_0 \ldots m_1 - 1, k \in \mathbb{Z}\}$. If the function f is localized to an interval the number of the fine resolution level scaling function coefficients from expansion (5) equals to the total number of the necessary coefficients from the expansion (6), since every refinement of the resolution level doubles the number of scaling function coefficients. The second type (6) of approximation of the function f is, however more favorable, because the most of the information about the function is in the scaling function coefficients (that are of rather small number), and the large number of wavelet coefficients contain only the fine details about the function. Lots of these wavelet coefficients are almost zero. In image compression the almost-zero coefficients are set to zero, and by this step, a lot of storing capacity can be saved. In Image processing the image itself is the set of the finest resolution level scaling function coefficients, and by consecutive applications of wavelet transform steps, the rougher level wavelet and scaling function coefficients are calculated. At one transformation step the higher level scaling function coefficients are "filtered" into scaling function and wavelet coefficients of the rougher resolution level, using the high-pass and low-pass analysis filter coefficients.

3. CALCULATIONS

As a first step we analyzed images of several localization types with the Structural entropy based method, introduced by Pipek and Varga [1] and used in image processing first by us [6]. The description of the structural entropy S_{str} and the filling factor q can be found in these articles. We recall only, that if the (1D) color coordinate values I_n of the pixel n are normalized so, that they form a probability distribution (i.e., $\sum_n I_n = 1$), the filling factor is

$$I_i \geq 0, \qquad \text{for } i = 1, \ldots, N \tag{7}$$

$$q = \frac{1}{N \sum_{i=1}^{N} I_i^2}, \tag{8}$$

where N is the total number of pixels, whereas the structural entropy is

$$S_{str} = S_1 - S_2 \tag{9}$$

with S_i being the ith Rényi entropy

$$S_i = \frac{1}{1-N} \ln \sum_{j=1}^{N} I_j^i. \tag{10}$$

The distributions of a given localization type (e.g., a Gaussian, or a power-law) form well characterized lines on the $S_{str}(q)$ plots.

The localization type changing according ti the MRA resolution levels of Gaussian, exponential and 2nd order power-law distributions are shown in Fig. 1. The resolution increases from left to right as $m = 0, \ldots, 8$. It is clearly visible, that as the resolution level decreases, the distributions tend to maps of slower decreasing functions, than the original one, and the smaller scale functions do it more rapidly.

As a second step, an additive superstructure of a large-scale Gaussian and several small-scale exponential functions is studied. The resulting images can be seen in Fig. 2. The differences of two scaling function levels are also plotted (the inverse transformed version of the rougher resolution level scaling function image with zero wavelet coefficients were applied). The fine resolution level images are similar for a long m range, the interesting thing happens in the low resolution part, where the support of the scaling functions are similar to the radius of the background pattern. The scaling function images lose the exponential peaks continuously, while the difference images keep these spots until the last image, where the tendencies of the background become dominant.

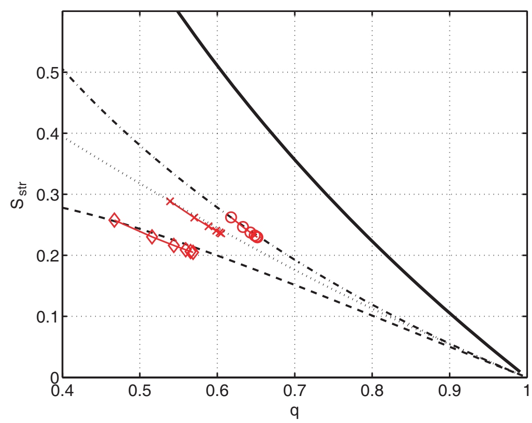


Figure 1: Structural entropy vs. filling factor maps of a Gaussian ($\diamond$), an exponential ($\times$) and a 2nd order power-law ($\circ$) distribution and their wavelet-transformed versions of resolution levels $m = 0, \ldots, 8$. The 2D Gaussian, exponential, 2nd order power-law and limiting curves are also plotted by lines.

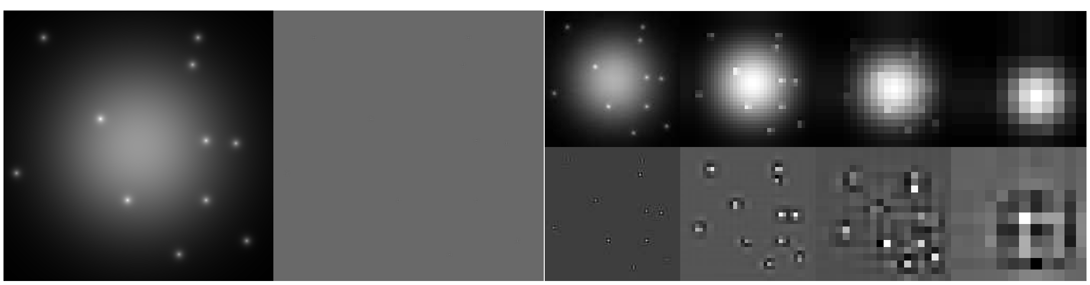


Figure 2: Superstructure of a Gaussian background a several small-scale exponential distributions and the Daubechies-6 wavelet transforms of the image. The upper images are the scaling function coefficients, the lower images are the difference of a finer level scaling function image and the inverse transform of the next, rougher level scaling function image. The images are for resolution levels $m = 8, 5, 3, 2, 1$ ($m = 8$ is the image itself).

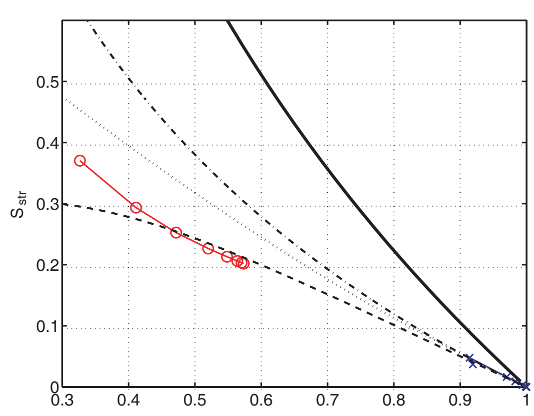


Figure 3: Structural entropy vs. filling factor maps of a superstructure of a large-scale Gaussian and more smaller scale exponential distributions ($\circ$) and their wavelet-transformed versions of resolution levels $m = 0, \ldots, 8$. The 2D Gaussian, exponential, 2nd order power-law and limiting curves are also plotted by lines. The points at the lower right corner ($\times$) correspond to the difference images of Fig. 2.

Figure 3 shows the appearance of the Gaussian behavior as the observing resolution level increases. The difference images at high resolution level are almost uniform gray, that is why they are at the ($S_{str} = 0, q = 1$) point. As the small-scale peaks start to dominate, the structural entropy moves from zero (the plot shows quicker decay, then the real exponential one) and when the background takes the lead, the tendency breaks down to below exponential. Together the two kinds of plots can detect the background pattern, but not accurately enough, thus further refinement of the method is needed. Similar, but not so expressed behavior can be seen on the opposite superstructure (small-scale Gaussians on large-scale exponential). The reason is probably the weakening effect of the tendency mentioned in the discussion of Fig. 1, namely, that all the structures tend to show a slower localization as the resolution level increases.

4. CONCLUSION

The different scale additive superstructures can be distinguished at the wavelet transformed series of images using structural entropy based localization analysis. However, Identifying the superstructure's parts are still not solved.

ACKNOWLEDGMENT

This work was supported by the Bolyai János Research Fellowship of the Hungarian Academy of Sciences and by the Széchenyi István University under project TáMOP 4.2.1.B.

REFERENCES

1. Pipek, J. and I. Varga, "Universal classification scheme for the spatial localization properties of one-particle states in finite d-dimensional systems," *Phys. Rev. A*, Vol. 46, 3148–3164, 1992.
2. Pipek, J. and I. Varga, "Scaling behavior of energy functionals of highly complex distributions," *Int. J. Quantum Chem.*, Vol. 70, 125–131, 1998.
3. Molnár, L. M., Sz. Nagy, and I. Mojzes, "Structural entropy in detecting background patterns of AFM images," *Vacuum*, Vol. 84, 179–183, 2010.
4. Varga, I. and J. Pipek, "Rényi entropies characterizing the shape and the extension of the phase space representation of quantum wave functions in disordered systems," *Phys. Rev. E*, Vol. 68, 026202, 2003.
5. Varga, I., E. Hofstetter, and J. Pipek, "One-parameter superscaling at the metal-insulator transition in three dimensions," *Phys. Lett.*, Vol. 82, 4683–4686, 1999.
6. Mojzes, I., Cs. Dominkonics, G. Harsányi, Sz. Nagy, J. Pipek, and L. Dobos, "Heat treatment parameters effecting the fractal dimensions of AuGe metallization on GaAs," *Appl. Phys. Lett.*, Vol. 91, No. 7, Article No. 073107, 2007.
7. Varga, B., A. Ürmös, Sz. Nagy, and I. Mojzes, "Fractal properties of gold, palladium and goldpalladium thin films on InP," *Vacuum*, Vol. 84, 247–250, 2010.
8. Varga, B., Molnár, L. M., Sz. Nagy, and I. Mojzes, "Fractal properties of AlGeNi layers on GaAs surfaces," *Vacuum*, Vol. 84, 251–253, 2010.
9. Varga, B., Sz. Nagy, and I. Mojzes, "Fractal properties, structural entropy and color of printed circuits boards processed by laser treatment," *PIERS Proceedings*, 281–285, Moscow, Russia, August 18–21, 2009.
10. Schuszter, M., L. Dobos, K. A. Nemcu, Sz. Nagy and I. Mojzes, "Analysis of morphology changes of heat treated metallization of compound semiconductors by the fast wavelet-transform based on B-spline," *J. Optelectron. Adv. Mat.*, Vol. 9, No. 7, 2241–2244, 2007.
11. Christopoulos, C., A. Skodras, and T. Ebrahimi, "The JPEG2000 still image coding system: An overview," *IEEE Trans. Consumer. Electron.*, Vol. 46, No. 4, 1103–1127, 2000.
12. MJ2 Format Standard, "Motion JPEG 2000 derived from ISO base media file format," ISO/IEC 15444-3:2002/Amd 2:2003, 2003.
13. Martucci, S. A., I. Sodagar, T. Chiang, and Y. Q. Zhang, "Zerotree wavelet video decoder," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 7, No. 1, 109–118, 1997.
14. Heil C. E. and D. F. Walnut, "Continuous and discrete wavelet transforms," *SIAM Review*, Vol. 31, No. 4, 628–666, 1989.
15. Daubechies, I., *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM Publications, Philadelphia, 1992.
16. Chui C. K., *An Introduction to Wavelets*, Academic Press, San Diego, 1992.

A New Spatio-temporal ICA for Multi-temporal Endmembers Extraction and Change Trajectory Analysis

S. Hmissi[1,2], **K. Saheb Ettabaa**[1,2], **I. R. Farah**[1,2], **and B. Soulaiman**[2]

[1]RIADI-GDL, ENSI, Tunis, Tunisia

[2]Téléom Bretagne, Technopôle Brest-Iroise-CS 83818, 29238 BREST CEDEX 3, France

Abstract— Independent component analysis (ICA) has been commonly applied to extract sources from hyperspectral images. The purpose of this paper is to study the ability of ICA in extracting spatio-temporal patterns from different hyperspectral sensors, or with different acquisition conditions and dates. Spatio-temporal ICA (*stICA*), which maximizes the degree of independence over space and time, is a suitable method for analyzing multi-temporal hyperspectral images and extracting local features.

1. INTRODUCTION

Multi-temporal hyperspectral images are fundamental appliances for the worldwide assessment and monitoring of natural resources. Such tools allow a better description of the seasonal dynamics of land cover types. Thus, the temporal variation can provide an insight in the remote sensing studies [1]. Several approaches have been developed to explore temporal information in order to magnify our knowledge about natural systems. The richness of multi-temporal hyperspectral data is only beginning to be understood. For instance, we have provided in [2] a new 3D model for images interpretation. It is now important to understand the properties of these signals when developing processing algorithms. ICA has recently demonstrated considerable promises in characterizing hyperspectral data, mostly due to its intuitive nature and potentiality for flexible characterization and extraction of endmembers. In this article, we first introduce ICA and its application to identify spatio-temporal patterns which are the endmembers in hyperspectral processing.

2. STATE OF THE ART

Following its first application, ICA has been successfully used in hyperspectral applications. Examples include the identification of endmembers (spectral unmixing), the analysis of multi-angular hyperspectral data, etc. A comprehensive review of ICAs approaches for hyperspectral data processing is given in [3]. To address this problem, many works have been done. Roberts et al. [4] proposed the multiple endmember spectral mixture analysis (MESMA) technique to extract materials in a hyperspectral image using endmembers and a spectral library. Fischer et al. present in [5] a new approach by using GIS and multi-temporal imaging data. In [6], Martnez et al. proposed a new algorithm based on mathematical morphology and spectral analysis of the hyperspectral spatial-temporal signature and endmembers identification. However, several weaknesses are percepted in the presented approaches. First, models are sensitive to data dimensionality, noise, and differences in illumination and atmospheric conditions. Second, usually temporal unmixing techniques address the change analysis problem by examining each pixel individually and to determine if it is an endmember. However, hyperspectral images reveal frequently the co-activation of spatially disparate endmembers, which cannot be rigorously investigated with the reviewed techniques since they ignore the spatial relationships between pixels. Then, most works focused on temporal dependence and ignored the fact that endmembers should be extracted by maximizing independency over both space and time. Moreover, the analysis of fractional abundances obtained by multi-temporal hyperspectral images unmixing allows a more refined and sophisticated interpretation [1].

3. PROPOSED APPROACH

The main objective of this paper is to propose a novel approach for multi-temporal hyperspectral data interpretation based on spatio-temporal ICA (*stICA*). It provides a robust framework for change analysis and multi-temporal endmembers extraction. Input dataset can be taken from images which are taken at different acquisition times and having different temporal/spectral resolutions. As we know, Independent component analysis (ICA) is a statistical technique which attempts to discover hidden sources or features from a set of observed data [7]. Typically, it represents a generative model where the sources are maximally independent, the observations are assumed to be linear mixtures of independent sources. As a part of our research, we have adapted the linear static ICA model for multi-temporal case. To achieve this goal, we propose a new data organization. These contributions are illustrated in the following sections.

3.1. Data Organization

Regularly, in widely applications, real-world data sets possess a particularprocessing scheme in addition to the necessary instantaneous independence required by ICA. If we consider a set of multi-temporal hyperspectral images,collected measurements contain both temporal and spatial indices. So, a data entry $x = x(x, y, \lambda, t)$ can depend on spatial position (x, y), *wavelength* λ as well as time t. Conventionally, we want to consider data sets $x(r, t)$ depending on two indices r and t, where $r \in \Re^n$ can be a multi-dimensional index and t indexes the time axis. For the case of hyperspectral images, we could assume $t \in [1 : T] := [1, 2, \ldots, T]$ and $r \in [1 : x] \times [1 : y] \times [1 : w]$, where T is the number of images, $x \times y$ is the images size and w is the bands numbers. So the number of spatial observations is ${}^s m := xyw$ and thenumber of temporal observations ${}^t m = T$.

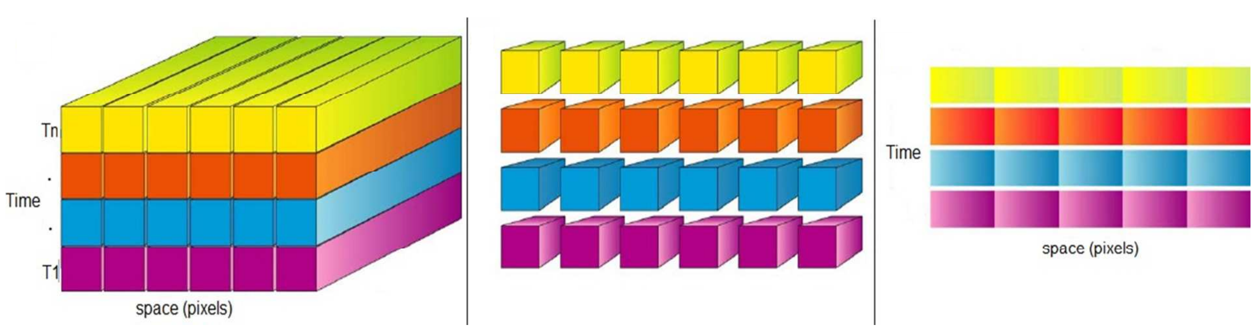
In the following, the multi-dimensional index r is contracted into an one-dimensional index by column or row concatenation. From that time, the data set $x(r, t) := x_{rt}$ can be represented by a data matrix X of dimension ${}^t m \times^s m$. The goal of the proposed spatio-temporal technique is to determine either a spatial sources matrix ${}^S S$ (endmembers) or a temporal source matrix ${}^t S$ (*time courses*). After mean removal wecan without loss of generality assume that the mixtures are spatiotemporallycentered.

3.2. Spatio-temporal ICA for Multi-temporal Endmembers and Change Analysis

Unlike principal component analysis (PCA) which uncorrelates with the data, ICA use higher-order statistics to achieve independence. Chang et al. [2] introduced ICA for hyperspectral images analysis, with the assumption that hyperspectral data is a mixture of spatially independent components. From which, ICA is becoming a very successful tool for endmembers identification and it gives very promising results. With the success of ICA in satellite images processing, there is a strong interest in this technique for the analysis of spatio-temporal data. Hyperspectral data sets contain mixtures of many different sources of variability. Endmembers may overlap spatially and temporally with errors due to atmospheric conditions, spectral distortions and other noise sources. In our research, we propose to extend individual ICA by analyzing time series hyperspectral images. Therefore, the proposed approach is performed by putting the sequence of hyperspectral images in one spatial-temporal multidimensional space as explain in previous section.

Properties such as non-Gaussianity and spatial/temporal independence of sources need to be addressed for the application of ICA to multi-temporal hyperspectral data. If the endmembers do not have a systematic overlap in time and/or space, then the distributions can be considered independent [5]. The signals of interest in hyperspectral data have typically a super-Gaussian spatial distribution. However, the noise will be more varied and potentially sub-Gaussian [2].

Given the input matrix: $X = [X_{i,j}] \in \Re^{r \times T}$ matrix containing a sequence of T images $X_1, \ldots, X_T, X_{ij}$ represents the reflectance of the ith pixel associated with the jth sample (time



(a) (b) (c)

Figure 1: Schematic diagram of the data structure for spatiotemporal ICA. (a) The data matrix in its initial form can be considered a cube with pixels represented in a full data matrix. For stICA, the cube is "broken up" (b) and the wavelength axe for each pixel is rotated into the same dimension as space. (c) The two-dimensional matrix now have the pixel values stacked end to end such that space and wavelength values are represented in the rows of the data matrix.

point), $(i = 1, \ldots, r; j = 1, \ldots, T)$ is modelled as:

$$X = SA \tag{1}$$

where $S \in \Re^{r\times T}$ and $A \in \Re^{k\times T}$ are respectively source and mixing matrix, k is the sources number. Thus, separating multi-date hyperspectral data into spatio-temporal independent components involves determining r-dimensional maps which are multi-temporal endmembers and their associated time courses of activation (abundance fractions). Therefore, the signal at each pixel can be shown as a mixture of underlying sources signals (Eq. (2)).

$$X_i = \begin{bmatrix} a_{11} \\ \ldots \\ a_{1T} \end{bmatrix} S_1 + \begin{bmatrix} a_{21} \\ \ldots \\ a_{2T} \end{bmatrix} S_2 + \ldots + \begin{bmatrix} a_{K1} \\ \ldots \\ a_{KT} \end{bmatrix} S_K \tag{2}$$

If the data is considered to be composed of T spatial patterns $\vec{x}_t(\vec{r}\,) := x(\vec{r}, t)$, this case is known as spatial ICA ($sICA$). Spatial independent component analysis (sICA)aims to decompose the spatial observation vector ${}^S\vec{x}(\vec{r}\,) := (x_{t1}(\vec{r}\,), \ldots, x_{tT}(\vec{r}\,))^T$ into ${}^SX(\vec{r}\,) := {}^SA^S\vec{S}(\vec{r}\,)$, where ${}^SA$ denotes the spatial mixing matrix and ${}^S\vec{S}(\vec{r}\,)$ the spatial sources. This is analogous to the following factorization: $X := {}^SA^SS$.

In the case of temporal ICA ($tICA$) the data is considered as a collection of time series $X_{\vec{r}}(t) := x(\vec{r}, t)$ for each spatial location $\vec{r}$. ICA is then applied to the temporal observation vector ${}^t\vec{x}(t) := (x_{\vec{r}_1}(t), \ldots, x_{\vec{r}\vec{s}_m}(t))^T$ that contains ${}^sm$ entries. The goal is to find a decomposition ${}^t\vec{x}(t) := {}^tA\,{}^t\vec{S}(t)$ with a temporal mixing matrix ${}^tA$ and temporal sources ${}^t\vec{S}(t)$. This is analogous to the following factorization: $X := {}^TA^TS$.

The choice of spatial or temporal independency is controvertible. Nevertheless,spatio-temporal ICA seeks the following factorization $X := {}^SS\,{}^tS$ with a spatial source matrix ${}^SS$ with column vectors ${}^S\vec{S}_i$ and a temporal source matrix ${}^tS$ with column vectors ${}^T\vec{S}_i$, which both have to fulfill the constraints imposed onto them.

For multi-temporal hyperspectral data processing, sources (endmembers) should be spatially independent and linearly mixed by time courses which are vectors of size $(1 * N)$. The first step in the proposed approach consists to perform a Singular Value Decomposition (SVD). Motivations for this choice are two-fold. In view of hyperspectral data dimensionality, SVD is used to provide a reduced rank data set as input to ICA. Moreover, SVD is adopted for ICA methods evaluation. Then, we define: $\tilde{X} = X$ by:

$$\tilde{X} = X = UDV^T == (UD^{1\backslash 2})(VD^{1\backslash 2})\tilde{U}\tilde{V}T^T \tag{3}$$

Given $X = \tilde{U}\tilde{V}T^T$, spatio-temporal ICA was used to find the following decomposition:

$$\tilde{X} = S_S \Lambda S_T^T \tag{4}$$

where S_S contains a set of k independent r-dimensional independent images, and S_T as a set of k independent temporal sequences of length T, and Λ is a diagonal scaling matrix. There exist two $k \times k$ mixing matrices, W_S and W_T such that $S_S = \vec{U}W_S$ and $S_T = \vec{V}W_T$.

As we can see, Spatio-temporal ICA enforces independence constraints over space as well as over time. In a linear decomposition of hyperspectral data, the data matrix can be transformed into a set of spatial independent components by taking linear combinations of multi-temporal endmembers.

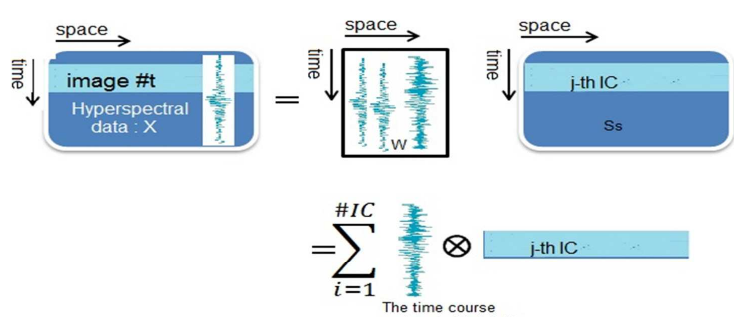


Figure 2: Spatio-temporal unmixing of hyperspectral images.

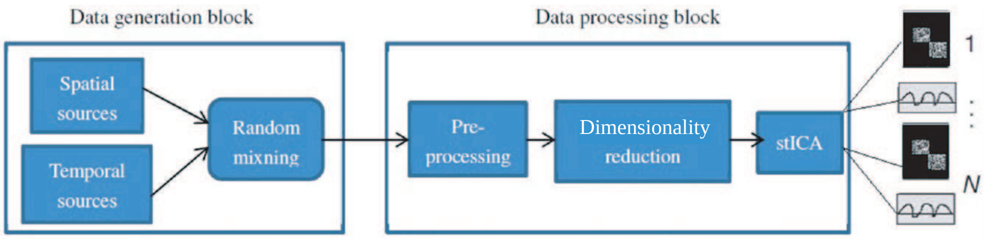
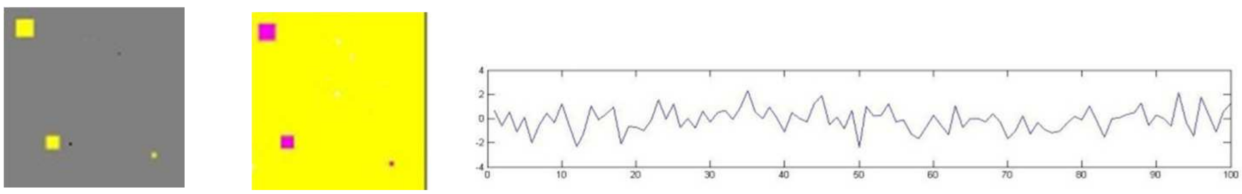


Figure 3: A model for applying stICA to multi-temporal hyperspectral data.



(a) (b) (c)

Figure 4: (a) Ground truth; (b) Changes detected by *stICA*; (c) Sample of extracted time courses.

4. RESULTS

To validate our approach, we have developed the model shown in Figure 3; it provides a framework containing various stages applied to the processing of multi-temporal hyperspectral data.

The data generation block consists of a set of statistically independent spatial endmembers (sources) which have been randomly mixed with some temporal sources. In order to assess the ability of *stICA* in multi-temporal endmembers recovering, we used the Matlab implementation of the *stJADE* algorithm proposed in [8]. The data processing block consists of a pre-processing step, including atmospheric correction, geometric correction, etc. It is also common to perform a data reduction. Finally, *stICA* was performed to generate multi-temporal sources which are the endmembers in our case. Figure 4(b) shows the detected endmembers (in pink) obtained by using *stICA* over the simulated temporal dataset. The associated time course is illustrated by Figure 4(c).

The obtained results correspond accurately with visual inspection of the simulated images.To validate this result, we have also compared the proposed approach with mono-temporal ICA performed for each image. Notably *stICA* across data performed slightly better than mono-temporal ICA. This suggests that the idea of imposing independence across times or images are not to be dismissed. Spatio-temporal ICA has made it possible to reveal multi-temporal endmembers that were not detected by purely spatial mono-temporal ICA. Therefore, the proposed framework provided very promising recognition performance even in noise presence and small sample size conditions.

5. CONCLUSIONS

The main purpose of this paper is the development and testing of new multi-temporal techniques for remote sensing data based on spatio-temporal ICA.Conclusively, understanding the data structure and analyzing a dimensional complex-valued data provides greater challenges for our study. Further work on the application of ICA to hyperspectral data includes improving the identification accuracy such that we can fully utilize the information contained in the data. This can be achieved by the incorporation of prior information about the sources and their nature without unnecessary constraining the ICA estimation result. Grateful selection and incorporation of prior information can prove to be even more beneficial for ICA of hyperspectral data.

REFERENCES

1. Farah, I. R., S. Hemissi, K. S. Ettabaa, and B. Soulaiman, "Multi-temporal hyperspectral images unmixing and classification based on 3D signature model and matching," *PIERS Online*, Vol. 6, No. 5, 480–484, 2010, doi:10.2529/PIERS091219165514.
2. Jin, H., P. Li, and W. Fan, "Land cover classification using multitemporal CHRIS/PROBA images and multitemporal texture," *IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, Vol. 4, IV-742–IV-745, 2008.

3. Van de Voorde, T., L. Demarchi, and F. Canters, "Multi-temporal spectral unmixing to characterise change in the urban spatial structure of the Greater Dublin area," *28th EARSeL Symposium and Workshops: Remote Sensing for a Changing Europe*, Istanbul, Turkey, June 2–5, 2008.
4. Freitas, R. M., Y. E. Shimabukuro, R. R. Rosa, and A. Huete, "Fraction images derived from EO-1 Hyperion multitemporal data for dry season green up analysis in Tapajós National Forest, Brazilian Amazonia," *IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, Vol. 4, IV-1015–IV-1018, 2009.
5. Martinez, P., R. M. Perez, J. J. Merino, D. Valencia, A. Plaza, and J. Plaza, "New spatial/spectral multitemporal endmember extraction algorithms," *ESA-EUSC Image Information Mining Workshop*, Madrid, Spain, 2006.
6. Hyväinen, A. and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks*, Vol. 13, No. 4–5, 411–430, 2000.
7. Chiang, S.-S. and C. I. Chang, "Unsupervised hyperspectral image analysis using independent component analysis," *IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, Vol. 7, 3136–3138, 2000.
8. Theis, F. J., P. Gruber, I. Keck, A. Meyer-Baese, and E. W. Lang, "Spatiotemporal blind source separation using double-sided approximate joint diagonalization," *EUSIPCO 2005*, Antalya, 2005.

Available Seat Counting in Public Rail Transport

P. De Potter[1], **C. Billiet**[1], **C. Poppe**[1], **B. Stubbe**[2], **S. Verstockt**[1,3], **P. Lambert**[1], **and R. Van de Walle**[1]

[1]Multimedia Lab, Department of Electronics and Information Systems, Ghent University-IBBT, Belgium
[2]Televic Group NV, Belgium
[3]ELIT Lab, University College West Flanders, Ghent University Association, Belgium

Abstract— Surveillance cameras are found almost everywhere today, including vehicles for public transport. A lot of research has already been done on video analysis in open spaces. However, the conditions in a vehicle for public transport differ from these in open spaces, as described in detail in this paper. A use case described in this paper is on counting the available seats in a vehicle using surveillance cameras. We propose an algorithm based on Laplace edge detection, combined with background subtraction.

1. INTRODUCTION

Video analysis has become a necessity because of the never ending growth of the amount of video surveillance cameras. Different intelligent video surveillance systems have already been developed for the detection of unwanted events [1–3]. In public transportation as well, video surveillance cameras have made their entrance, and thus video analysis is needed. However, because the conditions in vehicles for public transportation are different, a different approach is necessary.

The primary goal of video surveillance cameras is of course video surveillance, but if the cameras are installed, they can also be used for other video analysis tasks. One of these tasks is available seat counting, which makes it possible to refer passengers to cars with empty seats and to gather statistical information.

In this paper, we propose a system that combines an illumination change invariant edge detection technique with a background subtraction technique that can detect the blobs representing the moving objects.

The remainder of this paper is organized as follows. In Section 2, we discuss some of the related work in this area. We elaborate on the specific conditions in vehicles for public transportation in Section 3. In Section 4, an approach is presented to count the available seats in a vehicle. An evaluation of this implementation is given in Section 5. Finally, conclusions and future work are given in Section 6.

2. RELATED WORK

While a lot of recent research is done on the topic of video analysis, the number of publications in the area of analysis inside moving vehicles is quite limited.

In [4], Milcent et al. present a system to detect baggage in transit vehicles. They preprocess the video stream to correct the lighting. A light location mask, indicating reflecting metallic posts inside the vehicle, is used to gather the different parts of one object. To increase the speed of the segmentation algorithm, it is only applied on a region indicated by a probability location mask.

Several projects, such as PRISMATICA (pro-active integrated systems for security management by technological, institutional and communication assistance [5]) and BOSS (on board wireless secured video surveillance [6]) mention the transmission of video feeds upon the triggering of an alarm, but do not describe how the alarm is exactly triggered.

In [7], Vu et al. present an event recognition system based on face detection and tracking combined with audio analysis. 3D context such as zones of interest and static objects are recorded in a knowledge base and 3D positions are calculated for mobile objects using calibration matrices. Strong changes in lighting conditions occasionally prevent the system to detect people correctly.

Yayahiaoui et al. [8] and Liu et al. [9] report a high accuracy in passenger counting using a dedicated setup. Since the cameras used for this setup can not be used for other purposes, this solution is too expensive to be used in real life. Also, it is impossible to retrieve the location of the passengers.

3. CONDITIONS IN VEHICLES FOR PUBLIC TRANSPORTATION

In this section, we elaborate on the conditions that are specific for vehicles for public transportation.

One of the big challenges in video analysis is dealing with illumination variances and shadows. Inside moving and turning vehicles, this problem becomes even worse. The angle of the incoming sunlight changes, driving into a tunnel reduces the illumination drastically and static and moving objects almost always cast a dark shadow.

Another problem that is enlarged in vehicles is that of occlusion. Because of the limited space, moving objects easily occlude each other and the biggest part of seated passengers cannot be seen because of occlusion by the seats.

The windows pose other problems: during the day, a fast moving background can be observed through this windows. When it gets darker outside, objects are reflected.

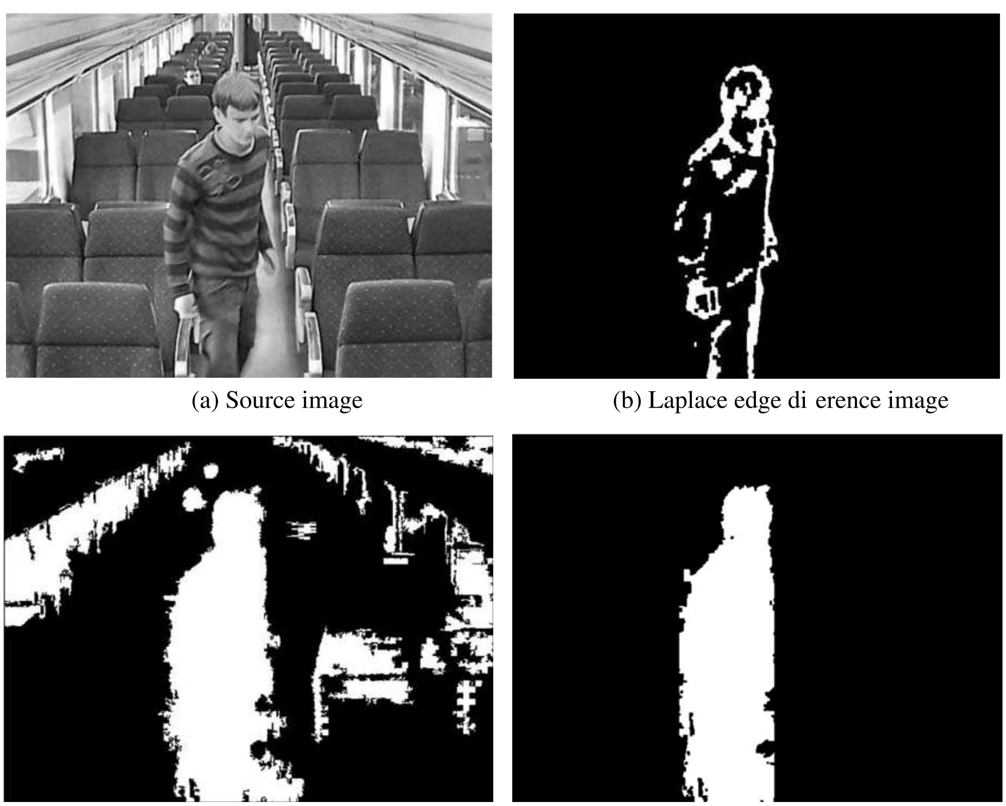
Since all the equipment has to be installed on the train, it has some limitations: the bandwidth between different pieces of equipment is low, the available processing power is limited and the installed cameras provide low quality video streams containing a lot of noise.

4. AN APPROACH FOR AVAILABLE SEAT COUNTING

In this section, our approach to count the number of occupied seats, from which the number of available seats can be derived, is described. In a first subsection, we describe how we detect moving objects using different methods. In a second subsection, we introduce the event detection that leads to the number of occupied seats.

4.1. Object Detection

The classification of pixels in foreground and background pixels is done in the object detection phase. We use different techniques for the object detection, as it has already been proven in the past that the combination of multiple techniques can reduce the individual weaknesses of these techniques [10]. The object detection consists of three consecutive steps: first, an edge detector is applied to discover the contours of moving objects. Secondly, a background subtraction method is used to retrieve blobs of potential foreground objects. A last step consists of merging the results



(a) Source image

(b) Laplace edge di erence image

(c) Background subtraction

(d) Combined result

Figure 1: Object detection.

of both techniques to obtain the blobs of the actual foreground objects. In this subsection, these three steps are described in more detail and illustrated using the image depicted in Figure 1(a).

To detect the contours of moving objects, a Laplace edge detector with a 7×7 pixel mask is used to extract the edges in a frame. The size of the mask is not too big, in order to keep the computation time within limits. The resulting edge image is then compared to the edge image of the previous frame to obtain a difference image containing the contours of moving objects. An example of such a difference edge image is given in Figure 1(b). The edge detector is quite independent on changes in illumination, since it only considers 2 frames.

For the background subtraction method, in a training phase the median value is calculated for every RGB channel of each pixel, as well as the average deviation from this median value. During the object detection phase, a pixel is then assumed to be a foreground pixel if its value differs more than 11 times the average deviation from the median value in each RGB channel. In a post processing phase, noise is eliminated by removing small connected areas. The areas that remain are interpreted as foreground areas and holes in these areas are filled up. The result of the background subtraction method is illustrated in Figure 1(c).

The results of both techniques are merged by taking the result of the background subtraction method and removing all the blobs that do not contain a significant amount of edges in the result of the edge detector method. By doing this, only the blobs corresponding to moving objects remain, as shown in Figure 1(d). It can be seen that only the moving, standing person is detected as a moving object, while the sitting persons in the upper, left-hand side of the image, are not detected as moving objects. This is not necessary, since they stay seated and thus no events need to be detected by the event detection mechanism.

4.2. Event Detection

The counting algorithm is based on the following principle: the number of seated passengers can only be increased or decreased when a passenger sits down or leaves a seat. The total number of seated passengers can thus be determined by counting the sit and leave actions.

In order to count these actions, multiple rectangular regions are identified on a camera view of the vehicle, as illustrated in Figure 2(a). These rectangular regions can be divided into three groups: the regions representing the left-hand side seats, the aisle and the right-hand side seats. The regions in the seats groups are split up in smaller rectangles, named tiles below, as illustrated in Figure 2(b), to improve the results.

The creation of the rectangles and tiles is done manually, but since it is a static environment and only a couple of seat configurations exist, this only has to be done a few times.

A tile is triggered when at least half of its pixels are detected as foreground pixels. When half of the tiles of a rectangle are triggered, the rectangle is triggered and sit action detection is started. The order in which the tiles were triggered is checked and is compared with previous presence of foreground pixels in either the aisle or an adjacent seat region, depending on the situation. If the combination of the order in which the tiles are triggered and previous presence in adjacent regions makes sense, then a sit activity is registered. E.g., If the tiles in the region from a seat near the aisle are triggered in an order from the the aisle to the window and previous presence was registered in

(a) Rectangular regions: seats and aisle (b) Further splitting up of the seat regions in tiles

Figure 2: Identification of different regions in a camera view.

(a) Group 1: One actor (b) Group 2: Multiple actors not acting in group (c) Group 3: Multiple actors acting in group

Figure 3: The different categories of test sequences.

Table 1: Performance for the different groups of video sequences.

Group	Precision	Recall
1	100%	100%
2	100%	57.14%
3	85.71%	85.71%

the aisle, then a sit action is registered for one of the seats near the aisle. If the comparison makes no sense, then no action is taken.

A rectangle can also become untriggered after being triggered before, when at least half of the tiles are untriggered, that is when at most half of its pixels are detected as foreground pixels, after being triggered before. Again, the order in which the tiles are untriggered is checked and compared to presence in adjacent regions. In this situation, a leave action is registered when the combination of the order and presence makes sense.

5. EVALUATION

The performance of our algorithm was tested on 78 acted sequences recorded in a passenger coach provided by the NMBS (national railway company of Belgium). The resolution of the sequences is 640×480 pixels and the framerate is 25 fps. The sequences are divided by difficulty into three groups: in the first group, only one actor is present. In the second group, multiple actors are present, but they do not perform group actions. In the third group, multiple actors are present and they do perform actions, such as entering the vehicle and sitting down, in group. An example of each of these categories can be seen in Figure 3.

The ground truth for these sequences is build up using the knowledge of the actual number of persons sitting down at a certain moment and can thus differ from the number of persons sitting down as observed at a certain moment because of occlusion.

The accuracy of the proposed algorithm is given in Table 1, by the values for precision and recall for the different groups of sequences. Unfortunately, currently no datasets and results are available to compare with.

The lower recall value for group 2 then for group 3 can be explained by the fact that, when people want to go sit together, they have to wait turns before moving into their seats. When people do not move in group, there is a bigger chance they perform a sit or leave action at the same time, which can lead to an occlusion.

The algorithm has not yet been optimized for speed, but manages to process one frame every 60 milliseconds on a 2.2 GHz processor. Some speed optimizations are thus required to meet real time constrains, for which a maximum processing time of 40 milliseconds is demanded for a single frame of the 25 fps sequences. Of course, more adaptations would be needed to be able to run this software on embedded hardware.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we described the conditions that are specific for vehicles in public rail transport and give an approach to count the number of seated passengers under these conditions. The proposed solution performs quite well in terms of precision and recall, but would need adaptations before being able to run on embedded hardware.

One thought that came into mind was that it might be useful not only to use the same hardware as the video surveillance system, but also a part of the software. For example the results of the object detection algorithm could be shared. This could lead to a minimal extra load for additional video analysis tasks.

Future work consists of trying other methods for a robuster object detection, so that also other events, such as fights, could be detected. Also, other camera configurations, with cameras in the middle of the aisle can be investigated. The dataset we used for the evaluation will be published to make a comparison of our results with other results possible.

ACKNOWLEDGMENT

The research activities as described in this paper were funded by Ghent University, the Interdisciplinary Institute for Broadband Technology (IBBT), the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT), the Fund for Scientific Research-Flanders (FWO-Flanders), and the European Union.

REFERENCES

1. Haritaoglu, I., D. Harwood, and L. S. Davis, "W4: Real-time surveillance of people and their activities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 8, 809–830, 2000.
2. Bremond, F., M. Thonnat, and M. Zuniga, "Video understanding framework for automatic behavior recognition," *Behavior Research Methods*, Vol. 36, No. 3, 416–426, 2006.
3. Adam, A., E. Rivlinand, I. Shimshoni, and D. Reinitz, "Robust real-time unusual event detection using multiple fixed-location monitors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 3, 555–560, 2008.
4. Milcent, G. and Y. Cai, "Location based baggage detection for transit vehicles," 15 p, Carnegie Mellon University, CyLab, Pittsburgh (PA), December 2005.
5. Velastin, S. A., L. Khoudour, B. P. L. Lo, J. Sun, and M. A. Vicencio-Silva, "PRISMATICA: A multi-sensor surveillance system for public transport networks," *Proceedings of 12th IEE International Conference on Road Transport Information and Control*, 19–25, London, UK, April 2004.
6. Jeney, G., C. Lamy-Bergot, X. Desurmont, R. Lopez da Silva, R. A. García-Sanchidrián, M. Bonte, M. Berbineau, M. Csapodi, O. Cantineau, N. Malouch, D. Sanz, and J.-L. Bruyelle, "Communications challenges in the celtic-BOSS project," *Lecture Notes in Computer Science*, Vol. 4712, 431–442, 2007.
7. Vu, V.-T., Q.-C. Pham, J.-l. Rouas, F. Bremond, N. Allezard, S. Ambellouis, G. Davini, P. Sayd, A. Flancquart, and M. Thonnat, "Audio-video event recognition system for public transport security," *Proceedings of IET Conference on Imaging for Crime Detection and Prevention*, 414–419, London, UK, June 2006.
8. Yahiaoui, T., C. Meurie, L. Khoudour, and F. Cabestaing, "A people counting system based on dense and close stereovision," *Lecture Notes in Computer Science*, Vol. 5099, 59–66, 2008.
9. Liu, N. and C. Gao, "Bi-directional passenger counting on crowded situation based on sequence color images," *Lecture Notes in Computer Science*, Vol. 4282, 557–564, 2006.
10. Poppe, C., G. Martens, S. De Bruyne, P. Lambert, and R. Van de Walle, "Robust spatio-temporal multimodal background subtraction for video surveillance," *Optical Engineering*, Vol. 47, No. 10, 107203, 2008.

Image Processing Methods for Evaluating Infrared Thermographic Image of Electrical Equipments

Mohd Shawal Jadin[1], **Soib Taib**[1], **Shahid Kabir**[2], **and Mohd Ansor Bin Yusof**[1]

[1]School of Electrical and Electronic Engineering
Universiti Sains Malaysia, Nibong Tebal, P. Pinang 14300, Malaysia
[2]Collaborative Micro-electronic Design Excellence Centre (CEDEC)
Universiti Sains Malaysia, Nibong Tebal, P. Pinang 14300, Malaysia

Abstract— Infrared thermography is well known as one of the effective tools in monitoring the condition of electrical equipments. It has the capability to detect the thermal abnormality in electrical equipments. The recent research in this field has shown the interest on an automatic diagnosis system. This is due to fast analysis and robust compared to manual inspection. Various techniques have been used to identify and classify the thermal anomalies in the infrared thermographic image of electrical equipments. The common method that normally used in analyzing infrared thermogram can be divided into four steps: image preprocessing, segmentation, classification and decision making. The majority of the techniques were implemented in bottom-up order. Conversely, another approach that is more top-down oriented has been introduced recently. This paper presents the review of image processing methods for both approaches in classifying the level of faults in electrical equipments. Some advantages and disadvantages of both approaches are also discussed.

1. INTRODUCTION

In 1800, William Hershel has discovered infrared radiation and it was the first experiment that showed there were forms of light not visible to the human eye [2, 3]. The infrared wavelength spectrum is ranged from about 1 mm down to 750 nm. All objects emit energy proportional to its surface temperature. As infrared energy functions outside the dynamic range of the human eye, special equipment is required to transform the infrared energy to another signal, which can be seen. For this purpose, infrared imagers were developed to see and measure this heat. Nowadays, various types of IR imager with more advanced and sophisticated features have been developed [4]. The basic concepts of IR imager, commonly known as the thermographic camera, is that it can capture an image of the thermal pattern and measures the emissive power of a surface in an area at various temperature ranges. The digital output image of IRT is called as thermogram. Each pixel of a thermogram has the specific temperature value, and the image's contrast is derived from the differences in surface temperature.

2. TYPICAL FAULTS AND DIAGNOSIS METHODS

Faults in electrical power systems can be classified into a few categories such as poor connection, short circuit, overloading, load imbalance and improper component installation [4–7]. In most cases, poor connections are among the more common problems in transmission and distribution lines of electrical power systems [8]. According to a thermographic survey conducted during the period of 1999–2005 [9], it was found that 48% of the problems were found in conductor connection accessories and bolted connections. This is mainly due to loose connection, corrosion, rust, and non-adequate use of inhibitory grease. On the other hand, 45% of the thermal anomalies appear in disconnector contacts. Most of the anomalies are due to deformations, deficient pressure of contact, incorrect alignment of arms, and accumulation of dirt. Only 7% of the problems were found in electrical equipment. By utilizing IRT, the thermal image will clearly indicates the problematic area.

There are two methods that are commonly used in diagnosing the conditions of electrical power equipments which are qualitative and quantitative. The qualitative measurement is widely used method by employing the ΔT criteria [1, 12]. It is also known as comparative thermography [3]. When the comparative technique is used appropriately and correctly, the differences between the two (or more) samples will often be indicative of their condition. Fig. 1 shows the example of hotspot and the reference point. Hot area is the suspected component and the reference must be another similar component which has the same condition. It could be also a similar component in other phases.

The advantage of this method is that it is a practical method to establish "failure" or "no failure" and the emissivity has only a minor impact on the result [1]. A drawback is that the ΔT criterion does not say anything about whether the equipment's temperature limits are actually exceeded. Furthermore, using the ΔT criteria will not expose systematic failures affecting all three phases [12]. In quantitative measurement, the observation is established by measuring the absolute temperature of electrical equipment under the same ambient conditions. As the reference temperature has to be measured, it is requires an even greater understanding of the variables influencing the radiometric measurement, as well as its limitations. It is vital to determine what margin of error is acceptable before beginning an inspection, and to work carefully to stay within those bounds [3]. This will include all the variables such as ambient condition, type of object to be inspected, thermal imager specification as well as thermographer itself. In this case, related data and information must be collected and adjustment should be made accordingly.

3. IMAGE PROCESSING METHODS

Research in image processing incorporated with an intelligent system for diagnosing the condition of electrical equipments by evaluating its IRT image is still at the early stages. Only few researches have been done with some different techniques. In evaluating the condition of electrical equipment, the image processing methods can be classified into bottom-up and top-down approaches. Generally, the common image processing method that is normally used in an infrared thermogram can be divided into few steps: image preprocessing, segmentation, feature extraction, classification and decision. The straightforward approach is by following the steps one by one in bottom-up order. Conversely, top-down oriented is more problem specific process of identifying the interesting image region. Both methods yield a region of interest (ROI). Decision will be made base on the information extracted and evaluated from these regions. This part will highlight some image processing methods that have been used for diagnosing the IRT image of electrical equipments based on the above-mentioned approaches.

3.1. Bottom-up Approach

To identify the hotspot within the IRT image, the simplest way is using the thresholding technique. Thresholding technique has been a common technique for image segmentation due to its intuitive and simple implementation. Based on the gray-level histogram of an image, the target object is separated from the background at a specific threshold, T. A. thresholded image, $g(x, y)$ of an image, $f(x, y)$ is defined as

$$g\left(x,y\right)=\begin{cases}1 & if \quad f(x,y)\leq T\\ 0 & if \quad f(x,y)\leq T\end{cases} \tag{1}$$

The hotspot region can be segmented through selecting a suitable threshold value [13]. For an automatic thresholding technique, Otsu [10] method is widely used in various applications. Chou and Yao [1] have proposed an algorithm for segmenting the hotspot region which was inspired by Otsu method. Morphological image processing was used to extract the hotspot where the maximum gray pixel value determines the maximum temperature of the hotspot region. The reference

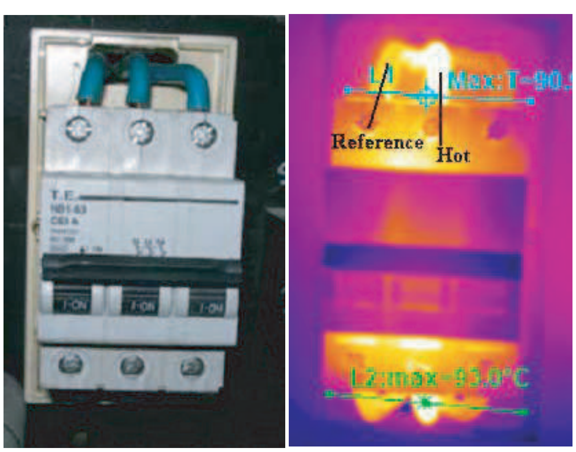


Figure 1: The hotspot and the reference point.

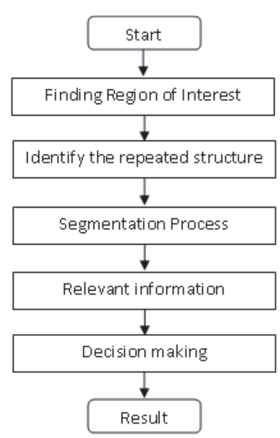


Figure 2: Flowchart.

temperature is derived from the average gray values of the object other than the hotspot region. The condition of the electrical equipment is calculated by comparing the hotspot temperature and the reference temperature. In another research, watershed transformation algorithm was successfully used for segmenting the thermogram of lightning arrester [13]. It is extremely noise and non uniform illumination robust procedure. However, the most critical problem with this technique is over-segmentation due to local minima attainment. The morphological gradient was used to eliminate the over-segmentation and produce a smooth edge of the lightning arrester.

Instead of using a gray level image, the original image has also been used for image segmentation and feature extraction processes. In this regard, the relevant information around the faulty point area is selected and all the pixel values outside the faulty area are set to zero [14]. Only fault region of gray level image is extracted using Zernike moment as the input features of support vector machine (SVM). It is also possible to extract features of a thermogram directly from its RGB data. This method is quite straightforward without implementing any applying advance image processing method. Nevertheless, the problem with this technique is high processing time due to the large feature vector to be computed by neural network algorithm [15].

3.2. Top-down Approach

Since the method of IRT evaluation of electrical equipment is using a qualitative measurement, all the similar and identical structures within the thermogram should be grouped together. Detecting a regular structure will be the first step in segmentation by grouping the repetitive structure. The common steps used in evaluating the thermal anomaly of electrical equipments based on top-down oriented can be summarized by a flowchart as shown in Fig. 2. In detecting the electrical equipment's structure, the region of interest (ROI) should be identified first. For finding the repeated objects or structures, the tasks can be broken down into two separate steps: (i) finding interesting features in the image and describing these using pre-specified descriptors, and (ii) comparing all the features and look for matches [16, 17]. In finding a repeated pattern in an image, one approach is to identify distinctive features in the image, describe the features, and compare them with each other to find similar regions within the image.

One of the most popular techniques to identify the features is using Scale Invariant Feature Transform (SIFT), which was introduced by David Lowe [18]. This method was implemented by [16, 17] for finding interest region of electrical installation. It first locates points of interest in a linear scale-space, and then assigns a descriptor vector constructed as local histograms of image gradient orientations around the point. The advantages of SIFT descriptor are invariant to image rotation and scale. Using the concept of feature matching in SIFT algorithm, the repeated object in the thermogram can be detected. The detected objects then segmented before enter the classification process.

There are three approaches were used to determine the thermal severity of an electrical equipment. The first is direct interpretation by identifying the maximum temperature for each region and evaluate the condition based on ΔT criteria. The maximum temperature is determined by finding the highest pixel value within the selected region. Calculating the histogram or histogram distance can also be used for finding the similarity between two objects. In this case, the histogram for each region is computed and compared with other region in order to get the ΔT. Another method is by analyzing the gradient of the segmented region. One of the advantages of utilizing the gradient analysis is that the source of the hotspot which was occurred in electrical equipment can be identified [16].

4. CONCLUSIONS

Both bottoms-up and top-down approaches were successfully implemented using various techniques in identifying fault in electrical equipment. Since the thermographic images of electrical equipment normally show a very high variation in appearance such as object distance from camera (scale change) and angle (view point change). The most suitable approach is describing the image using local features based on top-down approach, which are more robust to this kind of changes. Generally, local segmentation using top-down oriented has a great advantage over bottom-up approach. However, upgrading technique or introducing a new method could improve the diagnosing method of electrical equipments.

ACKNOWLEDGMENT

This research was supported by Fundamental Research Grant Scheme (FRGS), Universiti Sains Malaysia and Universiti Malaysia Pahang.

REFERENCES

1. Chou, Y. and L. Yao, "Automatic diagnostic system of electrical equipment using infrared thermography," *Proceedings of International Conference of Soft Computing and Pattern Recognition*, 155–160, 2009.
2. Braunovic, M., N. K. Myshkin, and V. V. Konchits, *Electrical Contacts Fundamentals, Applications and Technology*, CRC Press, 2007.
3. Chuck, H., *Handbook of Nondestructive Evaluation*, 1st Edition, McGraw-Hill Professional, 2001.
4. Epperly, R., G. Heberlein, and L. Ead, "A tool for reliability and safety: Predict and prevent equipment failures with thermography," *Proceedings of Petroleum and Chemical Industry Conference*, 59–68, 1997.
5. Cao, Y., X.-M. Gu, and Q. Jin, "Infrared technology in the fault diagnosis of substation equipment," *Proceedings of China International Conference on Electricity Distribution*, 1–6, China, 2008.
6. Dos Santos, L., E. C. Bortoni, L. E. Souza, G. S. Bastos, and M. A. C. Craveiro, "Infrared thermography applied for outdoor power substations," *Proceedings of SPIE*, 69390R-69390R-11, 2008.
7. Kregg, M. A., "Benefits of using infrared thermography in utility substations," *Proceedings of SPIE*, 249–257, 2004.
8. Azmat, Z. and D. Turner, "Infrared thermography and its role in rural utility environment," *Proceedings of Rural Electric Power Conference*, B2/1–B2/4, 2005.
9. Martínez, J. and R. Lagioia, "Experience performing infrared thermography in the maintenance of a distribution utility," *Proceedings of International Conference on Electricity Distribution*, Vienna, 2007.
10. Otsu, N., "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 9, No. 1, 62–66, 1979.
11. Li, B., X. Zhu, S. Zhao, and W. Niu, "HV power equipment diagnosis based on infrared imaging analyzing," *Proceedings of International Conference on Power System Technology*, 1–4, Chongqing, October 2006.
12. Lindquist, T. M., L. Bertling, and R. Eriksson, "Estimation of disconnector contact condition for modeling the effect of maintenance and ageing," *Proceedings of IEEE Russia Power Tech.*, 1–7, St. Petersburg, Russia, June 2005.
13. Laurentys Almeida, C. A., A. P. Braga, S. Nascimento, V. Paiva, H. J. A. Martins, R. Torres, and W. M. Caminhas, "Intelligent thermographic diagnostic applied to surge arresters: A new approach," *IEEE Transactions on Power Delivery*, 751–757, Vol. 24, No. 2, April 2009.
14. Rahmani, A., J. Haddadnia, and O. Seryasat, "Intelligent fault detection of electrical equipment in ground substations using thermo vision technique," *Proceedings of 2nd International Conference on Mechanical and Electronics Engineering*, Vol. 2, V2-150–V2-154, Kyoto, August 2010.
15. Shafi'i, M. A. and N. Hamzah, "Internal fault classification using artificial neural network," *Proceedings of 4th International Power Engineering and Optimization Conference*, 352–357, Shah Alam, June 2010.
16. Wretman, D., "Finding regions of interest in a decision support system for analysis of infrared images," Master Thesis, Royal Institute of Technology, 2006.
17. Smedberg, M., "Thermographic decision support — Detecting and classifying faults in infrared images," Master Thesis, Royal Institute of Technology, 2006.
18. Lowe, D. G., "Distinctive image features from scale-invariant key points," *International Journal of Computer Vision*, Vol. 60, No. 2, 91–110, 2004.

Fair-weather Atmospheric Electric Field Measurements at the Gaisberg Mountain in Austria

Helin Zhou[1], Gerhard Diendorfer[2], Rajeev Thottappillil[1], and Hannes Pichler[2]

[1]Division of Electromagnetic Engineering, School of Electrical Engineering
Royal Institute of Technology (KTH), Stockholm, Sweden
[2]Department of ALDIS (Austrian Lightning Detection & Information System)
Austrian Electrotechnical Association (OVE), Vienna, Austria

Abstract— A field mill (FM) has been permanently operated at a distance of about 170 m from the Gaisberg Tower (GBT) in Austria since several years. The electric field measurements suffer from field enhancement due to its location on a 4-m tall metal platform near the tower which itself is located on a 1280 meter high mountain. A special measuring campaign was conducted to determine the fair-weather atmospheric electric field at the Gaisberg Mountain on June 24th, 2010. The main objectives of this campaign were to calibrate the field mill in order to infer the relation between the electric fields at the tower tip and the ground level measured by the field mill under thunderstorm conditions. Besides the permanent field mill near the tower, two Campbell Scientific CS100 electric field meters were used during this campaign, and distances between each other were determined by using the Global Positioning System (GPS). Overall we determined an enhancement factor of 2.75 due to the mountain itself with reference to the mountains surrounding terrain. A field enhancement factor of 7.81 was obtained for the permanently installed field mill at the measurement platform next to the GBT with reference to the undisturbed electric field at the mountain top close to the platform at ground level. The electric field near the tower (distances about the tower height of 100 m) was smaller than the field measured at larger distance from the Tower. This observation was possibly caused by a shadowing effect of the tower.

1. INTRODUCTION

The amplitude of fair-weather atmospheric electric field is usually of the order of one to two hundred Volts per meter [1, 2]. It can be measured by a field mill on flat ground. If the field mill is mounted above the ground or installed at elevated objects, it will suffer from field enhancement. On the other hand, if the field mill is located close to tall objects (i.e., buildings, big trees, towers), the field mill reading will be lower due to the shadowing effect of these objects. Electric field enhancement factors at the top of buildings with different heights for lightning radiated fields are reported in the range of 1.5 to 2.3 compared with the measurements at ground level [3–6]. Next to the Gaisberg Tower (GBT) in Austria, a field mill (FM) is permanently operated since several years. The main purpose of those electric field measurements is to determine the background electric field that exists right at the time when upward lightning is triggered by the 100 m high tower. This radio tower on the mountain triggers about 60 upward initiated lightning flashes per year and is instrumented in order to measure the lightning current waveforms [7]. Knowledge of the background electric field at the mountain top will allow the estimation of the local electric field at the tower top by using a finite element model of the tower structure.

The FM is mounted on a 4-m tall metal platform at a distance of about 170 m from the GBT and hence the electric field measurements suffer from field enhancement. In order to get an approximate correction factor, simultaneous measurements are necessary. Besides the permanent field mill, a reference field mill is needed.

In this paper, we report the results of measured atmospheric electric fields under fair weather conditions at the Gaisberg Mountain in Austria. The enhancement factors due to the mountain itself and the elevated metal platform are investigated. Also the shadowing effect of the 100-m radio tower on the fair-weather electric field is presented.

2. INSTRUMENTATION

Two CS100 electric field meters manufactured by Campbell Scientific Inc. are employed for the simultaneous measurements. Unlike the traditional continuously rotating field mills, this type of electric field meter uses a reciprocating shutter, which reduces significantly power consumption. This feature makes it possible to operate the field mills powered by batteries and move them around easily. Each CS100 has been already calibrated in the factory when it is mounted at 2 m height tripod mast. An embedded CR1000M data-logger is used to record the data.

On the other hand, the permanently installed field mill was manufactured by Previstorm Inc. and is mounted on the 4-m tall metal platform at a distance of 170 m from the GBT. This field mill is in operation for several years. In order to synchronize the data measured by the mobile electric field meter, during the field campaign, the sampling rate of the permanent field mill was set to 1 Hz, the same sampling rate as used for the other two field mills.

3. MEASUREMENTS AND RESULTS

On June 24th, 2010, a field campaign was conducted in order to measure the fair-weather atmospheric electric field at the Gaisberg Mountain. During this campaign two mobile Campbell Scientific CS100 electric field meters and one permanent field mill were employed at different locations and distances between each other were determined by using the Global Positioning System (GPS).

Firstly, one CS110 Electric Field Meter was deployed and operated at a site of the surrounding area of the Gaisberg Mountain (786 m above sea level). After an initial test to verify that both CS110 field meters provided the same result when they were placed at the same location, the second CS110 was taken to measure the electric field simultaneously at sites of different altitudes along the way up to the top of the Gaisberg Mountain. Due to the mountain itself increasing field enhancement with increasing altitudes of the measuring sites were found. In Figure 1, we present as an example the simultaneously measured field from the reference site at the surrounding terrain and field at the top of the Gaisberg Mountain (site was at a distance of 249 m from the GBT). An arithmetic mean of 66 V/m at the surrounding terrain and 181 V/m at the top of the Gaisberg Mountain is obtained, respectively, corresponding to an enhancement factor of 2.75 due to the mountain itself.

In order to determine the enhancement factor for the field mill mounted on the 4-m tall metal platform at a distance of 170 m from the GBT, we deployed a CS100 electric field meter at a distance of 185 m nearby to the metal platform but at ground level. Simultaneous measurement results are shown in Figure 2. An electric field arithmetic mean value of 181 V/m at the ground level and 1406 V/m at the 4-m tall metal platform is obtained, corresponding to an enhancement factor of 7.81 due to the 4-m tall metal platform.

Similar to the measurements along the way up to the mountain top, at the local area of the Gaisberg mountain top, one CS110 was deployed at a distance of 249 m, assuming this was far enough from the 100 m high GBT to have negligible effect of the metal tower on the fair weather electric field. The second CS110 was used to measure simultaneously the electric field at different distances from the tower and ranging up to about the height (100 m) of the tower. Figures 3, 4 and 5 show the simultaneously measured data obtained from the two identical CS110 electric field meters. The mean ratios of the field values measured at 111 m, 85 m and 50 m to the reference values measured at 249 m from the GBT are 0.65, 0.73 and 0.71, respectively. We can see that

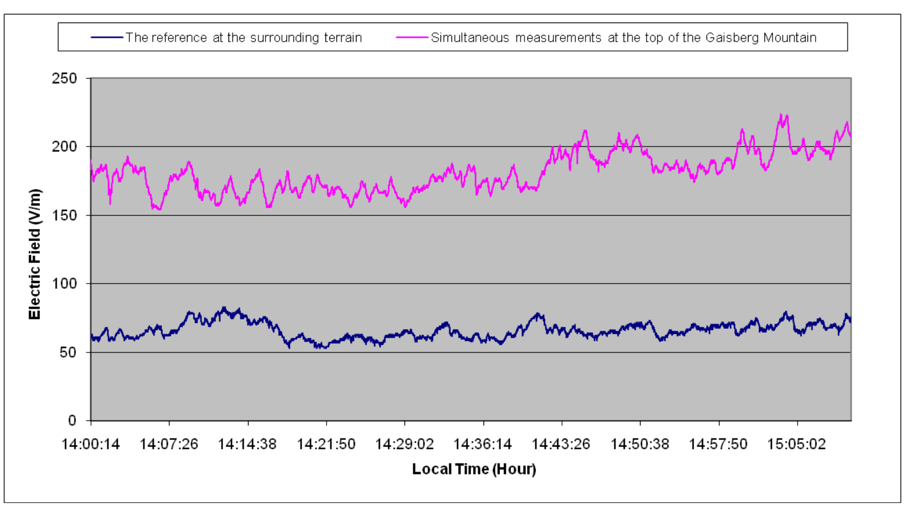


Figure 1: Simultaneously measured fields at the surrounding terrain and at the top of Gaisberg Mountain.

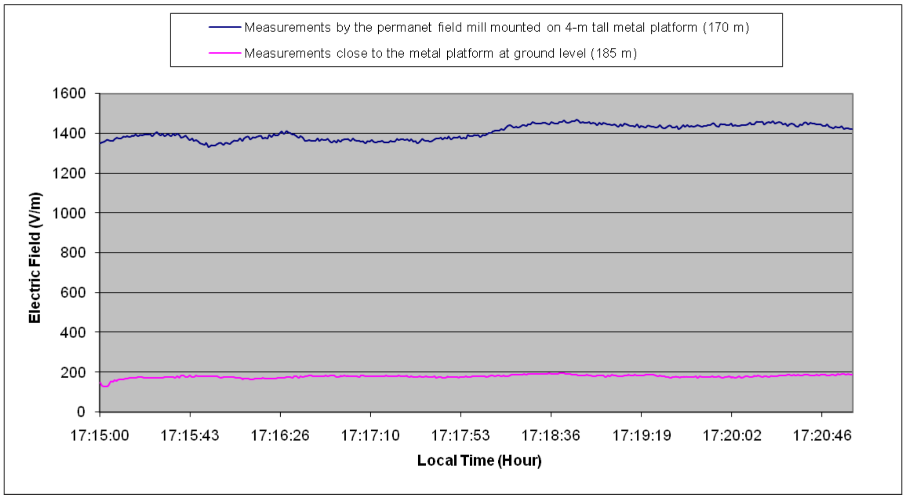


Figure 2: Simultaneously measured fields at the 4-m tall metal platform and at ground level.

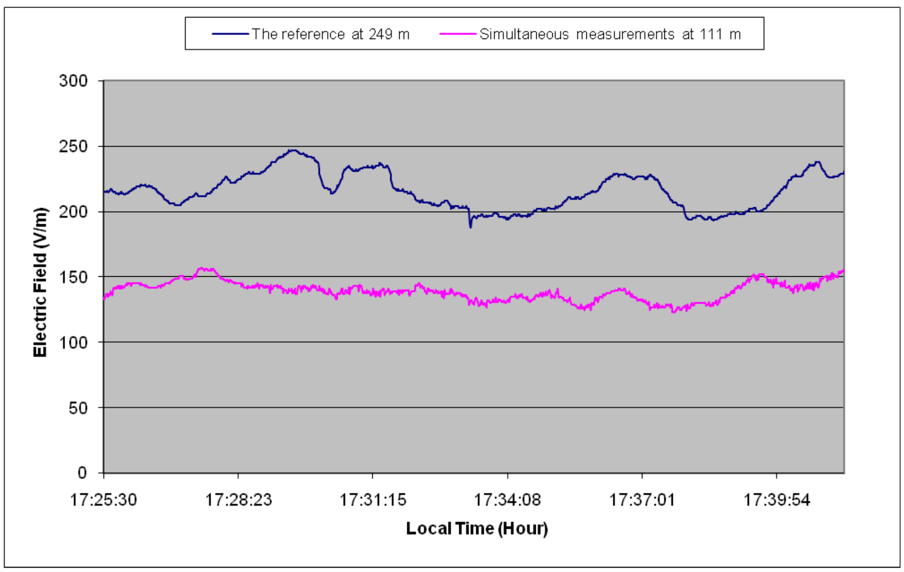


Figure 3: Simultaneously measured fields at distances of 249 m and 111 m from the GBT.

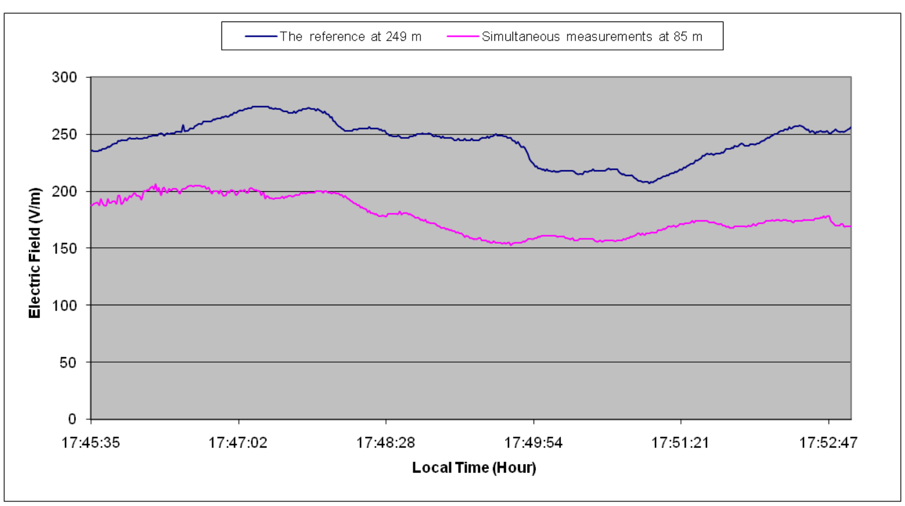


Figure 4: Simultaneously measured fields at distances of 249 m and 85 m from the GBT.

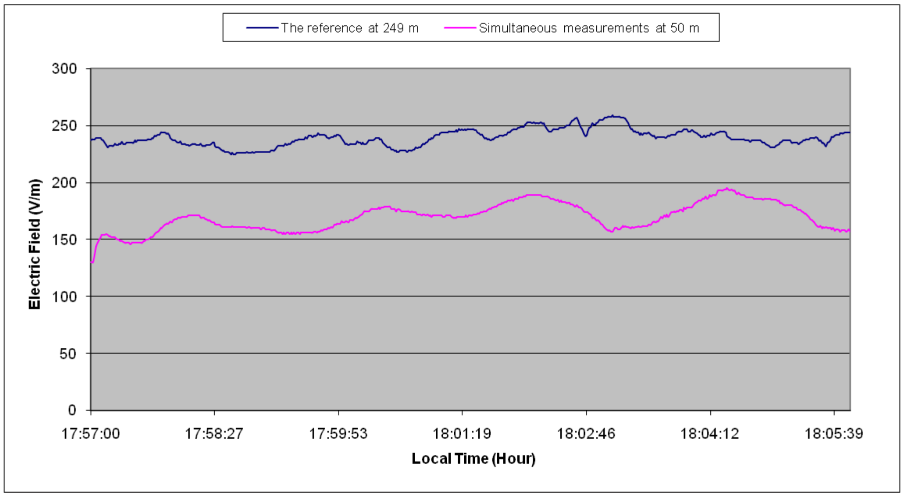


Figure 5: Simultaneously measured fields at distances of 249 m and 50 m from the GBT.

relatively smaller values were obtained at closer distance compared to the field measured at larger distance from the tower. This observation was possibly caused by a shadowing effect of the tower. The shadowing effect of the tower has also been observed in transient electric fields from lightning strikes to the GBT measured at very close distances [8].

4. CONCLUSION

In this paper, we present results of fair-weather electric field measurements during a campaign conducted at the Gaisberg Mountain in Austria at the end of June of 2010. By comparing the simultaneously measured electric fields from a reference field mill placed at the top of the mountain and the permanently installed field mill at the metal platform at a distance of about 170 m from the GBT, a field enhancement factor of 7.81 for the permanently installed field mill was determined. We also determined an enhancement factor of 2.75 due to the mountain itself with reference to the mountains surrounding terrain. A shadowing effect of the tower was observed when the electric field was measured at distances of up to about the tower height of 100 m compared to the reference field mill measured at a distance of 249 m from the tower.

ACKNOWLEDGMENT

The authors acknowledge the financial support received from the Swedish Research Council (VR Grant: 621-2005-5939), and the Swedish National Rail Administration (Banverket), Verbund (Contract 4500157745), the Austrian Science Fund (Project P17336-N07), Telekom Austria and Austrian Broadcasting Services (ORS). The authors also acknowledge the support of the European COST Action P-18 'Physics of Lightning Flash and its Effects'.

REFERENCES

1. Israelsson, S. and H. Tammet, "Variations of fair weather atmospheric electricity at Marsta Observatory, Sweden, 1993–1998," *J. Atmos. Sol. Terr. Phys.*, Vol. 63, No. 16, 1693–2003, 2001.
2. Guha, A., B. K. De, S. Gurubaran, S. S. De, and K. Jeeva, "First results of fair-weather atmospheric electricity measurements in northeast India," *J. Earth Syst. Sc.,* Vol. 119, No. 2, 221–228, 2010, Doi: 10.1007/s12040-010-0014-9.
3. Rubinstein, M., E. Montandon, and M. Ianoz, "Analysis of multi-station cloud lightning electric field pulses recorded with the swiss LPATS network," *Proceedings of 22nd International Conference on Lightning Protection*, Budapest, Hungary, 1994.
4. Bonyadi, R. S., R. Moini, S. H. H. Sadeghi, and A. Mahanfar, "The effects of tall buildings on the measurement of electromagnetic fields due to lightning return strokes," *Proceedings of 2001 IEEE EMC International Symposium*, Montreal, Canada, 2001.
5. Baba, Y. and V. A. Rakov, "Electromagnetic fields at the top of a tall building associated with nearby lightning return strokes," *IEEE Trans. Electromagn. Compat.*, Vol. 49, No. 3, 632–643, 2007.

6. Mosaddeghi, A., D. Pavanello, F. Rachidi, M. Rubinstein, and P. Zweiacker, "Effect of nearby buildings on electromagnetic fields from lightning," *Journal of Lightning Research,* Vol. 1, 52–60, 2009.
7. Diendorfer, G., H. Pichler, and M. Mair, "Some parameters of negative upward-initiated lightning to the gaisberg tower 2000–2007," *IEEE Trans. Electromagn. Compat.*, Vol. 51, No. 3, 443–452, 2009.
8. Mosaddeghi, A., A. Shoory, F. Rachidi, G. Diendorfer, H. Pichler, D. Pavanello, M. Rubinstein, P. Zweiacker, and M. Nyffeler, "Lightning electromagnetic fields at very close distances associated with lightning strikes to the Gaisberg tower," *J. Geophys. Res.*, Vol. 115, D17101, 2010, Doi: 10.1029/2009JD013754.

An Engineering Approach in Modeling Lightning Effects on Megawatt-class Onshore Wind Turbines Using EMTP and Models

Y. Méndez Hernández[1], **G. Drobnjak**[1], **A. Claudi**[2], **and M. Kizilcay**[3]

[1]GE Global Research Europe, Germany
[2]University of Kassel, Germany
[3]University of Siegen, Germany

Abstract— This publication addresses the topic of the effects of traveling waves on onshore wind parks caused by direct lightning strikes with main focus on predicting potential risk of damage due to overcurrents and overvoltages. The simulation program EMTP-ATP, programming language MODELS, pre-processing program ATPDraw and post-processing program PLOTXY were chosen for this study. This approach is an attempt to reflect an insight on this topic from the engineering point of view.

1. INTRODUCTION

With increasing nacelle heights and rotor blade lengths in modern wind turbines (WTs) the unexpected effects of lightning and electromagnetic traveling waves in form of overvoltages and overcurrents are brought to the attention of the scientific community, especially wind parks (WPs) with several WTs and rated power of several megawatts e.g., 100 MW and above. In some cases the earthing (grounding) systems of the WTs are galvanic connected to each other in order to ensure a common earth potential in the whole electric installation; however these techniques can lead to other effects such as overvoltages caused by direct lightning strikes.

2. APPROACH

The approach of modal components (surge impedance with a propagation velocity) was chosen for this work, this in order to model the fast electromagnetic transients and provide an acceptable frequency response of the models. The simulation is divided in the following components:

- Lightning source represented by a Heidler source.
- WT Rotor blades, which were modeled by surge impedances with a specific propagation velocity.
- Rotor blade and Nacelle Azimuth's bearings, that were represented as concentrated elements calculated at an electric frequency of 500 kHz.
- Steel tower, which was modeled as a surge impedance per segment.
- Earthing (grounding) system, modeled as a variable footing resistance.
- Low voltage cables, represented as a PI-equivalent diagram.
- Low voltage panelboard, modeled as a surge capacitance.
- Linear distribution transformer, including LV and MV bushings and coil surge capacitors.
- Medium voltage cables, represented as a JMarti model calculated at an electric frequency of 1 MHz.

3. WIND PARK MODELS FOR FAST TRANSIENTS EFFECTS

Figure 1 shows the general assumptions and references for the simulation, where, ii corresponds to the index of the conductor i; h_i the height of the conductor above ground and, r_i the radius of the conductor i. The surge impedances are assumed to have similar values in the rotor blades R, S and T.

3.1. Rotor Blades

The rotor blade lightning protection system (LPS) chosen for model was the receptor-based lightning protection system, which consists in this case of two metal air termination systems (so-called lightning receptors) embedded within the aerodynamic glass fiber composite (GFC) structure or skin of the rotor blade and galvanically connected to an internal down-conductor and widely explained in [1] and [2].

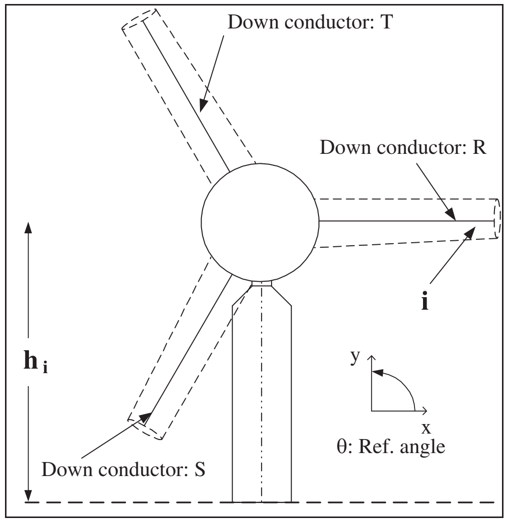


Figure 1: Surge impedance references.

Table 1: Parameter calculation for the rotor blades, using the theory in [4].

	μ_r	L [mH]	ε_r	C [µF]	Z_w [Ohm]	Length [m]	v [m/µs]
Blade Segment 1	1	4.14E-2	2.4	2.35E-5	4.20E0	18.75E0	1.50E2
Blade Segment 2	1	4.14E-2	2.4	2.35E-5	4.20E0	18.75E0	1.50E2

Table 2: Parameter calculation for the rotor blade and azimuth bearings.

	R_{bear} [Ohm]	L_{bear} [mH]	C_{layer} [µF]	Flashover Threshold [V]	C_{lub} [µF]
Rotor Blades Bearings	3.17E-6	7.67E-8	4.72E-2	1.00E-3	1.06E-4
Azimut Bearing	2.06E-5	7.59E-8	1.17E-4	1.00E-3	2.63E-4

The blades were modeled in two segments (representing each lightning receptor and its internal downconductor connection), each one with own characteristic surge impedance Z_w and a propagation velocity v, as explained in [4]. The magnetic and capacitive coupling between the rotor blades were neglected; due to the reason that special interest and focus was made on the effects of direct lightning; however, these couplings may influence the traveling waves patterns (reflection and transmission). Table 1 shows the parameters used for the calculation.

Special consideration to this first modeling of the rotor blades should be made, due to the reason that it was adopted from the transmission line theory, indeed this is an attempt to model the complex electromagnetic nature of the rotor blades during lightning strikes.

3.2. Rotor Blade and Azimuth Bearings

The bearing's parameters were calculated with a FEM-Program (Maxwell-3D Ver. 11.1.1) and imported as concentrated linear electric circuit elements to the main simulation program. The bearing's metal parts (flanges and metal balls mainly) were modeled as a resistance R_{bear} in series with an inductance L_{bear} calculated at an electric frequency of 500 kHz in order to consider the skin effect. The non-conductive parts such as the thin layer of lubricants between the bearing balls were modeled as a thin capacitor C_{layer} and the surrounding lubrication layers with a shunt capacitor C_{lub}, which are short-circuited with a voltage-controlled switch after a certain flashover voltage threshold has been reached. Table 2 depicts the parameters used for the bearings, which are connected in series.

3.3. Tower

The tower is usually manufactured in steel or reinforced concrete. The surge impedance model proposed in [5] was chosen for a steel tower, where the tower height and base radius of the equivalent

Table 3: Parameter calculation for the tower.

	Z_{Tower} [Ohm]	Length [m]	v [m/μs]
Tower Segment	1.98E2	2.90E1	2.50E2

Table 4: Parameter input values for the calculation for LV and MV cables.

	Radius [m]	Cross-Section [m]	Length [m]	ε_r	ρ [Ωm]	Cable Array
LV Cables	1.40E-2	6.16E-4	1.50E1	2.00E1	2.30E-8	3 Conductors/Phase
LV Earthing Cable	7.00E-3	7.00E-3	1.50E1	1.00E1	2.30E-8	1 Conductor
MV Cables	2.00E-2	1.26E-2	3.00E2	2.70E1	2.30E-8	1 Conductor/Phase
MV Earthing Cable	2.00E-2	1.26E-2	3.00E2	2.70E1	2.30E-8	1 Conductor

Table 5: Parameter input values for the distribution transformer.

	MV [kV]	LV [kV]	R_{sec} [Ω]	L_{sec} [mH]	R_m [Ω]
Transformer	33.00E1	0.69E0	2.00E-3	5.20E-2	1.00E-6

cone of the tower are used as input parameters to obtain the equivalent surge impedance Z_{tower}. For this calculation a tower of 3 segments is proposed, whereby every segment observes the same surge impedance value. Table 3 depicts the parameters.

In this publication the tower was assumed to be the only down-conductor for the lightning current; however some WT manufacturer install earthing down-conductors inside the tower, which are out of the scope of the study.

3.4. Earthing (Grounding) System and Earthing Rods

The earthing type B ring electrode system was chosen for this study and is usually installed in onshore WTs; it comprises a ring earth electrode external to the structure in contact with the soil or a foundation earth electrode [1]. The footing resistance R_T can be modeled from the guidelines proposed in [5] and widely explained in [2].

The soil resistivity assumed in this work for the calculation of the parameters was 500 Ωm. The earthing rods were modeled as a resistance in parallel with a capacitor as proposed in [6].

3.5. Low and Medium Voltage Cables

The non-shielded LV-cables connecting the WT panelboard to the LV side of the WT distribution transformer were modeled with a PI-equivalent model, three conductors/phase (equivalent to a 1000 Kcmil in AWG- notation) and one additional copper conductor for earthing were modeled.

The MV-cables shielded cables (sheath) connecting the primary side of the WT transformers were simulated and calculated with the JMARTI model of EMTP-ATP at 1 MHz (10 decades, 9 Points/Dec). Table 4 depicts the dimensions and lengths of the cables.

3.6. Three-phase Distribution Transformer

Each WT is connected to a distribution transformer, which was also modeled in order to address the overvoltage effects on the primary and secondary sides. Table 5 depicts the input parameters of the transformer; the impedances are referred to the secondary side. The surge capacitances were modeled with the guidelines proposed in [6].

4. SIMULATION RESULTS AND ANALYSIS

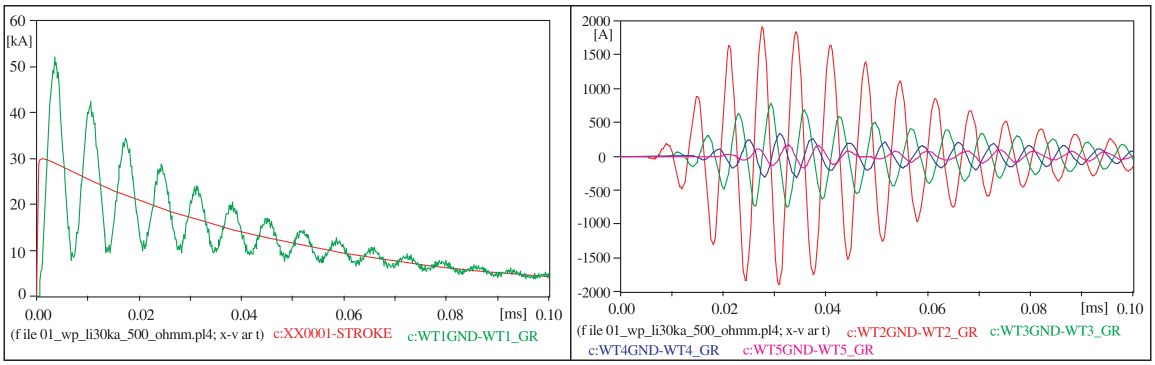
A total of five WTs were modeled, which are connected to a medium voltage Thevenin-equivalent network with a short circuit three-phase impedance (75 MVA, 33 kV Z_{base}); this in order to simulate the medium voltage connection of the WTs distribution transformers. The earthing systems of the WTs were galvanic connected between each other on the distribution transformer's LV-Side (star connection) and firmly earthed with earthing rods. Each MV cable sheath and earthing conductor were connected to the LV earthing system of the WT. The distance between the WTs is 300 m.

4.1. Case Study: 30 kA (1.2/50 µs) Lightning Impulse

A positive downward lightning stroke (30 kA, 1.2/50 µs) on the tip of the rotor blade R and the WT1 was simulated in order to explore the effects of the traveling waves across the WP. A simulation time step of 5 ns for a total simulation time of 100 µs was chosen.

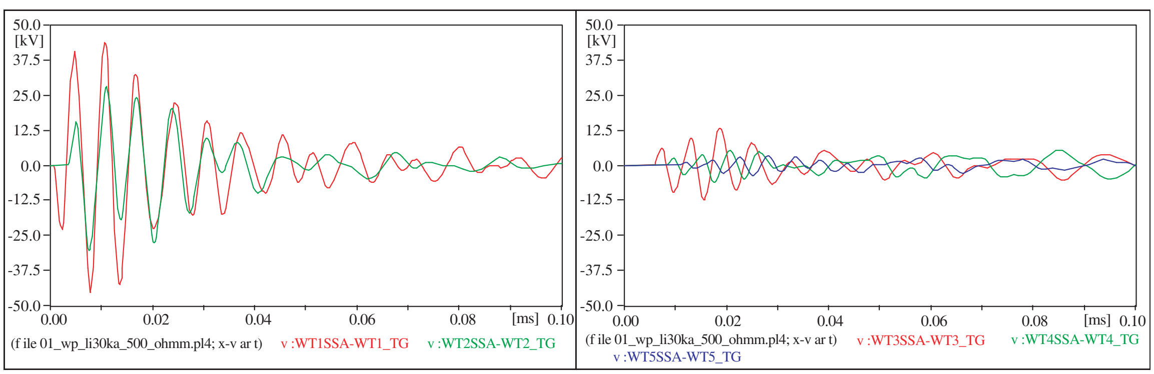
Figure 2 depicts the lightning current decay effects of the stroke. On the left figure a comparison of the stroke current injected at the tip of the blade (in red color) and the lightning current flowing across the earthing system of the WT1 is depicted; the effects of the lightning current traveling waves (reflection and transmission) across the earthing system with a high frequency over-imposed oscillation are showed. The segments of the rotor blades and tower, which may introduce additional transmission and reflection effects, may cause the superimposed high-frequency oscillations.

Figure 3 depicts the surge voltage on the transformer LV-side; WT1 and WT2 observe the highest overvoltage values, which could lead to flashover or similar effects on the LV coils and LV



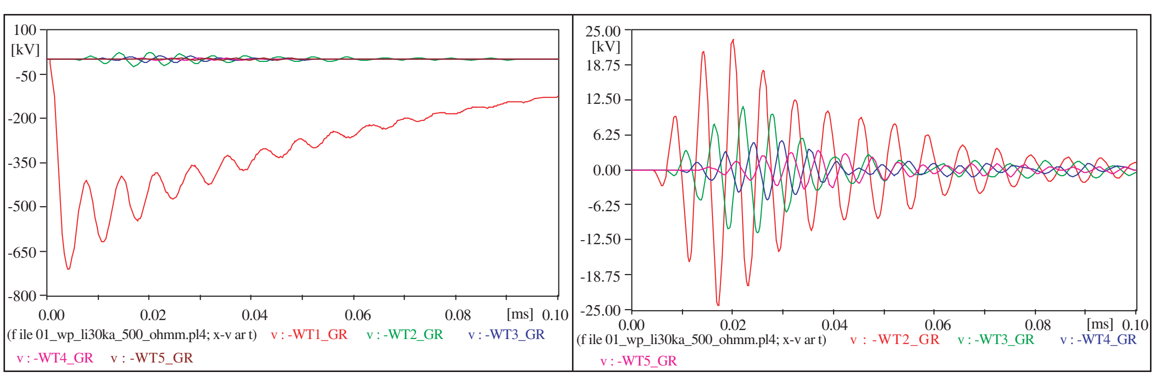
(a) (b)

Figure 2: Lightning current plot: (a) stroke and WT1, (b) WT2 to WT5.



(a) (b)

Figure 3: Transformer LV-side voltage: (a) WT1 and WT2, (b) WT3 to WT5.



(a) (b)

Figure 4: Voltage in the earthing system: (a) stroke and WT1, (b) WT2 to WT5.

grid connection. On the right figure, the delay in the formation of the overvoltages is observed, mainly caused by the traveling time required to reach the other WTs.

Figure 4 depicts the effects of the traveling waves on the footing resistance models described in Section 3.4 and [2]; the propagation of the traveling waves across the earthing system of the WTs shows the expected delay and remarkable overvoltages in the WTs of the WP.

5. CONCLUSIONS

The effects of direct lightning strikes and traveling waves on WPs in form of overcurrent were addressed using approximations and models for fast transients applications. The suggested rotor blade models are an attempt to establish a model for these complex structures; indeed the decoupled transmission line model approximation is the first approximation that needs to be further developed.

The methodology presented in this publication suggests the consideration of several WTs in order to address the possible negatives effects and their mitigation e.g., with proper shielding/earthing or the implementation of surge protection devices.

During a direct lightning strike to a wind turbine, the electromagnetic effects on the earthing system with higher overvoltages and overcurrents amplitudes are registered in the wind turbine affected and in the nearby located wind turbines, taking into consideration the travel time necessary to overcome the complex low and medium voltage network in the wind park.

Further work is required in order to improve the modeling of the different components, especially the rotor blades and bearings.

ACKNOWLEDGMENT

The authors would like to gratefully thank the University of Siegen, University of Kassel and the General Electric Company for the support and facilitation of this work.

REFERENCES

1. IEC 61400-24, "IEC 61400-24 Ed.1: Wind turbines — Part 24: Lightning protection," IEC — International Electrotechnical Commission, June 2010.
2. Méndez, Y., et al., "A first approach in the modelling of failures in rotor blades caused by lightning strikes," *Proceedings of The 30th International Conference on Lightning Protection*, Cagliari, Italy, September 2010.
3. Fast Front Transients Task Force of the IEEE Modeling and Analysis of System Transients Working Group, "Modeling guidelines for fast front transients," *IEEE Transactions on Power Delivery*, Vol. 11, No. 1, 1996.
4. Dommel, H. W., *EMTP — Theory Book*, Branch of System Engineering, Bonneville Power Administration, 1986.
5. Dudurych, I. M., et al., "EMTP analysis of the lightning performance of a HV transmission Line," *IEE Proc. — Gener. Transm. Distrib.*, Vol. 150, No. 41, 2003.
6. De Conti, A., et al., "Calculation of lightning-induced voltages on complex low-voltage distribution networks," *Proceedings of The 30th International Conference on Lightning Protection*, Cagliari, Italy, September 2010.

Consideration on Artificial Neural Network Architecture in Application for Microwave Filter Tuning

J. J. Michalski, T. Kacmajor, J. Gulgowski, and M. Mazur
TeleMobile Electronics, Poland

Abstract— This work focuses on the way in which the ANN should be designed considering the number of input layer neurons in the application for microwave filter tuning. Based on the Cauchy interpolation method we can decide, for a given filter topology, on the number of points at which we sample the reflection characteristic of the filter (obtained from the vector network analyzer). As a result, we can specify the number of ANN input neurons. Minimizing ANN topology, (the number of weights), results in better performance with regard to ANN learning time and ANN generalization errors. Experimental investigation performed on 6- and 11-cavity filters has verified correctness of the presented approach.

1. INTRODUCTION

Currently, after a filter is assembled on factory lines, it is necessary to tune each filter separately. Normally, this is a manual task performed by an operator, who checks appropriate parameters of a filter, e.g., the scattering parameters. In order to reach the specification goals, tunable elements have to be adjusted. Numerous tuning algorithms are published in literature. In recent years we introduced new methods based on heuristic optimization [2]. In this paper we prove that tuning methods based on optimization of well-defined cost function can be very effective. We have recently introduced a novel method of cavity filter tuning with the usage of Artificial Neural Network (ANN) [3]. We present the method for preparing learning and testing vectors containing sampled scattering characteristic and tuning screw errors corresponding to it. The ANN is trained on the basis of samples obtained from a properly tuned filter. It has been proved that the usage of multidimensional approximation ability of the ANN makes it possible to map characteristics of a detuned filter reflection in individual screw errors. The mechanism, based on using multiple golden filters in the process of preparing training vectors has been introduced in [4]. In our approach we have chosen the Feed-Forward ANN with one hidden layer (Figure 1).

We defined the number of output layer neurons as a number of tuning elements N, $\Delta Z \in R^N$. (So far we have considered only the tuning of cavities. We assumed that tuning elements for couplings are pre-tuned and do not change during the tuning process.) The number of neurons in the hidden layer was chosen experimentally to be a minimal one, but still guaranteeing sufficient learning and generalization error of the ANN. The input layer consisted of $2L = 512$ neurons for filter reflection characteristic sampled at $L = 256$ complex points. This number was chosen as a proper value for algorithm customization, in the sense of ANN learning and generalization error, but this value was not optimized in any way. This paper suggests how the ANN should be designed optimally, considering the number of input layer neurons, in accordance with the given filter topology.

2. GENERAL CONCEPT

First we have to ask one question: What is the smallest number of frequency points at which filter reflection characteristic should be sampled to allow its reconstruction in an analytical form? Generally, transmission and reflection characteristics of a two-port filter network, composed of a

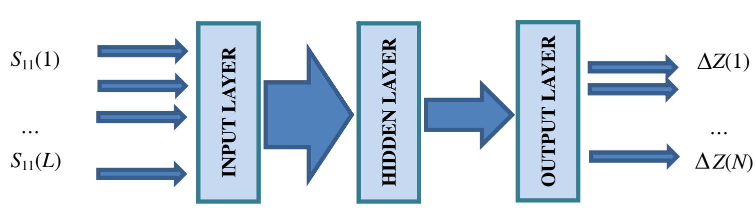


Figure 1. Artificial neural network (ANN), proposed in [3]. S_{11} represents reflection characteristic of a filter sampled at L frequency points and ΔZ represents the error for N tuning elements.

series of N inter-coupled resonators, can be defined as a ratio of two polynomials [1,7] $S_{11}(\omega) = \frac{F_N(\omega)}{E_N(\omega)}$, $S_{21}(\omega) = \frac{P_N(\omega)}{\varepsilon E_N(\omega)}$, where ω represents angular frequency and ε is a constant normalizing S_{21} to the equiripple level. The degree of a common denominator $E(\omega)$ and $F(\omega)$ is N, and the degree of polynomial $P(\omega)$ corresponds to the number of non-infinite transmission zeros. Based on the Cauchy interpolation method, a filter characteristic can be represented as [7]

$$S(\omega) = \frac{A(\omega)}{B(\omega)} = \frac{\sum_{i=0}^{M} a_i \omega^i}{\sum_{j=0}^{N} b_j \omega^j}, \tag{1}$$

Assuming that we have S characteristic sampled at L frequency points ω_l, $l = 1 \ldots L$, Equation (1) can be transformed into

$$X_{L\times(M+1)} a_{(M+1)\times 1} - Y_{L\times(N+1)} b_{(N+1)\times 1} = 0 \tag{2}$$

where

$$X_{L\times(M+1)} = \begin{bmatrix} 1 & \omega_1 & \omega_1^2 & \cdots & \cdots & \omega_1^M \\ 1 & \omega_2 & \omega_2^2 & \cdots & \cdots & \omega_2^M \\ \vdots & \vdots & \vdots & \square & \square & \vdots \\ \vdots & \vdots & \vdots & \ddots & \square & \vdots \\ \vdots & \vdots & \vdots & \square & \ddots & \vdots \\ 1 & \omega_L & \omega_L^2 & \cdots & \cdots & \omega_L^M \end{bmatrix},$$
$$Y_{L\times(N+1)} = \begin{bmatrix} S(\omega_1) & S(\omega_1)\omega_1 & S(\omega_1)\omega_1^2 & \cdots & \cdots & S(\omega_1)\omega_1^N \\ S(\omega_2) & S(\omega_2)\omega_2 & S(\omega_2)\omega_2^2 & \cdots & \cdots & S(\omega_2)\omega_2^N \\ \vdots & \vdots & \vdots & \square & \square & \vdots \\ \vdots & \vdots & \vdots & \ddots & \square & \vdots \\ \vdots & \vdots & \vdots & \square & \ddots & \vdots \\ S(\omega_L) & S(\omega_L)\omega_L & S(\omega_L)\omega_L^2 & \cdots & \cdots & S(\omega_L)\omega_L^N \end{bmatrix} \tag{3}$$

Equation (2) can now be written as the following system of linear equations

$$Z_{L\times(M+N+2)} d_{(M+N+2)\times 1} = 0 \tag{4}$$

where

$$Z_{L\times(M+N+2)} = \left[\left[X_{L\times(M+1)} \right] \; \left[-Y_{L\times(N+1)} \right] \right] \qquad d_{(M+N+2)\times 1} = \begin{bmatrix} \begin{bmatrix} \square \\ a_{(M+1)\times 1} \\ \square \end{bmatrix} \\ \\ \begin{bmatrix} \square \\ b_{(N+1)\times 1} \\ \square \end{bmatrix} \end{bmatrix} \tag{5}$$

This system can be solved if $L \geq N + M + 2$. If we take into consideration reflection characteristic, we have $M = N$ [7]. Taking this into account, we can conclude that $L = 2(N + 1)$ is the minimal number of complex points of sampled filter reflection characteristic S, which is necessary to reconstruct this characteristic in an analytical form. Therefore, we can conclude that, for $L \geq 2(N + 1)$, it is possible to have unambiguous mapping between filter characteristic $S \in C^L$ (sampled at L complex points) and positions of tuning elements $\Delta Z \in R^N$. Based on the above we can decide, for a given filter topology, on the number of points at which we sample the reflection characteristic (obtained from the vector network analyser). In consequence, we can specify the number of ANN input neurons.

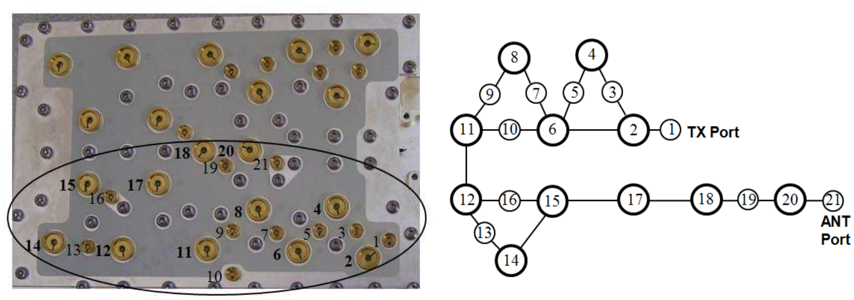
3. EXPERIMENT

In the experiment, performed on two filters of different topology, we have shown the influence of ANN topology on ANN generalization error. As presented in [3, 4] ANN generalization error describes the ability of ANN regarding filter tuning. The smaller the error, the higher the ANN tuning ability. Filters with their topology used in the experiment are presented in Figure 2 and Figure 3. The first filter is a TX filter of GSM duplexer (Figure 2) with self-locking screws. The second (Figure 3) is an RX filter of GSM duplexer with standard locking system (nuts). The ANN learning vectors were collected automatically using the IAFTT robot [5].

The first experiment was performed to examine the dependence between ANN generalization error in function of the number of neurons in the input layer In order to check the ANN generalization ability during the ANN training process we used the following ANN generalization error definition previously introduced in [3]

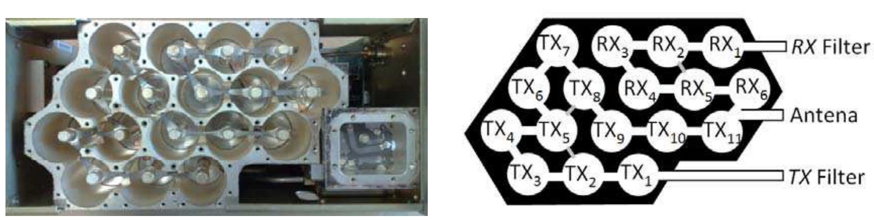
$$GE = 2K \sum_{a=1}^{t} \sum_{b=1}^{N} \left| \Delta Z_{0T}^{a}(b) - \Delta Z_{xT}^{a}(b) \right| / (t * N)[u] \tag{6}$$

where t — number of testing vectors, ΔZ_{0T}^{a} — screw position increment — known proper value, ΔZ_{xT}^{a} — screw position increment generated by ANN, for a given reflection characteristic. One screw-adjustment increment $1u$ is defined as tuning element resolution and K is the maximum screw deviation as a multiple value of u. During the ANN training process, for an 11-cavity filter (Figure 2), we had 1000 learning vectors in total, 500 for $K = 5$, 500 for $K = 10$, and 100 testing vectors, 50 for $K = 5$, 50 for $K = 10$. The tuning element change resolution was $1u = 9$ (deg). For a 6-cavity filter (Figure 3), we had 2000 learning vectors in total, 1000 for $K = 10$, 1000 for $K = 20$, and 100 testing vectors, 50 for $K = 10$, 50 for $K = 20$. The tuning element change resolution was $1u = 18$ (deg). Figure 4 presents the ANN generalization error in function of the number of input layer neurons WI for both filters (6-cavity filter — solid line, 11-cavity filter — dashed line). Each sampled complex point of reflection characteristic requires two input neurons, one for real and the second for the imaginary part. According to our findings, for the $N = 6$-cavity filter we need at least $L = 14$ complex samples of filter characteristic, which gives us $WI = 28$ input neurons. For the $N = 11$-cavity filter the minimal number of input layer neurons is $WI = 48$. These values



(a) (b)

Figure 2. (a) Layout and (b) topology of the TX filter used in the experiment. Small circles represent tunable coupling and cross-couplings. Bigger circles represent cavities. There are no tuning elements for coupling between cavities 15–17, 17–18. Fixed cross-coupling occurs between cavities 2–6.



(a) (b)

Figure 3. (a) Layout and (b) topology of the RX filter used in the experiment. There is one coupling element between cavities 2–5.

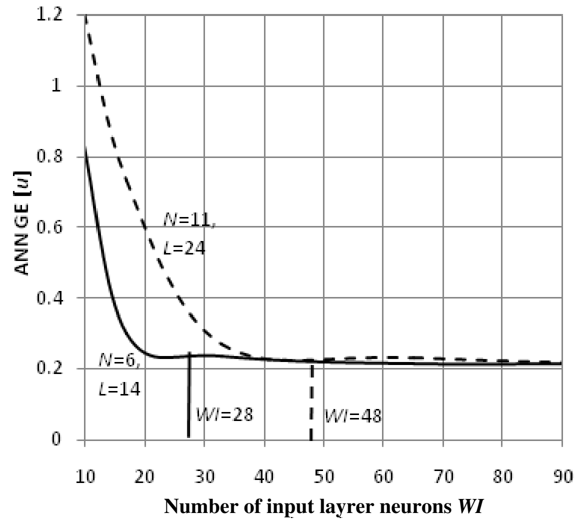


Figure 4. ANN generalization error in function of the number of input layer neurons. Dashed line — 11-cavity filter (Figure 2), solid line — 6-cavity filter (Figure 3).

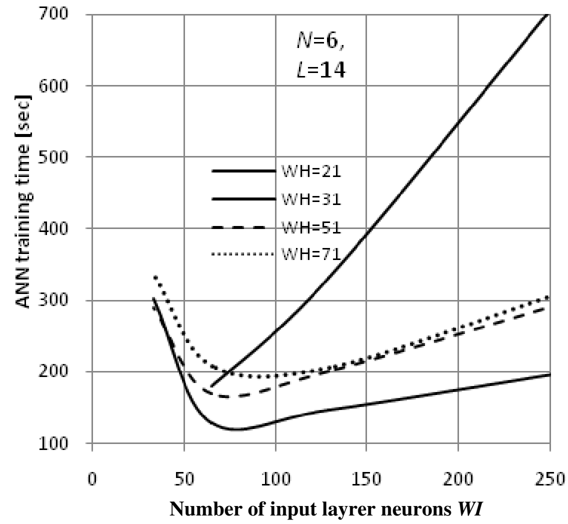


Figure 5. ANN training time necessary to achieve the generalization error level at the value of 0.035 [u] in function of the number of input layer neurons. Curves are calculated for a different number of hidden layer neurons WH.

are marked in Figure 4. We can observe that for $L \geq 2(N+1)$ the ANN is able to generalize at the efficient level and, consequently, is able to tune the filter. In general, minimizing the ANN topology (the number of ANN weights), results in better performance in the sense of both ANN generalization error and ANN learning time.

Exact dependencies between the number of training vectors and the number of neurons in a hidden layer are very difficult to define. Theoretical estimation can be performed using the Vapnik-Chervonenkis dimension [6], In practice this relation should be chosen experimentally. Figure 5 presents the ANN training time for a 6-cavity filter, in function of the number of input layer neurons. The number of neurons in the hidden layer is a parameter. In this experiment the ANN was trained until the generalization error reached the value of 0.035 [u]. Having analysed the curves, we can conclude that the number of neurons in the hidden layer (WH) has significant influence on ANN training time. Moreover, we can observe that, for $WH = 21$ neurons in the hidden layer and for the number of neurons in the input layer smaller than ~ 50, the neural network cannot reach the error value of 0.035 [u]. This means that neural network topology is not well chosen, as the number of network weights is too small to be trained for such set of training vectors.

4. CONCLUSIONS

While solving problems with the use of artificial neural networks it is always difficult to select optimal ANN topology. Very often the only one way to optimize the ANN topology is optimization by trial and error. In this paper we show how the number of input layer neurons can be specified and thus limited and how it influences ANN generalization ability and ANN learning time for the algorithm presented in [3]. The experiment performed on two filters of different topology proved correctness of our assumptions resulting from the Cauchy interpolation and regarding consideration on the dependence between the number of samples of reflection characteristic and the possibility of unambiguous mapping of these samples to errors of tuning elements.

ACKNOWLEDGMENT

This work was supported by National Ministry for Science & Higher Education (Decision nr. 736/N-COST/2010/0), under the project name “New optimization methods and their investigation for the application in physical microwave devices which require tuning”, performed within the COST Action RFCSET IC0803.

REFERENCES

1. Cameron, R. J., "General coupling matrix synthesis methods for Chebyshev filtering functions," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 47, No. 4, 433–442, April 1999.
2. Cegielski, T. and J. J. Michalski, "Heuristic methods for automated microwave filter tuning," *Proceedings of XVII International Conference on Microwave, Radar and Wireless Communications*, Vol. 3, 647–650, Poland, Wroclaw, May 19–21, 2008.
3. Michalski, J. J., "Artificial neural networks approach in microwave filter tuning," *Progress In Electromagnetics Research M*, Vol. 13, 173–188, 2010.
4. Michalski, J. J., "Artificial neural network algorithm for automated filter tuning with improved efficiency by usage of many golden filters," *Proceedings of XVIII International Conference on Microwave, Radar and Wireless Communications*, Vol. 3, 264–266, Lithuania, Vilnius, June 14–16, 2010.
5. Intelligent automatic filter tuning tool (IAFTT) is registered with "Priority certificate # 2516", (Hannover Cebit 2007), and patent pending "European Patent Application No. P382895 assigned by Polish National Patent Office".
6. Vapnik, V. N. and A. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities," *Theory of Probability Appl.*, Vol. 16, 264–280, 1971.
7. Lamperez, A. G., T. K. Sarkar, and M. Salazar-Palma, "Filter model generation from scattering parameters using the Cauchy method," *32th European Microwave Conference and Exhibition*, Vol. 1, 413–416, 1989.

Area of Phase Shifter Operation of the Azimuthally Magnetized Coaxial Ferrite Waveguide

Mariana Nikolova Georgieva-Grosse[1] **and Georgi Nikolov Georgiev**[2]
[1]Meterstrasse 4, D-70839 Gerlingen, Germany
[2]Faculty of Mathematics and Informatics
University of Veliko Tirnovo "St. St. Cyril and Methodius"
BG-5000 Veliko Tirnovo, Bulgaria

Abstract— Two iterative techniques are applied to compute the limits of the area in which the ferrite-loaded coaxial waveguide with azimuthal magnetization gives rise to differential phase shift (works as a digital phase shifter) for the normal TE_{01} mode. The first of them uses the peculiar phase behaviour of configuration at cut-off and the second one — the existence of a physical state for higher frequencies at which the propagation in case of negative (clockwise) magnetization stops. A graphical image of the functional dependence of this state on the structure parameters is a special envelope curve in the phase diagram, marking off the end of the characteristics for the magnetization referred to. Its equation, written through certain positive real numbers ($L(c, \rho, n)$ numbers) is employed to perform the relevant calculations. The influence of the relative thickness of central switching conductor with regard to that of the geometry on the size and shape of the area in question is studied.

1. INTRODUCTION

The circular waveguides, containing azimuthally magnetized ferrite, under normal TE_{01} mode excitation possess potentialities of nonreciprocal digital phase shifters with possible application in the electronically scanned antenna arrays [1–4]. Sophisticated iterative procedures have been elaborated to investigate their phase shifting properties, necessary for the development of these devices [1–3, 5–7]. It has been established that the simplest of the configurations brought up — the one, entirely filled with the anisotropic medium mentioned, exhibits the characteristics of interest in a restricted frequency domain [6]. It might be expected that the more complex geometries of the family treated would behave in a similar manner.

In this study the boundaries of the area in which the coaxial waveguide, completely filled with azimuthally magnetized ferrite produces differential phase shift for the aforesaid wave are examined as a function of the relative dimension ρ $(0 < \rho < 1)$ of its central switching conductor with respect to the one of the structure. For the purpose, the iterative schemes, suggested lately to trace the borderlines of the region of phase shifter operation of the circular ferrite geometry [6], are extended. Unlike before, the positive purely imaginary roots of the characteristics equation of the coaxial transmission line [2] and the positive real numbers $L(c, \rho, n)$, advanced recently [8, 9], are used. The main result of the analysis is that the width of area inspected expands when the parameter ρ grows. Simultaneously, the maximum value the phase shift might attain, decreases.

2. NORMAL TE_{0n} MODES IN THE COAXIAL FERRITE-LOADED WAVEGUIDE WITH AZIMUTHAL MAGNETIZATION

The propagation of normal TE_{0n} modes in the infinitely long, perfectly conducting coaxial waveguide, uniformly filled with lossless latching ferrite, magnetized in azimuthal direction to remanence, is nonreciprocal and is described by the equation [2, 11]:

$$\Phi\left(a, c; x_0\right) / \Psi\left(a, c; x_0\right) = \Phi\left(a, c; \rho x_0\right) / \Psi\left(a, c; \rho x_0\right) \tag{1}$$

in which $\Phi(a, c; x)$ and $\Psi(a, c; x)$ stand for the Kummer and Tricomi confluent hypergeometric functions [10] with $a = c/2 - jk$ — complex, $c = 3$, $x_0 = jz_0$, $k = \alpha\bar{\beta}/(2\bar{\beta}_2)$, $z_0 = 2\bar{\beta}_2\bar{r}_0$, $\rho = \bar{r}_1/\bar{r}_0$, k, z_0, ρ — real, $-\infty < k < +\infty$, $z_0 > 0$, $0 < \rho < 1$, $\mathrm{sgn}k = \mathrm{sgn}\alpha$ ($\alpha = \gamma M_r/\omega$ — off-diagonal ferrite permeability tensor element, $-1 < \alpha < 1$, γ — gyromagnetic ratio, M_r — ferrite remanent magnetization, ω — angular frequency of the wave, $\bar{\beta} = \beta/(\beta_0\sqrt{\varepsilon_r})$, $\bar{\beta}_2 = \beta_2/(\beta_0\sqrt{\varepsilon_r})$, $\bar{r}_0 = \beta_0 r_0\sqrt{\varepsilon_r}$, $\bar{r}_1 = \beta_0 r_1\sqrt{\varepsilon_r}$, β — phase constant, $\beta_2 = [\omega^2\varepsilon_0\mu_0\varepsilon_r(1-\alpha^2) - \beta^2]^{1/2}$ — radial wavenumber, r_0 and r_1 — outer and inner conductor radii, ρ — central conductor to waveguide radius ratio, $\beta_0 = \omega\sqrt{\varepsilon_0\mu_0}$ — free space phase constant, ε_r — ferrite relative permittivity). Eq. (1) holds, if its positive purely imaginary roots $\chi^{(c)}_{k,n}(\rho)$ in x_0 (in z_0) satisfy the relation $\bar{\beta}_2 = \chi^{(c)}_{k,n}(\rho)/(2\bar{r}_0)$ $(n = 1, 2, 3, \ldots)$ which determines the eigenvalue spectrum of the structure for the fields studied.

3. PHASE SHIFTER OPERATION OF THE WAVEGUIDE

For each set of parameters $\{\rho, |\alpha|, \bar{r}_0\}$, subject to the condition [11]:

$$\chi_{0,n}^{(c)}(\rho)/2 < \bar{r}_0\sqrt{1-\alpha^2} < L(c,\rho,n)/|\alpha| \tag{2}$$

the transmission line considered generates differential phase shift $\Delta\bar{\beta} = \bar{\beta}_- - \bar{\beta}_+$ for the relevant TE_{0n} mode. Here $L(c,\rho,n)$ is a general notation of certain positive real numbers — the finite limits of the sequences of numbers $\{K_-(c,n,\rho,k_-)\}$ and $\{M_-(c,n,\rho,k_-)\}$ where $K_-(c,n,\rho,k_-) = |k_-|\chi_{k_-,n}^{(c)}(\rho)$, $M_-(c,n,\rho,k_-) = |a_-|\chi_{k_-,n}^{(c)}(\rho)$, provided $k_- \to -\infty$ [8,9]. In particular, if $c = 3$, $n = 1$, in case $\rho = 0\,(0.1)\,0.9$, it is fulfilled: $L(c,\rho,n) = 6.59365$, 7.65009, 9.96386, 13.67913, 19.69958, 30.06288, 49.76036, 93.58730, 222.36897, 936.90302. The inequalities (2) constitute a mathematical representation of the physical criterion for phase shifter operation of the geometry [11]:

$$\bar{r}_{0cr} < \bar{r}_0 < \bar{r}_{0en-}. \tag{3}$$

It means that for a specific $|\alpha|$ the quantity $\Delta\bar{\beta}$ is called into being then and only then when the pertinent working normalized radius $\bar{r}_0$ is larger than the critical one $\bar{r}_{0cr}$ (connected with the cut-off frequency) [2,11]:

$$\bar{r}_{0cr} = \chi_{0,n}^{(c)}(\rho)/\left[2\sqrt{1-\alpha_{cr}^2}\right] \tag{4}$$

and less than the abscissa $\bar{r}_{0en-}$ of the point at a characteristic envelope En_{1-} — line in the phase portrait [8] for the same $|\alpha|$. The line mentioned appears at negative magnetization. It restricts the propagation of the modes thrashed out from the side of higher frequencies. Its equation $\bar{\beta}_{en-} = \bar{\beta}_{en-}(\bar{r}_{0en-})$ is written in parametric form as [2,8,11]:

$$\bar{r}_{0en-} = L(c,\rho,n)/\left[|\alpha_{en-}|\left(1-\alpha_{en-}^2\right)^{1/2}\right], \quad \bar{\beta}_{en-} = \left(1-\alpha_{en-}^2\right)^{1/2}, \tag{5}$$

($|\alpha_{en-}|$ is the parameter). (For $\alpha_+ > 0$ such boundary does not exist.) Assuming $c = 3$, $n = 1$, for $\rho = 0$ (0.1) 0.9, it holds: $\chi_{0,n}^{(c)}(\rho) = 7.66341\,19404$, 7.88188 32204, 8.47149 60889, 9.41155 10774, 10.78236 23991, 12.78631 35232, 15.86018 08438, 21.04406 45874, 31.47510 37789, 62.85832 85992. Obviously, if $|\alpha|$ varies from 0 to 1, $\bar{r}_{0cr}$ changes from $\chi_{0,n}^{(c)}(\rho)/2$ to infinity. For a given ρ, $\bar{r}_{0en-}$ has a minimum $\min\bar{r}_{0en-} = 2L(c,\rho,n)$ at $\alpha_{en-,\min} = -1/\sqrt{2}$ and $\bar{\beta}_{en-,\min} = 1/\sqrt{2}$. In case $\alpha_{en-} \to 0$ and $\alpha_{en-} \to -1$, then $\bar{r}_{0en-} \to +\infty$.

The analysis shows [11] that as when $\rho = 0$ (circular waveguide with an infinitely thin switching wire [6]), at cut-off, if $\alpha_+ > 0$, then $\bar{\beta}_{cr+} = 0$ (no propagation takes place and the corresponding number $k_{cr+} = 0$). The case $\alpha_- < 0$, however, is more complicated. The function $\bar{\beta}(\bar{r}_0)$ for $\bar{r}_0 \le \bar{r}_{0cr}$ is double-valued. Provided $\bar{r}_0 = \bar{r}_{0cr}$: i) there is no transmission, resp. $\bar{\beta}_{cr-} = 0$ (and the answering to it $k_{cr-} = 0$); ii) TE_{0n} mode with phase constant $\bar{\beta}_{c-} \neq 0$ is supported (and the pertinent to it $k_{c-} \neq 0$, $k_{c-} < k_{cr-}$). At a fixed ρ, if $|\alpha|$ and $\bar{r}_0$ change, acquiring consecutively all their allowable values, determined by the relation (2), the entire set of numerical equivalents of $\Delta\bar{\beta}$ which the structure might produce, is obtained. In a co-ordinate system $\bar{r}_0 - \Delta\bar{\beta}$ this set forms the area of phase shifter operation of the waveguide. Since the intervals of variation of both $\bar{r}_0$ and $|\alpha|$ are continuous, the area referred to, is single-connected. In the co-ordinate system regarded the abscissas of the cut-off points $\bar{r}_{0cr}$, conforming to any $0 < \rho < 1$ fix those of the lower limit of the domain in which the differential phase shift is afforded for given $|\alpha| \equiv |\alpha_{cr}|$. It is established that at the points considered $\Delta\bar{\beta}$ attains its maximum. Therefore, the ordinates of the border are specified by the equality $\Delta\bar{\beta}_{cr} = \bar{\beta}_{c-}$. The abscissas $\bar{r}_{0en-}$ of points from the envelope are the ones of the upper boundary of the domain in question. For their ordinates it is true: $\Delta\bar{\beta}_{en-} = \bar{\beta}_{en-} - \bar{\beta}_{e+}$. Finding the area of interest is equivalent to the tracing of its bounds. (The subscripts "+" ("−") distinguish the quantities, relating to the positive (negative) magnetization; and the ones "*cr*" ("*c*−") — those, coming up to the cut-off (to the transmission state for $\alpha_- < 0$, linked with the cut-off.) The subscript "*en*−" ("*e*+") is attached to the quantities, characterizing the envelope (the point from the $\bar{\beta}_+(\bar{r}_0)$ — curve for certain $|\alpha|$ of abscissa equal to that of the end point of the $\bar{\beta}_-(\bar{r}_0)$ — one for the same $|\alpha|$ at the envelope)) [11]. Below the discussion is confined to the normal TE_{01} mode ($n = 1$) only.

4. LIMIT OF THE AREA OF PHASE SHIFTER OPERATION, CONNECTED WITH CUT-OFF

4.1. Method for Computation of the Limit

First the root $\chi^{(c)}_{0,n}(\rho^{ch})$ of Eq. (1) is calculated for a fixed ρ^{ch}. Next, the critical radius $\bar{r}_{0cr}$ is counted from formula (4) for a chosen value of the off-diagonal tensor element $|\alpha^{ch}_{-}|$. Afterwards, Eq. (1) is solved consecutively for an arbitrary number of equidistant negative values of the parameter k, the first one of which is also arbitrarily picked out. At each cycle the numerical equivalent of k^{ch}_{-}, of the relevant root $\chi^{(c)}_{k^{ch}_{-},n}(\rho^{ch})$ and of $|\alpha^{ch}_{-}|$ are put in the expression [2, 11]:

$$\bar{r}_0 = \left(k\chi^{(c)}_{k,n}(\rho)/\alpha\right)\left\{\left[1+(\alpha/(2k))^2\right]/\left(1-\alpha^2\right)\right\}^{1/2}. \tag{6}$$

The outcomes are compared with that for $\bar{r}_{0cr}$. When an interval is located, containing it, the one for k which corresponds to it, is divided in parts. The procedure goes on, until the length of the interval of computed values of the normalized radius becomes less than the prescribed accuracy ε^{ch} (ε^{ch} — infinitesimal positive real number). Any numerical equivalent, belonging to the last interval for k, is taken as one of k_{c-}. Putting α^{ch}_{-} and k_{c-} in [3, 11]:

$$\bar{\beta} = \left\{\left(1-\alpha^2\right)/\left[1+(\alpha/(2k))^2\right]\right\}^{1/2} \tag{7}$$

yields $\bar{\beta}_{c-}$. Then α^{ch}_{-} is changed and everything begins over again. Finally ρ is altered. (The superscripts "*ch*" and "*comp*" are attached to the symbols, standing for the parameters chosen, resp. the quantities computed.)

4.2. Numerical Example

In case ρ^{ch} =0.2, $n = 1$, $|\alpha^{ch}| = 0.1$ and $\varepsilon^{ch} = 10^{-10}$, it is obtained: $\chi^{(c)}_{0,n}(\rho) =$ **8.47149 6088**9 and $\bar{r}_{0cr} =$ **4.25708 6960**6. Adopting $k^{ch}_{-} =$ **−0.001**9, **−0.0020**, **−0.002**1 and **−0.002**2 permits to size up for $\rho^{ch} = 0.2$: $\chi^{(c)}_{k^{ch}_{-},n}(\rho^{ch}) =$ **8.464**84 07155, **8.464**49 05621, **8.464**14 04217, **8.46**379 02942 and $\bar{r}_0 =$ **4.25**681 26038, **4.25**696 80441, **4.257**14 04363, **4.257**32 97741, resp. Obviously, the value of $\bar{r}_{0cr}$ lies between the second and third figured out ones of $\bar{r}_0$. Hence, k_{c-} looked for belongs to the interval $k^{ch}_{-} =$ (**−0.0020**, **−0.002**1). Dividing it to 100 parts and taking at each step that of them which results in an interval for $\bar{r}_0$, involving $\bar{r}_{0cr}$, gives consecutively: $k^{ch}_{-} =$ **−0.00207 0**, **−0.00207**1, $\chi^{(c)}_{k^{ch}_{-},n}(\rho^{ch}) =$ **8.46424** 19611, **8.464**46 25504 and $\bar{r}_0 =$ **4.25708 69**391, **4.25708** 86977, resp.; $k^{ch}_{-} =$ **−0.00207 001**, **−0.00207 00**2, $\chi^{(c)}_{k^{ch}_{-},n}(\rho^{ch}) =$ **8.46424 54**274, **8.46424 5**3924, $\bar{r}_0 =$ **4.25708 69**566, **4.25708 69**742, resp.; $k^{ch}_{-} =$ **−0.00207 00122**, **−0.00207 0012**3, $\chi^{(c)}_{k^{ch}_{-},n}(\rho^{ch}) =$ **8.46424 54197**, **8.46424 5419**4, $\bar{r}_0 =$ **4.25708 6960**5, **4.25708 6960**7, resp. Any of the last numerical equivalents of k^{ch}_{-} could be accepted as one of k_{c-}. Formula (7) enables to cast up: $\bar{\beta}_{c-} =$ **0.04115 7466**2, **0.04115 746**82. The upshot is **0.04115 74669**.

5. LIMIT OF THE AREA OF PHASE SHIFTER OPERATION, CONNECTED WITH THE ENVELOPE OF PHASE CHARACTERISTICS FOR NEGATIVE MAGNETIZATION

5.1. Method for Computation of the Limit

Assuming the same ρ^{ch}, as the one, taken in Section 4.1, the $L(c, \rho^{ch}, n)$ number is worked out. Next, a value of the element $|\alpha^{ch}_{en-}|$ is selected and those of the normalized guide radius $\bar{r}^{comp}_{0en-}$ and of the phase constant $\bar{\beta}^{comp}_{en-}$ are reckoned for it from formulae (5). After that the scheme, described above, is effectuated for positive numerical equivalents of the parameter k^{ch}_{+} to get $\chi^{(c)}_{k^{comp}_{+},n}(\rho^{ch})$, resp. $\bar{r}^{comp}_{0e+}$ and $\bar{\beta}^{comp}_{e+}$ sought. The computations continue, until it is fulfilled $|\bar{r}^{comp}_{0en-} - \bar{r}^{comp}_{0e+}| < \varepsilon^{ch}$. The quantity $\Delta\bar{\beta}_{en-}$ is counted as a difference between $\bar{\beta}^{comp}_{en-}$ and $\bar{\beta}^{comp}_{e+}$.

5.2. Numerical Example

For the values of parameters from Section 4.2 with $L(c, \rho^{ch}, n) = 9.96386$ [9], it pans out: $\bar{r}_{0en-} = 100.14056$, $\bar{\beta}^{comp}_{en-} = 0.99498\ 74371$, $k^{ch}_{+} = 0.84129$, $\chi^{(c)}_{k^{comp}_{+},n}(\rho^{ch}) = 11.82262$, $\bar{r}^{comp}_{0e+} = 100.14056$, $\bar{\beta}^{comp}_{e+} = 0.99323$, $\Delta\bar{\beta}^{comp}_{en-} = 0.00175$.

6. AREA OF PHASE SHIFTER OPERATION

Figures 1–4 show the influence of the switching conductor to waveguide radius ratio on the area studied. The LEn_1 — dashed and REn_1 — dotted curves represent the limits, connected with the cut-off and with the envelope, resp. For specific $\bar{r}_0 \in [\chi^{(c)}_{0,n}(\rho)/2, 2L(c,\rho,n)]$ differential phase shift is provided for all $|\alpha| \in [0, |\alpha_{cr}|]$, $|\alpha_{cr}| = \sqrt{1 - [\chi^{(c)}_{0,n}(\rho)/(2\bar{r}_0)]^2}$. Its magnitude ranges from the horizontal axis to the LEn_1 — line. If $\bar{r}_0 \in (2L(c,\rho,n), +\infty)$, there are three zones: lower, middle and upper. In the first (third) one $\Delta\bar{\beta}$ is obtained for $|\alpha| \in [0, |\alpha_2|]$ ($|\alpha| \in [|\alpha_1|, |\alpha_{cr}|]$), $|\alpha_{1,2}| = 0.5\sqrt{1 \pm \{1 - 4[L(c,\rho,n)/\bar{r}_0]\}^2}$. The relevant numerical equivalent of this quantity increases from the horizontal axis to the lower branch of the REn_1 — limit (from the upper branch of the REn_1 —

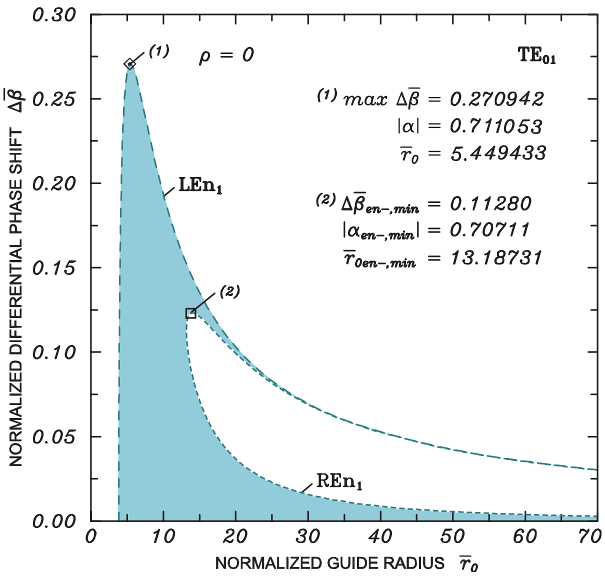


Figure 1: Area of phase shifter operation of the azimuthally magnetized coaxial ferrite waveguide for normal TE_{01} mode in case $\rho = 0$.

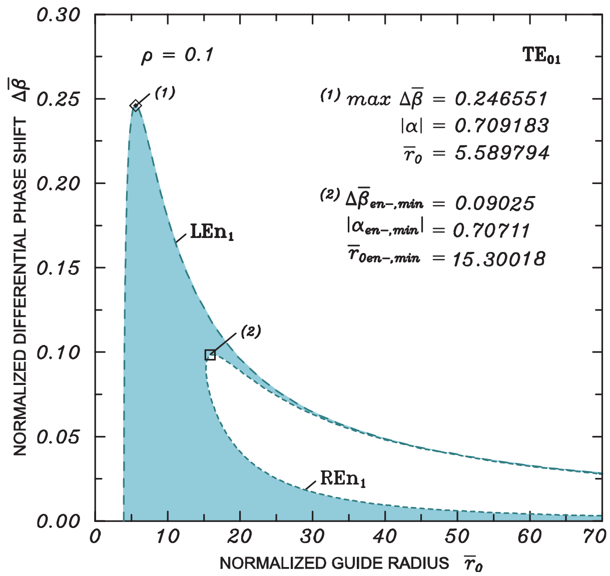


Figure 2: Area of phase shifter operation of the azimuthally magnetized coaxial ferrite waveguide for normal TE_{01} mode in case $\rho = 0.1$.

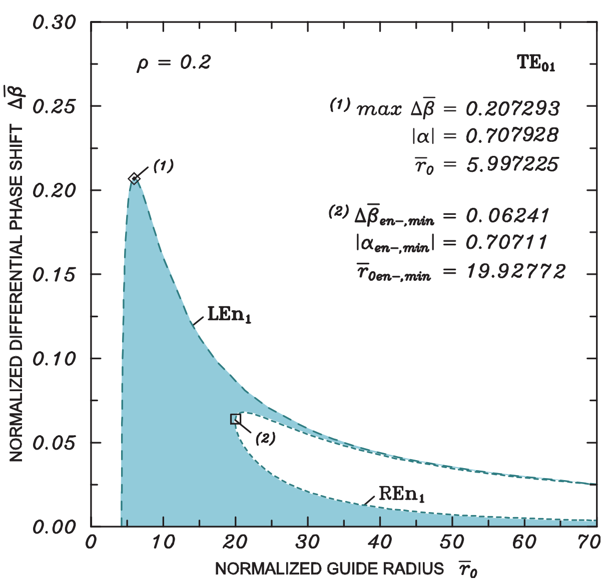


Figure 3: Area of phase shifter operation of the azimuthally magnetized coaxial ferrite waveguide for normal TE_{01} mode in case $\rho = 0.2$.

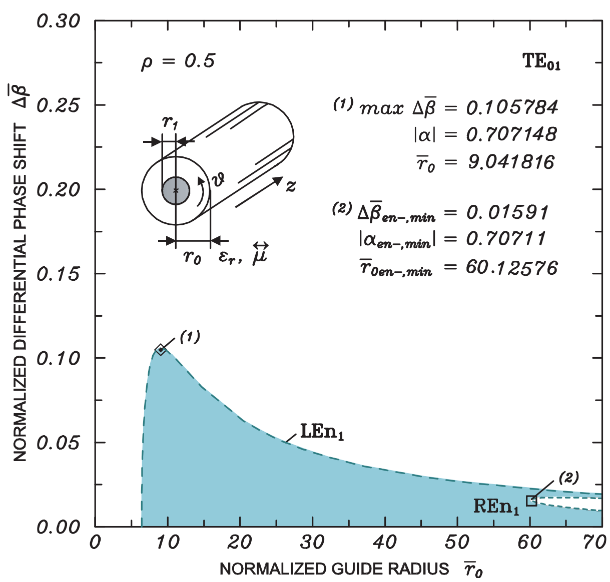


Figure 4: Area of phase shifter operation of the azimuthally magnetized coaxial ferrite waveguide for normal TE_{01} mode in case $\rho = 0.5$.

limit to the LEn_1 — line). On condition that $|\alpha| \in [|\alpha_2|, |\alpha_1|]$ (in the region, encircled by the REn_1 — curve) no phase shift is raised. There are two characteristic points in the Figures: the maximum of LEn_1 — limits (labeled by a rhomb and by the number 1) and that — denoted by a square and by the number 2. The parameters of the first point are determined through an iterative technique, changing the element $|\alpha|$, until the greatest value of $\Delta\bar{\beta}$ is found with the prescribed accuracy. The point 2 represents the minimum of the REn_1 — line. Here $|\alpha_{en-,\min}| = 1/\sqrt{2}$, irrespective of ρ. The enlargement of the latter broadens the area examined. At the same time it diminishes the maximum, the phase shift might acquire.

7. CONCLUSION

The borders of the area of phase shifter operation of the coaxial waveguide, entirely filled with azimuthally magnetized ferrite and propagating normal TE_{01} mode are counted. For the purpose two iterative approaches are employed. They are based on the special features of the phase characteristics of transmission line: double-valuedness at cut-off and availability of a distinctive envelope curve, representing the state in which the wave is no more supported. The basic idea in them is to solve numerically many a time the characteristic equation of configuration, written by the Kummer and Tricomi confluent hypergeometric functions, varying the imaginary part of their complex first parameter. The newly introduced positive real $L(c, \rho, n)$ numbers are used, too in the computations. The outcomes are presented graphically in normalized form and are discussed.

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

1. Georgiev, G. N. and M. N. Georgieva-Grosse, "Iterative method for differential phase shift computation in the azimuthally magnetized circular ferrite waveguide," *PIERS Online*, Vol. 6, No. 4, 365–369, 2010.
2. Georgieva-Grosse, M. N. and G. N. Georgiev, "Differential phase shift characteristics of normal TE_{0n} modes in an azimuthally magnetized coaxial ferrite waveguide," *Proc. 4th Europ. Conf. Antennas Propagat. EuCAP 2010*, article ID P4-53, 5 pages, Barcelona, Spain, April 12–16, 2010.
3. Georgiev, G. N. and M. N. Georgieva-Grosse, "Analysis of the differential phase shift in the circular ferrite-dielectric waveguide with azimuthal magnetization," *Proc. 2010 IEEE Int. Symp. Antennas Propagat. & CNC-USNC/URSI Radio Science Meeting AP-S 2010*, article ID 330.9, 4 pages, in CDROM, Toronto, ON, Canada, July 11–17, 2010.
4. Kashin, V. A. and A. P. Safonov, "A monopulse transmit-receive phased array with polarization discrimination of targets in the main beam," *J. Commun. Technol. Electron.*, Vol. 50, No. 8, 853–862, August 2005 (in Russian).
5. Georgiev, G. N. and M. N. Georgieva-Grosse, "An application of the complex Tricomi function," *Proc. Eleventh Int. Conf. Electromagn. Adv. Applicat. ICEAA'09*, 819–822, in CDROM, Turin, Italy, September 14–18, 2009.
6. 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.
7. Georgiev, G. N. and M. N. Georgieva-Grosse, "Numerical study of the differential phase shift in a circular ferrite-dielectric waveguide," *Proc. 5th Europ. Conf. Antennas Propagat. EuCAP 2011*, 1666–1670, Rome, Italy, April 11–15, 2011.
8. Georgiev, G. N. and M. N. Georgieva-Grosse, "A property of the $L(c, \rho, n)$ numbers and its application to waveguide propagation," *Proc. XXIX URSI General Assembly*, article ID BK.6(120), 4 pages, in CDROM, Chicago, IL, USA, August 7–16, 2008.
9. Georgiev, G. N. and M. N. Georgieva-Grosse, "Theorem for the $L(c, \rho, n)$ numbers," *PIERS Proceedings*, 1478–1482, Moscow, Russia, August 18–21, 2009.
10. Tricomi, F. G., *Funzioni Ipergeometriche Confluenti*, Edizioni Cremonese, Rome, Italy, 1954.

11. Georgieva-Grosse, M. N. and G. N. Georgiev, "Criteria for phaser shifter operation of an azimuthally magnetized coaxial ferrite waveguide," *Proc. 18th Int. Conf. Microwaves, Radar, Wireless Commun. MIKON-2010*, article ID C7–3, 7(4) pages, Vilnius, Lithuania, June 14–16, 2010 (Focused Session).

A Dual Linear Polarization Feed Antenna System for Satellite Communications

A. Tribak[1,2], **A. Mediavilla**[1], **Alicia Casanueva**[1], **and K. Cepero**[2]

[1]Department of Communication Engineering, University of Cantabria
Av. Los Castros s/n, Santander 39005, Spain
[2]Dpartment of Physique, University of Abdelmalek Essaadi, Tetouan, Morrocco

Abstract— A dual lineal polarization feed antenna system for satellite communication will be described. It consists of a turnstile-based orthomode transducer (OMT) and two identical duplexers formed by a plane T-junction and of two iris filters. This 4-ports subsystem transmits and receives radio frequency signals in double track in which the transmission Tx is made through two ports having as access the standard rectangular waveguide WR229 while the reception Rx is made through two ports having as access the standard rectangular waveguide WR159. The proposed subsystem overcomes the current practical bandwidth limitations by using a very compact octave bandwidth OMT along with two robust duplexers. The subsystem is working in the extended C-band, the 5.8–7.1 GHz range is used for the uplink, whereas the 3.6–4.8 GHz range is set to the downlink. The presented architecture exhibits a return loss better than 20 dB in all ports, an isolation between the different rectangular ports better than 70 dB and a transmission loss less than 0.15 dB in both frequency bands (3.6–4.8 GHz and 5.8–7.1 GHz) which represents state-of-the-art achievement in terms of bandwidth.

1. INTRODUCTION

With the increased communication capacity and versatility of the antenna systems, an orthogonal dual polarization operation is often required. One of the key components for orthomode operation is the orthomode transducer (OMT) [1–4]. The OMT is a RF device often used to combine or separate orthogonally polarized signals, thus providing polarization discrimination. The double polarization systems are normally used for satellite communication systems, where an orthomode polarization operation is used to increase the traffic capability of a given link. Another interesting application can be found in the recipients of radio astronomy [1–4]. Unfortunately, most OMTs today are not fully satisfactory. For example, they may not be effective in preventing the generation of undesired higher order modes or providing sufficiently high isolation between ports or their often excessive bulk and thickness may impede many important applications. In [2–5] several configurations of OMTs have been published focusing on the physical symmetries of the structure and making clear the distinction between OMTs of narrow- and broad- band. The criteria used to compare the performance of the different OMTs are the return losses, the insertion losses and the isolation between polarizations which is associated to the specifications of the cross-polarization for the system including the antenna and the feeders. In addition, the symmetry is a key question in the design of OMTs since it determines the high order mode generated in the structure and, therefore, the bandwidth of work.

Our main aim in this work is to increase in the quantity of information to be transmitted or received for solving the saturation problems of the available spectral bands. We solved this problem in two different ways: either by the widening of the frequency bands available of the subsystems, or by the use of the double ways for Tx and Rx by using the orthomode polarization operation to increase the traffic capability of the link, or by both of them at the same time. For this reason, we have used a waveguide OMT based in a turnstile junction [5, 6] which provides a good return loss (30 dB) along with an isolation of 60 dB across a very broadband frequency (2 : 1). The very wide bandwidth is achieved mainly due to the use of a double symmetry turnstile junction reduced height waveguide arms along with the use of a reduced height E-plane power combiners/dividers and simple mitered 90° bends. This facilitates the design and the construction of these OMTs. Moreover, two wideband diplexers have been designed and constructed to demonstrate the feasibility of using stepped waveguide junctions for duplexer designs. It consists of two channel filters and a Tee E-plane junction. The relative bandwidth of the whole diplexer is about 64%. The inputs and outputs return losses were better than 25 dB, the isolation is better than 80 dB and the transmission loss was less than 0.1 dB. Furthermore, experimental measurements of the whole 4-port subsystem (see Fig. 1 for the proposed 4-ports subsystem) exhibit very good agreement with the simulation results.

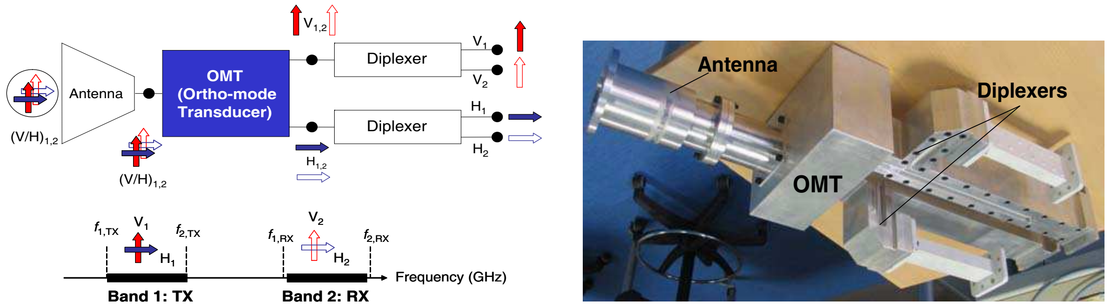


Figure 1: Extended C-band feed assembly and block diagram: dual-polarization horn antenna, proposed OMT and two diplexers.

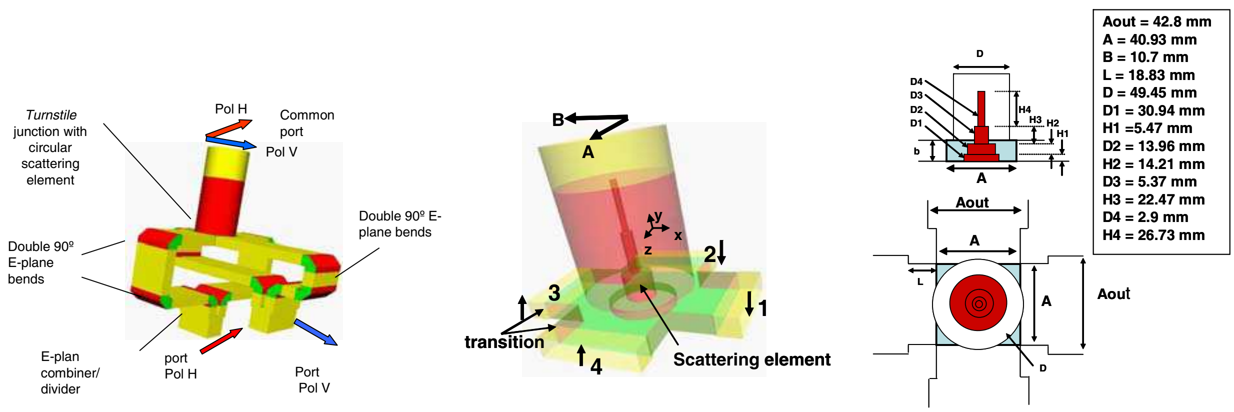


Figure 2: Turnstile junction OMT: internal structure of the complete OMT (left), turnstile junction side view and top view (sketch and dimensions).

2. THEORY AND DESIGN

2.1. A-Turnstile Junction Orthomode Transducer (OMT)

The turnstile junction OMT is working according to the following manner: signals split by the turnstile junction exit in opposite waveguides 180° out-of-phase. These signals are recombined using a power combiner that is also 180° out-of-phase; implemented using a single step E-plane divider assures 3 dB/180° power division in more than one octave bandwidth operation along with −50 dB return loss at the input port. In the first design we have used one input step divider which allows to increase the waveguide height in order to obtain reduced height waveguide ports in the outputs divider [5]. We have used the easiest form divider which provides an extremely good result in contrast to other [2–4] full height combiners in which multi-step transition is necessary to obtain a full height at the output ports combiner. For each polarization, a pair of identical 90 E-plane single mitred waveguide bends in reduced height configuration link, from the turnstile junction outputs to the E-plane junction inputs, can be easily designed for return loss better than 48dB in more the 68% bandwidth. A matching element may be provided at the base of the cavity formed by the rectangular waveguide ports and the main waveguide to enhance a very wide band operation of the turnstile junction with a low reflection coefficient.

Figure 2 shows an internal view of a four-section reduced- height turnstile junction. The internal diameters and heights of the scattering element are optimized for a specific frequency band as a function of the rectangular input waveguide dimension "a". The number of cylindrical sections can be reduced if the application needs a restricted band. The optimisation procedure aims for good return losses for the two orthogonal modes TE_{11} at the common circular waveguide. Two fundamental modes (designated as A and B) can propagate in this circular guide as independent orthogonal linear polarization states when the frequency is above the cutoff frequency value (frequencies above 3.5 GHz). The turnstile junction splits A equally between the reduced-height rectangular waveguide outputs 2 and 4, but does not couple it to ports 1 and 3. Similarly, B is split equally between ports 1 and 3, but does not couple it to ports 2 and 4. The operating mode is explained according to Fig. 3. Taking as example the vertical polarization with the polarized electric field according to

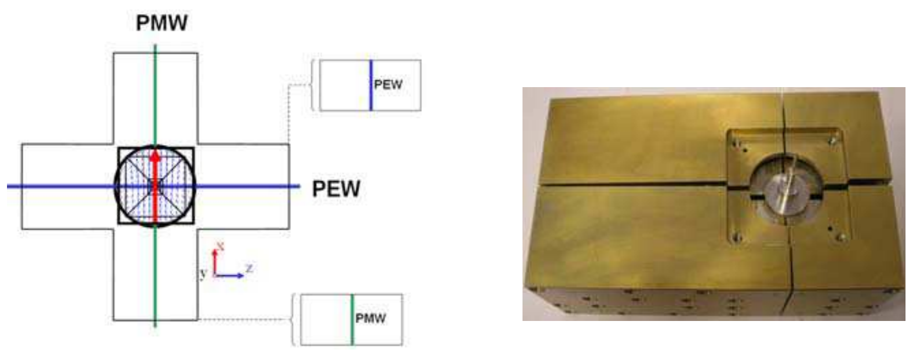
x-axis, the only modes generated in the junction are those with symmetry of the perfect magnetic wall (PMW) in the plane xy and those with symmetry of the perfect electrical wall (PEW) in the plane yz.

In the PMW rectangular ports propagate the TE_{10} fundamental mode over a bandwidth of almost 76% before the excitation of the first higher order TM_{11} mode. The first higher order mode excited in PEW rectangular ports is the TE_{20} mode (which appears in 66.66% of bandwidth from the fundamental mode TE_{10}).

If we can match the incident TE_{11} mode, from which the electric field is polarized according to the x-axis Fig. 2, all the power will be divided between ports 2 and 4 with no power coupled to ports 1 and 3. This operation is valid between the cutoff frequency of the fundamental TE_{10} mode of the rectangular waveguide and the first higher order mode TE_{20} excited in the rectangular ports 1 and 3 (PEW). Therefore, the maximum bandwidth that a turnstile junction can achieve is an octave bandwidth $((F_{c\mathrm{TE}_{20}}/F_{c\mathrm{TE}_{10}}) = 2)$. Due to the symmetry of the turnstile structure, the behaviour is the same for the other incident TE_{11} mode, in which the electric field is polarized according to z-axis. Apart from the fundamental TE_{11} mode propagated in the circular common port, the first higher order mode propagated in the common circular port is the TM_{11} mode, which existed in 72.6% bandwidth from the fundamental TE_{11} modes.

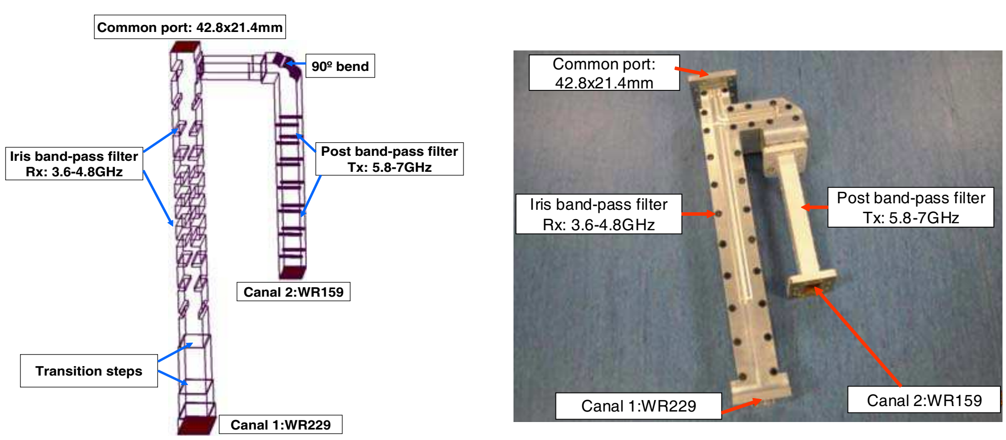
2.2. B-Diplexers

Basically the structure of the diplexer consists of a T junction E-plane, and two band-passes filters one for the Tx and the other for the Rx. Fig. 4 shows an internal view of the total diplexer. The rectangular common port dimensions of the diplexer are $42.8 \times 21.4\,\mathrm{mm}$ ($fc = 3.5\,\mathrm{GHz}$) which allows the propagation of TE_{10} mode in the whole frequency band 3.6–7 GHz, where the Tx port is in standard waveguide WR159 ($fc = 3.7\,\mathrm{GHz}$) and the Rx one is in standard waveguide WR229



(a) (b)

Figure 3: Operating mode of the (a) turnstile junction and (b) designed OMT. External dimensions are $240 \times 140 \times 94\,\mathrm{mm}^3$.



(a) (b)

Figure 4: (a) Internal structure and (b) designed diplexer.

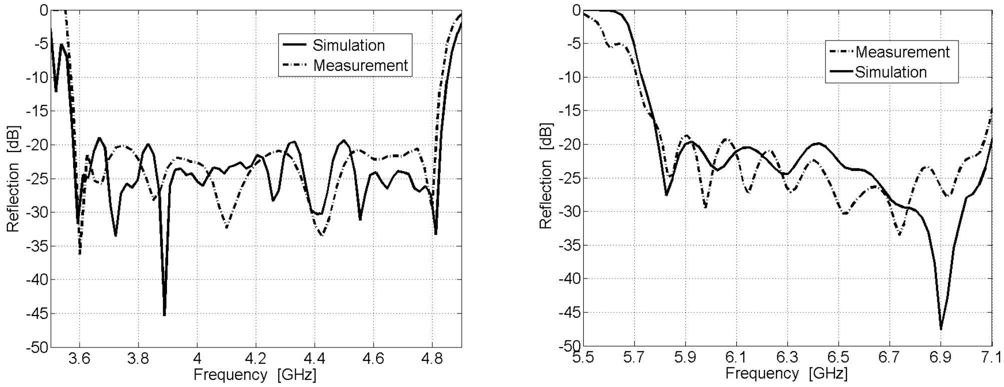


Figure 5: Measured and simulated reflection coefficients of the designed 4-ports subsystem in both Rx band (3.6–4.8 GHz) and Tx band (5.8–7 GHz).

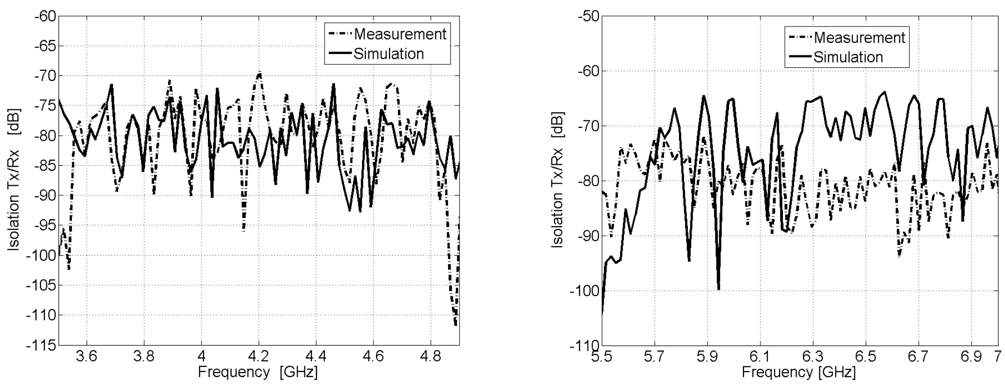


Figure 6: Measured and simulated isolation Tx/Rx in both bands.

($fc = 2.57$ GHz). The relative bandwidth of the whole diplexer is about 64%. The channel of filter 1 (Tx) has a relative bandwidth of 19% and the channel of filter 2 (Rx) has a relative bandwidth of 29%. The filter of the high frequency band (Tx) is based in rectangular waveguide with inductive cylindrical posts and the other of the low frequency band is made in rectangular waveguide with inductive irises. Both filters are well matched; provide a return loss better than 25 dB in case of the low band and 35 dB in the high band along with an insertion loss less than 0.05 dB in both bands. In order to have a TX and RX access in the same plane, we have used a 90° E-plane bends with small curves that exhibit a return loss better than 42 dB.

3. SIMULATED AND MEASURED RESULTS

In Figs. 5 and 6, we illustrate the comparison between the measured and simulated results of the whole 4-ports subsystem depicted in Fig. 1. In these graphs, we can observe that the return loss in the various ports of the system is better than 20 dB and the isolation is better than 70 dB. The measured insertion loss does not exceed 0.15 dB in the two frequency bands of interest.

4. CONCLUSIONS

The electrical and mechanical design along with the measurements of very broadband 4-ports feed antenna is presented. Experimental data shows a return losses better than 20 dB in both Tx and Rx bands included in the 3.6–7 GHz (64% bandwidth) and an isolation between Tx and Rx ports better than 70 dB along with an insertion losses lower than 0.04 dB in the two separated bands of interest. To date, the designed OMT has the best measurement performance reported in literature along with a reduced size (very challenging for this type of structures), being a more robust design than devices existing previously. These OMTs are well-suited to very wide bandwidth and very separated dual band applications with good performances and good manufacturability mechanical tolerances.

ACKNOWLEDGMENT

This work was supported by the Ministerio de Educacion y Ciencia (Spain), Plan Nacional de I+D+I, Astronomy and Astrophysics program AYA2007-68058-C03-03, and the FPI grant BES-2005-6730. The authors like to thank Spanish MICINN and AECID through their codes TEC-2008-06684-C03-01, CSD2008-68, and D/018607/08 for their human capital support. The authors wish also to thank the staff of the Spanish Agency for International Development Cooperation (AECID).

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 diffractional 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 metalsemiconductormetal 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.

Study of a Coplanar Circulator Based on a Barium Hexaferrite Nanocomposite

T. Boyajian[1], **D. Vincent**[1], **M. Le Berre**[2], **and S. Neveu**[3]

[1]DIOM Laboratory, Jean-Monnet University, 25 Dr. Rémy Annino Street, Saint Etienne 42000, France
[2]Lyon Institute of Nanotechnology (INL), NSA of Lyon, 7 Jean Chapelle Avenue
Villeurbanne 69621, France
[3]PECSA-CNRS/UPMC/ESPCI Laboratory, 4 Place Jussieu, Paris 75252, France

Abstract— This paper presents the design and analysis of a Y-junction coplanar circulator operating in the frequency range of 40–60 GHz. The magnetic material used in this circulator's numerical study is a Barium Hexaferrite with a magnetization lower than the usual ferrite. Its purpose is to assume the performance of such circulator when its conventional ferrite is replaced by a magnetic nanocomposite. Magnetization depends on the volume of the magnetic material; therefore, the volume fraction of the barium hexaferrite nanoparticles in the final nanocomposite is introduced. A three dimensional finite element method was used to evaluate numerically the circulator's performance. Preliminary experiments on the nanocomposite magnetic material showing its orientation and non-reciprocity are presented.

1. INTRODUCTION

Circulators belong to a large family of nonreciprocal devices widely used in microwave components relying on magnetic materials. The operation of a stripline Y-junction circulator was first explained by Bosma [1, 2], Fay and Comstock [3], and other authors [4, 5]. Their work became a general basic study for all other circulator designs such as stripline, microstrip and coplanar.

For commercial applications, besides a wide frequency band, low reflection levels at all ports, low insertion loss (< 1 dB) in the forward direction (from port 1 to port 2, from port 2 to port 3, and from port 3 to port 1), high isolation (> 20 dB) in the reverse direction are required. In addition to its performance, a circulator should be easily implemented in new technology components.

Circulators play important roles in different telecommunication systems, especially when one antenna is used for transmission and reception. Mobile phone systems, satellite links, radar duplexers, and phased-array antennas [6] could be good examples.

This paper deals with the operation of a coplanar circulator based on a Barium Hexaferrite thin film. It describes the design and the simulations results. First trials and microwave characterizations are exposed concerning the nanocomposite material.

2. FUNDAMENTALS AND FORMULATIONS

The magnetic material integrated in such devices guarantees its non reciprocity. The signal is maximal from the transmitter to the receiver, but the signal cannot propagate in the reverse direction. The Barium Hexaferrite ($BaFe_{12}O_{19}$) in its bulk state has a saturation magnetization M_s of 382 kA/m. Having a high anisotropy (1.7 T), the microwave excitation and the magnetization of such material are related by a permeability tensor $\bar{\bar{\mu}}$:

$$B = \bar{\bar{\mu}} H \tag{1}$$

Polder [7] was the first to theoretically approach the problem of the tensorial analysis of the permeability in a uniformly magnetized single-domain anisotropic magnetic particle. The permeability is described by:

$$\vec{\vec{\mu}} = \begin{bmatrix} \mu & j\kappa & 0 \\ -j\kappa & \mu & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

$\mu = \mu' - j\mu''$ is the diagonal term and $\kappa = \kappa' - j\kappa''$ is the off-diagonal term. This latter term determines the non reciprocity. In case of an isotropic material, it is equal to zero. μ and κ are given by:

$$\mu = 1 + \frac{(\omega_0 + j\alpha\omega)\omega_m}{(\omega_0 + j\alpha\omega)^2 - \omega^2} \tag{3}$$

$$\kappa = \frac{\omega\omega_m}{(\omega_0 + j\alpha\omega)^2 - \omega^2} \tag{4}$$

Here

$$\begin{aligned} \omega_m &= \gamma M_s \\ \omega_0 &= \gamma H_i \end{aligned} \tag{5}$$

γ is the gyromagnetic ratio, α is the damping factor and H_i is the internal field of the particle considered uniform and given by:

$$H_i = H_0 - N_Z M_s \tag{6}$$

where N_z is the demagnetizing factor which depends on the geometric shape and dimensions of the ferrite. Polder's approach supposes the material to be saturated and the external excitation perpendicularly applied. Anisotropic magnetic materials have a high remanent magnetization when no field is applied. This means they don't need to be externally biased by permanent magnets, they are self-biased. The applied field H_0 is replaced by H_a which is the anisotropy field of the ferrite.

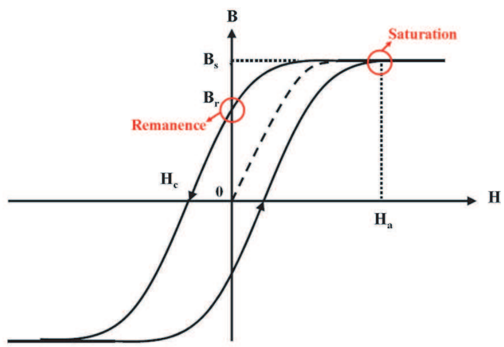


Figure 1: Hysteresis loop of hard magnetic materials.

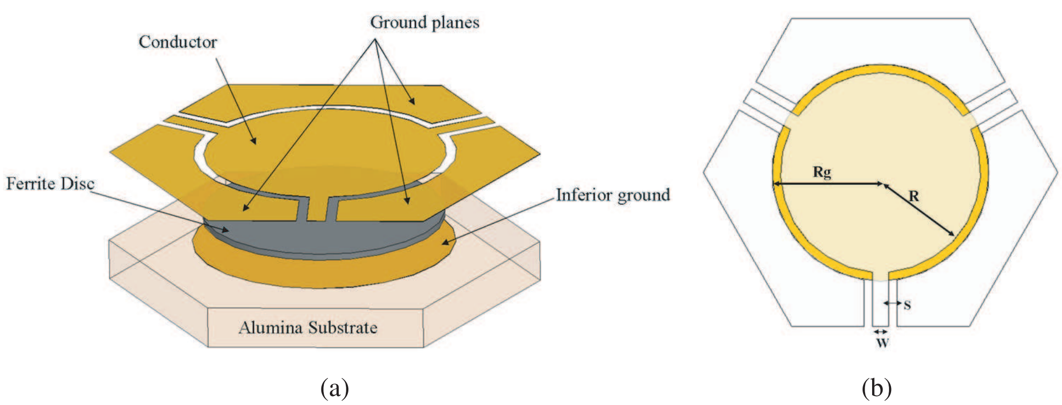


Figure 2: (a) Coplanar circulator's structure and (b) geometric dimensions.

3. DESIGN AND NUMERICAL STUDY

3.1. Circulator Structure

The configuration of the proposed circulator is shown in Figure 2. The ground plane and the access lines are in the same plane, thus the use of the term "coplanar". The metallization is made of 1 µm gold film. Another ground plane in a circular shape is placed on the Alumina dielectric substrate (Al_2O_3). This inferior ground plane ensures the transition of the signal from the access lines to the central circular conductor part. The ferrite disc separates the dielectric substrate and the signal plane.

3.2. Geometric Dimensions

The thicknesses and dimensions of the structure's different parts are resumed in the following table:

The signal lines are CPW lines with a width $W = 72\,\mu\text{m}$ and longitudinal spacing $S = 40\,\mu\text{m}$, resulting a characteristic impedance around $50\,\Omega$.

Table 1: Coplanar circulator's geometric dimensions.

Structure Part	**Proposed Circulator**		
	Type	**Thickness**	**Radius**
Substrate	Alumina (Al_2O_3)	635 µm	Hexagonal form
Ferrite Disc	Barium Hexaferrite ($BaFe_{12}O_{19}$)	100 µm	Disc of radius = 465 µm
Metallization	Gold (Au)	1 µm	Circular Inferior Ground Rg = 505 µm
			Circular conductor R = 465 µm

3.3. Magnetic and Dielectric Characteristics

The magnetic material is defined by the following characteristics using a simulator based on a three dimensional finite element method.

- $\varepsilon_f = 14.2$, the relative permittivity of the ferrite
- $\alpha = 0.0175$, the damping factor, or $\Delta H = 500$ Oe at a frequency of 40 GHz
- $\tan\delta_e = 10^{-2}$, the dielectric loss tangent.

The dielectric substrate's characteristics are:

- $\varepsilon_d = 10$, the relative permittivity of Alumina
- $\tan\delta_e = 10^{-2}$, the dielectric loss tangent.

The gold metallization has a relative permittivity of 1, relative permeability of 0.99996 and a conductivity of 41.10^6 S/m.

3.4. Numerical Results

The saturation magnetization is directly related to the volume fraction of the magnetic material. Our starting value of the saturation magnetization was 0.48 T, considering the case of bulk material (100%). By decreasing this concentration, we find that the circulator remains operational but with higher insertion losses, lower isolation and higher reflections. Figure 4 shows the operation of our designed circulator with different concentrations of 100%, 50% and 10%. We note that the S parameters shown here are those between the ports 1 and 2, and evidently we have the same results for the other couples of ports (1 and 3, 2 and 3).

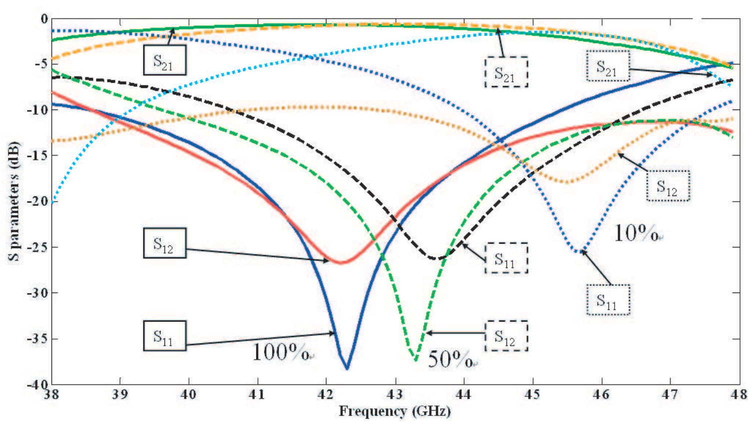


Figure 3: Circulator's performance with different values of M_s (Tesla).

4. MAGNETIC MATERIAL

Practically, our purpose is to replace the conventional barium hexaferrite thin film with a nanocomposite magnetic material. The magnetic thin film, of a thickness less than 100 µm, does not provide optimal results; this is mainly caused by the difficulty of orienting this solid film after deposition.

4.1. Barium Hexaferrite Nanoparticles

Our nanoparticles of barium hexaferrite have flattened hexagonal form with an average size of 200 nm. The synthesis was done in autoclave at 200°C. They are single domain particles. Magnetic measurements show that these particles have a saturation magnetization around 0.47 T and a coercivity of 25 kA/m. These values match those of the bulk Barium hexaferrite.

4.2. Mixture and Nanocomposite

Many experiments of mixtures were done: the resin SU-8 is the dielectric matrix. A volume fraction of 0.1 was reached.

This mixture was spread out on a coplanar line CPW used for characterization. A magnetic field of 0.7 T was applied during this process in the direction of the layer plane, in order to align the nanoparticles inside the nanocomposite. Then, the whole sample was left to dry.

4.3. Microwave Characterization

When an oriented magnetic material is deposited on a line, without any external applied field, the propagation becomes non reciprocal and provides first indications on the material efficiency. The transmission scattering parameters obtained from the microwave characterization, show that non-reciprocity is of 0.562 dB (Figure 5). It is still a low value due to the small quantity of the magnetic material and/or to the insufficient orientation of the particles within the nanocomposite. However, this phenomenon indicates the global composite's orientation which is not observed in the case of non-oriented nanocomposite.

Further rigorous characterization measurements will be made to complete this first analysis and to introduce a permeability tensor model in the design 3D software.

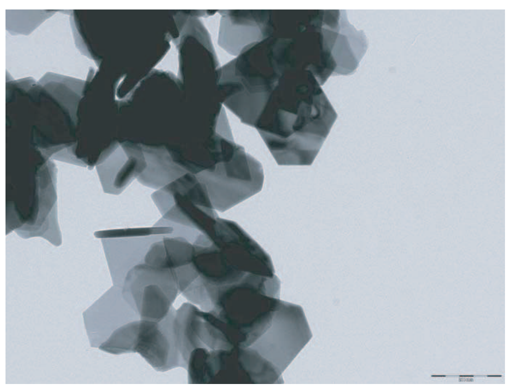
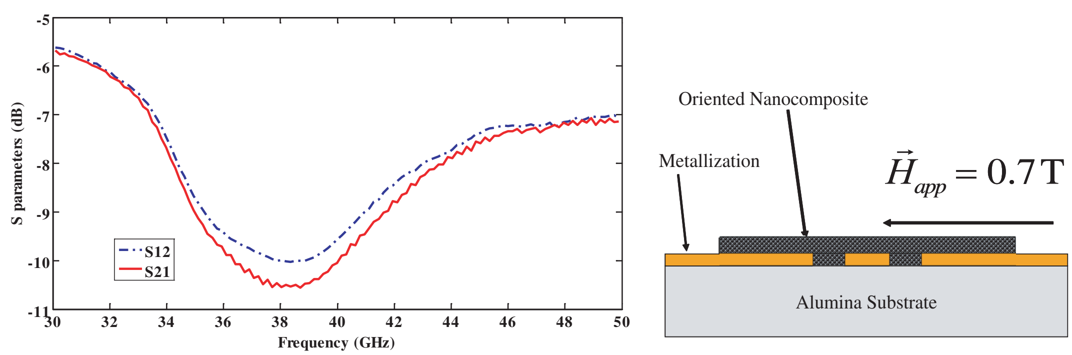


Figure 4: Barium Hexaferrite Nanoparticles.



(a) (b)

Figure 5: Transmission S parameters showing (a) the non-reciprocity of (b) the oriented nanocomposite .

5. CONCLUSION

This paper discusses a numerical study of a circulator, a microwave non-reciprocal device, based on a magnetic material. Structure, geometric dimensions and different characteristics were presented.

In addition, the magnetic ferrite was also described. Measurements were shown to highlight the non-reciprocal effect of the magnetic nanoparticles used to create a nanocomposite magnetic material. Our future aim is to reach higher volume fractions of magnetic material in the composite, consequently, higher non-reciprocal effects.

ACKNOWLEDGMENT

We wish to acknowledge the Rhône-Alpes Region for supporting this research.

REFERENCES

1. Bosma, H., "On the principle of stripline circulation," *The Institution of Electrical Engineers*, 137–146, Jan. 1962.
2. Bosma, H., "On stripline Y-circulation at UHF," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 12, 61–72, Jan. 1964.
3. Fay, C. E. and R. L. Comstock, "Operation of the ferrite junction circulator," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 13, 15–27, Jan. 1965.
4. Wu, Y. S. and F. J. Rosenbaum, "Wide-band operation of microstrip circulators," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 22, 849–856, Oct. 1974.
5. Schloemann, E. and R. E. Blight, "Broad-band stripline circulators based on YIG and Li-ferrite single crystals," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 34, 1394–1400, Dec. 1986.
6. Schloemann, E. F., "Circulators for microwave and millimeter-wave integrated circuits," *Proceedings of the IEEE*, Vol. 76, 188–200, Feb. 1988.
7. Polder, D., "On the theory of ferromagnetic resonance," *Philos. Mag.*, Vol. 40, 99–115, 1949.

A Linear Ultrasonic Motor Using a Quadrate Plate Transducer

Jiamei Jin and Chunsheng Zhao
Precision Drive Laboratory, Nanjing University of Aeronautics and Astronautics
Nanjing 210016, China

Abstract— A linear ultrasonic motor using a quadrate plate transducer has been developed for precision positioning. This motor consists of two pair of Pb(Zr,Ti)O_3 piezoelectric ceramics elements, which are piezoelectrically excited into the second-bending mode of the motor stator's neutral surface in two orthogonal directions, on which the tops of four projections move along an elliptical trajectory, which in turn drive a contacted slider into linear motion via frictional forces. The stator is obtained easily coincident frequency for its coincident characteristic dimension in two orthogonal directions. The performance characteristics achieved by the motor are: (i) a maximum linear speed of > 60 mm/s; (ii) a stroke of > 150 mm; (iii) a driving force of > 5.0 N; (iv) a response time of ~ 2 ms.

1. INTRODUCTION

An ultrasonic motor is different from a coil motor and has many features such as high driving torque at low speed, quick response, zero backlash, self-braking without power. These merits are available for the precise positioners [1–3]. The USM featuring high holding torque and high response characteristics are promised to be used as a precise and accurate positioning actuator [4]. Several linear ultrasonic motors found in literature base on the use of two different vibration modes. Most often flexural and longitudinal modes are combined to achieve an elliptic micro-motion of surface points. This micro-motion is converted to direct linear (or translatory) motion of a slider. To gain high amplitudes of the micro-motion and thus having a powerful motor, the ultrasonic vibrator should be driven near the eigen-frequency of its modes. Additionally, low mechanical and electric losses lead to increased efficiency and large amplitude magnification in resonance. This demands a geometrical design that fits the eigen-frequency of the two different modes. A frequency deviation of only a few percent leads to non-acceptable disturbance of the elliptic motion. Thus, the mechanical design of the vibrator has to be done very carefully. The stator presented in this paper employed a quadrate configuration easy to obtain technologically coincident eigen-frequencies of two same shape modes in two orthogonal directions.

Bolted Langevin- or "sandwich"-transducers are commonly used in many different technical field. For the set-up of a low-cost motor, it seems helpful to use this technology, because its structure is rather simple and production processes are well known [5]. Unfortunately, resonance frequency of this kind of transducers strongly depends on production accuracy and pre-stress. The transducers bonded piezoelectric ceramics with epoxy are more Consistent in output performance. Piezoelectric ceramic plates used by Nanomotion and Physikinstrumente can be produced with high accuracy at low price for large numbers. Therefore, the piezoelectric ceramic plates of stator presented in this paper are bonded by epoxy glue to the quadrate substrate.

The stators of ultrasonic motors operating in intermittent contact scheme drive the rotors (or sliders) when the stators contact the rotors, and the rotors (or sliders) will move in itself inertia after the stators left the rotors. The duty cycle of the contact and the "flight" manage the output force and the velocity of motor. The pattern using alternative work of multi-stator or multi-driving-end of single-stator is promising a large thrust and output velocity [6]. We prepared vibrator that consist of four piezoelectric plates that are bonded by epoxy glue to a quadrate substrate with four projections is a multi-driving-end of single-stator.

2. VIBRATOR DESIGN

The configuration of the vibrator is illustrated in Fig. 1. Two piezoelectric ceramic plates on the cross were poled in reverse directions with respect to each other. Upon application of drive voltage to the two piezoelectric ceramic plates, a bending vibration is generated in one-direction, one expands, whereas the other contracts.

Two pair such piezoelectric ceramic plates were assembled orthogonally in the stator. This results in two orthogonal bending vibration modes, as shown in Fig. 2. The four projections were located at the crest and trough of B_{03} in the y-direction, and at the middle point between the crest

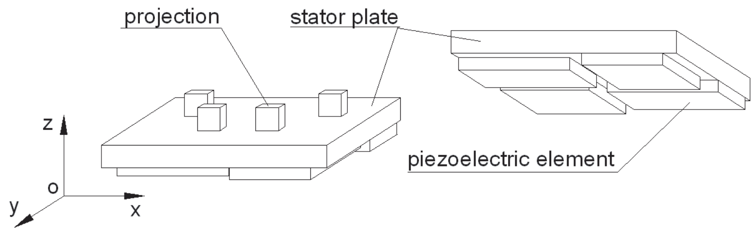


Figure 1: Stator.

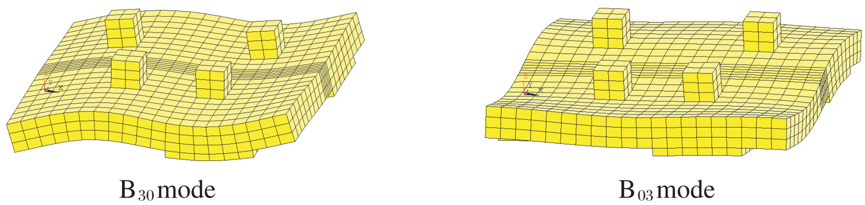


Figure 2: Operating modes.

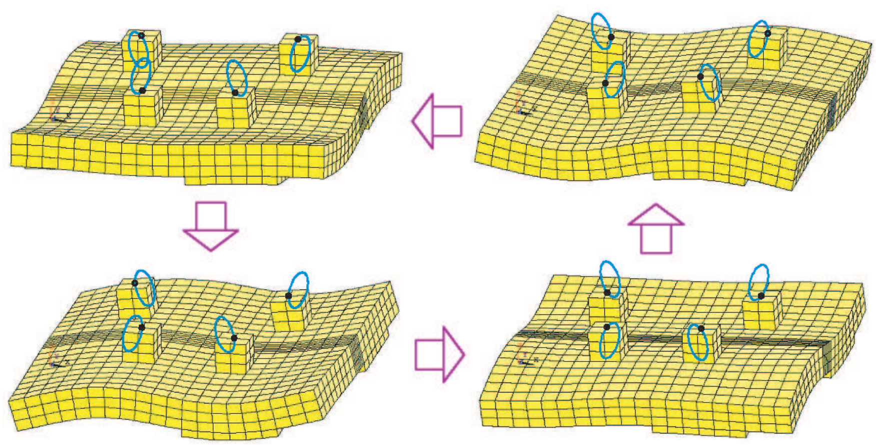


Figure 3: Harmonic analysis.

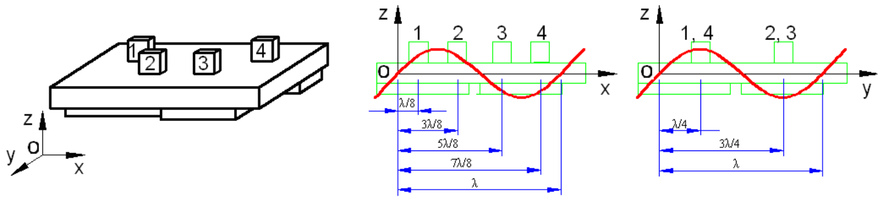
(or trough) and nodal line of B_{30} in the x-direction, respectively. Upon application of two sine drives to the two pair of the piezoelectric ceramic plates (phase shifted by 90°), a vibration can be excited at each projection. The tops of the projections move in ellipses, but with different phase. Analysis using ANSYS finite element method software illustrates the ellipse movement of the tops, as shown in Fig. 3. This square stator operated in its B_{30} and B_{03} bending modes is obtained easily coincident frequency, and could potentially be used for a new slider drive mechanism.

3. MOTION OF DRIVING ENDS

To simplify problem, the deformations of the vibrator during a wavelength λ in two directions are described as sinusoids, as shown as Fig. 4. we have made the following assumptions,

a. The vibrations of B_{30} and B_{03} have coincident amplitude ξ when the vibrator is driven by the same driving signal.
b. The distance between the ends of the projections of the stator and the neutral plane of the vibrator is denoted with h.
c. The rotary angles α of the projections are very small when the stator vibrates. Then, $\sin\alpha \approx \tan\alpha$.

Now, the vibration displacements of the tops of the projections can be described.



(a) (b) (c)

Figure 4: Projections of stator.

(1) In B_{30} mode, as shown as Fig. 4(b), the motions of the tops of the 4 projections can be written as, respectively,

$$\begin{cases} x_{1,(30)}(t) = -\sqrt{2}\pi\frac{h}{\lambda}\xi\sin\omega t \\ z_{1,(30)}(t) = \frac{\sqrt{2}}{2}\xi\sin\omega t \end{cases} \tag{1}$$

$$\begin{cases} x_{2,(30)}(t) = \sqrt{2}\pi\frac{h}{\lambda}\xi\sin\omega t \\ z_{2,(30)}(t) = \frac{\sqrt{2}}{2}\xi\sin\omega t \end{cases} \tag{2}$$

$$\begin{cases} x_{3,(30)}(t) = \sqrt{2}\pi\frac{h}{\lambda}\xi\sin\omega t \\ z_{3,(30)}(t) = -\frac{\sqrt{2}}{2}\xi\sin\omega t \end{cases} \tag{3}$$

$$\begin{cases} x_{4,(30)}(t) = -\sqrt{2}\pi\frac{h}{\lambda}\xi\sin\omega t \\ z_{4,(30)}(t) = -\frac{\sqrt{2}}{2}\xi\sin\omega t \end{cases} \tag{4}$$

(2) In B_{03} mode, as shown as Fig. 4(c), the motions of the tops of the 4 projections can be written as, respectively,

$$z_{1,(03)}(t) = z_{4,(03)}(t) = \xi\cos\omega t \tag{5}$$

$$z_{2,(03)}(t) = z_{3,(03)}(t) = -\xi\cos\omega t \tag{6}$$

Let $\varphi = \tan^{-1}\sqrt{2}$, when the B_{30} and B_{03} modes are excited with phase shifted by $\pi/2$, the motions of the tops of the 4 projections can be written as, respectively,

$$\begin{cases} x_1(t) = x_{1,,(30)}(t) = \sqrt{2}\pi\frac{h}{\lambda}\xi\sin(\omega t+\pi) \\ z_1(t) = z_{1,(30)}(t) + z_{1,(03)}(t) = \sqrt{\frac{3}{2}}\xi\sin(\omega t+\varphi) \end{cases} \tag{7}$$

$$\begin{cases} x_2(t) = x_{2,(30)}(t) = \sqrt{2}\pi\frac{h}{\lambda}\xi\sin(\omega t) \\ z_2(t) = z_{2,(30)}(t) + z_{2,(03)}(t) = \sqrt{\frac{3}{2}}\xi\sin(\omega t-\varphi) \end{cases} \tag{8}$$

$$\begin{cases} x_3(t) = x_{3,(30)}(t) = \sqrt{2}\pi\frac{h}{\lambda}\xi\sin(\omega t) \\ z_3(t) = z_{3,(30)}(t) + z_{3,(03)}(t) = \sqrt{\frac{3}{2}}\xi\sin(\omega t+\varphi+\pi) \end{cases} \tag{9}$$

$$\begin{cases} x_4(t) = x_{4,(30)}(t) = \sqrt{2}\pi\frac{h}{\lambda}\xi\sin(\omega t+\pi) \\ z_4(t) = z_{4,(30)}(t) + z_{4,(03)}(t) = \sqrt{\frac{3}{2}}\xi\sin(\omega t-\varphi+\pi) \end{cases} \tag{10}$$

The tops of the projections move in ellipses, but with different phase in φ. The tops move to their max positions in z-direction at different moment, and the pattern of alternative work of multi-drive-end of single-stator is available when the driving ends are used to drive a slider.

4. PROTOTYPE AND PERFORMANCES

Figure 5 illustrates the construction and working principle of the quadrate linear ultrasonic motor. The stator's projections were elastically pressed together in order to ensure frictional contact with the slider. The slider was then excited into linear motion, by contact to the projections of the stator. The operating pattern, four projections drive the slider by turn in one vibration cycle of the stator, is promising a large thrust and a quick velocity.

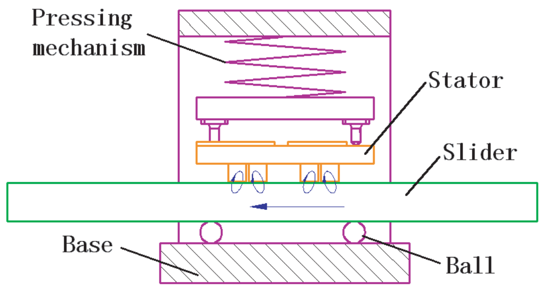


Figure 5: Operating principle of linear ultrasonic motor.

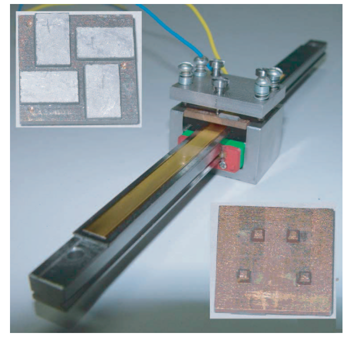


Figure 6: Prototype of linear ultrasonic motor.

The stator was a square plate with a thickness of 2 mm and a length of 20 mm, and its each projection has a dimension of $2 \times 2 \times 2\,\text{mm}^3$. Each piezoelectric ceramic plate coated silver electrode has a dimension of $10 \times 6 \times 0.5\,\text{mm}^3$. A grating with a resolution of 40 nm was fixed on the slider in order to investigate experimentally the velocity and displacement of the motor, as shown in the Fig. 6. The resonant modes of the assembled stator were measured using a Doppler Laser Vibrometer OFV056(Polytec). A distinct B_{30} bending mode with resonant frequency of 57.31 kHz was found and the resonant frequency of B_{30} bending mode was 57.36 kHz. The main performance characteristics of the prototype were experimentally determined. Using a pair of drive voltages of 150–200 Vpp (phase shifted by 90°) that were applied to the stator, the traveling speed of the slider was found to be 120–180 mm/s. By applying weight to the rotor, the maximum load (driving force) of the motor was determined to be ∼5.3 N.

5. CONCLUSION

A square plate, quadrate projections, linear-type ultrasonic motor has been developed with characteristics of (i) a maximum linear speed of > 60 mm/s; (ii) a stroke of > 150 mm, which in principle is not limited from being higher; (iii) a driving force of > 5.0 N; (iv) a response time of ∼ 2 ms. The coincident frequency of two operating modes is obtained easily because its stator has coincident characteristic dimension in two orthogonal directions. The tops of four projections move along an elliptical trajectory, which in turn drive a contacted slider into linear motion via frictional forces. The work pattern is promising a large thrust and output velocity.

ACKNOWLEDGMENT

The work is supported by National Natural Science Foundation of China (No. 50735002).

REFERENCES

1. Ueha, S. and Y. Tomikawa, *Ultrasonic Motors: Theory and Applications*, 4–6, Clarendon, Oxford, 1993.
2. Sashida, T. and T. Kenjo, *An Introduction to Ultrasonic Motors*, 6–8, Clarendon, Oxford, 1993.
3. Uchino, K., *Piezoelectric Actuators and Ultrasonic Motors*, Kluwer Academic, Boston, 1997.
4. Senjyu, T., T. Kashigawi, and K. Uezato, "Position control of ultrasonic motors using MRAC and dead-zone compensation with fuzzy inference," *IEEE Trans. Power Electr.*, Vol. 17, 265, 2002.
5. Zhao, C., *Ultrasonic Motors Technologies and Applications*, Science Press, Beijing, 2007.
6. Jin, J. and C. Zhao, "A method for improving the performances of ultrasonic motors using intermittent contact scheme," *IWPMA*, Vol. 4, 307, 2007.

A Novel LLCC Resonant Network for Ultrasonic Motor

Hua-Feng Li and Wei-Qing Huang
Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 210016, China

Abstract— High driven voltage is required for ultrasonic motor, so the driving circuit needs transformer to boost voltage. Since the USM is a nonlinear capacitive load, a matching circuit is needed to keep the output voltage and the phase shift constant. For the switching circuit with push-pull topology, the switching loss is less under inductive load rather than capacitive load, which will increase its efficiency and reliability. A novel LLCC resonant network was proposed and its parameters were given. Theoretical analysis and simulation were presented, which demonstrated that the output voltage and phase shift of the resonant network was independent of driving conditions, such as driving frequency and load, and the load characteristics can be changed from capacitive to inductive. This novel driver is applied to the TRUM-60 type ultrasonic motor, and the experimental results are presented to verify the effectiveness of the proposed circuit.

1. FOREWORD

The Ultrasonic Motor (USM) is a new concept of new motor. It utilizes the reverse piezoelectric effect of piezoelectric materials and ultrasonic vibration to convert the micro-deformation of elastic materials (stator) to the macro movement of rotor or slider by the resonance amplification and friction coupling [1–4]. Compared with the traditional electromagnetic motors, ultrasonic motors have characteristics which are small inertia, fast responses, no electromagnetic interference, no producing magnetic fields and so on. Ultrasonic motors have been used in aerospace engineering and military fields such as spacecraft, missiles and nuclear warheads. They also have been used in civilian areas such as auto-focus lens and quartz watches [5–8].

Although the USM has the above advantages, its control circuit becomes very complicated because of the nonlinear and time varying of the internal parameters. The driving circuit which is designed based on these characteristics requires that the amplitude and the phase of the output voltage do not change with the frequency and load change [9, 10]. When the driving voltage is sine wave, the mechanical vibration characteristics of USM are the best. So it is very common to use the resonant circuit to produce sine wave. Each kind of resonant circuits has its advantages and disadvantages. When they are applied to USM, the amplitude and phase of output voltage are easily changed with the load changes. The common method to keep the amplitude of output voltage constant is using the output voltage feedback to adjust the input voltage. But this method will increase the complexity of the driver.

The supply voltage of driving circuit of USM is generally low, but the voltage amplitude for USM is high. So the driving circuit of USM needs transformer to achieve isolation and boost. We can choose push-pull, half-bridge and full-bridge circuits for this purpose. Literature [11] proposed a LLCC resonant network based on half-bridge topology. In the push-pull circuit, the MOSTFETs have the common ground so that they do not need isolation. The push-pull circuit has simple structure which is especially suitable for the USM driver at low supply voltage. It is the most common topology circuit at present. This paper uses the push-pull circuit as the driving circuit of USM, designs and analyzes the LLCC resonant circuit to make the voltage amplitude and the phase not change with the load and the driving frequency. The output voltage of motor does not need feedback loop control so that the circuit structure is simple.

2. PUSH-PULL DRIVING CIRCUIT OF USM

Near the resonance frequency, the USM can be represented by an equivalent circuit in Fig. 1(a) [12], where C_d is the static capacitance of PZT, L_m is equivalent inductance, C_m and R_m are capacitance, resistance related to the stator's mass, elasticity and mechanical loss, respectively, which construct a dynamic branch. They determine the electromechanical coupling characteristics of the motor. The values of these parameters are determined not only by the stator itself, but also by the pre-pressure rotor acting on the stator. With the condition that not change the external characteristic of the circuit, the equivalent circuit of the motor can be transformed to a RC parallel circuit, shown

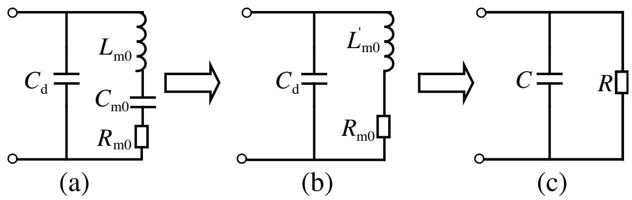


Figure 1: Transform of USM equivalent circuit.

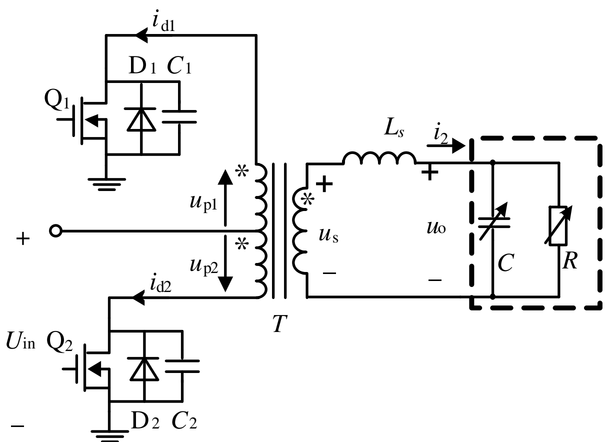


Figure 2: Push-pull driving circuit for USM (single phase).

in Fig. 1(c). The transformation relation of the parameters are derived as:

$$\begin{aligned} L'_{m0} &= L_{m0} - \frac{1}{\omega^2 C_{m0}} \\ C &= C_d - \frac{L'_{m0}}{R_{m0}^2 + (\omega L'_{m0})^2} \\ R &= R_{m0} + \frac{(\omega L'_{m0})^2}{R_{m0}} \end{aligned} \tag{1}$$

The Push-pull driving circuit for USM (single phase) is shown in Fig. 2, where D_1 (D_2), C_1 (C_2) are the body diode and the junction capacitance of MOSTFET Q_1 (Q_2), respectively. The equivalent circuit of USM is shown in the dashed box of the figure. Both the equivalent capacitance C and the equivalent resistance R change with load and frequency. L_s is a series matching inductor.

As shown in the Fig. 2, the amplitude and phase of the output voltage change with different C because it is obtained by the resonance of the series matching inductance L_s and the equivalent capacitance C of USM [13]. Fig. 3 is the waveforms of the two-phase voltages of TRUM-60 (its electrical parameters are shown in Table 1) which is drove by the series inductance of 1.5 mH. It is clearly seen from the figure that the amplitude of the single-phase voltage varies widely and both the amplitude and phase shift of the two-phase voltages are different. This is the inherent weaknesses of using single series inductance resonant circuit. So it is necessary to study other forms of resonant circuits.

3. LLCC RESONANT NETWORK

The functions of the resonant network which is between the driving circuit and the USM are as follows: voltage amplitude and phase of the USM does not change with the load and driving frequency change; The load of the driving circuit is resistive. Therefore, this article proposes a LLCC resonant network shown in Fig. 4 where the equivalent capacitance and resistance of Z_1 which represents USM are variable; L_m, L_s and C_s are the external components, which constitute the LLCC resonant network with the capacitance C. The input and output of the network are the high-voltage square wave U_s and the voltage of USM, respectively.

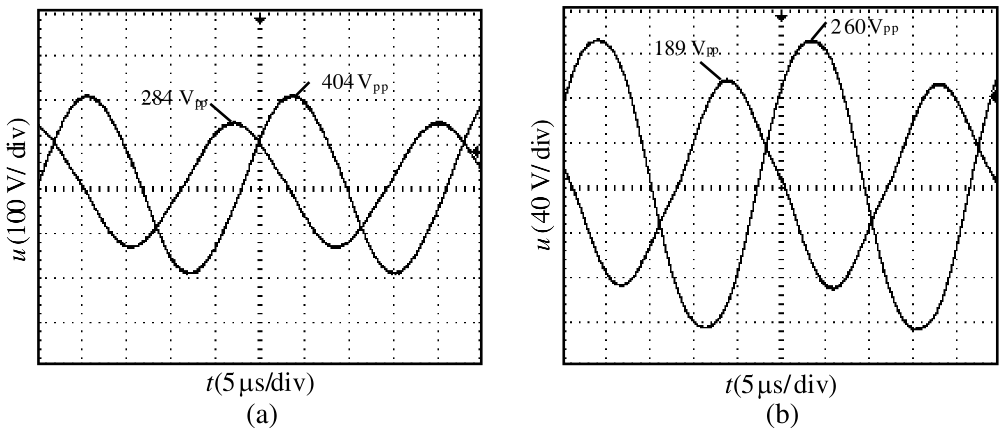


Figure 3: Output voltages vs. frequency with L_s resonance. (a) $f = 43.42$ kHz. (b) $f = 41.19$ kHz.

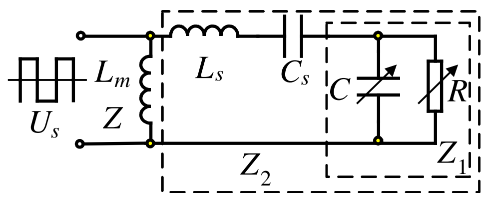


Figure 4: LLCC resonant network.

Thus the voltage gain and phase of resonant network can be obtained.

$$G_V = \frac{Z_1}{Z_2} = \frac{1}{(1 + \frac{C}{C_s} - \omega^2 L_s C) + j(\frac{\omega L_s}{R} - \frac{1}{\omega R C_s})} \tag{2}$$

$$\angle G_V = \text{arctg}\frac{\frac{\omega L_s}{R} - \frac{1}{\omega R C_s}}{1 + \frac{C}{C_s} - \omega^2 L_s C} \tag{3}$$

It can be seen from Eq. (2) that when the imaginary part of the denominator is 0 and the real part is 1, the voltage gain is 1 and the phase of the resonant network is 0 in Eq. (3). Thus we can get

$$\begin{cases} \frac{1}{R}(\omega L_s - \frac{1}{\omega C_s}) = 0 \\ 1 + C(\frac{1}{C_s} - \omega^2 L_s) = 1 \end{cases} \tag{4}$$

Because generally R is high (several kΩ) and C is small (several nF), when the L_s and C_s are in series resonant, the voltage amplitude and phase of USM do not change with the frequency and load.

The total impedance of resonant network and USM is

$$Z = \frac{j\omega L_m Z_2}{j\omega L_m + Z_2} = \frac{Z_{up}}{Z_{down}} \tag{5}$$

when $\angle Z_{up} = \angle Z_{down}$, the imaginary part of the total impedance is zero and the driving load is resistive, then we have

$$C + C_s - \omega^2 C C_s (L_m + L_s) = 0 \tag{6}$$

It can be obtained from (4) and (6) that the parameters of the network which meet the functions of the above resonant network are

$$\begin{cases} \omega^2 L_s C_s = 1 \\ C + C_s - \omega^2 C C_s (L_m + L_s) = 0 \end{cases} \tag{7}$$

In order to verify the correctness of the proposed network, we use the parameters in Table 1 to make a simulation analysis of the resonant network. Fig. 7 is the relationships between the voltage gain as well as phase and the driving frequency.

Table 1: Parameters for USM and resonant network.

f (kHz)	C (nF)	R (kΩ)	L_m (mH)	L_s (mH)	C_s (nF)
$38 \sim 42$	$6 \sim 7$	$2 \sim 6$	2.6	1.0	15

Table 2: Output voltage vs. frequency.

f (kHz)	41.49	41.67	42.48	42.7	42.99	43.22	43.59
V_o (V_{pp})	496	492	496	496	492	496	496

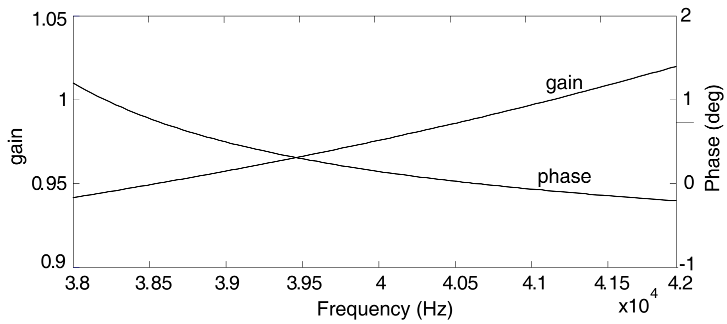


Figure 5: Gain and phase of resonant network vs. frequency.

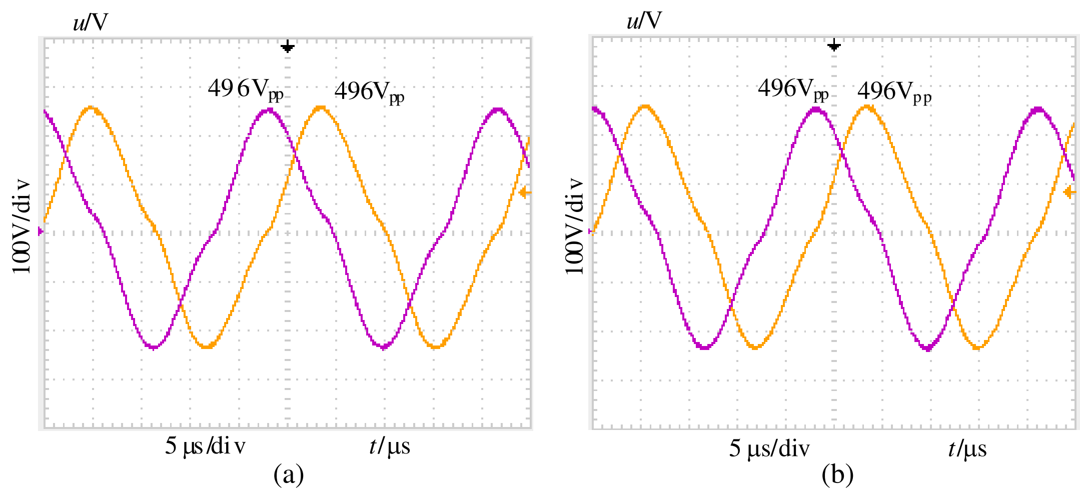


Figure 6: Output voltages vs. frequency with LLCC resonance. (a) Driving frequency is 42.48 kHz. (b) driving frequency is 43.59 kHz.

Figure 5 shows that the voltage gain and the phase shift of the resonant network are very small and it meets the design requirements.

4. EXPERIMENTAL RESULTS

To verify the rationality of the above analysis, we do some experiments by using the USM of TRUM-60 made by the institute. The parameters of USM and resonant network are shown in Table 1. The input voltage of the driving circuit is 12 V and the transformer ratio of the push-pull circuit is 1 : 15.

Figure 6 are the waveforms of the two-phase voltages at different frequencies. Table 2 is the single-phase voltage at different frequencies. It can be seen from the results that the amplitudes and phase shift of USM do no change with the driving frequency and the two phase voltages can maintain the same, so that the LLCC resonant network meets the design requirements.

Figure 7 is the relationship between the output voltage of USM and the DC input voltage at a fixed frequency. The figure shows that they have a very good linear relationship.

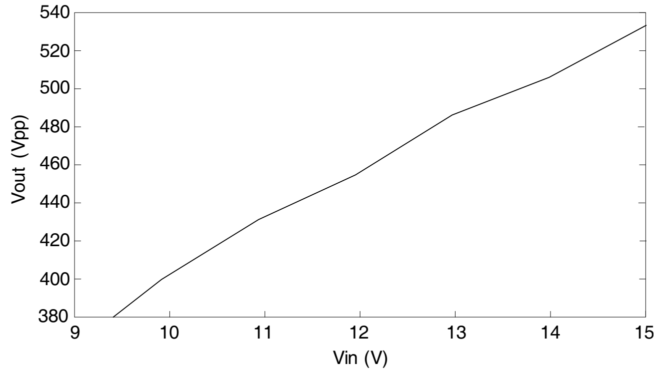


Figure 7: Output voltages vs. input voltage with LLCC resonance.

5. CONCLUSION

In this paper, to overcome the inherent shortcomings of traditional resonant network, a new LLCC resonant network is proposed and the calculating method is given. The LLCC resonant network achieves the decoupling of motor voltage and phase shift and driving conditions. The experimental verification by using the USM of TRUM-60 is given. The experimental results show that the proposed method is correct and the LLCC resonant network lays a good foundation for the development of high-performance driver of USM.

ACKNOWLEDGMENT

This project is supported by National Natural Science Foundation of China (50875123, 50735002) and National Natural Science Foundation of Jiangsu Province (Bk2009373).

REFERENCES

1. Sashida, T., "Trial construction and operation of an ultrasonic vibration–driven motor [J]," *Japan Applied Physics*, Vol. 51, No. 6, 713–720, 1982.
2. Sashida, T. and T. Kenjo, *An Introduction to Ultrasonic Motors [M]*, 4–10, Oxford, London, UK, 1993.
3. Xia, C., Y. Zheng, and T. Shi, "FEM analysis stator vibration of traveling-wave type contact ultrasonic motor [J]," *Proceedings of the CSEE*, Vol. 21, No. 2, 25–32, 2001 (in Chinese).
4. Li, H. and C. Gu, "Research on large torque travelling wave ultrasonic motor [J]," *Proceedings of the CSEE*, Vol. 22, No. 8, 67–70, 2002 (in Chinese).
5. Ueha, S. and Y. Tomikawa, *Ultrasonic Motors Theory and Applications [M]*, Oxford Science Publications, London, UK, 15–18, 1993.
6. Zhao, C., "Ultrasonic motor techniques for 21st century [J]," *Engineering Science*, Vol. 4, No. 2, 86–91, 2002 (in Chinese).
7. Zhe, X., W. Chen, and Z. Chen, "Development of ultrasonic motor in USA [J]," *Piezoelectrics & Acoustooptics*, Vol. 1, No. 21, 37–40, 1999 (in Chinese).
8. Hosoe, K., "An application of an ultrasonic motor to automatic focusing lenses [C]," *25th Tohoku Univ. Tsuuken Symp.*, Japan, 1989
9. Lin, F.-J., R.-J. Wai, and H.-H. Lin, "An adaptive fuzzy-neural-network controller for ultrasonic motor drive using the LLCC resonant technique [J]," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 46, No. 3, 715–727, 1999.
10. Li, R., N. Frohleke, and J. Bocker, "LLCC-PWM inverter for driving high-power piezoelectric actuators [C]," *2008 13th International Power Electronics and Motion Control Conference*, 159–164, 2008.
11. Lin, F., R. Wei, and R. Duan, *Driving and Intelligent Control of Ultrasonic Motor [M]*, 73–89, Sea Publisher, Central Taiwan, 1999.
12. Zhao, C., *Ultrasonic Motors Technologies and Applications [M]*, 226, Science Press, Beijing, 2007.
13. Shi, J. and H. Wang, "Analysis of push-pull convertor used in ultrasonic motor drive," *Electrical Machine and Control Application*, Vol. 36, No. 1, 1–4, 2009.

Theory and Experiment of the Valveless Piezoelectric Pump with Rotatable Unsymmetrical Slopes

Jianhui Zhang[1], **Qixiao Xia**[2], **Chunsheng Zhao**[1], **and Jiamei Jin**[1]

[1]Precision Driving Laboratory, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
[2]College of Mechanical & Electronic Engineering, Beijing Union University, Beijing 100020, China

Abstract— With no relative motion at joint, no internal contamination, low manufacturing cost, and easy to microminiaturization, the valveless piezoelectric pump has unrivaled application in one-time use and miniaturization. However, the microminiaturization of the pump itself cannot microminiaturize the whole system including the pump. Thus, this paper focuses on how to integrate mixing and transport of piezoelectric liquid. Furthermore, the difficult and challenging issues of the ratio of mixing fluid and composition control are also discussed. The valveless piezoelectric pump with rotatable unsymmetrical slopes is put forward under this background. Its major features are setting drive and transfer in one, also setting transfer and mixing in one. Firstly, the design of the valveless piezoelectric pump with rotatable unsymmetrical slopes is proposed and the one-way flow principle is analyzed. Then, the fluid mechanics model of the valveless piezoelectric pump with rotatable unsymmetrical slopes is established. Meanwhile, the flow numerical analysis and calculation of the pump cavity are done. Finally, the experiments on relationship between the rotation angles of the slope and the flow rate and slope angles of rotation and the two inlets flow ratio are conducted. The experimental results show that: the maximum flow reaches 32.32 ml/min. The maximum relative error between the theoretical results and the experimental ones is 14.59%. According to the curve indicating the rotating different angles of rotatable unsymmetrical slopes to the flow ratio of the two inlets, the experimental and theoretical maximum relative error is 3.75%. Thus, the principle feasibility of the valveless piezoelectric pump with rotatable unsymmetrical slopes and the theory reliability are verified and proven.

1. INTRODUCTION

By integrating the driving into the transmission, the valveless piezoelectric pump eliminates the traditional mechanical transmission chain. What is more, this type of pump shares the features of no relative motion at joint of drive part, no internal contamination caused by abrasive wear and lubrication. So, it has the advantages: low processing costs, be easy to microminiaturize. This valveless piezoelectric pump owns unrivaled application in one-time use and miniaturization [1–7]. However, only miniature pump also does not necessarily in system miniaturization including pump. Therefore, the research valveless piezoelectric pump is continually expanding functional integration, particularly in the integration of piezoelectric liquid mixing and transportation functions and applications [8].

J. C. Rife, etc. proposed the piezoelectric devices designed for liquid mixture in 2000 [8]. The essence is that piezoelectric vibrator drives obstacles in the mixing tank to produce vortex, to achieve uniform mixing effect. However, it only can do liquid mixture and cannot transport.

However, no variability and non-regulatory of unsymmetrical slopes, cone block, and staggered convex determine non-regulatory of ratio of mixture. Thus, these previous results only involves simple mixing and transport integration, and does not take into account the liquid mixing ratio and composition control needed transport.

Utilize underlying mechanisms of unsymmetrical slopes valveless piezoelectric pump working principle, this paper provides a valveless piezoelectric pump with rotatable unsymmetrical slopes (VPPRUS) which is easy to mixing and adjustable mixing ratio control.

Firstly, the design of the VPPRUS is proposed and the one-way flow principle is analyzed. Then, the fluid mechanics model of the VPPRUS is established. Meanwhile, the flow numerical analysis and calculation of the pump cavity are done. Finally, the experiments of pumping capacity are conducted. Thus, the principle feasibility and the theory reliability are proven.

2. VALVELESS PIEZOELECTRIC PUMP WITH UNSYMMETRICAL SLOPES

Unsymmetrical slopes are the key component of the valveless piezoelectric pump with unsymmetrical slopes. They are arranged of several groups of slopes whose slope angles in the same direction are equal, while unequal in the different direction ($\alpha_1 \neq \alpha_2$). As shown in Figure 1.

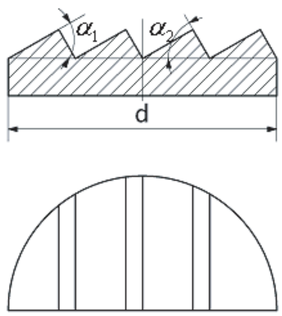


Figure 1: Unsymmetrical slopes.

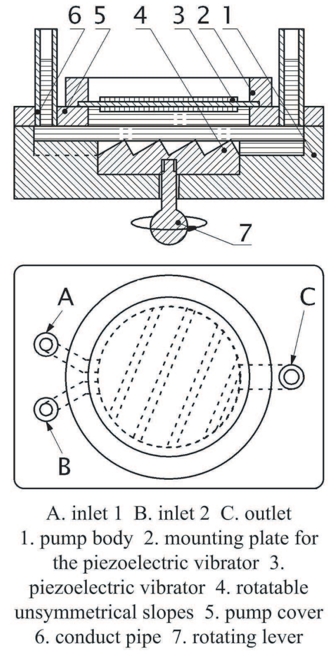


Figure 2: Valveless piezoelectric pump with rotatable unsymmetrical slopes.

3. THE STRUCTURE OF THE VPPRUS

Figure 2 shows the structure of the VPPRUS. To facilitate the flowing of different mixing fluid, two inlets and one outlet are designed; In order to achieve mixing ratio can be controlled, the unsymmetrical slopes is rotatable.

When pump working, the two inlets are filled with two different fluids. Rotating unsymmetrical slopes can change the flow resistance of each inflow and outflow, thereby achieving control of the flow proportion of two inlets. Direct-flow is caused by one-way mechanism of fluid through the unsymmetrical slopes. Meanwhile, the liquids mixed by the complex turbulence and vortex through unsymmetrical slopes are transported to the outlet, and then the functions of simultaneous mixing and delivery are completed.

Flow field and vortex in cavity will be change after the changing of geometrical position of the rotatable unsymmetrical slopes. When the piezoelectric vibrator working, because the slopes are rotated, the fluid can flow into pump chamber through two inlets with needed proportion, and can mix with each other under the effect of rapid turbulent in cavity. The vortex is the driving force of liquid mixing. The higher swirl intensity, the more complex form, the greater ability of mixing and stir are.

4. THE PRINCIPLE OF THE VPPRUS

When pump working, driven by the alternating voltage, the center of the piezoelectric vibrator warps upward and downward periodically caused by the peripheral fixing, so that the volume of the pump chamber periodically changes. This change cycle T is the reciprocal of the driving frequency f $(T = 1/f)$. The flow principles of suction and discharging cycle are indicated in Figure 3 and Figure 4 respectively. To illustrate conveniently and simply, the piezoelectric vibrator model of peripheral fixing and center warping is reduced to a type of piston in average change.

5. EXPERIMENTS

To verify the effectiveness of the theoretical analysis and simulation, this research develops the pump for experimental verification. The design parameters of the pump chamber, the flow channel and the rotatable unsymmetrical slopes are the same with that of the theoretical analysis. Table 1 shows the geometrical parameters and the drive parameters of the pump. Figure 5 is the photograph of the VPPRUS for the experiments.

Figure 6 shows when the slope angle $\alpha_1 = 30°$, the theoretical and experimental relationship curves between rotating different angles of the rotatable unsymmetrical slopes and the flow rates.

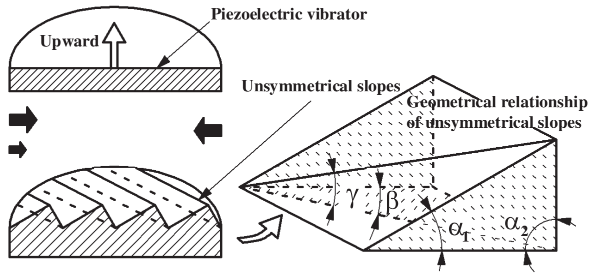


Figure 3: The flow principle of suction cycle.

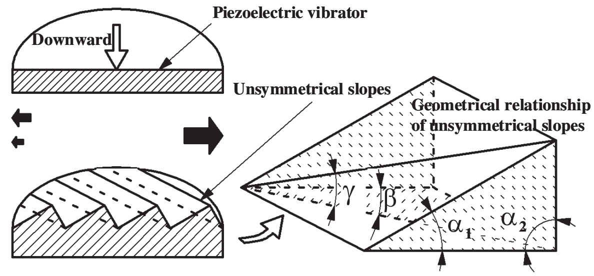
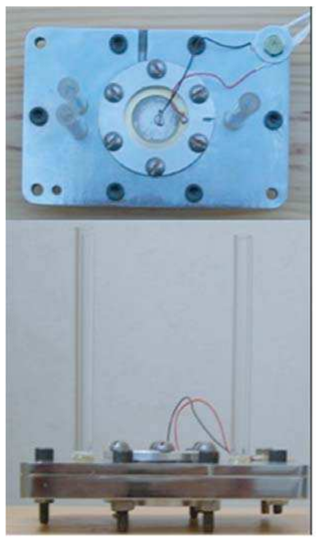
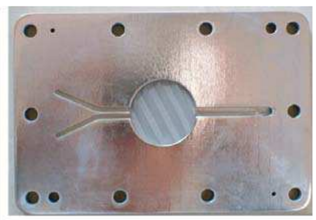


Figure 4: The flow principle of discharging cycle.



(a) Experimental pump



(b) The pump chamber

(c) The rotatablc unsymmetrical slopes

Figure 5: The photograph of the VPPRUS.

The data of the theoretical curve are from the simulation and calculation of last chapter. The results show that: the data of experimental results and that of theoretical analysis are the same trend, but the error also increases with the flow rate increases. The maximum experimental flow rate $Q_S = 32.32\,\mathrm{ml/min}$, and the maximum theoretical flow rate $Q_L = 37.31\,\mathrm{ml/min}$. The maximum relative error of the pump rate between experiment and theory is 14.59% when $\beta = 9^\circ$.

Figure 7 shows when the slope angle $\alpha_1 = 30^\circ$, the relationship curve of the ratio n_L between the theoretical flow rate of inlet 1 Q_{1L} and the one of inlet 2 Q_{2L}, the ratio n_S between experimental flow rate of inlet 1 Q_{1S} and the one of inlet 2 Q_{2S}, and rotating different angle β of rotatable unsymmetrical slopes. The specific ratio is given from formula (1).

$$n_i = Q_{i1}/Q_{i2} \qquad i = L, S \tag{1}$$

The results show that: the experimental ratio between the slopes rotation angle and the flow rate of two inlets is only slightly lower than the theoretical value. From the curve between the rotating different angles of the rotatable unsymmetrical slopes and the ratio between the inlet 1 and inlet 2, the experimental and the theoretical maximum relative error is 3.75%, when $\beta = 14^\circ$.

A systematic error exists in the theoretical and experimental values. The roughness error of the flow passage in pump chamber during theoretical calculation, and the multiple turbulence error, etc, have not been considered in the numerical modeling. The fluid loss on the wall during experimental measurement, the impurities in fluids, and other factors will all affect the accuracy of the experiment.

See references [9–12], the non-regulatory changes of the unsymmetrical slopes, the tapered block and the staggered convex determine non-regulatory issues of the mixing ratio. However, it is the VPPRUS that overcomes the above-mentioned characteristics. To achieve the integration of

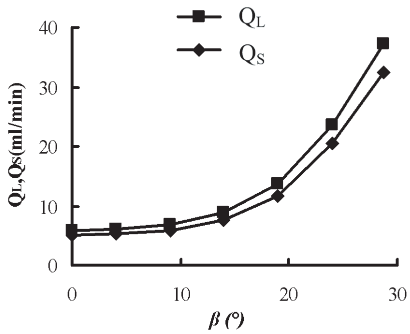


Figure 6: The curve between the slope rotation angles and the flow rates.

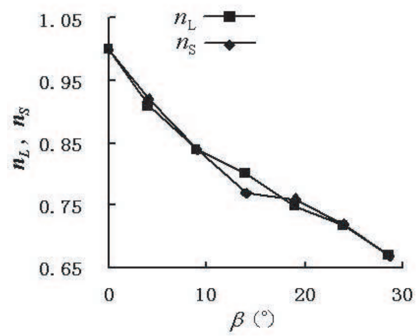


Figure 7: The curve between the slope rotation angles and the flow rate ratios of two inlets.

piezoelectric liquid mixing and delivery, and make the overall system miniaturized possible. To achieve setting drive and transfer in one, also setting transfer and mixing in one.

6. CONCLUSIONS

(1) This paper presents the VPPRUS. This pump achieves the integration of piezoelectric liquid mixing and delivery, and makes the overall system miniaturized possible. To achieve setting drive and transfer in one, also setting transfer and mixing in one.

(2) The working principle of the VPPRUS is analyzed, the formula for pump flow rate relationship is established, and simulation and numerical methods are used to verify the above theories.

(3) The flow field and the flow rate of the VPPRUS are simulated, and the flow fields after the inlet 1, inlet 2 and outlet are separated alone are simulated. In the role of the rotatable unsymmetrical slopes, pressure in pump chamber is uneven, and the multiple turbulences are resulted. The turbulences generated in inlet 1 and 2 are different, then the flow rates are also different, and the proportional control of flow rate can achieve fluid mixing.

(4) The experimental pump is designed. The relationships between the slope rotation angles and the flow rates, and the relationship between the slope rotation angles and the flow rate ration of two inlets are all experimented. The results of slope rotation angle and flow rate show that: the data of experimental results and that of theoretical analysis are the same trend, but the error also increases with the flow rate increases. The maximum experimental flow rate $Q_S = 32.32\,\text{ml/min}$, and the maximum theoretical flow rate $Q_L = 37.31\,\text{ml/min}$. The maximum relative error of the pump rate between experiment and theory is 14.59% when $\beta = 9°$.

The results of the slopes rotation angles and the flow rate ratios of two inlets show that: From the curve between the rotating different angles of the rotatable unsymmetrical slopes and the ratio between the inlet 1 and inlet 2, the experimental and the theoretical maximum relative error is 3.75%, when $\beta = 14°$. The experimental value is only slightly lower than the theoretical value.

ACKNOWLEDGMENT

This project is supported by National Natural Science Foundation of China (Grant No. 50775109), the National Science Foundation of China important project (Grant No. 50735002) and National Natural Science Foundation of China (Grant No. 51075201).

REFERENCES

1. Erik, S. and S. Göran, "A valve-less diffuser/nozzle-based fluid pump," *Sensors and Actuators A*, Vol. 39, No. 12, 159–167, 1993.
2. Gerlach, T., "Microdiffusers as dynamic passive valves for micropump applications," *Sensors and Actuators A*, Vol. 69, 181–191, 1998.
3. Ivano, I., A. Dino, et al., "Modelling and experimental validation of a piezoelectric micropump with novel no-moving-part valves," *Sensors and Actuators A*, Vol. 133, 128–140, 2007.
4. Ribeiro, M. F. T., J. L. M. Santos, and J. L. F. C. Lima, "Piezoelectric pumping in flow analysis: Application to the spectrophotometric determination of gabapentin," *Analytica Chimica ACTA*, Vol. 600, 14–20, 2007.

5. Morris, C. J. and F. K. Forster, "Low-order modeling of resonance for fixed-valve micropumps based on first principles," *Journal of Microelectromechanical Systems*, Vol. 12, No. 3, 325–334, 2003.
6. Nguyen, N.-T. and X. Huang, "Numerical simulation of pulse-width-modulated micropumps with diffuser/nozzle elements," *Technical Proceedings of the 2000 International Conference on Modeling and Simulation of Microsystems*, 636–639, San Diego, California, USA, March 27–29, 2000.
7. Zhang, J., Y. Li, and Q. Xia, "Analysis of the pump volume flow rate and tube property of the piezoelectric valveless pump with Y-shape tubes," *Chinese Journal of Mechanical Engineering*, Vol. 43, No. 11, 136–141, 2007.
8. Rife, J. C., M. I. Bll, et al., "Miniature valveless ultrasonic pumps and mixers," *Sesors and Actuators*, Vol. 86, 135–140, 2000.

Design of a Multilayer Composite-Antenna-Structure by Spiral Type

D. Kim[1], **J. Kim**[1], **J. Kim**[3], **W.-S. Park**[2], **and W. Hwang**[1]

[1]Department of Mechanical Engineering, POSTECH, Republic of Korea
[2]Department of Electronic and Electrical Engineering, POSTECH, Republic of Korea
[3]Antenna R&D Group, Mobile Comm. Samsung Electronics Co., Ltd., Republic of Korea

Abstract— We study a composite-antenna-structure (CAS) having high electrical and mechanical performances that we have designed and fabricated. The CAS, consisting of a glass/epoxy face sheet and a honeycomb core, acts as a basic mechanical structure, in which a spiral antenna type is embedded. To increase the intensity, a carbon fiber plate is used as a bottom sheet. This structure of the $0.5 \sim 2\,\text{GHz}$ band has a gain of $5 \sim 9\,\text{dBi}$ with circular polarization characteristics and reflection loss below $-10\,\text{dB}$ within the desired frequency band.

1. INTRODUCTION

In the last 15 years there has been much research into the embedding of antennas in load-bearing structural surfaces of aircraft, so as to improve both structural efficiency and antenna performance [1–3]. Structural, material and antenna designers have collaborated to develop a novel high-payoff technology known as a Conformal Load-bearing Antenna Structure (CLAS) [3]. This technology shows great promise for enhancing the performance and capability of aircraft, by reducing weight, improving the structural efficiency of airframes that contain antennas, and improving the electromagnetic performance of antennas. To develop the load-bearing antenna structure, we proposed the use of antenna-integrated composite structures of sandwich construction, specifically the surface-antenna-structure (SAS) [4–6] and the composite-smart-structure (CSS) [7–9]. In those studies, we designed and fabricated a microstrip antenna structure which implemented satellite communication in the X band ($8.2 \sim 12.4\,\text{GHz}$) and the Ku band ($12.4 \sim 18\,\text{GHz}$). At such high frequencies the microstrip antenna has only a small bandwidth. In the present paper we report a new CAS based on a spiral antenna type giving good performance in a low frequency band with higher bandwidth [10, 11]. A sandwich composite consisting of a glass/epoxy face sheet and honeycomb core is used as a basic mechanical structure, in which a spiral antenna type is embedded. To increase the intensity, a carbon fiber plate is used as the bottom sheet.

2. STRUCTURE AND MATERIALS

The basic design concept of the CAS panel is an organic composite multi-layer sandwich panel into which spiral antenna elements are inserted. This concept originates mechanically from a composite sandwich structure, and electrically from a spiral antenna, as shown in Figure 1. The sandwich structure consists of two thin load-bearing facesheets, bonded to either side of a moderately thick and lightweight core that prevents the face sheets from buckling. The sandwich structure gains its bending rigidity mainly by separating the facesheets, and has very high structural efficiency (ratio of strength or stiffness to weight). The SAS panel consists of several basic layers. Each layer must meet its own combination of structural and electrical design requirements, as well as the manufacturing and assembly requirements. The basic panel layers are: an outer facesheet, antenna element, honeycomb core, and supporter elements. These are shown in Figure 2 in an exploded view, which also specifies the materials chosen in each layer. The layers are bonded by adhesive to form the final assembly. The outer facesheet must carry a significant portion of the in-plane loads, since it contributes to the overall panel buckling resistance, and it also provides low velocity impact and environmental resistance. This outer facesheet must also permit the transmitting and receiving of RF signals. The facesheet material must be low loss and only weakly dielectric in order to minimize signal attenuation and reflection loss. The honeycomb cores transmit shear loads between layers induced by bending loads in the panel, and support the facesheet against compression wrinkling. They also provide impact resistance and increase the overall panel buckling resistance. The thickness of the honeycomb cores contributes significantly to the overall rigidity, and is involved in the balance between panel thinness and structural rigidity. The supporter also carries a significant portion of the in-plane loads together with the outer facesheet, as well as supporting the whole structure. It can be selected without need to consider electrical performances, and therefore has the best mechanical properties of any layer in the CAS construction. Spiral antennas [10, 11]

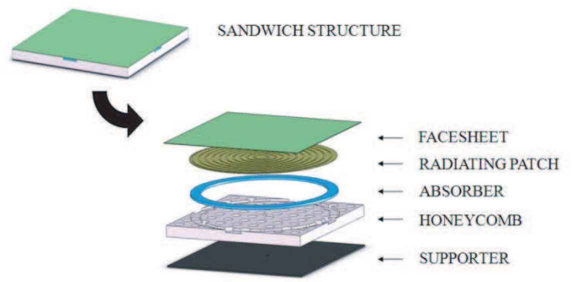


Figure 1: Design concept of the composite antenna structure.

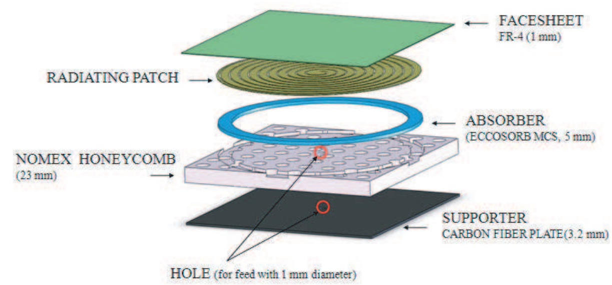


Figure 2: Structure and materials of the composite antenna structure.

can be used in high-performance aircraft, spacecraft, and in satellite and missile applications, where constraints include size, weight, cost, performance, ease of installation, and aerodynamic profile. These antennas are low-profile, conformable to planar and nonplanar surfaces, simple and inexpensive to manufacture using modern printed-circuit technology, and compatible with MMIC designs. Our CAS design is based on a spiral antenna type with a bottom layer of carbon fiber plate. The antenna does not work well if the the spiral antenna current is interrupted. To overcome these problems, we used an electromagnetic wave absorber on a hole in honeycomb core. The spiral antenna is placed on the absorber, which absorbs electromagnetic waves. Unwanted radiation cannot pass through the absorber, reducing the effect of the reflector.

3. DESIGN AND EXPERIMENTAL PROCEDURE

The antenna is to be designed for low frequency and broadband communication. The antenna requirements are: frequency range 0.5 to 2 GHz (bandwidth 1.5 GHz), and gain at least 10 dBi with circular polarization. In designing the antenna elements, a computer-aided design tool (CST Microwave Studio) is used to select a large number of strongly interacting parameters by means of integrated full-wave electromagnetic simulation. The resulting antenna elements and their dimensions are shown in Figure 3. The facesheet is used to a FR-4 glass/epoxy radiating patch comprising a spiral antenna, 1 mm thick. This spiral antenna, set in a circle of diameter 30 cm, is 2 mm thick and has 10.5-turns. FR-4 glass epoxy is a popular and versatile high pressure thermoset plastic laminate grade with good strength to weight ratio. FR-4 undergoes negligible water absorption and is commonly used as an electrical insulator possessing considerable mechanical strength. The main objective is to obtain a good impedance match as seen by the feedline, in the range of frequencies from 0.5 to 2 GHz. The feedline connected to the input port has characteristic impedance 50 Ω, chosen for impedance matching at the port. Coaxial cable is used to feed the antenna in the center. Manufacture of the CAS is a sequential process. The facesheet, including antenna elements, are first prepared by a photolithographic process. The honeycomb cores and each layer must be aligned prior to permanent bonding, in order to give precise electromagnetic coupling. For alignment, four guide holes are made near the edge of all layers. These are confirmed to have no effect on antenna performance. The CAS is assembled by aligning these holes using a plastic nut and bolt. Each layer is bonded to the top and bottom of its neighbors in the designed sequence, using epoxy film adhesive. The assembly, covered by a vacuum bag, is then cured in an autoclave according to the recommended curing cycle for this adhesive (125°C for 90 minutes at a pressure of 3 kg/cm^2). Figure 4 shows the appearance of each layer and the top view of the final assembly after fabrication. The size of the CAS is $300 \times 300 \times 27$ mm. Antenna performance of the fabricated CAS is determined by electrical measurements. The return loss characteristic, which measures the mismatch or the ratio of the reflected power to the incident power at the input port, is measured using a Network Analyzer 8510 under laboratory conditions. The radiation patterns are measured in an anechoic chamber at four frequencies, 0.5, 1, 1.5 and 2 GHz, in order to show the bandwidth patterns. Gains and axial ratios are calculated by comparing the magnitude of the electric field against a standard-gain horn antenna.

4. RESULTS AND DISCUSSION

We have studied an antenna embedded in a structural surface, which provides good structural and good electrical efficiencies at the same time. The design is a composite sandwich structure

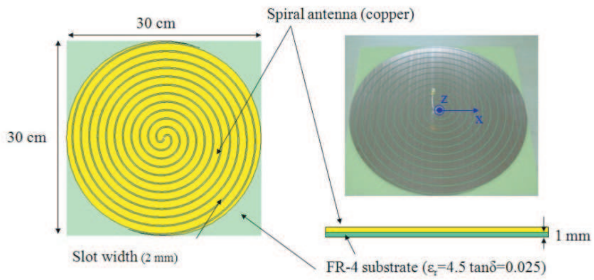


Figure 3: Design of the spiral antenna element.

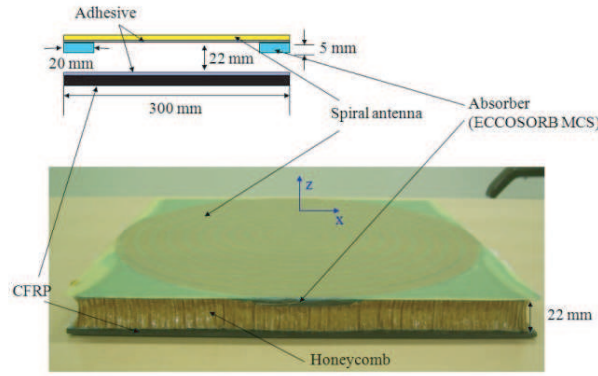
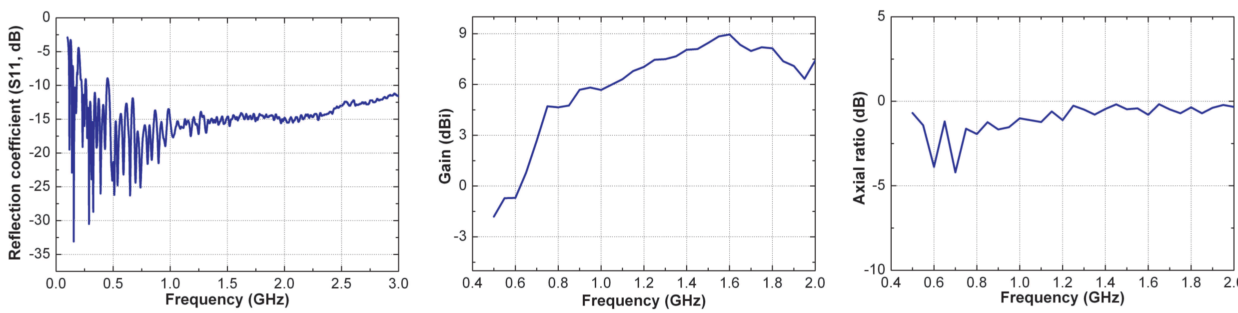


Figure 4: Structure and fabrication of the composite antenna structure.



(a) Reflection coefficient (b) Gain (c) Axial ratio

Figure 5: Electrical performance of composite antenna structure.

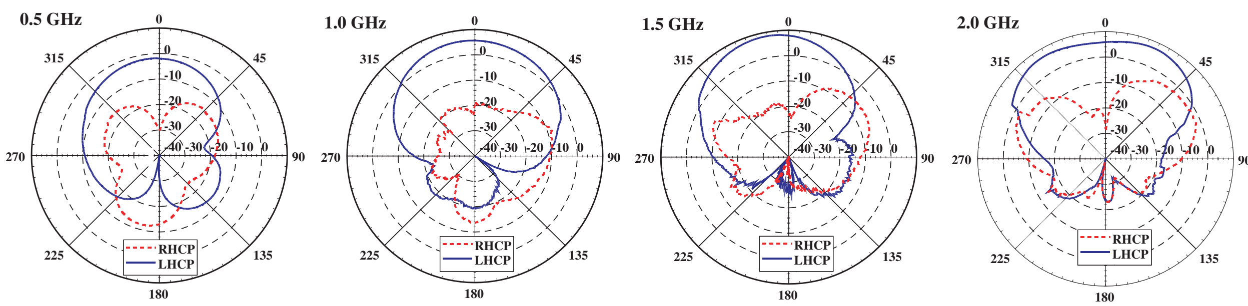


Figure 6: Radiation patterns in bandwidth.

in which a spiral antenna element has been inserted. This design provides antenna performances that meet our requirements. Figure 5 shows the electrical performance of the composite antenna structure. Figure 5(a) shows the return loss characteristic; a bandwidth of approximately 1.5 GHz is seen, corresponding to the frequency range of interest (0.5–2 GHz). Figure 5(b) shows the antenna gain and the gain reduction occurs at low frequencies and is caused energy loss due to absorption. Figure 5(c) shows the axial ratio, and measuring less than 3 dB within the band and get the value of circular polarization is well formed. Figure 6 shows the radiation patterns at 0.5, 1, 1.5 and 2 GHz. The radiation pattern indicates that emissions from the front of the spiral antenna is LHCP (Left Hand Circularly Polarized), RHCP (Right Hand Circularly Polarized) include coming out the back. The back radiation is low because of the carbon fiber plate, and front radiation is large with a LHCP characteristic.

5. CONCLUSIONS

In this paper, we have designed and fabricated a spiral antenna of honeycomb sandwich construction. The final demonstration article is a $300 \times 300 \times$ 27 mm flat antenna panel with an antenna element. Electrical measurements of the fabricated structure show that is satisfies the design requirements, with a bandwidth above 1.5 GHz and a high gain with circular polarization. The design concept can be extended to give a useful guide for manufacturers of structural body panels as well as antenna designers, and promises to be an innovative future communication technology.

ACKNOWLEDGMENT

This research was financially supported by the Ministry of Education, Science Technology (MEST) and Korea Institute for Advancement of Technology(KIAT) through the Human Resource Training Project for Regional Innovation.

REFERENCES

1. Lockyer, A. J., K. H. Alt, D. P. Coughlin, M. D. Durham, and J. N. Kudva, "Design and development of a conformal load bearing smart-skin antenna: Overview of the AFRL smart skin structures technology demonstration (S3TD)," *Proc. SPIE*, Vol. 3674, 410–424, 1999.
2. Lockyer, A. J., J. N. Kudva, D. Kane, B. P. Hill, and C. A. Martin, "Qualitative assessment of smart skins and avionics/structures integration," *Proc. SPIE*, Vol. 2189, 172–183, 1994.
3. Lockyer, A. J., K. H. Alt, J. N. Kudva, R. W. Kinslow, and A. C. Goetz, "Conformal load-bearing antenna structures (CLAS): Initiative for multiple military and commercial applications," *Proc. SPIE*, Vol. 3046, 182–196, 1997.
4. You, C. S., W. Hwang, H. C. Park, R. M. Lee, and W. S. Park, "Microstrip antenna for SAR application with composite sandwich construction: Surface-antenna-structure demonstration," *Journal of Compos. Materials*, Vol. 37, No. 4, 351–364, 2003.
5. Jeon, J. H., W. Hwang, H. C. Park, and W. S. Park, "Buckling characteristics of smart skin structures," *Composite Structures*, Vol. 63, No. 3–4, 427–437, 2004.
6. Kim, D. H., W. Hwang, H. C. Park, and W. S. Park, "Fatigue characteristics of a surface antenna structure designed for satellite communication," *Journal of Reinforced Plastics and Composites*, Vol. 24, No. 1, 35–51, 2005.
7. You, C. S., W. Hwang, and S. Y. Eom, "Design and fabrication of composite smart structures for communication, using structural resonance of radiated field," *Smart Materials and Structures*, Vol. 14, 441–448, 2005.
8. You, C. S. and W. Hwang, "Design of load-bearing antenna structures by embedding technology of microstrip antenna in composite sandwich structure," *Composite Structures*, Vol. 71, 378–382, 2005.
9. You, C. S. and W. Hwang, "Design and fabrication of composite smart structures with high electric and mechanical performances for future mobile communication," *Mechanics of Composite Materials*, Vol. 40, No. 3, 237–246, 2004.
10. Kramer, B. A., S. Koulouridis, C.-C. Chen, and J. L. Volakis, "A novel reflective surface for an UHF spiral antenna," *IEEE Antennas and Wireless Propagation Letters*, Vol. 5, 32–33, 2006.
11. Hertel, T. W. and G. S. Smith "Analysis and design of two-arm conical spiral antennas," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 44, No. 1, 25–37, 2002.

Impact Behavior of Composite-Surface-Antenna Having Dual Band

J. Kim[1], **D. Kim**[1], **D. Shin**[2], **W.-S. Park**[2], **and W. Hwang**[1]

[1]Department of Mechanical Engineering, POSTECH, Republic of Korea
[2]Department of Electronic and Electrical Engineering, POSTECH, Republic of Korea

Abstract— Embedding of an antenna within a structural surface is an excellent way to improve structural efficiency and antenna performance. We study the impact characteristics of a Composite-Surface-Antenna (CSA) and the degradation in the performance of the antenna after impact testing. We used an annular ring patch antenna designed to operate in two bands: GPS (1.575 GHz) and DMB (2.62 GHz). The CSA also included composite materials and adhesive film. When 20 J of impact energy was applied to the CSA surface, the maximum contact force was 4 kN, and the CAS proved resistant to impact. The antenna performance, measured by the return loss and radiation pattern, remained excellent after the impact tests.

1. INTRODUCTION

A number of wireless communication systems and broadcasting services have been integrated into modern vehicles for transport. Conventional antennas have protruded aerials, which are structurally weak and liable to increase drag in aerodynamic applications [1]. Researchers have therefore sought to embed antennas in structural surfaces of aircraft in order to improve both structural efficiency and antenna performance [2–4]. The performances of those antennas have verified that suitably chosen composite materials, honeycomb core, and adhesive film simultaneously satisfy the electrical and structural requirements without degradation in performance [5]. A dual band antenna can reduce the installation space, and provides advantages in cost and size. The system nevertheless has problems reconciling structural and electrical performance. The choices of structural material for the face sheets and core are limited to low dielectric systems that transmit RF signals with minimum loss [2]. Also, the adhesive that bonds the different layers and provides mechanical strength reduces antenna performance. The effect of the adhesive film should be considered at the design stage, using an experimental correction [6]. Furthermore, since it acts as the primary load-carrying component in vehicles, composite construction experiences impacts such as tool drops, hail, bird strikes, and runway debris. Prediction of the effect of low-velocity impact damage is difficult [7]. In the present paper we study the impact behavior characteristics and the electrical performance of the CSA, focusing on the effect of composite materials and adhesive film. After determining the electrical performance and verifying the corrections, we measured and analyzed the impact behavior characteristics and the electrical performance of the CSA.

2. STRUCTURE DESIGN

The CSA is a sandwich panel with inserted antenna elements having GPS and DMB bands. It is composed of composite laminates, honeycomb core, dielectric substrates and adhesive films as shown in Fig. 1. In the top facesheet, glass/epoxy laminate is used for the antenna functions of radiating or receiving radio waves. The bottom facesheet is also glass/epoxy laminate. Mechanically, the facesheets carry a significant portion of the in-plane load, and provide low velocity impact resistance and environmental resistance. The core is Nomex honeycomb, to support various loads; it provides an air-gap to improve the performance of antenna. The first dielectric substrate between the top

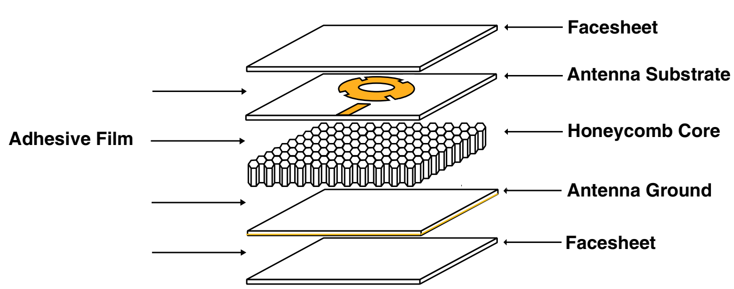


Figure 1: Composite surface antenna structure.

facesheet and honeycomb is FR4 plate, from which a dual-mode annular ring antenna of GPS and DMB is designed. The second dielectric substrate, between the honeycomb and bottom facesheet, is FR4 plate acting as the antenna ground. Adhesive films are inserted between the layers to bond them together and provide mechanical strength. Table 1 sets out the physical properties of the materials used in the CSA structure.

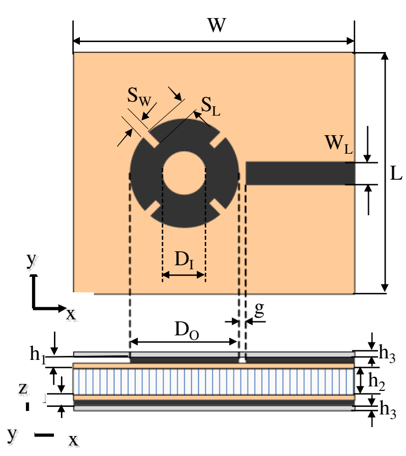
The CSA design is based on two different resonant frequencies of 1.575 GHz (GPS) and 2.6 GHz (DMB). An antenna design tool (CST MWS 2006b) was used for simulation, and the effects of various parameters were investigated. The proposed antenna comprises a coupling feed line which has impedance 50 Ω and four slots on the annular ring patch. To excite the two resonant frequencies, a gap is used between the feed line and the annular ring patch for input impedance matching. The GPS resonant frequency is attained by adjusting the inner and outer diameters of the annular ring. Four slots were made in the patch, to allow matching to the DMB resonant frequency [8]. A frequency shift is observed after deploying the glass/epoxy and adhesive films in the antenna. Compensation for this shift is therefore designed in. The resulting design of the antenna elements is shown in Fig. 2.

3. EXPERIMENTAL PROCEDURE

Impact testing was carried out using a drop weight impact test machine (Dynatup drop weight impact test machine). A high speed camera was used to capture the antenna behavior immediately after impact, as shown in Fig. 3. The impactor has a hemispherical tip of 12.7 mm diameter and the

Table 1: Properties of materials used in the composite surface antenna structure.

Materials	Properties
Composite Laminates (Glass/Epoxy$[0/90]_{2S}$)	Elastic Modulus (GPa) 25.4 Tensile Strength (MPa) 573.6 Dielectric Constant 4.0 Loss Tangent 0.03
Honeycomb Core (Nomex Honeycomb)	Compressive Modulus (MPa) 414 Compressive Strength (MPa) 7.76 Shear Strength (MPa) 88.6 Dielectric Constant 1.1 Loss Tangent ≈ 0
Dielectric Substrate (FR4 plate)	Dielectric Constant 4.5 Loss Tangent 0.002
Adhesive Film Dielectric Constant 2.87 **(Epoxy Polymer)**	Loss Tangent 0.028



Parameters	Dimensions (mm)
Antenna width (W)	140
Antenna Length (L)	140
Antenna substrate (h_1)	0.4
Honeycomb core (h_2)	2.61
Glass/Epoxy (h_3)	1.0
Gap between annular ring and feed line (g)	0.5
Outer diameter of annular ring (D_O)	86
Inner diameter of annular ring (D_I)	26
Slot width (S_W)	2
Slot length (S_L)	11
Feed line width (F_L)	13

Figure 2: Dimensions of dual band antenna.

impact energy is 20 J. Impacts took place at various critical points on the antenna surface as shown in Fig. 4. The test specimens were of two types: SNF (Specimen of No composite Facesheets) which does not have glass/epoxy laminate facesheets, and SCF (Specimen of Composite Facesheets) which has facesheets on the outer side.

4. RESULTS AND DISCUSSIONS

Upon impact, some of the energy was transformed into elastic energy and the rest was absorbed by the specimen, causing structural damage [1]. Fig. 5 shows the relation between time, load, and energy history for our drop-weight test apparatus with impact energy 20 J. The elastic energy is 8 J, and the energy absorbed is 12 J. The maximum contact force in the case of SCF is three times greater than for SNF. The impacter bounced elastically off the SCF surface, but penetrated the SNF surface. The SCF surface shows damage in the shape of a cross.

The return loss was measured using a Hewlett Packard HP8510 network analyzer. Fig. 6 shows the input return loss of the composite surface antenna. The feed network is a critical point in the antenna, because the input impedance is mismatched. For SCF the impedance was mismatched, but it operates in both the GPS and DMB bands because of its 6 dBi value. The compact range

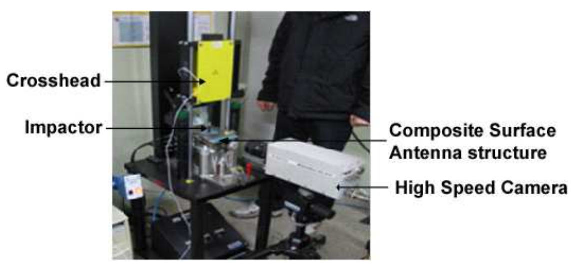


Figure 3: Set-up for impact test.

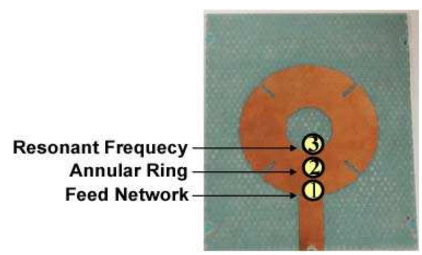
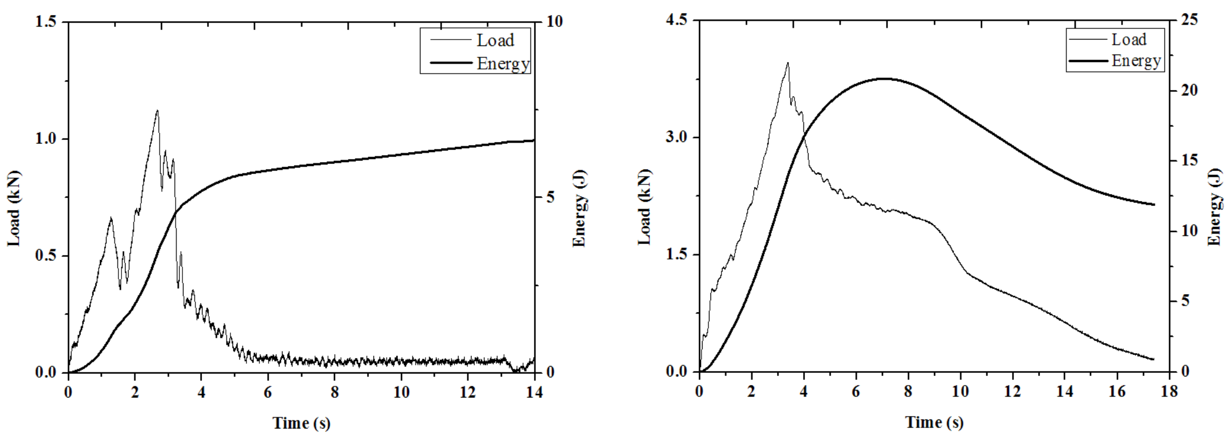
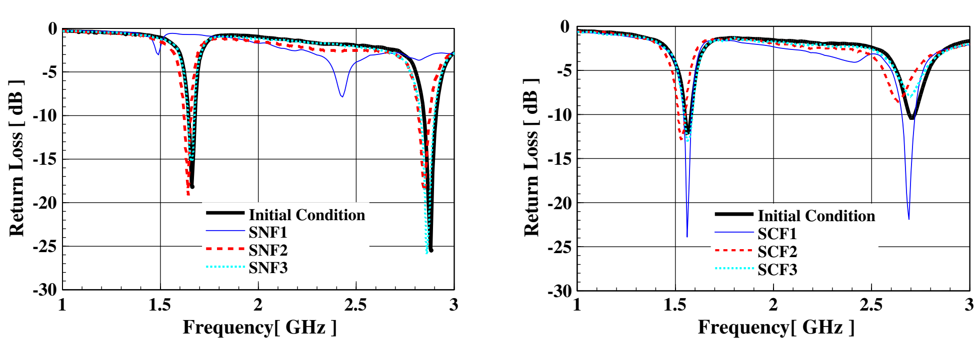


Figure 4: Impact points on the antenna surface.



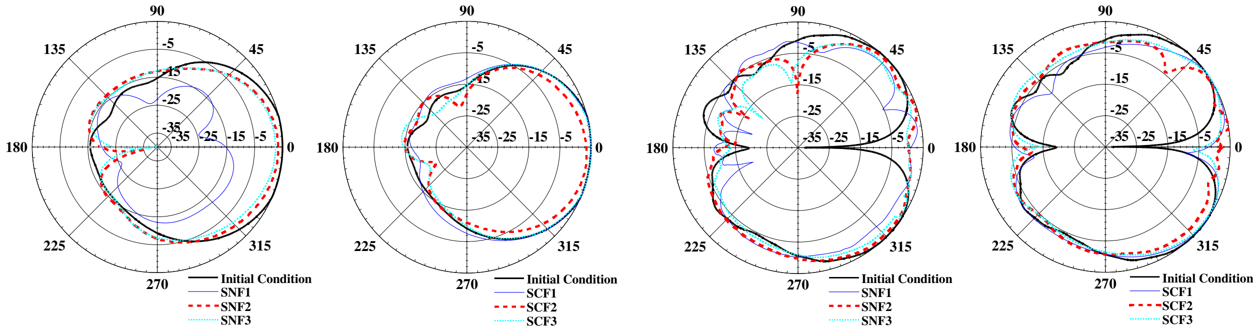
(a) (b)

Figure 5: Impact energy and impact load history. (a) Test results for SNF. (b) Test results for SCF.



(a) (b)

Figure 6: Return loss of composite surface antenna. (a) Return loss of SNF. (b) Return loss of SCF.



(a) (b)

Figure 7: Radiation pattern of composite surface antenna. (a) Radiation patterns at GPS. (b) Radiation pattern at DMB.

provides an efficient means of obtaining the antenna patterns (which normally involve the far field) at a relatively short distance. Fig. 7 shows the radiation patterns for DMB and GPS frequencies, measured after the impact tests. A different current distribution is excited in the CSA upon comparing the distribution for dual band before and after impact. The radiation patterns for SNF clearly show the effect of the dent; the gain is decreased below 3 dBi after the impact. For SCF there is little change and the maximum gain is constant. The figures show that, despite the local damage to the SCF surface, the antenna performs well.

5. CONCLUSIONS

The CSA is a composite sandwich structure, integrated with an annular ring patch antenna designed for dual band performance. Glass/epoxy and adhesive film having structural function can be used as the antenna materials without reducing the antenna efficiency. The CSA design is based on two different resonant frequencies, 1.575 GHz (GPS) and 2.6 GHz (DMB). After impact tests, the antenna continued to perform well both mechanically and electrically. The CSA concept facilitates the design of antennas with structurally effective materials for body panel communication to and from mobile vehicles.

ACKNOWLEDGMENT

This research was financially supported by the Ministry of Education, Science Technology (MEST) and Korea Institute for Advancement of Technology(KIAT) through the Human Resource Training Project for Regional Innovation.

REFERENCES

1. Kim, C., L. Lee, W. Hwang, and W. S. Park, "Impact damage and antenna performance of conformal load-bearing antenna structures," *Smart Materials and Structures*, Vol. 12, 672–679, 2003.
2. Lockyer, A. J., K. H. Alt, D. P. Coughlin, M. D. Durham, and J. N. Kudva, "Design and development of a conformal load-bearing smart-skin antenna: Overview of the AFRL Smart Skin Structures Technology Demonatration (S3TD)," *Proceedings of SPIE*, Vol. 3674, 410–424, 1999.
3. Lockyer, A. J., J. N. Kudva, D. Kane, B. P. Hill, and C. A. Martin, "Qualitative assessment of smart skins and avionics/structures integration," *Proceedings of SPIE*, Vol. 2189, 172–183, 1994.
4. Tuss, J., A. J. Lockyer, K. Alt, F. Uldirich, R. Kinslow, J. Kudva, and A. Goetz, "Conformal loadbearing antenna structure," *37th AIAA Structural Dynamics and Materials Conference*, Vol. 1415, 836–843, 1999.
5. Lockyer, A. J., K. H. Alt, J. N. Kudva, and R. W. Kinslow, "Conformal load-bearing antenna structures (CLAS): Initiative for multiple military and commercial applications," *Proceedings of SPIE*, Vol. 3046, 182–196, 1997.
6. Kim, D., C. S. You, and W. Hwang, "Effect of adhesive bonds on electrical performance in multi-layer composite antenna," *Composite Structures*, Vol. 90, 413–417, 2009.

7. You, C. S., D. Kim, S. H. Cho, and W. Hwang, "Impact behavior of composite antenna array that is conformed around cylindrical bodies," *Composites Science and Technology*, Vol. 70, 627–632, 2010.
8. Kim, J. H., D. W. Woo, and W. S. Park, "Design of a dual-mode annular ring antenna with a coupling feed," *Microwave and Optical Technology Letters*, Vol. 51, 3011–3013, 2009.

From Piezoelectric Actuator to Piezomotor

Y. Bernard, J.-L. Christen, C. Hernandez, and A. Razek
Laboratoire de Génie Électrique de Paris (LGEP)
Supelec, Université Paris VI, Université Paris XI, CNRS UMR 8507, France

Abstract— This paper discusses the ways to achieve a macroscopic motion from the very small strain generated in basic piezoelectric actuators. The first part reviews the operating principle and force-displacement characteristics of some types of piezoelectric actuators, bulk-type, multilayer, benders, flextensional and Langevin transducers. The second part focuses on the different ways to convert this small displacement into macroscopic sliding motion, usually through friction forces. Regardless of the actual use of the active piezoelectric elements, it is shown that the characteristics of each motor can be derived from the trajectory of the contact point between the active stator and the slider, either with a reversible or hysteretic motion. Finally, the third part shows two recent prototypes of piezomotors that have been built and tested in our lab: an Inchworm and a linear-traveling-wave motor.

1. INTRODUCTION

Piezoelectricity is a well known phenomenon that allows a direct generation of motion when an electrical field is applied with a rather high power density. However, the resulting strain has an order of magnitude of some micrometers for centimeter-sized piezoelectric elements. This particularity creates the need for complementary mechanical systems that allow obtaining suitable strokes for general applications.

Piezoactuators are devices that amplify the motion of active piezoelectric elements to the range of millimeters. They are the brick elements of piezomotors which are devices that transmit motion with potentially unlimited stroke.

A review of the main types of commercially available piezoelectric motors is presented in [1]. Several surveys over specific types of piezomotors are available in [2, 3].

Piezoelectric motors show characteristics such as high actuation forces, low speeds, which often render the use of gear unnecessary and high efficiency at a small scale. These features make of them an interesting solution to many applications in which the available space is limited. Additionally, the presence of friction forces in the transmission of motion often ensures a blocked position at rest. As for the small displacement amplitudes of some kinds of motors, it may lead to a very accurate motion capability. Application fields include watches, camera zoom actuators, valves for precise injection systems, and micropositionning devices.

2. PIEZOELECTRIC ACTUATORS

Piezoelectricity can be described as a linear phenomenon linking mechanical properties such as stress (T) and strain (S) to electric properties such as electric field (E) and electrical displacement (D) in some particular classes of materials, the most widely used being lead-zirconate-titanate (PZT) ceramics. When first poled, PZT ceramics are subject to longitudinal, transverse and shear efforts depending on the direction of the applied electrical field. Considering strain and electrical displacement as independent variables, the equations of piezoelectricity read:

$$\begin{cases} S = s^E T + d^t E \\ D = dT + \varepsilon^T E \end{cases}$$

where d is a $3 * 6$-matrix with only 5 non-zero coefficient when computed in the frame where the third direction is that of poling. The coefficient d_{33} then accounts for longitudinal displacement, $d_{31} = d_{32}$ for transverse displacement and $d_{15} = d_{24}$ for shear strain. It is to be noted that longitudinal and transverse displacements always occur simultaneously when the electrical field is applied along the poling direction, whereas shear strain appears alone when the field is applied orthogonally. The order of magnitude of these coefficients is $10^{-10}\,\mathrm{m \cdot V^{-1}}$.

The reverse piezoelectric effect is used in actuators that provide a limited displacement with potentially high forces. Piezoelectric ceramics can be used directly when they are sandwiched between two electrodes, as bulk-type piezoactuators. In order to achieve displacements of some tens of nanometers, the voltage applied to the piezoelement must be high (hundreds to thousands of Volts). This voltage has an upper limit imposed by the disruptive field of the material and a lower negative limit at which the natural polarization of the ceramic is reversed. Due to these constraints, several techniques are often employed to amplify the movement and reduce the applied voltage.

Most piezoelectric actuators can be characterized by their free stroke and blocking force at a given voltage. In this case, a linear behavior between the maximum values of these parameters is observed.

2.1. Multilayers

The multilayer actuators are composed of stacked layers of piezoelectric ceramic of some tens of micrometers sandwiched between very thin electrodes, mounted electrically in parallel and mechanically in series. Compared to the bulk-type actuator, this configuration allows an increment in the electric field applied to a single layer resulting in higher strains for lower applied voltages. Typical strokes of these actuators are in the range of the tens of micrometers. Blocking forces can reach the 10 kN range. Due to layers bonding characteristics, these devices are much weaker when they are subject to tensional forces than bulk piezoelectric ceramics. Therefore, they cannot be used at resonance. In addition, the capacitance of a multilayer actuator is rather high (up to tens of μF) leading to important currents in dynamic operation.

2.2. Benders

The benders use piezoelectric transverse (d_{31}) mode. The extension or contraction of the ceramic occurs along one of the surfaces of a beam, causing a bending motion. Benders can be manufactured in different ways: unimorphs (a single piezoelement glued to a passive beam), parallel bimorphs (two piezoelements poled in the same direction, glued on each side of a passive grounded electrode, and powered with opposite voltages) and serial bimorphs (two piezoelements poled in opposite directions, glued to each other). The latter bimorphs present a smaller stiffness due to the absence of a passive beam. Consequently, they can provide higher strokes. The piezoelements can also be made of multilayer piezoelectric ceramics in order to limit the input voltage. Depending on the stiffness and the length of the beam, the achieved free stroke of this kind of actuators can reach several millimeters, but with very low actuation forces (less than 2 N).

2.3. Flextensional

The flextensional actuators work with the piezoelectric multilayer stacks clamped in a passive elastic structure, whose deformation amplifies that of the ceramic in longitudinal mode. There are several designs of flextensional structures, most of which are elliptic or rhomboidal. The amplification factor in terms of displacement ranges between 2 and 6 according to their axis ratio. This allows higher strains in the direction of motion in detriment of mass and size in the orthogonal directions. Depending on the dimension of the flexible structure and the stack, the developed force can reach the kN range.

2.4. Langevin Transducers

The Langevin transducers [4] are made of one piece of massive piezoelectric ceramic sandwiched between two masses whose shape and acoustic impedance are chosen to amplify the motion at the resonance. Due to the piezoelectric ceramic's weak resistance to tensional forces a prestress is applied by bolting all three parts together. This kind of actuator is optimized for amplification purposes at a specific resonant frequency and can only be used in longitudinal (d_{33}) mode. As they work at resonance, the blocked force and free displacement cannot be defined for the Langevin transducers, which are instead characterized by their operating frequency (in the 10–100 kHz range) and output power, that can reach the 100 W range for use in acoustic devices.

All the actuators presented in this part show interesting characteristics in terms of actuation force, but suffer from very limited strokes when it comes to macroscopic applications. They are either used alone (e.g., in adaptive optics), but more often assembled in more complex structures intended to provide high displacements.

3. DIFFERENT STRATEGIES TO ACHIEVE A SLIDING MOTION

Converting microscopic displacements into macroscopic ones can be done by transmitting periodically the movement generated in a stator to a slider through friction forces. The different types of piezomotors can be classified according to the movement of the contact point between the slider and the stator, which can be hysteretic or reversible. The latter category can be subdivided in permanent or discrete motion. The "contact point" is defined here as one particle in the stator that has had contact with the slider and has transmitted its movement to it, at a given instant in the past.

For non-hysteretic movements of the contact points, there must be a dissymmetry between backward and forward phases of the movement, in order not to cancel the transmitted motion.

Motors with permanent contact points rely on the stick-slip effect, which is based on the difference between static (μ_s) and dynamic (μ_d) friction coefficients. The principle is shown on Figure 1. The piezoelement is driven slowly during the forward phase and rapidly during the backward phase. The forward force applied to the slider is given by:

$$F_F = \mu_s N$$

where N is the normal preload. The backward force is given by:

$$F_B = \mu_d N.$$

Due to the difference between μ_d and μ_s the movement can be achieved either in the backwards or in the forward direction.

Non-permanent-contact-point motors such as the impact-drive motor or the ultrasonic standing-wave motor use a geometric dissymmetry to achieve motion. The operating principle of this motor is shown on Figure 2. On the forward phase, the slider is driven by the stator with an angle α that is inside the friction cone of the slider-stator interface, so the stator sticks to the slider while imparting motion. On the backward phase, the stator tip drives the slider with an angle of $\frac{\pi}{2} - \alpha$, which is outside the friction cone, so that the stator slips and similarly transmits a lower force in the backward direction. The main inconvenient of that kind of motors is the fact that it can only be driven one way.

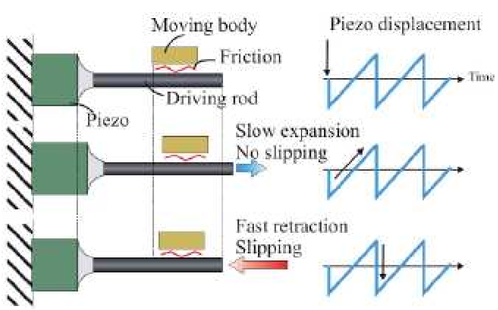


Figure 1: Stick-slip motor ([6]).

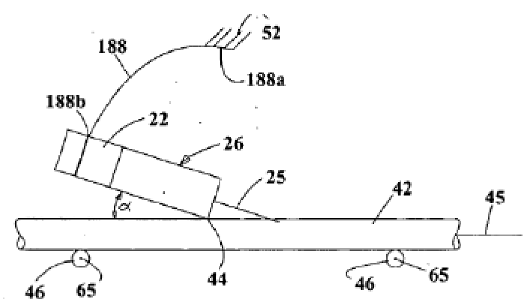


Figure 2: Impact-drive motor (US Patent 20050127789).

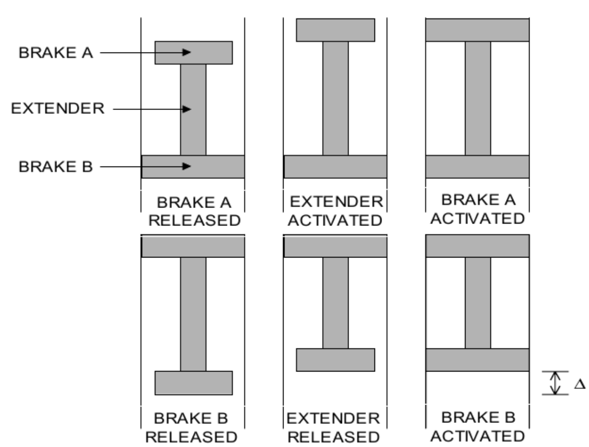


Figure 3: Return loss of a circular corrugated horn.

Finally, motors with hysteretic contact-point trajectory use this phenomenon to avoid exerting a force during the backward phase. One example of such a structure is the Inchworm-type motors, which use two piezoelements to clamp alternatively the slider and one between them to extend and contract. This results in a rectangular cycle during which one step is done. The process is explained on Figure 3. This motor works in a quasi-static regime, but many other motors use resonance to achieve an elliptic motion of the contact point, for example in the famous traveling-wave ultrasonic motor.

4. EXAMPLES OF REALISATIONS AT LGEP

The LGEP has designed two piezomotors for different applications.

4.1. Inchworm

This work concerns the study of tissue regeneration that takes place between the existing gap of a cut bone while it is progressively separated. The actuation principle used to move one part of the bone with respect to the other the displacement has to be discontinuous performing small steps at a time in order to maintain a daily fixed gap between the two bone parts. The forces required are very weak because they only stretch the tendons that are elongated in a short distance. An inchworm (Figure 4) has been designed and tested using multilayer actuators [6].

4.2. Travelling Wave Motor

This work was aimed to perform a micropump. Here, Langevin transducers have been used to create a travelling wave on a beam, as shown in Figure 5. The wave's energy has the ability to move a slider. Different strategies to create the travelling wave have been tested [7].

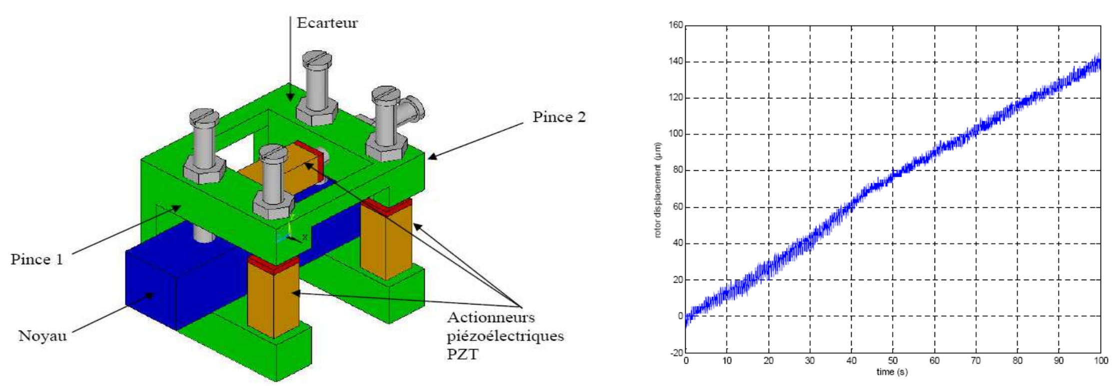


Figure 4: Designed structure and test (displacement versus time).

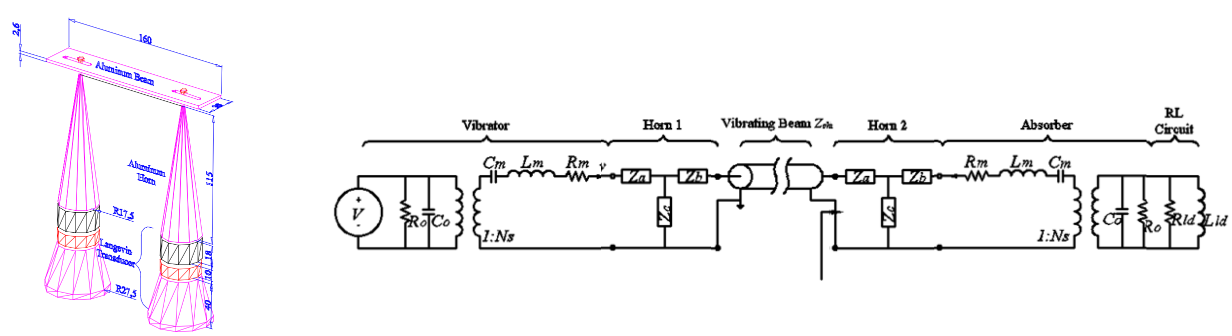


Figure 5: Designed structure and modeling.

5. CONCLUSION

In this article, the evolution of the basic piezoelectric ceramic and its extraordinary ability to generate microscopic displacements into a more elaborated structure capable of producing large or even unlimited strokes is presented.

As first stage element in the evolution there is the piezoelectric actuator. Here, different geometries have been reviewed (multilayer, benders, flextensional and Langevin). According to their principle, they favour certain aspects of the piezoelectric actuation like the force or displacement. For example, when high forces are required, it is recommended to use piezostacks whereas for important strokes benders can be the solution.

The linking between piezoelectric actuators and piezoelectric motors which compose the last element of the evolutionary chain has been effectuated by studying the motion of the contact point between the motion-generating element and the moved one. In this manner, it has been possible to classify piezomotors. The classification unveiled in this paper presents the advantage of embracing all sorts of piezoelectric motors. In addition, with the proposed categorization it is possible to establish easily differences and relationships between motors presenting very different operating principles.

REFERENCES

1. Spanner, K. and B. Koc, "An overview of piezoelectric motors," *Proceedings Actuator 10*, 167–176, Bremen, Germany, June 14–16, 2010.
2. Uchino, K., "Acoustic transduction — Materials and devices," Annual Report Vol. 4, Office of Naval Research, 1999.
3. Morita, T., "Miniature piezoelectric motors," *Sensors and Actuators A*, Vol. 103, 291–300, 2003.
4. Sashida, T., *An introduction to Ultrasonic Motors*, Oxford Science Publication, Oxford, 1993.
5. Rasmussen, T., O. Hansen, E. V. Thomsen, P. Sommer-Larsen, K. Uchino, S. Ural, and S. Tuncdemir, "Easy manufacturing of compact SIDM based motor," *Proceedings Actuator 08*, 644–647, Bremen, Germany, June 9–11, 2008.
6. Sunyoto, S., Y. Bernard, and A. Razek, "Design and realization of a linear piezoelectric actuator for orthopaedic applications," *Journal of Advanced Science*, Vol. 18, Nos. 1–2, 162–165, 2006.
7. Hernandez, C., Y. Bernard, and A. Razek, "Theoretical and experimental analysis of a two modes excitation linear motor using piezoelectric actuators," *Proceedings Actuator 10*, 652–656, Bremen, Germany, June 14–16, 2010.

A Comparison on the Radioelectric Propagation along Grasslands and Scrublands at Wireless Frequency Bands

J. A. Gay-Fernández, I. Cuiñas, and P. Gómez
University of Vigo, Spain

Abstract— There is a deficit of propagation studies at wireless frequency bands in a peer to peer configuration at rural environments, which could be really interesting when designing wireless sensor networks. This paper tries to improve the knowledge of the propagation in these ambiances, in order to optimize the distances between devices in an actual network deployment.

A large amount of received power samples were gathered in scrublands and grasslands, in order to obtain the general propagation equations for these scenarios.

1. INTRODUCTION

The use of wireless sensor networks is nowadays in an exponential growing. Although the initial applications were home automation and industrial control [1], or medical applications [2], the use of these wireless sensor/actuator networks in open rural areas is a new innovative application that was not considered at the beginnings. Technical agriculture and ecological farming are now new scenarios where sensor networks could constitute useful tools.

The research results provided by this study try to exploit the use of these wireless networks in these outdoor environments. Two different environments have been considered in this study: grasslands and scrublands. Within these scenarios, various propagation experiments have been deployed in order to analyse the radio channel characteristics at 2.4, 3.5 and 5.8 GHz, which are the most used frequency bands within wireless networks world.

There are some different studies analysing the propagation across vegetation [3, 4], and even the ITU-R published a Recommendation, 833 [5], which is related to this outdoor environments. However, the new rural applications mentioned above need also new radio propagation models. This is necessary because the peer to peer configuration usually employed in these wireless networks is not considered in this recommendation. ITU-R 833.6 text provides attenuation predictions for master-slave scenarios, as the typical base station to mobile terminal communication, and it is focused on propagation across the canopies of the trees. The environments considered in this research are only composed by grass, with a mean height of 12 cm and by weeds and scrubs with a mean height of 1.8 meters. Due to these differences the ITU-R recommendation does not work as well as expected in such situations. So, new data is needed to develop new propagation models.

The paper has been organised as follows. The second section is intended of the measurement system employed during the campaigns. The third section contains the description of the environments where the measurements were performed. The fourth section is devoted of the results, which are analysed in section fifth. Finally, the conclusions occupy the sixth section.

2. MEASUREMENT CAMPAIGNS

2.1. Measurement Setup

A separate transmitter and receiver configuration has been used during both measurement campaigns. Thus, large distances between transmitter and receiver could be accomplished in order to check how the signal strength attenuation with distance is.

The transmitter equipment consists of a signal generator Rohde-Schwarz SMR and an omnidirectional wide band antenna, Electrometrics EM-6865. This system is capable to generate pure tones at the frequencies under test: 2.4 (+18 dBm), 3.5 (+19 dBm) and 5.8 GHz (+15 dBm).

A portable spectrum analyzer Rohde-Schwarz FSH-6 is used at the receiver system with an omnidirectional antenna, similar to the transmitter end. When the tone under study was captured, the spectrum analyser was changed to time domain configuration, in order to gather time series of the received power.

2.2. Measurement Procedure

The data was collected around two different radials at each environment. Each radius consists of 25 points and 150 meters at the grassland environment, and 16 points and 32 meters at the scrubland one. The number of power samples gathered at grass and scrub lands is 301 and 3010 respectively.

Table 1: Distance (m) from transmitter to receiver at each measurement point.

Environment	**Measurement point**																								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
Grassland	1	2	4	6	8	10	14	18	22	26	30	35	40	45	50	55	60	65	70	80	90	100	115	130	150
Scrubland	1	2	3	4	6	8	10	12	14	16	18	20	23	26	29	32									

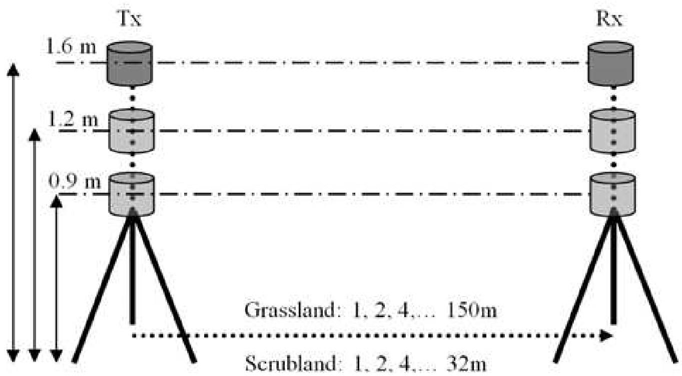


Figure 1: Configuration of the antennas.

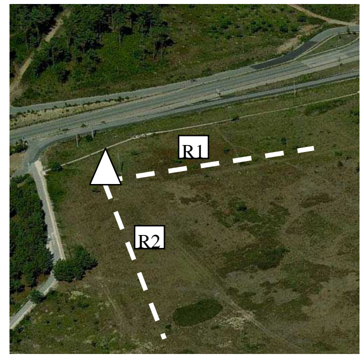


Figure 2: Grassland environment.

Three different heights were analyzed for the transmitting and receiving antennas: 0.9, 1.2 and 1.6 meters. Both antennas were placed at the same height in our analysis, in order to simulate the best conditions for a peer to peer propagation. Figure 1 depicts the configuration of the transmitting and receiving antenna used in these campaigns.

As shown in Figure 2, the receiver was moved along the defined radial. The distance between consecutive points of measurement is higher as the separation between transmitter and receiver is. Table 1 displays the distances in meters of the measurement points in both environments. The maximum distance is limited by the noise level.

3. MEASUREMENT ENVIRONMENTS

Two different environments have been taken into account in this study. The first one is a low height grassland, located near the campus of the University of Vigo. It was composed basically of grass and weeds with a mean height of 12 cm. Figure 2 shows the grassland environment, the transmitter location (triangle) and the two 150 meter long radii (white dashed lines). The second environment under study is a scrubland, located at 10 km far from the University of Vigo. It was composed basically of undergrowth with a mean height of around 180 cm. Thus, both the transmitter and the receiver were always completely shadowed by vegetation.

4. RESULTS

903 power samples per frequency were collected at each one of the 50 points under measure at the grassland environment. The power samples per frequency at each point were 9030 at the scrubland environment, because there was observed a large time-variance of the received power. Thus, the complete campaign represents a total of more than 1 million power samples. The processing of such amount of data, as well as the results, are the contents of this section

4.1. Data Processing

The objective of the data processing is the analysis of the results by means of a regression to know how the power decays with distance. Depending on the selected frequency, the antennas height and the environment, the attenuation of the received power seems to be quite different. Some vectors seem to fit a linear equation of the form $P = P_0 - n \cdot 10 \cdot \log_{10}(d)$, where d is the distance between transmitter and receiver in meters, P_0 is the received power, in dBm, at 1 meter from the transmitter, P is received power, in dBm too, at a distance d from the transmitter and n is a factor that shows the rhythm of the power decay with distance.

However, the behavior at other environments suggests that the best choice is a double linear regression, where each one of the regression sections fits an equation similar to that mentioned above.

4.2. Regression Fitting

When the previously explained regression fitting is applied to the collected samples, data from Table 2 is obtained for grassland and scrubland. These tables show the attenuation factors "n_1" and "n_2", obtained for the first and second regression section respectively; the mean error produced with this estimation; and the cut-off point of the two regressions. Rows with a dash in "n_2" and "Cut-off point" columns indicate that in these cases a single regression seems to fit the data better.

All the power values that are shown in the following figures have been normalized to a transmission power of 0 dBm, in order to easily use with another transmitting power value.

Figures 3 and 4 depict an example of the regressions obtained at 2.4 GHz in grasslands and at 5.8 GHz in scrublands respectively. These images show the double regression fitting that exists in both environments and how the cut-off point increases with increasing antenna height in grasslands, but not in scrublands. In this last scenario, this especial point is quite similar for all heights and frequencies.

5. ANALYSIS OF THE RESULTS

Data from Figure 3 and Table 2 show that attenuation increases with decreasing antenna heights in the grassland scenario. For instance, comparing the ranges where −90 dBm are achieved at 2.4 GHz depending on the antenna height, we obtain 55 m, 70 m and 100 m for 0.9 m, 1.2 m and 1.6 m antenna height respectively. This effect could be caused by the obstruction of the first Fresnel zone, which has a greater radio with increasing the antenna height.

Regarding the scrubland environment, the effect of the double regression seems not to be caused by the obstruction of the first Fresnel Zone, since the path obstruction is always complete due to the high vegetation located in the vicinity of the antennas. With minor differences, all the points where the attenuation slope change, are quite similar regardless the antenna heights and frequency. Even so, the ranges achieving −90 dBm in Figure 4 are between 15 and 25 meters for 0.9 and 1.6

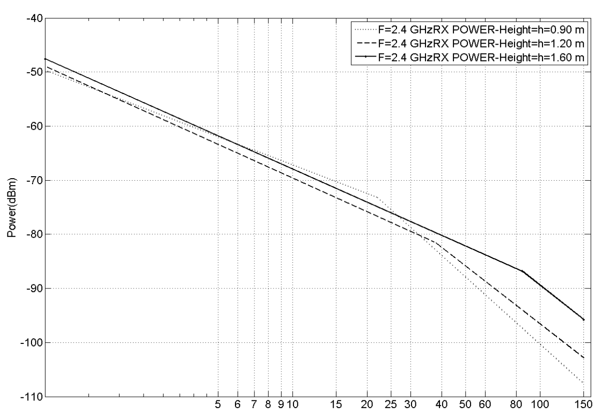


Figure 3: Regression fitting at 2.4 GHz in grasslands.

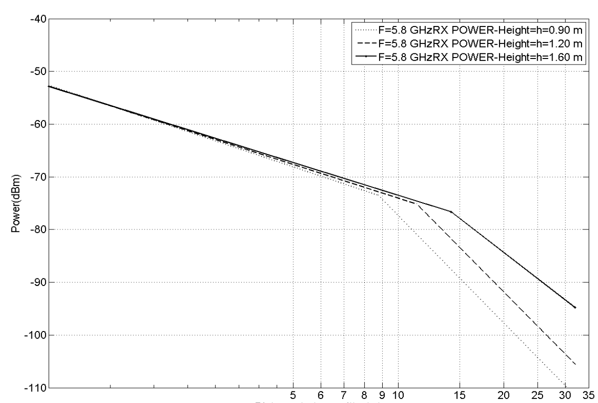


Figure 4: Regression fitting at 5.8 GHz in scrublands.

Table 2: Regression data.

Frequency [GHz]	Height [m]	Grasslands				Scrublands			
		n_1	n_2	Error [dB]	point [m]	n_1	n_2	Error [dB]	point [m]
2.4	**0.90**	1.75	4.13	1.47	22	2.63	4.63	2.61	13
	1.20	2.07	3.55	1.20	37	2.20	5.18	1.23	13
	1.60	2.04	3.61	1.70	85	1.88	5.58	1.60	13
3.5	**0.90**	1.97	3.13	1.41	27	2.19	5.37	2.69	7
	1.20	1.94	2.40	1.15	57	1.92	6.28	1.15	12
	1.60	1.90	-	0.943	-	1.98	5.53	1.03	13
5.8	**0.90**	2.14	2.78	0.871	48	2.22	6.76	1.77	9
	1.20	1.93	2.84	0.885	84	2.13	6.72	2.58	11
	1.60	1.98	-	1.08	-	2.07	5.12	0.768	14

meters high respectively.

6. CONCLUSION

The results of a large measurement campaign at 2.4, 3.5 and 5.8 GHz performed in grasslands and scrublands have been presented. Although the propagation across vegetation appears to be widely studied, the peer to peer applications need new radio propagation models.

This paper shows how the power decays depending on the type of environment, the frequency and the height of the transmitter and receiver antennas. The final results seem to indicate that the attenuation in scrublands is much greater than in grasslands at both frequencies.

ACKNOWLEDGMENT

This work was supported by Xunta de Galicia, Spain, under project InCiTe 08MRU045322PR.

REFERENCES

1. Egan, D., "The emergence of ZigBee in building automation and industrial control," *Computing & Control Engineering Journal*, Vol. 16, No. 2, 14–19, 2005.
2. Timmons, N. F. and W. G. Scanlon, "Analysis of the performance of IEEE 802.15.4 for medical sensor body area networking," *First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, 16–24, Oct. 4–7, 2004.
3. LaGrone, A. and C. Chapman, "Some propagation characteristics of high UHF signals in the immediate vicinity of trees," *IRE Transactions on Antennas and Propagation*, Vol. 9, No. 5, 487–491, 1961.
4. Richter, J., R. F. S. Caldeirinha, M. O. Al-Nuaimi, A. Seville, N. C. Rogers, and N. Savage, "A generic narrowband model for radiowave propagation through vegetation," *Vehicular Technology Conference*, 39–43, 2005.
5. Int. Telecommun. Union (ITU), "Attenuation in vegetation," ITU-R Recomm, 833-6, 2007.

Design and Development of an Electronic Cowbell Based on ZigBee Technology

J. A. Gay-Fernández, I. Cuiñas, and M. G. Sánchez
University of Vigo, Spain

Abstract— Nowadays, the ecologic food products consumption is in an exponential growing. People are realizing that meat of a cow living in an extensive farm is much better and healthier than the meat of a cow fattened up in a static industrial place. So, farmers are concerned to develop better techniques to extensive farming cattle, but maintaining the veterinary control of the animals. The aim of this contribution is the hardware and software design and test of an electronic cowbell to allow farmers to track the cows that live freely in extensive farms. Some veterinary data, like body temperature, or environmental data, like ambient humidity and temperature, could be gathered with the aid of different sensors plugged into the cowbell.

1. INTRODUCTION

Ecologic products are more and more important every new day in our modern life. The beef meat that can be found at the supermarkets came usually from intensive farms, where cows were fattened up in almost static places. Nowadays, things are changing, and many farmers are realizing that allowing their cows to move freely along large areas provides them better final products, with less fatty and better flavor. Due to these benefits, the creation of extensive farms is increasing every day.

One of the main problems in these extensive farms could be the cows tracking control, and how to gather some veterinary parameters, like, for example, the body temperature. This paper tries to propose a solution to this farmer's headache, allowing the cattle to move freely but keeping the cows' health controlled at any moment. The idea is the use of ZigBee [1] networks to provide both features: tracking and data sensing.

Some ZigBee motes of the wireless network could be installed in trees or bushes, at static positions, and other motes would be placed at the cow collars (cowbell). The static motes, usually called reference nodes, provide the location of the dynamic motes (blind nodes), by means of triangulation algorithms, and finally, the position is optimized with the aid of a Kalman filter.

Furthermore, the electronic cowbell would be equipped by body temperature sensors to obtain veterinary data, and some other environmental sensors, to gather, for instance, ambient humidity and temperature.

The static motes would be involved in the connectivity between the cattle and ambient data and the farmers. Thus, these data could be available at the farm, in a laptop or even in a PDA device.

2. ELECTRONIC COWBELL

The electronic cowbell is based on the chip CC2430 from Texas Instruments. This chip allows establishing an IEEE 802.15.4/ZigBee network, at 2.4 GHz ISM band, with very low power consumption. This last feature is very important for a system that is supposed to work for long time without changing batteries. Plugged to this central device, there are two data sensors providing body temperature, humidity and ambient temperature. This data could help veterinarians to keep animals under control while they are living in freedom.

2.1. CC2430 Chip

The ZigBee CC2430 is an IEEE 802.15.4 and ZigBee System-on-Chip solution developed by Chipcon from Texas Instruments. Its 8051 microcontroller core provides two USARTS where several sensors like contact temperature and ambient humidity and temperature can be plugged. Furthermore it has 21 general I/O pins, which can allow us the connection of any other sensor/actuator.

In addition, these wireless devices are capable of measuring digital RSSI, so with an exponential power decay estimation is quite easy to estimate the distances between nodes. Thus, with the aid of some fixed nodes, with a known position, the free animals with this tailor made electronic cowbell on them could be tracked along the grasslands or forests where they usually expend their time. Please refer to [2] for more details.

Figure 1: IEEE 802.15.4 Chipcon CC2430 EM module.

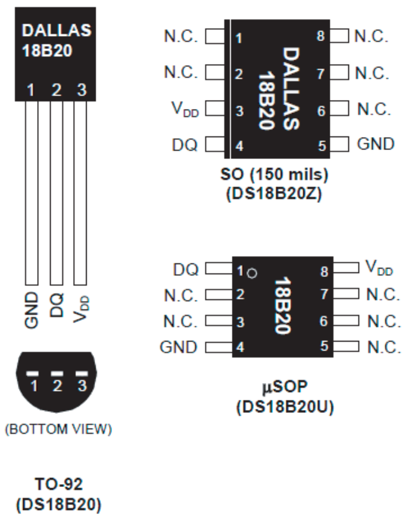


Figure 2: DS18B20 pin configurations.

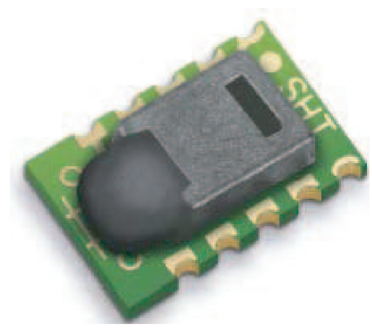


Figure 3: SHT10 humidity and temperature sensor.

2.2. Body Temperature Sensor

The body temperature sensor selected for this device is the DS18B20 from Dallas Semiconductor. This device may measure temperatures from $-55°$C to $+125°$C, which seems to be enough for body temperature measuring. Furthermore, it is able to transmit data only by one wire. From the point of view of hardware, this means that this device is quite simple to plug. However, its software configuration appears to be more difficult than in other devices with multiple transmission wires. For more details please see Figure 2 and [3].

2.3. Environmental Sensors

Two environmental sensors have been plugged into the CC2430 chip: one for humidity measurement and the other for ambient temperature. In fact, both parameters can be obtained with only one integrated sensor, the SHT10 ambient humidity and temperature sensor, from Sensirion. This sensor is able to measure temperatures from $-40°$C to $123.8°$C and the whole range of relative humidity with only $90\,\mu$W of power consumption. Please refer to [4] for more details. Figure 3 depicts this sensor device.

3. TRACKING SYSTEM

3.1. Introduction

The sensor data is only an extra application for the whole system. The main aim of this wireless network is to track the position of the animals, when they wear our tailor made cowbell. Cowbells are usually called Blind Nodes (BN) in these location systems, because they do not know their position.

In order to locate a BN along a rural outdoor environment, we need a grid of Reference Nodes (RN) at known positions. The way to estimate the position of the BN is based on triangulation.

We have to estimate the distance between our BN and at least three RN's in order to get a right position. As much RN's we have, as more precision we will get in the final location measure.

The method to estimate these distances based on the measured RSSI is shown in Section 3.2. Once these distances have been estimated, the location algorithm, based on a standard triangulation algorithm, is run. Finally, these location estimations have been improved with the aid of a Kalman filter, as shown in [5].

3.2. Distance Estimation

As we have said before, the distance estimation is based on the RSSI measured in the Reference Nodes. The BN that wants to be located sends a broadcast message to all RN's in its coverage area. Then, each RN can evaluate the RSSI of the received message and send it back to the BN. Once the BN has all these RSSI values, it can estimate the distance to each RN based on the Equation (1).

$$d = 10^{\frac{RSSI - A_0}{10 \cdot N}} \tag{1}$$

where d is the distance to estimate in meters, $RSSI$ is the measured received signal strength indicator, in dBm, A_0 is the power received one meter from the transmitter, in dBm, and N is the power decay rate (without units). Values for A_0 and N are very different for each environment so they need to be estimated in each case. In our test environment, a classroom of the School of Telecommunication Engineering of Vigo, it has been estimated a value of -32 dBm for A_0 and 2.875 for N. These parameter values have been obtained after developing a small measurement campaign at 2.4 GHz in this classroom.

3.3. Location Algorithm

Once distances between BN and RN's have been estimated, the location algorithm needs to be run. The way to estimate the position is shown below:

The estimated distances between the BN and at least three RN are known. The position of each RN is known, so an equation system with only 2 unknowns and number of RN's equations can be created. The best way to solve this kind of systems appears to be the minimum mean square solution, and it can be easily solved by an algorithm like that described in [6].

When the position of the cowbell (BN) is known, this information, together with sensor humidity and temperatures data, are sent to a ZigBee gateway, in order to be incorporated to a database or to rendering software.

4. RESULTS

Results presented in this section have been recorded in a classroom. A lot of possible paths were analyzed, but only the more representative ones have been included in this chapter. A six RN's network has been deployed in the classroom, and one cowbell has been moving along the corridors between desks.

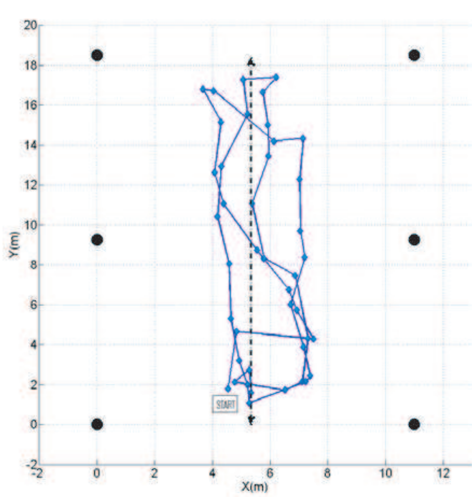


Figure 4: Path 1 with Kalman Filtering.

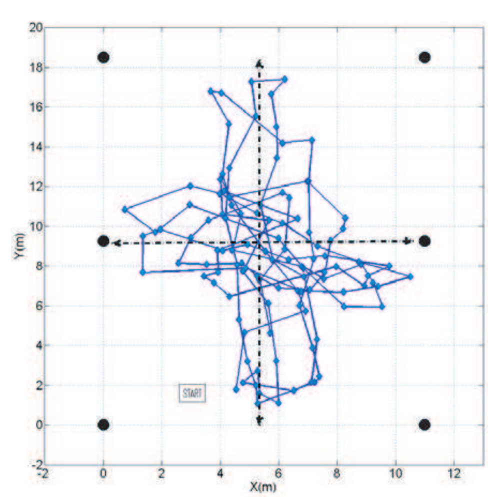


Figure 5: Path 2 with Kalman Filtering.

The first test performed was a linear motion at the coordinate $x = 5.5$ meters, while coordinate Y was changing between 0 and 19 meters. Figure 4 shows the path estimated with the aid of a Kalman Filter. The double black dash-dotted arrow depicts the actual motion of the cowbell. Black circles are the Reference Nodes. Blue lines interconnecting stars provide the predicted motion of the cowbell. The mean estimated error is around 2 meters, which seems to be enough for locating a cow in grasslands or forests.

The second test involves a change in the motion direction. The test consisted on moving the cowbell at fixed $X = 5.5\,\mathrm{m}$ and varying Y between 0 and 19 meters. Approximately after 2 min, we began a new motion direction, varying X between 0 and 11 meters, with Y fixed at 9.5 meters.

5. CONCLUSIONS

A design of a ZigBee cowbell to help farmers in extensive cow exploitations has been presented. This electronic cowbell provides, with the aid of static motes, tracking. In addition, some veterinary constants as body temperature, and some environmental data, as ambient temperature and humidity could be gathered with the aid of this device.

Finally, some test results show that the mean error when locating the animal appears to be acceptable, and the trajectory is tracked with a quite good performance.

ACKNOWLEDGMENT

This work was supported by Xunta de Galicia, Spain, under project InCiTe 08MRU045322PR.

REFERENCES

1. Adams, J. T., "An introduction to IEEE STD 802.15.4," *Aerospace Conference*, Vol. 8, 2006.
2. "CC2430 data sheet (rev. 2.1) SWRS036F," Texas Instruments, 2007.
3. "DS18B20 programmable resolution 1-wire digital thermometer datasheet," Maxim Integrated Products, 2010.
4. "SHT1x (SHT10, SHT11, SHT15) temperature and humidity sensor datasheet," Sensirion, 2010.
5. Nabaee, M., A. Pooyafard, and A. Olfat, "Enhanced object tracking with received signal strength using Kalman filter in sensor networks," *International Symposium on Telecommunications*, 318–323, Aug. 27–28, 2008.
6. Chen, J., X.-J. Wu, F. Ye, P. Song, and J.-W. Liu, "Improved RSSI-based localization algorithm for park lighting control and child location tracking," *International Conference on Information and Automation, ICIA'09*, 1526–1531, Jun. 22–24, 2009.

RFID from Farm to Fork: Traceability along the Complete Food Chain

I. Cuiñas[1], **L. Catarinucci**[2], **and M. Trebar**[3]

[1]Dept. Teoría do Sinal e Comunicacións, Universidade de Vigo, Spain
[2]Dept. of Innovation Engineering, University of Salento, Lecce, Italy
[3]Faculty of Computer and Information Science, University of Ljubljana, Slovenia

Abstract— The project "RFID from Farm to Fork" looks for the extension of RFID technologies along the complete food chain: from the farms where cows, fishes, sheep, grapes, etc. grow; to the final consumer at the supermarkets, including all intermediate stages: transports, factory processes, storage. The paper is intended to show the project objectives and concerns, as well as it highlights the main radio propagation problems detected within a RFID system installed in a food factory. The paper also shows a proposal of using RFID traceability in different study cases.

1. INTRODUCTION

People need to eat! For this reason, the production and distribution of food is the largest and the most important activity in each country all over the world. Because of this extension of food economy, different new proposals appear to improve the quality of the products and the information received by the consumers.

This paper has the aim to present the works developed along the project "RFID from Farm to Fork", a CIP-Pilot action involved within the 7th Frame Work of the European Union. Our proposal looks for the extension of RFID technologies [1] along the complete food chain: from the farms where cows, fishes, sheep, grapes, etc. grow; to the final consumer at the supermarkets, including all intermediate stages: transports, factory processes, storage. The main objective is the use of only one system to perform the complete traceability, recording data at each stage. These data could be useful to determine the perfect condition of the final product, but also to control the process during the elaboration. Thus, both final consumers and producers would take advantage of such systems.

The final consumers could obtain different data above the whole process suffered by the product they are buying, just by moving the object (labeled with a RFID tag) in the surroundings of a RFID reader, which can be installed in the supermarket or even as an application at each personal smart phone. The individual identification of the product allows the software to obtain a complete traceability report from a central database, and to bring the consumer this information. Each of the producers along the chain could use the identification by radio frequency to control his production and storage, and to know some previous information of his ingredient matters. The project involved both the design of the complete system and its tests at different stages of the chain: fishing companies, wine producers, food transporters, and final users, in order to define the actual interest of the system, its performance, and its advantages and disadvantages.

The following sections show the proposal of the project, the radio propagation difficulties in factory ambient, as well as an introduction to some of the pilot tests.

2. THE RFID F2F PROPOSAL

The project will showcase the ability of RFID technologies to make a return on investment for SMEs in the food industry, as well to provide large information to the consumers [2]. The opportunities for such a return on investment arise from the following:

- Opportunities to create markets for premium products (organic, etc.) if technology can address authentication, condition monitoring and quality control.
- New opportunities are created to increase quality, reduce wastage, reduce energy used for refrigeration, reduce chemical usage for preservatives, optimize carbon use, etc.
- Impact on competitiveness and productivity gains.
- Potential for new markets for food producers in the regions.
- Increased productivity through reduced wastage.

- Authentication of origin, process and transport of products.

These advantages have been realized in large concerns, which have control over most or all of the value chain and are in a position to make an end-to-end investment. However, they are not available to independent SMEs, which only participate in one stage of the value chain. By linking RFID and sensor network technologies with a Europe wide database, which can store the exact history of any food product, SMEs will be given the opportunity to optimize their own business process to maximize return. In addition, a pan-union resource will be created which will allow producers to demonstrate unequivocally the quality and freshness of their product, which will have the effect both of increasing consumer confidence, and increasing producer margins.

The emphasis of the current project is somehow advanced form previous projects, in that it proposes:

- Integration of RFID and sensor networks to provide end to end monitoring.
- Demonstration of end-to-end traceability for the whole value chain in a number of different product areas.
- Integration with an international database of food ingredients.

3. RADIO WAVE PROPAGATION IN FOOD FACTORY AMBIENTS

The use of RFID and/or sensor networks in the food chain has been previously tested, and it could obviously provide a competitive advantage to the involved companies. However, the radio propagation within a factory presents important differences compared to other more friendly environments: there are lots of metallic elements: machinery, store installations, transports, and even walls and ceiling, which limits the performance of the radio systems. When the factory belongs to the food sector, the situation becomes more and more complicated, as most of the food products present a large content in water, which is not the best friend of the radio waves. Even more, some of the packages are metallic (cans or metallic sheets), which are also limitative for the good performance of a radio system.

3.1. Fish Factories

The fish factories, both for freezing or for canning processing, resulted to be problematic for radio electric propagation due to the presence of large amounts of water and ice within the tags and the readers. The fresh fish is commonly transported inside plastic boxes containing the fish and a lot of ice around. This ice content makes really difficult to establish a radio link if the RFID tag is glued directly to the plastic: a solution consists of separating the tag away from the box (i.e., from the ice) to improve the propagation conditions.

Some tests have been performed using plastic boxes with frozen fish, or just fish with water, or even a simple wet box. When the box is dry, the tag can be read from several meters, but in wet conditions the coverage distance is reduced to several centimeters. Once the fish is frozen and packed, the installation of the tag is again a problem: if it is too close to the ice, the performance of the radio link could be reduced. And it is not easy to install a tag far from the ice inside a freezer!

When the final product is the canned fish, the metallic surfaces of the cans act as isolators for the radio links. The installation of the individual tags in a separate section of the carton package could be enough to solve the situation in a supermarket. But it is almost impossible to manage a store of cans, by means of individual RFID tags. Thus, a separate tagging (individual can vs. lots of cans in larger batches) would be needed to control the storage. This solution could also present some problems depending on the side of the box the tag is glued: if the tags fall in the middle of a package island, perhaps the RFID reader is unable to detect the tag information.

3.2. Wineries

When dealing with wine factories, there are different limitations to the propagation along the process from the vineyards to the bottled wine. The grapes are commonly carried from the vineyards to the winery in plastic boxes of around 20 kg. Again, the water contents could be a difficulty, although less important than in fishery.

Part of the process of manufacturing the wine is performed in steel barrels. The building where these recipients are placed does not present good propagation conditions, as a lot of reflectors and scatters are present in such indoor environment. In this situation, the performance of the physical layer of a RFID system could be affected, as various fading events could appear due to the multipath phenomena.

Once the wine is bottled, the location of the tags on the bottle surface is also problematic: the liquid content becomes a good conductor, and so the transmission capabilities of the tag are reduced. A good selection of the type of tags, as well as its location in the body of the bottle, is a key factor to achieve valid performances.

4. CASE I: WINE PRODUCTION

The quality and the taste of a wine depend on many factors: type of grape, soil, percentage of precipitation and temperature during grapes growing, eventual use of pesticides, types of harvest (day or night), storage conditions of grape before of wine making, and so on.

Experts wine consumers, therefore, are becoming more discerning and pretend to know as much data as possible in order to enlarge their knowledge, be sure of the authenticity of a wine and spend consciously.

One of the several pilots related to the wine supply chain is a small winery named "Vigne Mastrodomenico" located in Basilicata, a Region of Southern Italy, and producing no more than 10000 bottles per year of "Aglianico del Vulture". The vines grow on a hilly volcanic soil as large as eight ectars with different degrees of sun exposure depending on the vineyard zone. In some cases, hence, it could be necessary a selective forced irrigation. Once the harvest is done, the grapes are processed in a small winery carved into the rock, creating a natural habitat (for instance, a constant temperature of almost 15°C is naturally achieved) for this kind of wine. The presence of the wine and of steel drums, though, complicates the scenario from an EM point of view.

Two aspects have been considered so far. The first one regards the wireless sensor network (WSN) at the vineyard site, useful to collect both meteorological and soil status data to be associated with the single wine bottle and to automatically activate the forced irrigation. The second one, more critical, regards the RFID-related hardware investigation to be used in the winery. As previously stated, in fact, the presence of electromagnetically hostile materials, such as liquids and metals, makes challenging the single bottle traceability in the different steps of the wine supply chain.

For such a purpose, a very complex RFID-based test bed environment has been assembled. The test bed is capable to evaluate the tag performance in realistic situations and in each step of the wine supply chain. Results in terms of successful read rate as well as of RSSI (Received Strength Signal Indication) have been obtained. These results demonstrate that, among the possible RFID frequency bands, UHF has to be preferred to HF for several reasons. For instance, HF tags do not work properly in Far Field conditions, thus implying the impossibility to be used in some critical steps of the supply chain requiring a far field interrogation. Moreover, they are not EPCGlobal compliant, with relevant consequences in terms of multiple readings, generality of the implemented solution and many others. Consequently, two different UHF readers, three different reader antennas, two different Near Field UHF tags and six different Far Field UHF tags have been tested. With each tag attached to a wine bottle, the reading range has been firstly measured both with and without wine inside, so to characterize each tag on the bases of its tolerance to the liquid presence. Then, the problem of the tag performance degradation when the traceability system is implemented at items level in case of presence of liquid has been investigated. Results in terms of successful read rate have been obtained by varying the mutual orientation between tag antenna and reader antenna (i.e., 0° and ±90°). In Figure 1, for instance, the successful read rate of two near field tags (Cube 2 and Paper Clip) and of a far field one (Jumping Jack) are reported. It is quite evident that in critical conditions far field tags perform better than near field ones. Nevertheless, when different supply chain steps based on longer interrogation distances are considered, the effect of the performance degradation is more marked. The individuation of the best tag positioning on the wine bottle is the next step.

5. CASE II: FISH MANUFACTURING

A fish pilot in Slovenia deals with traceability of seabass grown in northern part of Adriatic sea. At farm, a young fish brought from hatchery is breeding for four years in cages under similar conditions as other fish. On the day of catch, fish is transferred to the processing factory, sorted by weight category and packed into styrofoam boxes depending on orders. In some cases fish is also cleaned or filleted. After that, boxes are covered with ice and transported to cold store. The next day, fish is delivered to retailers, restaurants, fish market and even private consumers.

In the supply chain is very important to secure required temperatures of fish between 0°C and 4°C. This can be easily done by using active RFID data logging device with temperature sensor.

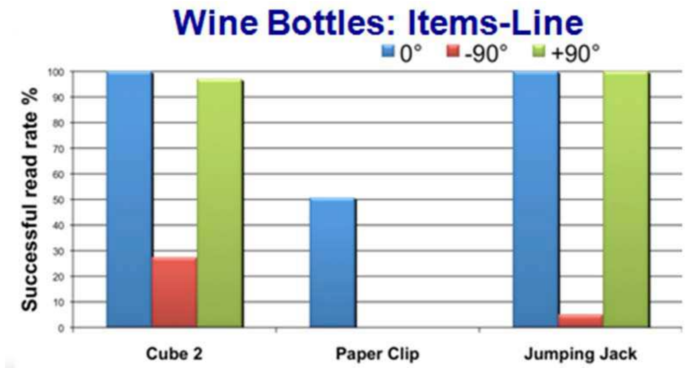


Figure 1: Successful read rate of three commercial tags in the wine items line.

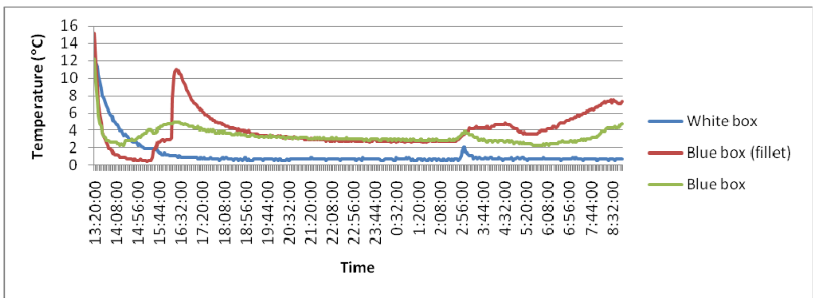


Figure 2: Temperatures measured inside three boxes of fish.

To perform tests in real environment we used two RFID development kits [3], one with IDS-SL900A chip working in UHF frequency bands and the other one with IDS-SL13A in HF frequency bands. Temperatures are automatically measured and recorded, and can be used in a system to automatically notify users about an event.

For the pilot implementation, RFID data loggers will be used to measure the temperature of fish which means that they are placed in the middle of box under the ice before they leave a processing room. By this kind of a solution we start measuring temperatures when fish is packed and read temperatures only at the delivery place. To show a temperature graph during complete supply chain a mobile RFID reader will be used when RFID data loggers are removed from the box.

An example of measured temperatures can give the consumer a clear picture what happened with fish he ordered. Figure 2 shows a temperature graph of three different packages of fish: a) White box — temperatures inside a box where fish was covered by ice; b) Blue box (Fillet) — the process of filleting took place after the sorting process and fish was stored to the blue box and covered by ice packing sheets; c) Blue box — fish was stored to blue box and covered by ice packing sheets. At the end, by the delivery only temperatures of fish covered with ice was more or less the same while in blue boxes temperature starts to increase due to the transport environmental conditions.

6. CONCLUSIONS

This paper presents the objectives and initial work of the project "RFID from Farm to Fork", highlighting the radio propagation concerns of a RFID system within an industrial environment. Some advances on the application to winery and fish manufacturing processes have been also commented.

Taking into account the electromagnetically complex propagation scenario of a winery and of a wine supply chain, an exhaustive tag characterization has been carried out and the performance degradation due to the presence of liquid has been confirmed. Once proved that, in the wine production case, UHF tags must be preferred to HF ones. The performance comparison of both near field and far field UHF tags in a practical scenario has been performed and the better suitability of far field tags has been pointed out.

Vice versa, as for the fish manufacturing case, RFID data loggers were used to measure and store temperatures inside a box of fish covered with ice. This data can be used to control the conditions of perishable food in the supply chain. The proposed implementation is based on radiofrequency identification and does not require the use of any additional sensor networks communication technology. The results were very similar for both systems (HF and UHF) used in the described scenario.

ACKNOWLEDGMENT

This work has been supported by European Union (CIP-Pilot Actions), under the project "RFID from Farm to Fork", grant agreement number 250444.

REFERENCES

1. *RFID Journal*, http://www.rfidjournal.com.
2. RFID-F2F From Farm to Fork, http://www.rfid-f2f.eu/.
3. IDS-SL13A Smart Label Chip with Sensor & SL 900A, EPC Class 3 Chip with Sensor, http://www.ids.si/.

Recent Evolution of ITU Method for Prediction of Multipath Fading on Terrestrial Microwave Links

Basile L. Agba[1], **Olfa Ben-Sik-Ali**[1], **Robert Morin**[2], **and Germain Bergeron**[2]

[1]Institut de Recherche d'Hydro-Québec, 1800 Lionel-Boulet, Varennes, Québec, Canada
[2]Hydro-Québec, 15000 University, Montreal, Québec, Canada

Abstract— Three methods are commonly used to predict multipath fading on terrestrial line-of-sight (LOS) links namely Barnett-Vigants and Morita methods used respectively in North America and Japan, and the worldwide ITU method. This later gains in importance even in North America and in Japan, due to the regular updates proposed by Study Group 3 of the ITU Radiocommunications sector. The ITU recommendation, ITU-R P.530 provides guidelines based on fading measurements of 251 links in various geoclimatic regions. Since 1978, thirteen revisions were proposed. The paper reviews the evolution of ITU-R P.530 from 1997 to 2009 with focus on the difference between the last two revisions (Rev. 12 versus Rev. 13). These differences concern mainly the geoclimatic factor, which depends on the refractivity gradient and the terrain roughness, and the multipath fade occurrence factor that affects worst month outage probability. The paper also presents parametric studies carried out using our prediction tool, Microwave link simulator, to determine the critical parameters affecting link performance. In our studies, 20 links located in Quebec, Canada, are designed based on real-world parameters. The detailed link profiles are known and refractivity gradients are extracted from ITU database according to ITU-R P.453-9. Climatic conditions are also taken into account using local rain database. The results of this work show that three climatic parameters, namely rain intensity, refractivity gradient and annual mean temperature are critical. Likewise for equipments, radio signature and antenna cross-polarization discrimination highly reduce the overall performance in terms of link availability. Finally, two year measurements data of five links are analysed. The results show a good agreement with ITU-R P.530-13.

1. INTRODUCTION

Microwave prediction methods used in links design are intensively investigated since many years. Nevertheless, the unpredictable variation of waves propagation remained a challenging issue, that Scientifics try to handle by proposing prediction methods (such as Barnett-Vigants [1] in North America, Morita in Japon [2]...) that take into account several climatic and environmental parameters. The Study Group 3 of ITU has proposed and maintained actively a microwave prediction method defined in the ITU-R P.530 recommendation. This method gains in importance even in North America and in Japan, due to its regular updates. ITU-R P.530 provides guidelines based on fading measurements of 251 links in various geoclimatic regions. Since 1978, thirteen revisions were proposed. This paper addresses the recent evolution of this method and presents some experimental validation examples.

Section 2 of this paper reviews the evolution of ITU-R P.530 from 1997 to 2009 with focus on the difference between the last two revisions (Rev. 12 [3] versus Rev. 13 [4]). Section 3 presents parametric studies carried out using our prediction tool, Microwave link simulator, to determine the critical parameters affecting link performance. This section also provides comparison between some experimental measurements and the predicted values.

2. EVOLUTION OF ITU-R P.530

It is interesting to look at the significant changes in the evolution of different versions since the 1978 original one. For practical reasons, we will limit ourselves to versions from version 7 (1997) to version 13 (2009, currently in use). Note that there are only minor differences between versions 7 and 8 and also between versions 11 and 12.

- From versions 7/8 to 9, the major changes concern the geoclimatic K factor and consequently the worst month unavailability calculation. The attenuation by hydrometeors is specified by long-term statistics of rain attenuation and includes the duration and the number of events.
- From 9 to 10, techniques for alleviating the effects of multipath propagation (with and without diversity use) are introduced and included in the calculation of the probability of unavailability.

- From 10 to 11/12, the main improvement concerns the development of a method that takes into account the combined effect of rain and wet snow.
- From 12 to 13, there is a radical change in the expression of the geoclimatic K factor and the worst month probability of unavailability.

The regular review of ITU-R P.530-13 allows the adaptation of the theoretical model to the real world experimental measurements on a regular basis. It also represents a challenging task in terms of updating the simulation and design tools to reflect theses changes. The examination of version 13 shows that there are two major modifications, namely in the expressions of the geoclimatic K factor and the "worst month" outage probability, P_w. To better understand the impact of these modifications, the corresponding formula proposed in versions 12 and 13 are presented and a comparative study of K and P_w is performed. The main changes are highlighted in the equations below:

ITU-R P.530-12 [3]

$$K = 10^{-3.9-0.003dN_1} S_a^{-0.42}, \tag{1}$$
$$p_w = Kd^{3.2}(1+|\varepsilon_p|)^{-0.97} \times 10^{0.032f-0.00085h_L-A/10}. \tag{2}$$

ITU-R P. 530-13 [4]

$$K = 10^{-4.4-0.0027dN_1}(10+S_a)^{-0.46}, \tag{3}$$
$$p_w = Kd^{3.4}(1+|\varepsilon_p|)^{-1.03} \times f^{0.8} \times 10^{-0.00076h_L-A/10}. \tag{4}$$

where, dN_1 is the point refractivity gradient in the lowest 65 m of the atmosphere not exceeded for 1% of an average year, S_a is the area terrain roughness defined as the standard deviation of terrain heights (m) within a 110 km × 110 km area, f is the frequency (GHz), d is the path length (km), $|\varepsilon_p|$ is the magnitude of the path inclination (mrad), h_L is the altitude of the lower antenna and A is the fade depth (dB). P_w is expressed in % and can be converted in seconds.

The modifications in Equations (1) to (4) are broadly reflected by a significant decrease of the K factor. Although the K-factor decreases significantly (Figure 1), the "worst month" unavailability increases as shown on Figure 2. To confirm this observation, we applied the two versions on 20 real world links located throughout Quebec, Canada. The location and the main parameters are summarized in Table 1 and the comparison results are presented on Figure 3 in terms of "worst month" unavailability. All the links operate at 7.425 GHz with length of 70 km and frequency and space diversity (2 receivers) is applied. The increasing of the "worst month" outage probability is equivalent to an additional loss of about 2 dB in system gain.

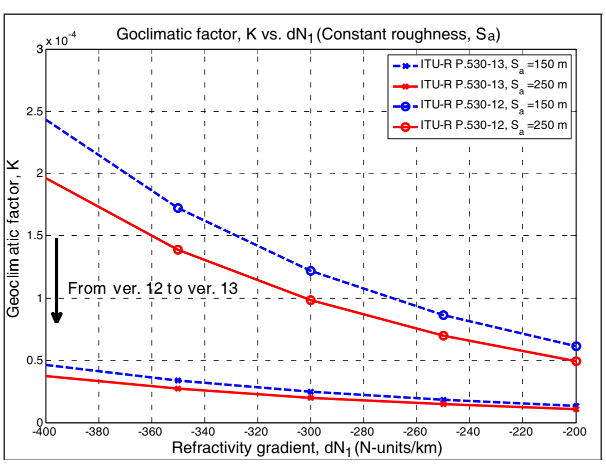


Figure 1: K factor as function of dN_1 (P.530-12 vs. P.530-13).

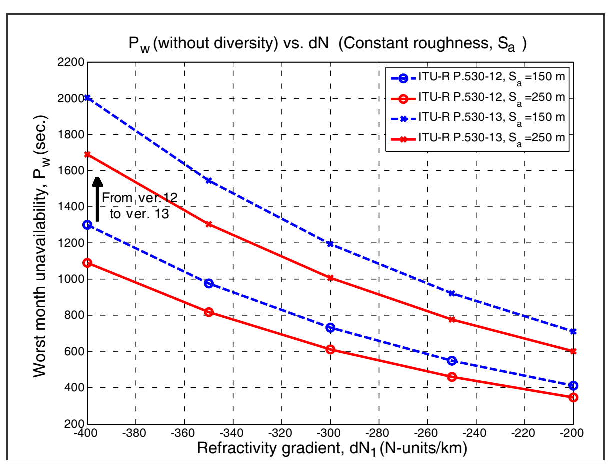


Figure 2: Unavailability as function of dN_1 (P.530-12 vs. P.530-13).

Table 1: Locations and main parameters of links.

No.	h_L (m)	$\|\varepsilon_p\|$ (mrad)	Lat. ξ (^{o})	dN_1 (N-units /km)	S_a (m)
BV1	477	100	53.14	–258.72	68.50
BV2	639	0.00	51.91	–249.23	68.67
BV3	57	0.03	50.30	–295.99	161.91
BV4	151	1.17	50.26	–273.64	210.62
BV5	108	2.59	49.18	–286.79	282.64
BV6	122	2.59	48.11	–273.64	168.29
BV7	74	2.57	49.19	–285.00	161.04
BV8	109	0.81	48.56	–243.42	56.77
BV9	152	0.99	46.84	–269.80	235.67
BV10	403	1.70	46.11	–259.81	152.84
BV11	254	0.74	45.28	–275.44	141.03
BV12	59	0.44	45.60	–274.74	58.48
BV13	89	0.44	45.55	–291.78	86.42
BV14	395	1.94	47.46	–299.12	66.89
BV15	401	0.33	48.48	–297.92	142.79
BV16	334	0.21	48.57	–326.57	46.10
BV17	399	0.59	49.79	–323.88	34.11
BV18	299	0.61	51.71	–335.89	35.47
BV19	332	0.07	53.61	–292.98	37.71
BV20	182	0.01	53.72	–273.00	53.31

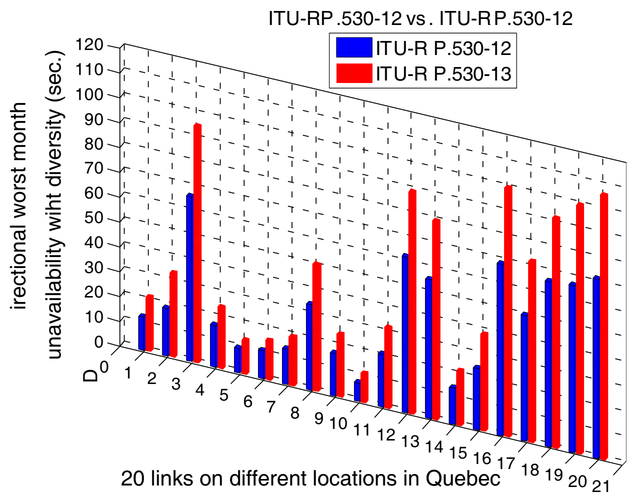


Figure 3: Unavailability with diversity over 20 links in Québec (P.530-12 vs. P.530-13).

3. PARAMETRICAL STUDIES USING ITU-R P.530-13 AND VALIDATION BY MEASUREMENTS

3.1. Impact of Geoclimatic Parameters

The parametrical studies are useful to give a better understanding of the impact of a single parameter or a group of parameters. For example, the combined effect of dN_1, S_a, $|\varepsilon_p|$ and h_L is included in the multipath fade occurrence factor, P_0 which indicates the global behavior of the considered link. None of the parameters mentioned above can neither explain alone the variation of the outage probability, nor be considered as the most dominant. Indeed, the effects of each one can be annihilated by the others. For example a high roughness can reduce the K factor and consequently the outage probability. But this potential reduction will be canceled or completely reversed if the link is located at low altitudes and has a low path inclination.

3.2. Influence of Radio Signature

Until recently, frequency selective fading are only taking into account by applying an additional mitigating factor on the flat fade margin (FFM), to obtain the effective fade margin (EFM). This mitigating factor is calculated using the radio equipment dispersive fade margin. ITU introduced in ITU-R P.530-7 an alternative method based on radio signature.

As it might be expected, taking into account the radio signature greatly degrades the total outage probability if no diversity technique is applied. However, applying the diversity techniques (in particular space diversity) effectively mitigates selective fading. The radio signature is useful to quantify the impact of selective fading in broadband systems.

3.3. Influence of Cross-polarization Discrimination, XPD

The cross polarization acts directly on the outage probability because the XP outage probability is added to the "worst month" outage. Three main parameters are used, namely the guaranteed maximum cross-polarization, XPDg, the cross-polarization improvement factor, XPIF and the carrier-to-noise ratio, C/I. If XPDg exceeds 35, then it has no impact on the XP outage. This value is usually specified by antennas' manufacturers. We then present only the outages as a function of XPIF and C/I. The XP outage probability for clear air increases with the increase of C/I and decreases with XPIF. These observations indicate that both parameters have an opposite influence. A tradeoff for equipment selection is necessary to minimize the XP effects.

3.4. Comparison between Prediction and Measurements

Two years' measurements data (April 2009–August 2010) of five links are analysed. The comparison results with ITU-R P.530-13 show a good agreement as confirmed on Figure 6. The parameters used are $dN_1 = 297.96$ N-units/km, $S_a = 140.63$ m, $|\varepsilon_p| = 0.285$ mrad and $h_L = 399$ m. For space

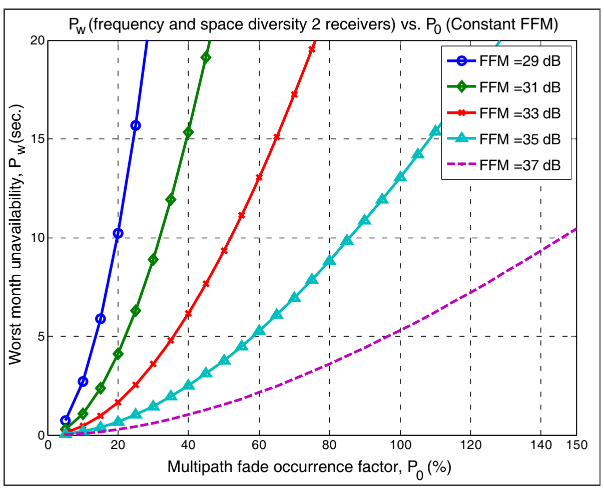


Figure 4: Worst month unavailability as function of multipath fade occurrence factor.

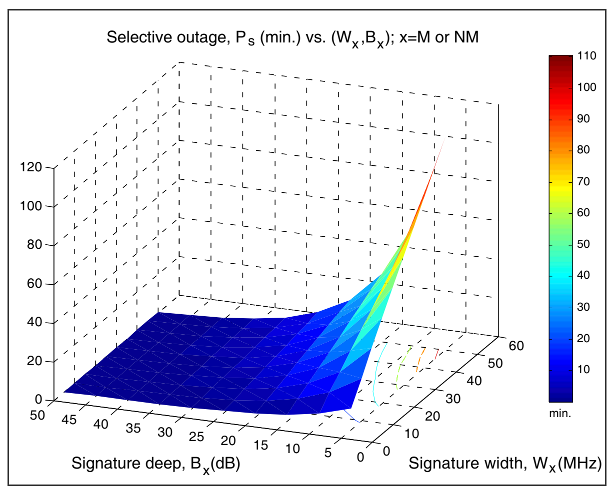


Figure 5: Selective outage probability as function of radio signature parameters.

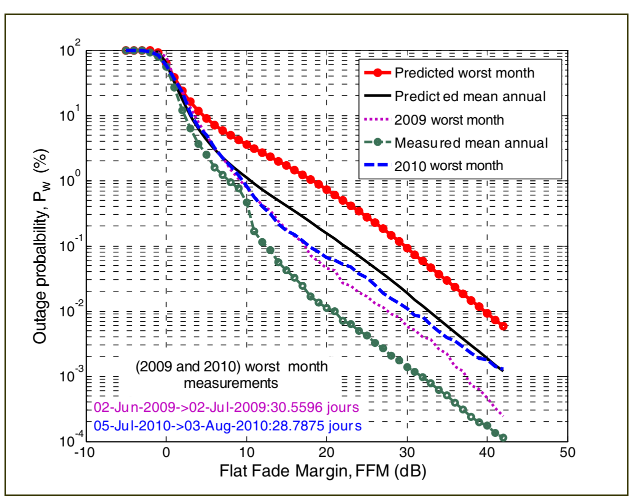


Figure 6: Prediction vs. measurements ("Worst month" and annual outage probabilities).

Table 2: Mean error between prediction and measurements (measured on the receiver site).

	site 1	site 2	site 3	site 4	site 5
2009 $< A_m >$ (dB)	28	27	31	26	33
2010 $< A_m >$ (dB)	26	31	28	29	35
A_{th} (dB)	32	32	34	40	42
$E = \lvert A_{th} - < A_m > \rvert$	4–6 dB	1–5 dB	3–6 dB	11–14 dB	7–9 dB

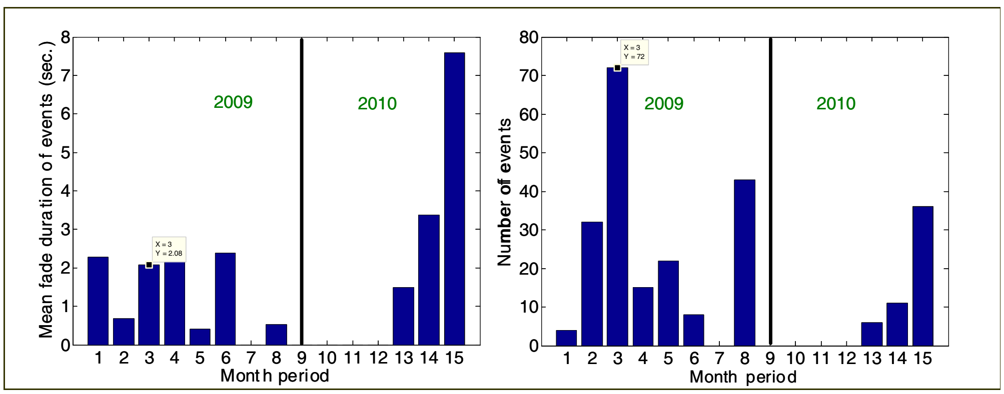


Figure 7: Statistics of fading events (mean fade duration and number of fading events) Fade margin, $A = 30\,\text{dB}$.

reasons, results over one link is presented with some statistics such as number of fading events and mean fade duration.

The analysis of all measurements over the six links during two years (2009–2010) leads to the following observations and conclusions.

- Measurement and prediction curves have similar tendency. Theoretical Prediction curve is an upper bound which is validated during the link measurement duration (2 years).
- There is a huge dispersion of the distribution of fading from month to month. The month to month cap varies from 1 dB to 20 dB.
- The error on the prediction of the "worst month" outage, (for an outage of 10^{-4}) is about 14 dB for 2009 and 11 dB for 2010. This result is consistent with the results published by [5] which are based on 239 link statistics used in the definition of the ITU-R P.530-8 recommendation.
- Finally, the analysis of diversity improvement factors show a good agreement between measurements and prediction for frequency and space diversity using 2 receivers but the improvement factor when 4 receivers are applied is low compared to ITU prediction.

4. CONCLUSION

The evolution of ITU-R P.530 and its validation by experimental measurements carried out in Quebec, Canada are the two main elements addressed in this paper. The results can be summarized as follows. The evolution of ITU-R P.530 is very dynamic and a new version is proposed approximately every two years. The validation of each new revision is confronted to measurements data collected over the world and available in SG3 database. A parametric study made on the latest revision (P.530-13) shows that the multipath fade occurrence factor mainly affects the "worst month" outage. Likewise for equipments, radio signature and antenna cross-polarization discrimination highly reduce the overall performance in terms of link availability. Finally, our own measurements confirm the accuracy of the prediction based on ITU-R P.530-13 with acceptable error margin.

REFERENCES

1. Barnett, W. T., "Multipath propagation at 4, 6 and 11 GHz," *Bell System Technical Journal*, Vol. 51, No. 2, 311–361, February 1972.
2. Morita, K., "Prediction of rayleigh fading occurrence probability of line of sight microwave links," *Rev. Elec. Comm. Lab. (Japan)*, Vol. 18, 310–321, November 1970.
3. International Telecommunication Union (ITU-R), "Propagation data and prediction methods required for the design of terrestrial line-of-sight systems," Recommendation UIT-R P.530-12, 2005.
4. International Telecommunication Union (ITU-R), "Propagation data and prediction methods required for the design of terrestrial line-of-sight systems," Recommendation UIT-R P.530-13, 2009.
5. Olsen, R. L. and T. Tjelta, "Worldwide techniques for predicting the multipath fading distribution on terrestrial L.O.S links: Comparison with regional techniques," *IEEE Transactions on Antenna and Propagation*, Vol. 51, No. 1, 23–30, January 2003.

Ultra-wideband Spatio-temporal Channel Sounding with Use of an OFDM Signal in an Indoor Environment

Daisuke Sugizaki, **Naohiko Iwakiri**, **and Takehiko Kobayashi**
Wireless Systems Laboratory, Tokyo Denki University, Japan

Abstract— Microwave vector network analyzers (VNAs) have been widely used as ultra wideband (UWB) channel sounders. However, channel sounders using VNAs have some limitations: there is only one snapshot and incoming waves tend to be strongly correlated due to sounding signals of sine waves. To overcome the above issues, a UWB spatio-temporal channel sounder, employing an orthogonal frequency division multiplexing (OFDM) signal and virtual array antenna, was developed along with a proposed technique for estimating time-of-arrival (TOA) and angle-of-arrival (AOA) of UWB OFDM signals. In the prototype sounder, the bandwidth of the OFDM signal was 500 MHz (3.25–3.75 GHz). The received collinear virtual array antenna consisted of 8 elements with spacing of 40 mm. The TOAs were calculated by correlating the received waveforms with a transmitting template waveform. The AOAs were derived by extending a multiple signal classification (MUSIC) algorithm and an angular histogram method. Comparisons were carried out between the VNA- and the OFDM-based channel sounders in terms of experimental estimation of TOAs and AOAs in an indoor three-path environment. Delay profiles and AOA estimation results agreed reasonably well with ray tracing calculations and the proposed design of the OFDM-based channel sounder was validated.

1. INTRODUCTION

Representative results of channel sounding have been presented in [1–4] for wideband and [5–9] for ultra-wideband (UWB). A microwave vector network analyzer (VNA) and a single-element antenna with a spatial scanner, which conforms to a synthetic aperture, are widely used structures to measure the spatial distribution of transfer functions [9]. However, channel sounders using VNAs have some limitations: there is only one snapshot, incoming waves tend to be strongly correlated due to sounding signals of sine waves, and the sounding signals are do not correspond to actual signals of wireless communications. To overcome the above issues, a UWB spatio-temporal channel sounder, employing an orthogonal frequency division multiplexing (OFDM) signal and a virtual array antenna, was developed along with a proposed technique for estimating time-of-arrival (TOA) and angle-of-arrival (AOA) estimation technique of UWB OFDM signals. Experiments were carried out [10] in a radio anechoic chamber to compare this OFDM-based channel sounder with a conventional VNA-based and a pseudo-noise-based channel sounders to verify the effectiveness of the proposal. This paper presents another comparison in an indoor three-path environment between VNA- and OFDM-based channel sounders in terms of experimental estimation of TOAs and AOAs.

Our underlying motivation is to apply it to a detect-and-avoid cognitive radio system [11]. The proposed sounder is capable of estimating the spectrum and AOA of an in-band (narrowband) signal as well as sounding the desired UWB signal itself [12]. To avoid interference, a spectral hole can be created within an OFDM signal by nulling subcarriers corresponding the estimated spectrum; the antenna beam can be steered to direct the null(s) to the AOA.

2. VNA- AND OFDM-BASED UWB CHANNEL SOUNDERS

The block diagrams of the VNA- and OFDM-based sounder are depicted in Fig. 1. In the VNA-based sounder, the VNA acts as a transceiver, as shown in Fig. 1(a). Omnidirectional antennas yielding a low voltage standing wave ratio (< 1.3) and very short group delay (< 0.1 ns) [13] were employed for transmission and reception in both systems. The receiving antenna was a virtual array using the omnidirectional antenna as an element that was scanned with a precise mechanical scanner.

In the prototype sounder, the transmitter of the sounder consisted of an arbitrary waveform generator (AWG), an orthogonal modulator, a microwave local oscillator, and a power amplifier, as shown in Fig. 1(b). A 500-MHz baseband OFDM signal was generated with a 2.5-Gsamples/s, 8-bit AWG. Next, a 3.5-GHz carrier frequency was modulated with the OFDM signal by using the orthogonal modulator, and the resulting radiofrequency UWB OFDM signal was amplified to feed a transmitting antenna. The received signal was amplified with a low noise amplifier, demodulated with an orthogonal demodulator, and analog-to-digital converted and recorded with a 2.5-Gsamples/s, 8-bit digital storage oscilloscope (DSO). The typical signal-to-noise ratio of this system was 30 dB at a distance of 3 meters between the transmitting and the receiving antennas. The whole system was controlled by a personal computer (PC). The data recorded in the DSO was sent to and analyzed off-line by the PC.

3. EXPERIMENTAL SETUP

The sounding signals of the VNA-based sounder are discrete sine waves in a frequency domain. The signals occupied a bandwidth of 500 MHz and its center frequency was 3.5 GHz. The number of points and IF bandwidth were set to 501 pts and 100 Hz, as listed in Table 1.

The sounding signals of the OFDM-based sounder complied with one of the subbands specified in Multiband OFDM (MB-OFDM) [14] — one of the UWB communication standards. Signal processing software, SPW® [15], numerically generated the OFDM signal data, from which the AWG generated the actual signal. The bandwidth of the OFDM signal was 500 MHz and its center frequency was 3.5 GHz. The number of subcarriers was 128; the lengths of a guard interval, the OFDM symbol, and the total symbol were 32 chips, 32 OFDM symbols, and 160 chips, respectively. The length of one OFDM burst was 5120 chips, which corresponded to 10.24 µs (= 2 ns/chip × 5120 chips). The OFDM signal parameters are listed in Table 2.

Propagation measurement was carried out in an indoor three-path environment. Two planar reflectors were placed in the room to create reflected waves, as shown in Fig. 2. The antennas used for transmission and reception were vertically polarized. The received collinear virtual array

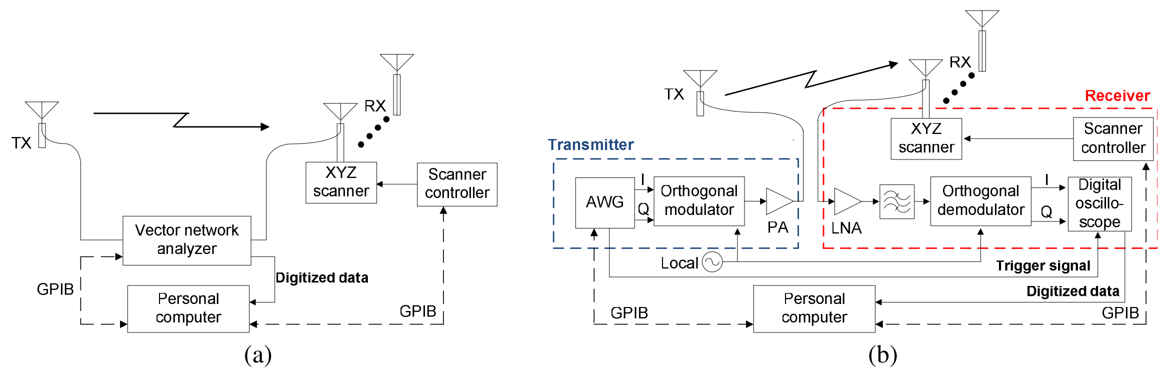


Figure 1: The block diagrams of the (a) VNA- and (b) OFDM-based UWB spatio-temporal channel sounder.

Table 1: VNA signal parameters.

Center Frequency	3.5 GHz
Bandwidth	500 MHz
Number of points	501 pts
IF bandwidth	100 Hz

Center Frequency	3.5 GHz
Bandwidth	500 MHz
Modulation scheme	QPSK-OFDM
PN code length	32 OFDM symbols
Number of subcarriers	128 chips
Guard interval length	32 chips
Burst length	5120 chips
Smapling rate	2.5 Gsps (5-times oversampling)

Table 2: OFDM signal parameters.

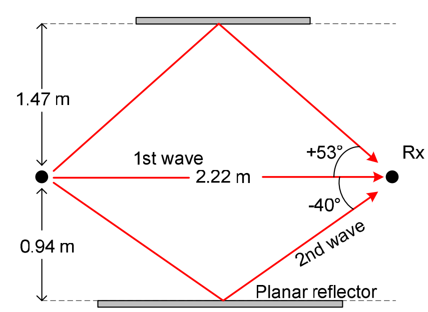


Figure 2: The measurement environment.

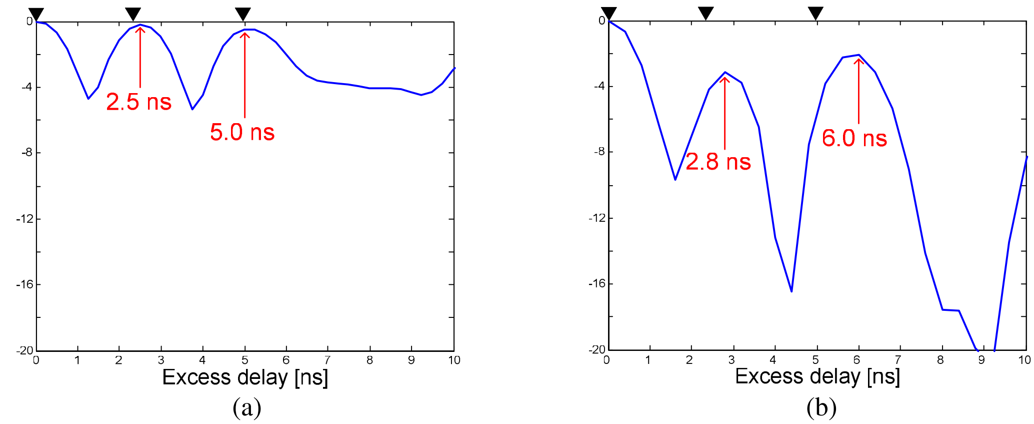


Figure 3: Delay profiles derived from measured data using: (a) VNA- and (b) OFDM-based channel sounders. Results of ray tracing calculation are indicated by inverted triangles.

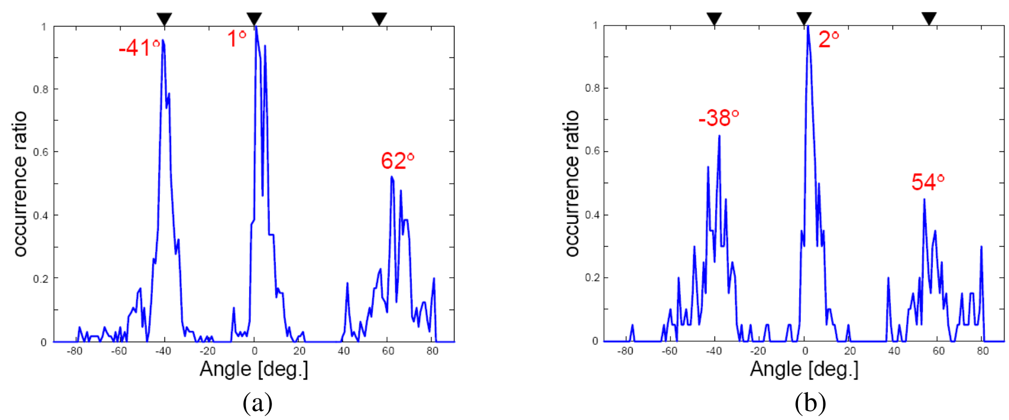


Figure 4: AOA estimation results derived from measured data using: (a) VNA-, (b) OFDM-based channel sounders. Results of ray tracing calculation are indicated by solid inverted triangles.

antenna consisted of 8 elements, and the element spacing was 40 mm.

4. RESULTS

The TOAs were calculated by correlating the received waveforms with a transmitting template waveform. The AOAs were derived by extending a multiple signal classification (MUSIC) algorithm and an angular histogram method [5]. Comparison was made between the VNA- and the OFDM-based channel sounders in terms of experimental estimation of TOAs and AOAs in an indoor three-path environment. The excess delay times and AOAs of the second and third wave were calculated with a ray tracing method at 2.3 ns and 4.9 ns, and $-40°$ and $53°$, respectively. In the VNA-based and OFDM-based channel sounder, the delay time of the second and the third waves were respectively estimated at 2.5 ns and 5.0 ns, and 2.8 ns and 6.0 ns, as shown in Fig. 3. Considering the time resolution of these systems = 0.4 ns, both delay profiles agreed reasonably well with the ray tracing calculation. Estimation results of the AOA (assuming three waves) also agreed reasonably well with ray tracing, as shown in Fig. 4. The design of the proposed OFDM-based channel sounder was validated.

5. CONCLUSIONS

Comparison was carried out between VNA- and OFDM-based channel sounders in terms of experimental estimation of TOAs and AOAs in an indoor three-path environment. The results of TOA and AOA estimation agreed reasonably well with ray tracing calculations. The proposed OFDM-based channel sounder was demonstrated to nearly equal the VNA-based channel sounder. Future works include replacement of the virtual array by a multiple antenna systems and experimental validation of its DAA function.

ACKNOWLEDGMENT

This work has been supported in part by the Research Institute for Science and Technology, Tokyo Denki University, under the grant Q06J-02. The authors would also like to express special thanks to Hiroto Iizuka for assistance in prototyping the sounder and Synopsys, Inc. for providing the simulation tool.

REFERENCES

1. Fleury, B. H., M. Tschudin, R. Heddergott, D. Dahlhaus, and K. I. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algolithm," *IEEE J. Select. Areas Commun.*, Vol. 17, No. 3, 434–449, Mar. 1999.
2. Kalliola, K., H. Laitinen, L. I. Vaskelainen, and P. Vainikainen, "Realtime 3D spatial-temporal dual-polarized measurement of wideband radio channel at mobile station," *IEEE Trans. Instrum. Meas.*, Vol. 49, No. 2, 439–448, Apr. 2000.
3. Sakaguchi, K., J. Takada, and K. Araki, "A novel architecture for MIMO spatio-temporal channel sounder," *IEICE Trans. Electron.*, vol. E85-C, No. 3, 436–441, Mar. 2002.
4. Bethel, R. E. and K. L. Bell, "Maximum likelihood approach to joint array detection/estimation," *IEEE Trans. Aerosp. Electron. Syst.*, Vol. 40, No. 3, 1060–1072, Jul. 2004.
5. Iwakiri, N. and T. Kobayashi, "Ultra-wideband indoor double-directional channel estimation using transformation between frequency and time domain signals," *IEICE Trans. Fundamentals*, Vol. E92-A, No. 9, 2159–2166, Sep. 2009.
6. Cramer, R. J.-M., R. A. Scholtz, and M. Z. Win, "Evaluation of an ultra-wide-band propagation channel," *IEEE Trans. Antennas Propag.*, Vol. 50, No. 5, 561–570, May 2002.
7. Carbonelli, C. and U. Mitra, "Clustered channel estimation for UWB signals," *IEEE Int. Conf. Commun.*, 2432–2436, Jun. 2004.
8. Haneda, K., J. Takada, and T. Kobayashi, "Double directional ultra wideband channel characterization in a line-of-sight home environment," *IEICE Trans. Fundamentals*, Vol. E88-A, No. 9, 2264–2272, Sep. 2005.
9. Haneda, K., J. Takada, and T. Kobayashi, "Cluster properties investigated from a series of ultrawideband double directional propagation measurements in home environments," *IEEE Trans. Antennas Propag.*, Vol. 54, No. 12, 3778–3788, Dec. 2006.
10. Iizuka, H., M. Takahashi, N. Iwakiri, and T. Kobayashi, "A prototype of ultra wideband spatio-temporal channel sounder and its performance evaluation," *Progress In Electromagnetics Research Symposium*, Moscow, Russia, Aug. 18–21, 2009.
11. Weiss, T. A. and F. K. Jondral, "Spectrum pooling: An innovative strategy for the enhancement of spectrum efficiency," *IEEE Commun. Mag.*, Vol. 42, No. 3, S8–13, Mar. 2004.
12. Iwakiri, N. and T. Kobayashi, "OFDM interference detection using flexible subcarrier channel estimator," *4th European Conference on Antennas and Propagation (EuCAP 2010)*, Barcelona, Spain, Apr. 2010.
13. Taniguchi, T., A. Maeda, and T. Kobayashi, "Development of an omni-directional and low-VSWR ultra wideband antenna," *Int. J. Wireless* & *Optical Commun.*, Vol. 3, No. 2, 145–157, Feb. 2006.
14. "Multi-band OFDM physical layer proposal for IEEE 802.15 task group 3a," IEEE P802.15-03/268r3, Mar. 2004.
15. http://www.synopsys.com/Tools/SLD/DigitalSignalProcessing/Pages/Signal-Processing.aspx.

Densitometry of Electromagnetic Field Exposure Due to Wi-Fi Frequency

A. D. Usman[1], **W. F. Wan Ahmad**[1], **M. Z. A. Ab Kadir**[1], **M. Mokhtar**[2], **and M. A. Zainal Abidin**[1]

[1]Centre of Excellence on Lightning Protection (CELP), Faculty of Engineering
Universiti Putra Malaysia, UPM Serdang, Selangor 43400, Malaysia
[2]Wireless and Photonic Networks (WiPNET), Research Center of Excellence
Faculty of Engineering, Universiti Putra Malaysia, UPM Serdang, Selangor 43400, Malaysia

Abstract— The world is undergoing electromagnetic revolution with many frequencies used for a variety of wireless devices. Recently, the emergence of Wi-Fi technologies deployed in schools, offices and other public places raised a lot of concern by the general public on the magnitude and safety of this exposure. To this effect, in situ measurements were conducted at various Wi-Fi access points at both academic and students' residential areas in Universiti Putra Malaysia, Serdang, Selangor, Malaysia. An RF survey meter with tri-axis probe and a spectrum analyzer were used to carry out measurements at far field distance for 6 minutes average time. Various locations were investigated and it is found that the highest and the lowest readings obtained during the downloading and uploading activities were 5.63 V/m and 0.37 V/m, respectively. These readings are far below the ICNIRP reference guidelines for exposure to general public, i.e., 61 V/m for 2.46 GHz Wi-Fi operating frequency.

1. INTRODUCTION

Wireless Fidelity or popularly known as Wi-Fi is one of the new technologies that are widely used by both the young and the aged. This new technology enable users to communicate to internet from any place such as offices, homes, in space and any place without the hassles of plugging wires.

Locations where users can connect to wireless networks are called Wi-Fi hotspots. The most popular Wi-Fi technology such as 803.11b operates at frequency range of 2.40 GHz to 2.4835 GHz. The network devices providing the Wi-Fi air interface between the broadband network and the user terminal is called Wi-Fi access point (Wi-Fi AP), which is a radio transceiver itself. The use of Wi-Fi services for private purposes is covered by an exemption order, but the provision of it to public in some countries is permissible under a class of license provided that the services do not cross the public streets or unleashed government lands. For these public purposes a fixed carrier license is required.

Regulations were given by international bodies and scientific communities in terms of maximum exposure allowed [1–4]. However, even with much lower exposure, the energy emitted from radio frequency (RF) and microwave frequencies has unwanted side effects [5–9]. To date, there are no studies adequately done on effect of exposure to Wi-Fi. However, there are similar studies on the frequencies around the frequencies of operation of Wi-Fi resulting on complicating conclusions. Yet attempts were made by some countries to turn their cities to e-cities with Wi-Fi technologies thereby exposing the greater populace to radio frequency radiation (RFR). Few studies were performed on the RFR emitted by Wi-Fi hotspots as in [10–12], with none is done in Malaysia being one of the leading nation in South East Asia that employ the use of this technology in almost all facet of life, has motivated this research. Another factor is the public concern and the exposure of the members of staff of university community and students spending extensive amounts of time in buildings in which Wi-Fi networks has been established made us to investigate the RFR levels at both the Faculty of Engineering and the 10th college student hostel putting into consideration the activity levels, RFR emission from both the routers and the Wi-Fi APs during the time of measurements. A controlled environment was also used to carry out the experiment during the downloading and uploading activities to validate the practical measurements results. Note that both Faculty of Engineering and the 10th college student hostel are located at Universiti Putra Malaysia, Serdang, Selangor, Malaysia.

2. MATERIALS AND METHOD

Measurements were performed at 10 different points within the Faculty of Engineering and 10th college student hostel, following the experimental protocol as in [13]. The instruments used were the

PMM8053A Broadband RF survey meter with a frequency range of 100 kHz to 3 GHz and electric field measurement range from 10 mV/m to 100V/m. The RF survey meter has a tri-axis probe. A laptop is used to retrieve the field strength information from the meter. A spectrum analyzer SA 9270A of frequency range 9 kHz to 2.9 GHz was used to determine the dominant operating Wi-Fi frequency. This meant for the dominant exposure level to be found. Having determined that, the RF survey meter operating in broadband is used to measure the field strength. Figure 1 shows the set-up of the measurement procedure at Wi-Fi APs 1 to 9, while Figure 2 illustrated the set up for a controlled experiment carried out in the 8th floor Laboratory at Wi-Fi AP 10. The far field approximate points were determined numerically using the far field equation as in (1), where r is the distance from observation point, r_n is the distance in m, D is the maximum dimension of the antenna and λ is the operating wavelength.

$$r > r_n = \frac{D^2}{2\lambda} \tag{1}$$

Note that the Wi-Fi APs used in Universiti Putra Malaysia campuses were Omnidirectional of two types, i.e., ceiling and wall mounted. The antenna were made from Cisco with operating frequency of 2.46 GHz, a gain of 5.2 dBi and approximate length of 15 cm. Hence, the far field was found to begin at 0.1 m. Therefore, electric field measurements were performed at 0.2 m, 0.4 m, and 1m distances. Equation (2) was also used to validate the field measurements at various distances, were E is the magnitude of the electric field strength in V/m, P is the power of the source in W, and r is the radial distance.

$$E = \frac{1}{2r}\left(377 \times \frac{P}{\pi}\right)^{\frac{1}{2}} \tag{2}$$

The measurements from the controlled experiment were validated by estimating the polarization of the signals from the Wi-Fi router. This was achieved by repeatedly engaging the network with downloading and uploading activities while the antenna of the router was set at different angles of rotation. The orientation of the antenna with maximum readings was noted to indicate that the antenna was aligned with the polarization of the emission. The measurements were taken by keeping the Wi-Fi router and the computer at a fixed point of a reasonable distance apart. Note that the computer was used for downloading and uploading activities while measurements were taken with the RF survey meter at 0.2 m, 0.4 m, and 1 m distances each for 2 minutes. Then, the router and the RF survey meter were kept fixed and the laptop was moved at a distances of 0.2 m, 0.4 m and 1 m, respectively. In both cases the readings were taken during downloading and uploading activities.

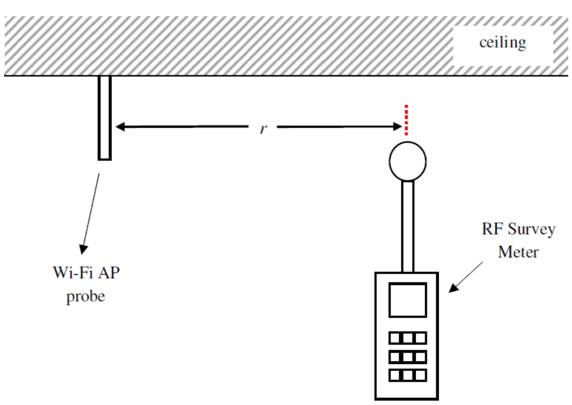


Figure 1: Experimental Set-Up for RF measurements at Wi-Fi APs 1-9.

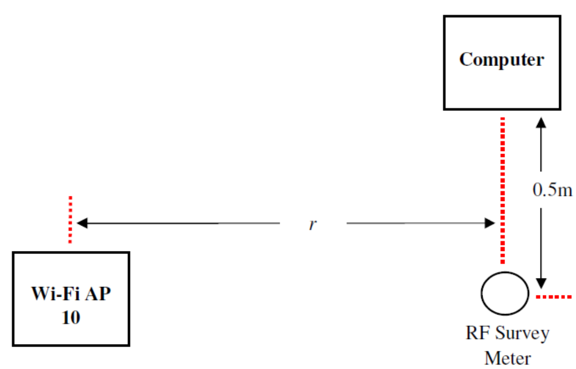


Figure 2: Top view of the experimental Set-Up for RF measurements at Wi-Fi AP 10.

3. RESULTS AND DISCUSSIONS

Based on the far field determination using (1) and (2), measurements were conducted at 0.2, 0.4 and 1 m distances at all the 10 Wi-Fi APs, where the results are as in Tables 1 to 11. At each location, the maximum, minimum and average values of the electric field magnitude are determined against time. Also, their percentage differences with ICNIRP standard for general public exposure are calculated, i.e., 61 V/m for 2.46 GHz Wi-Fi operating frequency.

At the Faculty of Engineering, it is found that the highest field strength of 5.16 V/m was obtained at Wi-Fi AP 3, situated between the two Lecture Halls 6 and 7. This corresponds to 8.5% of ICNIRP reference level for exposure to general public. The lowest readings of 0.37 V/m corresponding to 0.6% ICNIRP guidelines were obtained at Wi-Fi AP 1.

At the student's hostel, it is demonstrated that the highest reading is obtained at 4.51 V/m, i.e., at Wi-Fi AP 8, Wing C in 10th college which represents 7.4% of the ICNIRP reference level for exposure to general public. While, the lowest reading of 1.3 V/m at a distance of 1 m away from Wi-Fi AP 7, Wing B in 10th college was recorded where this is only 2.1% ICNIRP reference level.

The controlled environment results obtained at Wi-Fi AP 10, 8th Floor Laboratory, Faculty of

Table 1: Results obtained for Wi-Fi AP 1 at Cafeteria, Faculty of Engineering.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.00	1.49	0.67
0.4	0.00	0.87	0.43
1.0	0.00	0.37	0.09

Table 2: Results obtained for Wi-Fi AP 2 in front of Lecture Hall 4, Faculty of Engineering.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.00	1.1	0.54
0.4	0.00	0.69	0.33
1.0	0.00	0.43	0.02

Table 3: Results obtained for Wi-Fi AP 3 between Lecture Halls 6 and 7, Faculty of Engineering.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.00	5.16	1.59
0.4	0.00	2.69	1.29
1.0	0.00	0.78	0.27

Table 4: Results obtained for Wi-Fi AP 4 behind the Conference Room, Faculty of Engineering.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.00	0.98	0.33
0.4	0.00	0.78	0.18
1.0*	0.00	1.35	0.35

*Note that, the downloading and uploading activities were simulated at this distance at the time of measurement.

Table 5: Results obtained for Wi-Fi AP 5 between Lecture Hall 8 and Research Division, Faculty of Engineering.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.00	3.38	0.51
0.4	0.00	1.13	0.25
1.0	0.00	0.43	0.01

Table 6: Results obtained for Wi-Fi AP 6 at Wing A of 10th college student hostel.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.00	4.21	1.19
0.4	0.00	1.75	0.34
1.0#	0.00	1.36	0.47
1.0$	0.00	2.45	1.27

#Note that, the measurement is done at empty room 310.

$Note that, the measurement is done at room 309 during two students are inside the room, and downloading and uploading activities are available at that time.

Table 7: Results obtained for Wi-Fi AP 7 at Wing B of 10th college student hostel.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.00	2.64	1.16
0.4	0.00	1.98	0.82
1.0	0.00	1.30	0.45

Table 8: Results obtained for Wi-Fi AP 8 at Wing C 10th of 10th college student hostel.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.58	4.51	2.72
0.4	1.15	3.66	2.52
1.0	0.00	1.52	0.99

Table 9: Results obtained for Wi-Fi AP 9 at Wing D of 10th college student hostel.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.92	3.44	1.98
0.4	0.76	3.02	1.82
1.0	0.00	2.33	1.17

Table 10: Results obtained for Wi-Fi AP 10 at 8th Floor Laboratory when in NON-ACTIVE mode at Faculty of Engineering.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.00	5.63	1.07
0.4	0.00	2.16	0.70
1.0	0.00	0.54	0.03

Table 11: Results obtained for Wi-Fi AP 10 at 8th Floor Laboratory when in ACTIVE mode at Faculty of Engineering.

Distance from APs to the RF Survey Meter, r(m)	**Measured Electric Field Strength (V/m)**		
	Minimum	**Maximum**	**Average**
0.2	0.00	4.09	3.62
0.4	0.00	2.49	1.30
1.0	0.00	1.13	0.81

Engineering while keeping the computer and the Wi-Fi router at constant position and moving the RF survey meter away from the router are as listed in Tables Xa and Xb for non-active and active modes, respectively. Note that there is no downloading and uploading activity during the non-active mode and vice versa during the active mode. The highest electric field strength is obtained at 0.2 m, i.e., 5.63 V/m, and this field strength decreases as one move away from the AP with low readings of 1.13 V/m found at 1 m, corresponding to 5.9%, and 1.3% respectively, of the ICNIRP guidelines for public exposure. While keeping the Wi-Fi router and RF survey meter at 1 m apart ($r = 1$) and moving the computer, it is found that the highest electric field strength is obtained at 0.2 m, i.e., 4.09 V/m, and again the field strength decreases to 0.54 V/m at 1 m away where these are 6.7% and 0.1% respectively, of the ICNIRP guidelines for exposure to general public at 2.46 GHz operating frequency of Wi-Fi technology.

The electric field magnitudes measured in this work are found to be much lower than the ICNIRP standard general public exposure where the calculations are done in a fraction of the standard and presented in percentage values. These regardless the maximum, minimum and average values of the measured electric field magnitudes. The proximity of the Wi-Fi APs and the Wi-Fi routers to the point of downloading and uploading activities should be as far as possible as results obtained demonstrated that, the magnitude of the electric field generated at these APs are highly depending on the downloading and uploading activities levels as well as the distance one person is from an AP. It was also found that the electromagnetic field strength was higher adjacent to the Wi-Fi router rather than near to the computer, as been illustrated by a controlled experiment at Wi-Fi AP 10. This is due to the higher duty cycle at the base station as it serves as the hub in communications with many computers. These results are comparable to the findings in [10, 11] and will no doubt help in various campaign carried out about the fears of citing Wi-Fi hotspots in schools and residential areas.

4. CONCLUSION

Wi-Fi RF exposure assessment was done at Faculty of Engineering and 10th college student hostel, where both are located in Universiti Putra Malaysia, Serdang, Selangor, Malaysia. The frequency of operation of the Wi-Fi is 2.46 GHz which belongs to the IEEE 802.11g unlicensed frequency. The measured results obtained demonstrated that the highest readings to be 5.63 V/m and the lowest was 0.37 V/m. These readings were however much below the 61 V/m ICNIRP reference guidelines for general public. However care should be taken as these APs are very close to the students rooms at about 1.5 m away and the exposure are for 24 hours a day. While, in the case of academic areas people are only exposed to about 8 hours per day where the ICNIRP guidelines adopted by Malaysian government are meant to safeguard against short term thermal effect only. Generally, the electric field intensity is highly depending on the downloading and uploading activity levels, and the distance one person is from the Wi-Fi APs and the Wi-Fi routers. The extent of the exposure at any point was found to be a function of both the field emission and the distance of the transmitters.

REFERENCES

1. IEEE Std C95.1, "Standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz," Institute of Electrical and Electronics Engineers, New York, USA, 2005, ISBN: 0-7381-4835-0, http://ieeexplore.ieee.org.
2. International Commission on Non-Ionizing Radiation Protection, "Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz), *Health Phys.*, Vol. 74, 494–522, 1998, http://www.icnirp.de.
3. Federal Communication Commission, Office of Engineering and Technology, 1998, http://www.fcc.gov.
4. Australian Radiation Protection and Nuclear Agency, "Maximum exposure levels to radiofrequency fields-3 kHz to 300 GHz," 2002, http://www.arpansa.gov.au.
5. Michaelson, S. M. and J. C. Lin, *Biological Effects and Health Implications of Radio-Frequency Radiation*, 1st Edition, Springer, New York, 1987, ISBN: 10: 0306415801.
6. Golberg, B. R., "Literature resources for understanding biological effect of electromagnetic fields," 2002, http://infoventures.com.
7. Usman, A. D. and W. F. Wan Ahmad, "RF EM exposure: The issues, the debate and the realities," *Proceedings of the 2008 IEEE International RF and Microwave Conference (RFM*

2008), 41–45, Kuala Lumpur, MALAYSIA, Dec. 2–4, 2008, ISBN: 978-1-4244-2866-3, Digital Object Identifier: 10.1109/RFM.2008.4897381.
8. Charles, P. and P. Elliot, *Handbook of Biological Effects of Electromagnetic Fields*, 2nd Edition, Boca Raton, CRC Press, Florida, 1995, ISBN: 10: 0849306418.
9. Valberg, P. A., T. E. van Deventer, and M. H. Repacholi, "Workgroup report: Base stations and wireless networks-radiofrequency (RF) exposures and health consequences," *Environ. Health Phys.*, Vol. 115, 416–424, 2007, http://www.ehponline.org.
10. Office of the Telecommunication Authority, "Radiofrequency radiation measurements of public Wi-Fi installations in Hong Kong," 2007, http://www.ofta.gov.hk.
11. Stephenson, G., "Wireless local area network assessment at Methodist First School, Somerset Services, County Hall, Taunton," 2003, www.six.somerset.gov.uk.
12. Kühn, S., U. Lott, K. Axel, and K. Niels, "Assessment of human exposure to electromagnetic radiation from wireless devices in homes and office environments," 2006, http://www.who.int.
13. ECC recommendation (02) 04 Revise, "Measuring non-ionizing electromagnetic radiation (9 kHz–300 GHz)," 2003, http://www.erodocdb.dk.

Channel Characterization Techniques for Wireless Automotive Embedded Systems

M. Cheikh[1, 2], J. David[2], J. Tartarin[3], S. Kessler[1], and A. Morin[1]

[1]Continental Automotive Systems, 1, Av Paul Ourliac, Toulouse 31100, France
[2]LAPLACE, Av Charle Camichel, Toulouse 31000, France
[3]Université de Toulouse — LAAS, 7, Av Colonel Roche, Toulouse 31100, France

Abstract— The number of wireless systems developed for automotive applications increases for security and comfort considerations. Among these applications, the wireless automotive access systems (Wireless Access Systems: WAS) and the direct intelligent tire information system (Tire Pressure Monitoring System: TPMS) have expanded considerably. The emitted (from the badge or the wheel units) waves change in phase and power according to the Multi-path propagation and produces both destructive and constructive behaviors all along the channel and particularly at the receiver side. These environment and operating considerations increase the radio-link budget complexity, and contribute to degrade the global transmission quality of the systems. It is then necessary to develop non disturbing measurement techniques to avoid interaction between the measurement setup and the measured fields. The purpose of this study is to provide an innovating space characterization technique for TPMS and WAS, to characterize the received power distribution inside and around the vehicle.

1. INTRODUCTION

The high potential of accident prevention by using an intelligent tire system can be clearly seen through the different accident analysis. It has been shown that adverse road conditions, tire defects or their combination play an important role in road accidents. Moreover the decrease in the number of fatalities, provided that the entire car fleet is equipped with intelligent tire systems, could be significantly improved by preventing at least 10% of accidents. This would mean that over 4 000 life's could be saved every year in European countries. As shown in Figure 1(a) TPM System corresponds to a wireless communication between a transmitter module (T_X) fixed in each tire of the car, next called "Wheel unit" and a receiver (R_X) [1]. For the detection of the tire inflation problems, the "Wheel Unit" includes different electronic sensors (temperature, pressure, acceleration ...). The data are collected by the receiver where the different wheel unit frames from each tire are decoded by the control unit. Then a graphical display informs the driver with the pressure and temperature variations.

The car access systems operate the bilateral link LF/RF, by sending the different commands at 315 MHz or 434 MHz [2]. The signals are related to an electronic code featuring the key to control the vehicle (lock/unlock of doors and the trunk release, start of the engine, ...). As shown in Figure 1(b), automotive PAssive Start and Entry system (PASE) module generates a low frequency wake-up message (at 125 kHz) from the car towards the badge, and a RF challenge signal communicates back from the badge to the car at triggering event. The free radio license frequency of 434 MHz is chosen for this study [3].

2. RF SOURCE CHARACTERIZATION FOR TPMS IN NEAR FIELD

The radiofrequency source contains different elements which affect significantly the overall performance of the RF radio-link budget [4, 5]. Therefore, it is compulsorily to take in account the electromagnetic influence of each component influencing the field distribution of a radiofrequency source for TPMS (the antenna pattern, the rim, the tire and the ground). The main elements of the near field probing system are shown in Figure 2.

The full wheel with the ground influence is estimated in near field and reported in Figure 3. The near field measurements are performed for two different positions of the sensor at 90° (middle height of the wheel) and 180° (bottom of the wheel). The field distribution using a metallic plate as a ground put forward a distributed radiating source, whatever the position of the sensor. The presence of metallic walls close to the emitting module alters its radiating pattern. Other studies at higher RF frequencies show a larger number of distributed sources around the wheel unit, according to a distance close to the half-wavelength of the carrier $\lambda/2$.

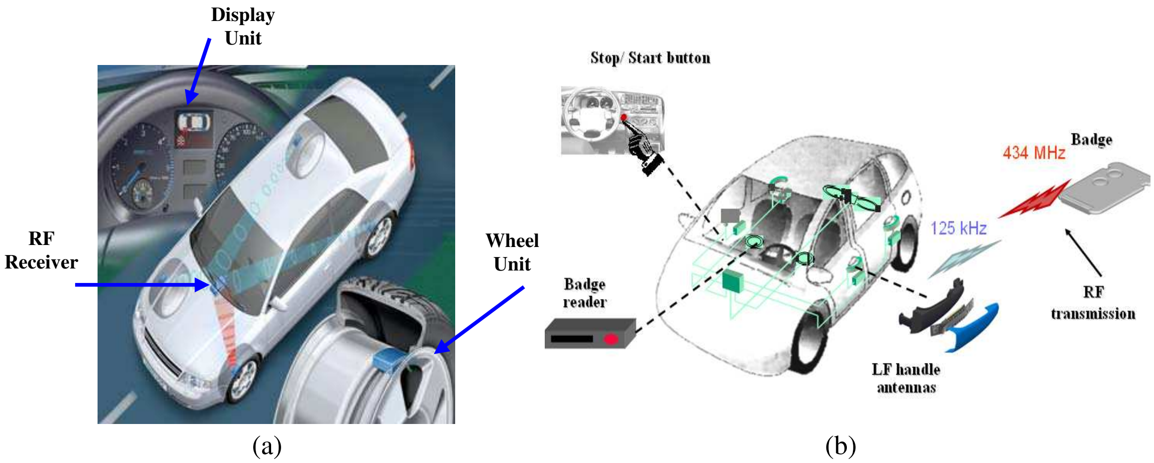


Figure 1: (a) Tire pressure monitoring systems. (b) Wireless car access systems.

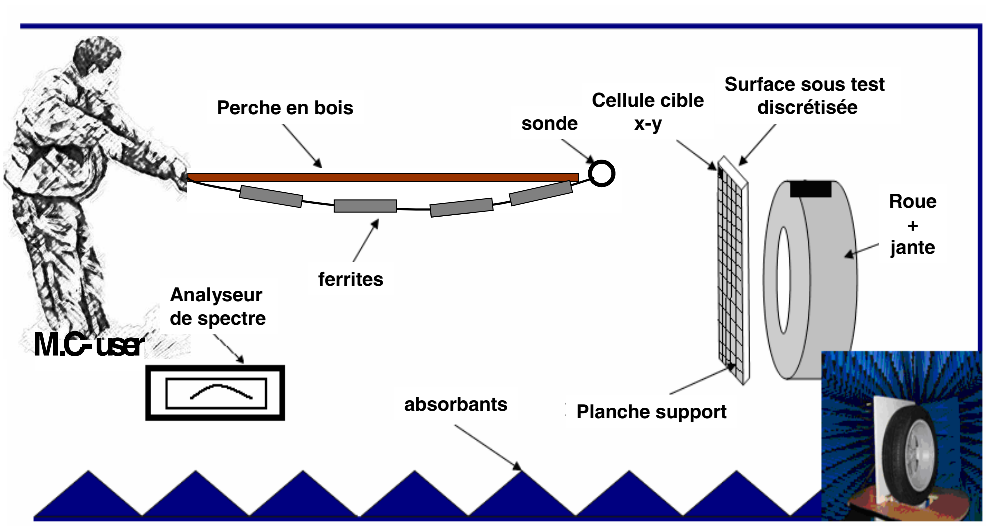


Figure 2: Near field measurement setup.

The measured power in near field varies between -67 dBm and -46 dBm (power dynamic range of 21 dB). Furthermore, the lumped source is divided into four main zones due to the reflections generated by the ground.

3. RADIO COVERAGE CHARACTERIZATION FOR TIRE PRESSURE MONITORING SYSTEMS

An efficient probing system for TPMS must provide a complete and reliable analysis of the radiofrequency channel. Thus, it is necessary to implement non disturbing measurement techniques to avoid harmful effects and to ensure an efficient field probing.

The principal elements constituting the characterization system for TPMS are:

- RF Transceiver system: is activated by LF commands, the wheel unit generates a pure carrier at 433.92 MHz during four minutes to ascertain a stable signal level for one wheel rotation.
- Reception system: is moved in eight different positions inside the sounded zone. For each receiver location we collect the angular power variation over three wheel turns. The receiver computes the RSSI values (Received Signal Strength Indicator).
- Transmission link: the harnesses have harmful effects on the electromagnetic measurements. Hence, a fiber optic link is used to connect the probed zone to the acquisition system without perturbations.

- Encoder angle: attached on a wheel and allows measuring precisely the angular variation of the probed wheel.
- Winch: pulls the probed vehicle with a constant speed (0.017 m/s) to ensure a complete wheel rotation with a high angular resolution (10 RSSI per degree).
- Acquisition system: developed software acquires the measured field from the receiver versus the rotation.
- Post processing: interpolate various measurement data with MATLAB and establish a spatial mapping of the internal coverage versus the angular sensor position and versus the space variation.

The field variation of the channel is measured versus space and wheel rotation angle for TPMS. The high resolution sounding system allows to track the channel variations and to locate fading effects. Optical components are used to get an accurate non disturbing field measurement technique. The RSSI (Received Signal Strength Indicator) collected at each reception cell is emitted through an optical fiber system which constituted with an optical transmitter, a receiver and a fiber link. The sounded zones are then presented versus a two dimensions diagram inside the vehicle close to the expected reception zone. The absence of any metallic cable (as usually used in dedicated setups) prevents from any interaction of the probing technique with the EM distribution close to the receiver or in the channel of propagation. Figure 5 shows a three dimensional representation of the collected radiofrequency power for a given TPMS emitter (wheel unit). The power variations

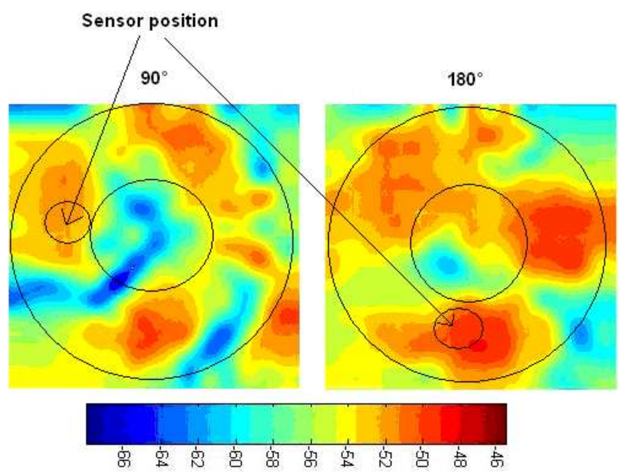


Figure 3: Near ground influence in near field measurement conditions (the sensor location is given by the black circle).

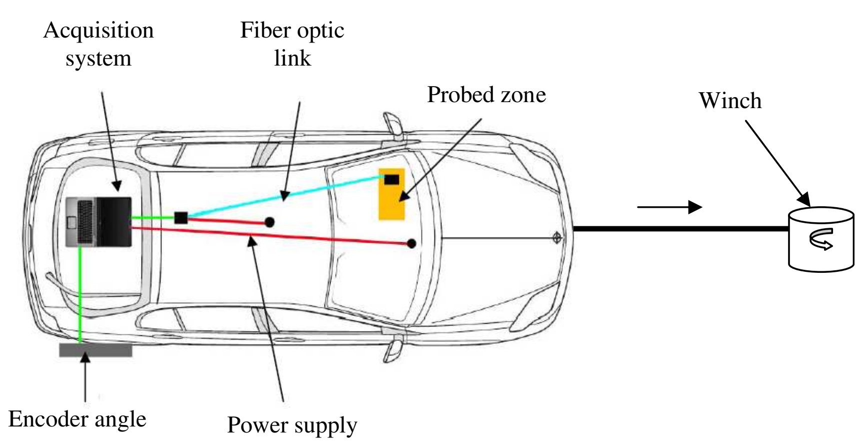


Figure 4: TPMS measurement setup for the inside coverage.

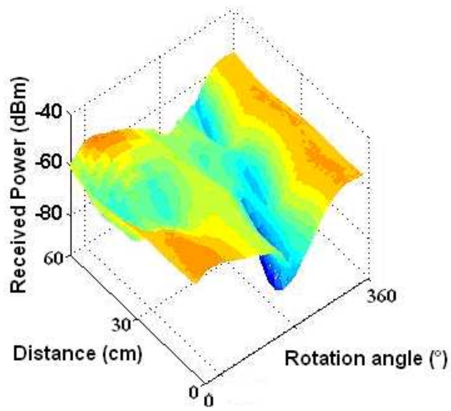


Figure 5: Radiofrequency channel variation in function of the distance and the wheel rotation.

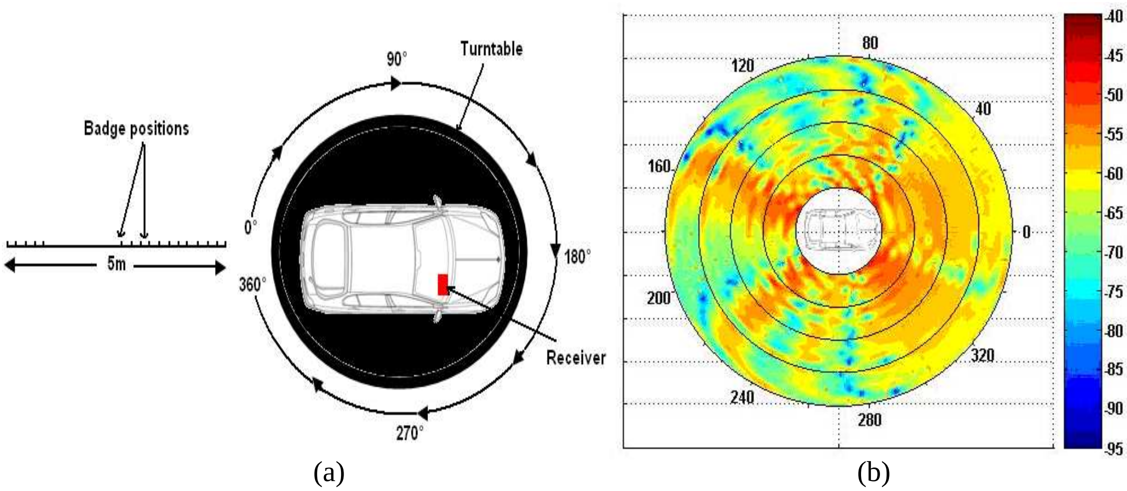


Figure 6: (a) WAS measurement setup for the outside coverage. (b) Measured power around the car for the horizontal polarization. $R_{\min} = 2\,\mathrm{m}$ and $R_{\max} = 6\,\mathrm{m}$.

are measured versus the rotation angle (wheel unit's emitter) and the antenna position (receiver location). The channel profile is unique for each wheel position and vehicle type.

It can be seen in Figure 5 that the multi-paths phenomena generates a stronger field variation. Further, some signal fading (zone with power is lower than the sensitivity threshold: $-90\,\mathrm{dBm}$) can be observed in the measured channel profile.

4. RADIO COVERAGE CHARACTERIZATION FOR WIRELESS ACCESS SYSTEMS

Various environmental parameters affect considerably the signal reception for WAS (Wireless Access Systems):

- Ground: the overall performance of an antenna system is extensively modified by the presence of the ground beneath it (power fall is observed for the horizontal polarization at 1.7 m in our systems).
- Path Loss: the measured path loss value is approximately 30 dB at 16 m from the transmitter.
- Vehicle structure: affects significantly the radiation patterns for both polarizations.
- Human body: decreases the antenna efficiency and deforms its pattern.

An effective measurement setup for the Wireless Access System is used to characterize the RF coverage around the vehicle and to track all space-angular fading. The measurement setup consists of five aim elements (shown in Figure 6(a)).

- Round table: turns the tested vehicle with a constant angular speed.

- Transceiver system: the badge is moved in 28 different positions form 2 to 6 m (the shift of the badge is about $\lambda/4$). For each badge location, we collect the angular power variation (0° to 360°).
- Receiver and post processing: collect the RSSI values and establish a 2D polar mapping of the external coverage.

A no-disturbed RF coverage in a defined zone outside vehicle (2 to 6 m) is shown in Figure 6(b). We notice that the received power varies importantly as function of the badge position around the vehicle.

5. CONCLUSIONS

Full innovating probing techniques were applied to characterize the field distribution in and around the vehicle. The presented procedures are used to model effectively the RF channel behavior inside the vehicle (versus the space location and the angle rotation) for TPMS and outside the vehicle for WAS (versus the badge position). An optical link is used to connect the probed zone to the acquisition system without creating any perturbation, and enabling to get an accurate field value. Based on the high measurement resolution, this technique allows to locate all signal fading effects. The probed zones are then mapped versus two dimensions polar diagram to define the suitable receiver position in order to improve the RF coverage. On the basis of the different experiments versus several car configurations (presence of many passengers, tire manufacturer, rim sizes . . .), we have proposed matched strategies for improving the reliability of PASE and TPMS systems [6]: time diversity, space diversity or polarization diversity techniques have provided improvements up to 12 dB.

REFERENCES

1. Fischer, M., “Tire Pressure Monitoring, design and functionality of direct measuring systems,” Siemens VDO Automotive, Die Bibliothek der Technik, Vol. 243.
2. Chakam, G.-A. and M. Brzeska, “Achieving long range communication in cars,” *EuCAP*, November 11–16, 2007.
3. Chakam, G.-A., M. Brzeska, A. Schafer, M. Weinberger, and A.-S. A. Arango, “Comparison of key-fob antennas for vehicle access system,” *EuCAP*, November 11–16, 2007.
4. Wiese, R. W., H. J. Song, and H. P. Hsu, “RF link budget analysis of a 315 MHz wireless link for automotive tire pressure monitoring system,” *SAE*, 2005-01-1532, 2005.
5. Cheikh, M., J. David, J.-G. Tartarin, S. Kessler, and A. Morin, “RF source characterization for TPMS,” *EuMC*, 176–179, 2009.
6. Cheikh, M., J. David, J.-G. Tartarin, S. Kessler, and A. Morin, “Efficient time diversity evaluation for direct tire pressure monitoring system,” *VTC*, 1–5, 2010.

Performance Comparison of OFDM, MC-CDMA and OFCDM for 4G Wireless Broadband Access and Beyond

Syed M. Zafi S. Shah, A. W. Umrani, and Aftab A. Memon
Institute of Information and Communication Technologies
Mehran University of Engineering and Technology
Jamshoro 76062, Pakistan

Abstract— The performance comparison off the three most feasible multiple access techniques proposed for fourth generation wireless communication systems, i.e., Orthogonal Frequency and Code Division Multiplexing (OFCDM), Orthogonal Frequency Division Multiplexing (OFDM), and Multi-Carrier Code Division Multiple Access (MC-CDMA) is evaluated. The objective is to find the most suitable technique for implementation in 4G Communication Systems. The modems of OFDM, MC-CDMA and OFCDM have been redesigned and performance is analyzed in terms of probability of error. The results reveal that OFCDM provides the lowest BER for a given SNR.

1. INTRODUCTION

The two main candidates for fourth generation (4G) mobile communication systems are WiMAX 802.16e and Long Term Evolution (LTE) Advanced initially endorsed by ITU-R enabling true broadband services with transmission rates up to 100 Mbps with full mobility and 1 Gbps with limited mobility [1–3]. However, none of these systems support such higher transmission rates. The major hindrance is the weak 'Resource Accessing' techniques used at the physical layer of the incumbent systems, which either rely on FDMA, TDMA and DS-CDMA or combination of them, to provide access to multiple users. There have been several proposals including Orthogonal Frequency Division Multiplexing (OFDM), Multi-Carrier Code Division Multiple-Access (MC-CDMA) and Orthogonal Frequency & Code Division Multiplexing (OFCDM) for the adoption as multiple access techniques [4–6]. MC-CDMA has been thoroughly explained in [4, 6]. The motivation was to combine the advantages of multicarrier transmission with the ability of CDMA to accommodate a greater number of users. Later, a combination of multicarrier transmission with two-dimensional spreading named Orthogonal Frequency & Code Division Multiplexing (OFCDM) is introduced by [5]. The authors in [5] have advocated the use of OFCDM for future 4G mobile communications by providing its basic structure and main functions over OFDM system. However, the authors do not take into account the consideration that the MC-CDMA is also a potential technique for 4G mobile communications. A comparison of OFCDM with MC-CDMA would have been more appropriate. Motivated by their work we compare, in this paper, the performance of OFCDM, OFDM and MC-CDMA systems, keeping the aim to find the suitable technique for implementation in 4G systems. The modems of OFDM, MC-CDMA and OFCDM have been redesigned and the performance is simulated under a practical 6-tap frequency-selective Rayleigh fading channel.

2. SIGNAL STRUCTURE COMPARISON

The structure for an OFDM symbol is shown in Figure 1, where it is assumed that the multicarrier transmission consists of four subcarriers, i.e., $N_c = 4$. Multiple access can be provided to users by either allocating each user a particular subcarrier for transmission, i.e., FDMA, illustrated in Figure 1(a), or by allocating all subcarriers to a user for a particular yet a reasonable amount of time, i.e., TDMA, illustrated in Figure 1(b). The signal structure for an MC-CDMA symbol accommodating a single user is illustrated in Figure 2(a). The MC-CDMA symbol is generated as follows: A block of real-valued BPSK modulated symbol, assuming $\{+1\}$ here from $\{+1, -1\}$, is repeated onto the subcarriers. The data is then spread by a frequency-domain spreading sequence, $\{+1, -1, +1, -1\}$ with the spreading factor $SF_{\text{freq}} = 4$. The number of subcarriers N_c and SF_{freq} have been chosen to be equal, however, $SF_{\text{freq}} < N_c$ is also possible. A key observation in the signal structure of MC-CDMA is that spreading in frequency domain does not require additional bandwidth, which is in contrast to DS-CDMA [4]. Figure 2(b) shows the signal structure of an MC-CDMA symbol that accommodates four users over the same number of subcarriers; by multiplexing the data of each user over the frequency-power axis. Since there is a requirement for a synchronous

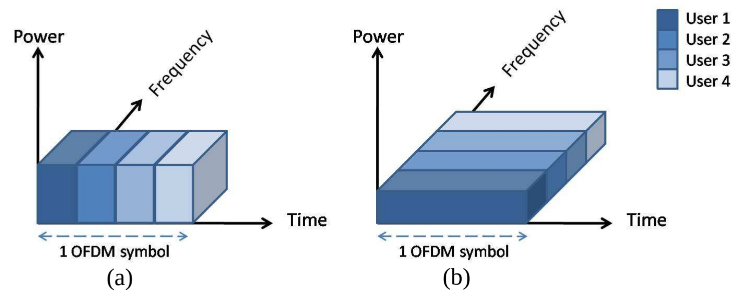


Figure 1: (a) OFDM-TDMA signal structure. (b) OFDM-FDMA signal structure.

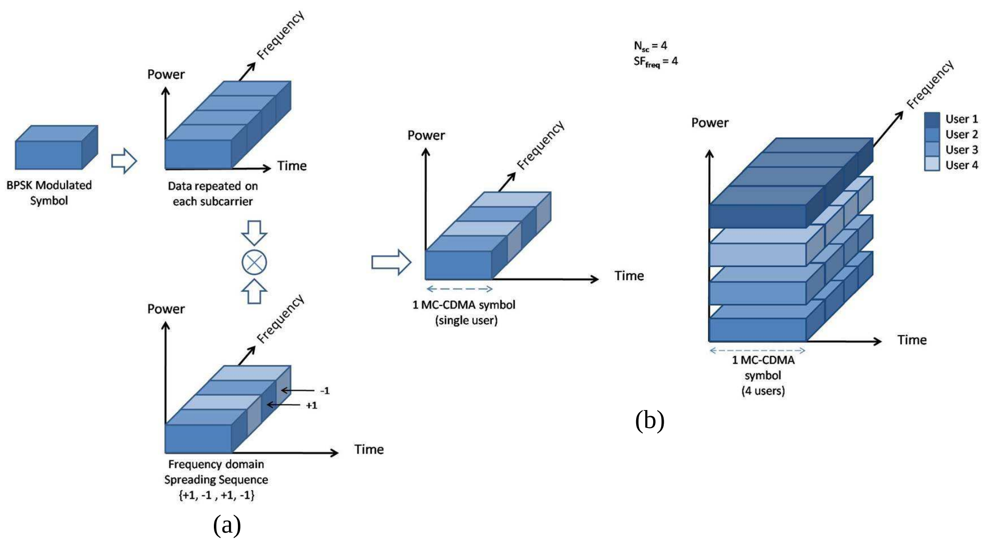


Figure 2: (a) MC-CDMA single user signal structure. (b) MC-CDMA multi-user signal structure.

link for MC-CDMA [4,6], orthogonal sequences like Walsh, OVSF and Zadoff-Chu are preferred over non-orthogonal sequences [4].

An illustration of signal structure of an OFCDM symbol is shown in Figure 3(a). OFCDM can be implemented in two basic ways, i.e., by first implementing DS-CDMA followed by MC-CDMA or vice versa. For our implementation we have chosen the former, with the generation of symbol as follows: A block of BPSK modulated symbol, assuming $\{+1\}$ here from $\{+1, -1\}$ is spread in time domain, over the power-time axes, by a spreading code $\{+1, -1\}$, with a spreading factor $SF_{\text{time}} = 2$. Next, every chip of the spread sequence is repeated onto the four subcarriers. This data is subsequently spread in the frequency-domain by a spreading sequence $\{+1, -1, -1, -1\}$ with the spreading factor $SF_{\text{time}} = 4$ to yield a signal whose power has been spread over two dimensions. OFCDM provides access to users by allocating a block of the 2D symbol. The spreading factor for a user is therefore the product of the factors of both domains, given as $SF = SF_{\text{time}} \times SF_{\text{freq}}$. In Figure 3(b), the signal structure for an OFCDM symbol is shown, which accommodates 8 users with $SF_{\text{time}} = 2$ and $SF_{\text{freq}} = 4$ ($SF = 8$). It can be inferred by comparison between the signal structures of the three techniques that OFCDM provides the best utilization of physical layer resources by accommodating a greater number of users. Another significant revelation is that an OFCDM symbol can be downgraded to either an MC-CDMA symbol by using $SF_{\text{time}} = 1$ and even to an OFDM symbol by using $SF_{\text{time}} = 1$ and $SF_{\text{freq}} = 1$. For comparison between OFDM and OFCDMA the reader is advised to [6, Figure 1].

3. SYSTEM COMPARISON

The transmitter structure for OFDM is shown in Figure 4(a), where user data, as a serial stream of bits, is modulated by a suitable digital modulation technique and then converted from a serial stream in to N_c parallel sub-streams. Each stream will be up-converted by carriers that are orthogonal to each other. This is performed by applying the IDFT algorithm onto the parallel sub-

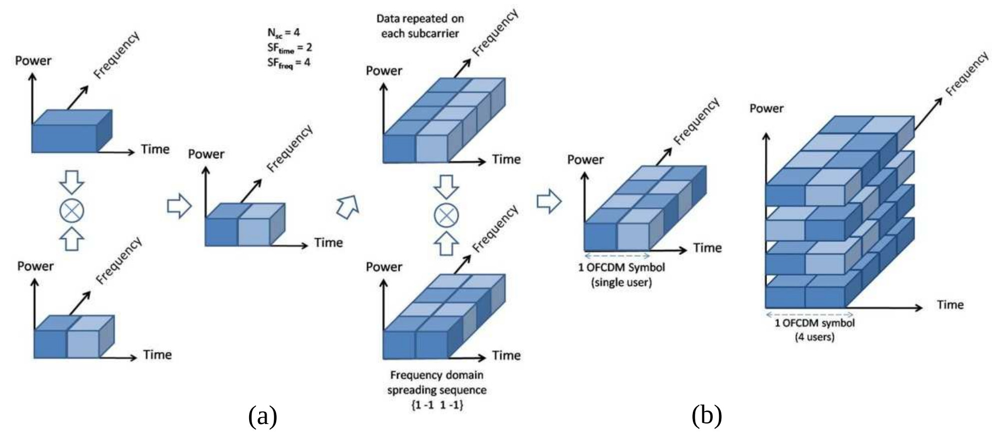


Figure 3: (a) OFCDM single user signal structure. (b) OFCDM multi user signal structure.

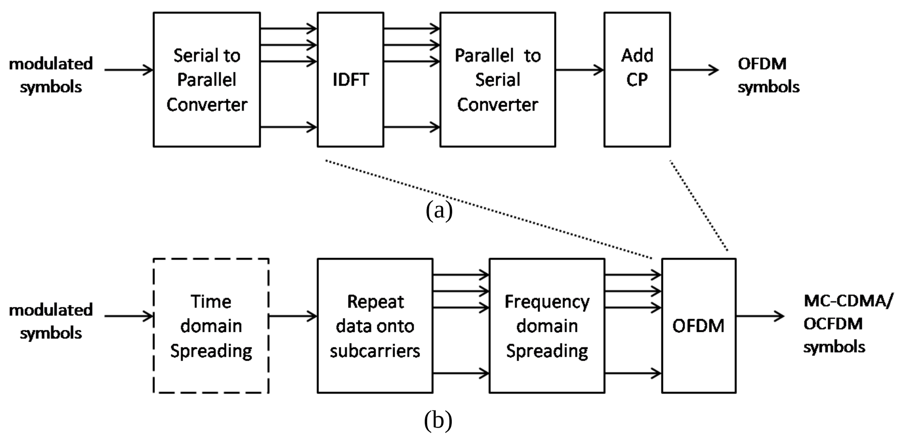


Figure 4: (a) Block diagram of an OFDM transmitter. (b) Block diagram of an MC-CDMA transmitter (solid line); block diagram of an OFCDM transmitter (solid and dotted line).

streams. A part of the signal's tail is copied to its front, therefore prefix, before sending it through the channel. This helps in creating a guard space in the time-domain for delayed multipath signals and prevents successive symbols interfering with each other. The transmitter for MC-CDMA, illustrated in Figure 4(b), is similar to that of OFDM. Data bits are modulated and repeated onto N_c parallel streams. After that, each sub-stream is multiplied by a chip of a unique spreading sequence. This step would perform the objective of multiplexing in the frequency-domain. The transmitter for OFCDM is illustrated in Figure 4(b). Data bits are modulated and spread in the time domain by a sequence of spreading factor SF_{time}. This step would perform the objective of multiplexing in the time domain. The second part of 2D spreading is in the frequency domain, which can be implemented by repeating each chip of the spread sequence onto N_c parallel streams. Similar to MC-CDMA, each stream is multiplied by a chip of a spreading sequence. The final part of the OFCDM transmitter is to implement OFDM transmission described in the previous section.

4. SIMULATION RESULTS

The bit error rate was computed by varying the Energy-of-Bit-to-Noise-ratio (E_b/N_o) over a 6-tap Rayleigh distributed static channel with taps, $[0.89, 0.71, 0.56, 0.45, 0.35, 0.28]$. Perfect channel estimation was assumed for equalization at the receiver so that the BER performance of the access techniques alone could be observed. The performance comparison for the equalizers has been illustrated in Figure 5, where OFDM, MC-CDMA and OFCDM transceivers have been designed with both types of equalizers. The spreading factor for MC-CDMA and OFCDM was kept constant at 8.

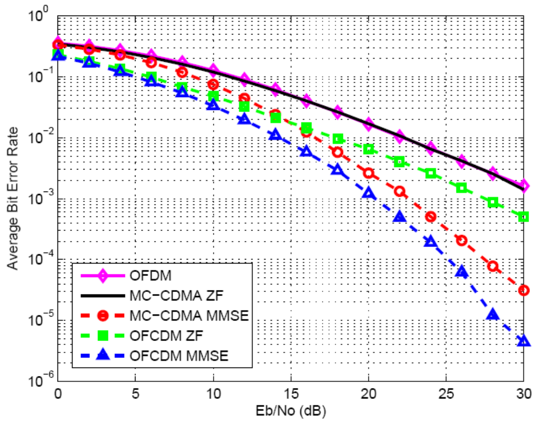


Figure 5: Comparison of frequency-domain Zero Forcing (ZF) and Minimum Mean Square Error (MMSE) equalizers.

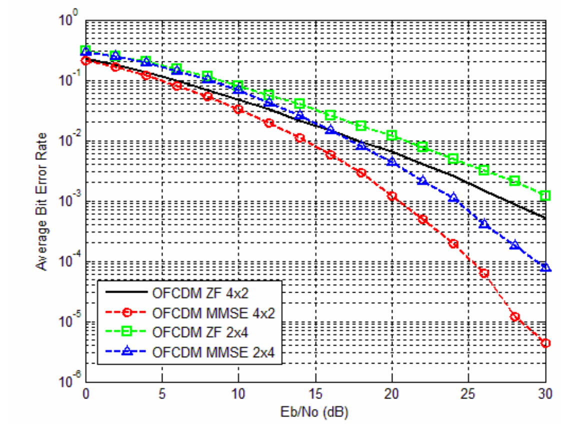


Figure 6: Comparison of the effect of time-domain and frequency-domain spreading factor on bit error rate.

It was observed that MMSE equalizer in general performs better than ZF equalizer. Interestingly, OFCDM with ZF equalizer had a better performance than MC-CDMA MMSE at low values of E_b/N_o, but beyond 12 dB, MC-CDMA MMSE's BER fell drastically with respect to E_b/N_o. However, OFCDM MMSE continues to perform better than MC-CDMA MMSE. With the Zero Forcing equalizer, the performance of MC-CDMA is similar to that of the performance of OFDM.

5. CONCLUSIONS

We elaborated in detail the signal structure of OFDM, MC-CDMA and OFCDM and it was established that OFCDM is in fact, a superset for OFDM and MC-CDMA. The techniques were then compared in terms of bit-error-rate performance, which was estimated over a 6-tap frequency-selective Rayleigh fading channel using the Monte Carlo randomization method. The results demonstrated that OFCDM provides the lowest BER for a given E_b/N_o as compared to the other two techniques. Results also show that the selection of a higher spreading factor for frequency-domain spreading led to an increase in the BER when the channel chosen to be highly frequency-selective and a higher spreading factor for time-domain spreading led to an increase in the requirement for the number of subcarriers.

REFERENCES

1. Report, "Requirements related to technical performance for IMT-Advanced radio interface(s)," ITU-R Report M. 2134, http://www.itu.int/pub/R-REP-M.2134-2008/en, 2008.
2. T-Mobile USA, *4G Speeds From T-Mobile Now Broadly Available in the Northeastern US and Other Major Cities*, T-Mobile USA Press, May 20, 2010.
3. Fu, I. K., et al., "Multicarrier technology for 4G WIMAX system," *IEEE Comm. Mag.*, Vol. 48, No. 8, Aug. 2010.
4. Adachi, F., et al., "Broadband CDMA techniques," *IEEE Wireless Comm.*, 8–18, Apr. 2005.
5. Zhou, Y., T.-S. Ng, et al., "OFCDM: A promising broadband wireless access technique," *IEEE Comm. Mag.*, Vol. 46, No. 3, 38–49, Mar. 2008.
6. Fazel, K. and S. Kaiser, *Multi Carrier and Spread Spectrum: From OFDM and MC-CDMA to LTE and WiMAX*, 2nd Ed., Wiley, 2008.

Influence of the Design of Resistance Welding Equipment on the Evaluation of Magnetic Field Exposure of Operators

R. Doebbelin[1], **T. Winkler**[2], **and A. Lindemann**[1]

[1]Institute of Electric Power Systems, Otto-von-Guericke-University, Magdeburg, Germany
[2]Fraunhofer-Institut für Fabrikbetrieb und-automatisierung, Magdeburg, Germany

Abstract— High-amperage welding currents of resistance welding equipment cause intense magnetic fields. As a consequence, the occurring magnetic field exposure of operators has to be assessed. In the paper existing procedures to evaluate the magnetic field exposure of resistance welders and to prove compliance with given limits are illustrated taking exposure-influencing design aspects into account. The evaluation procedures include the consideration of basic restrictions established for current density values occurring inside a human body and reference levels established for magnetic flux density values and other quantities occurring in the workplace area.

1. INTRODUCTION

Considering the evaluation of occurring exposures of the general public or occupational field exposures to electric, magnetic or electromagnetic fields, relevant regulations have to be taken into account and compliance with given limits has to be ensured to avoid any health hazards for humans. The limits are based on scientifically proven effects of field exposure. Effects, which are only assumed or not proven, do not provide a basis for limits. In the low-frequency range the stimulation of excitable body tissues is a proven effect which may occur as a consequence of exposure to extremely intense electric or magnetic fields [1].

Resistance welding is a widespread joining process in automotive industry and other fields of manufacturing. In resistance welding, e.g., in spot welding processes, welding currents in the kA range are applied resulting in intense magnetic fields in the vicinity of the equipment. Besides welding robots also man-operated welding machines and hand-held welding guns are widely used. Despite the intense magnetic fields in the working area of operators of resistance welding equipment, there is no information about occurrence of stimulation effects and incidence of diseases among resistance welders caused by magnetic field exposure due to statements of occupational physicians. Nevertheless, compliance with the existing field exposure limits has to be guaranteed by the employers. Hence, the procedures which are given in the relevant regulations have to be applied to carry out the inevitable exposure assessment.

2. RELEVANT GUIDELINES AND REGULATIONS

In 1998, the International Commission on Non-Ionizing Radiation Protection (ICNIRP) published its "Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields" [1], which have been regarded as the fundamental regulation since then. In the ICNIRP guidelines frequency-dependent exposure limits are given in form of unperturbed RMS values relating to sinusoidal fields. Limit values of the electric current density occurring in the tissues of a human body exposed to an electric or a magnetic field are defined as basic restrictions. Since it is not possible to measure the current density inside the body, a current-density based field exposure assessment requires calculations based on adequate modeling of the effects of field exposure which occur inside a human body. To enable an alternative evaluation, limit values (named reference levels) for measurable quantities like electric and magnetic field strength and magnetic flux density have been derived from the basic restrictions.

From ICNIRP guidelines, a directive of the European Parliament concerning the protection of employees against risks arising from electromagnetic fields [2] was deduced. In this directive the frequency-dependent exposure limits of ICNIRP guidelines were adopted. The concrete exposure assessment, measurement and calculation shall be ruled by European Standards. The directive should come into force in the EU member states by adoption into national law until 2008. Meanwhile the ICNIRP guidelines are subject of revision (so, there is a new draft concerning exposure to time-varying fields in the frequency range 1 Hz to 100 kHz [3]) and the EU directive shall become effective in an amended version in 2012. Nevertheless, the process of establishing related standards is under way.

According to the EU directive, member states may employ other scientifically-based standards or guidelines for the evaluation of field exposure of workers until harmonized standards from CENELEC cover all relevant situations. The German safety regulation on occupational field exposure BGV B11 "Electromagnetic Fields" [4] represents such a scientifically-based guideline. In this regulation additionally to exposure limits for continuous sinusoidal fields also instructions for assessment of exposure to time-varying fields with pulsed waveforms are given.

3. EXPOSURE ASSESSMENT CONSIDERING REFERENCE LEVELS

Since electric field strength and current density inside a human body cannot be measured, the field exposure can also be evaluated based on the quantities electric and magnetic field strength and magnetic flux density outside the body, as mentioned above. Performing the evaluation procedure, the values of the relevant quantity occurring in the area of the considered workplace have to be compared to the respective reference level. These reference levels are derived from the basic restrictions for worst-case exposure conditions. Consequently, compliance with the reference levels will ensure compliance with the basic restrictions. In the case of low-frequency magnetic fields the magnetic flux density B is mostly used as the basis for evaluation of field exposure.

As a special aspect, it has to be considered that, according to BGV B11, in the case of pulsed magnetic fields the concrete limit values depend on the actual waveform of the field parameters and therefore the respective limit values have to be determined following a procedure which is given in this regulation. Due to the proportionality of the magnetic flux density B to the welding current i_w, the determination of the permissible flux density values can also be performed using the waveform of the welding current instead of the waveform of the magnetic flux density. The results shall be the same in both cases because, considering the waveform, only time parameters and certain values related to the amplitude of the waveform have to be used in the given calculation procedure. Choosing the welding current is more convenient, since the magnetic flux density generally exists in form of three space-vectors corresponding to the axes of a three-dimensional coordinate system.

Exposure assessment considering reference levels shall be illustrated by the example of a resistance welding inverter (Fig. 1).

Figure 2 shows the welding current of a resistance welding inverter with 1 kHz switching frequency of the inverter unit (this frequency is used in widespread designs of resistance welding inverters).

Performing a BGV B11 based exposure evaluation, the single components which exist in the magnetic flux density and in the welding current waveform as well have to be considered separately. In the case of a resistance welding inverter, additionally to the basic shape of the waveform (the d.c. pulse showing exponential slopes) superimposed ripple components occur according to the operation of an inverter power source. So, a 300 Hz component may occur due to imperfect smoothing of the voltage of the d.c. link. In connection with output rectification a ripple component showing the doubled value of the switching frequency appears in the welding-current waveform (a 2 kHz component in the case of 1 kHz inverters). This superimposed ripple proves to be the most relevant component concerning field exposure because, in comparison with the other waveform components, it gives rise to the highest values of minimum distance to the resistance welding equipment [5]. Therefore, the following explanations of the procedure focus on the 2 kHz ripple component.

In BGV B11 different exposure areas are defined. Under controlled conditions and taking into account, that the exposure times are short in most resistance welding applications, the so-called

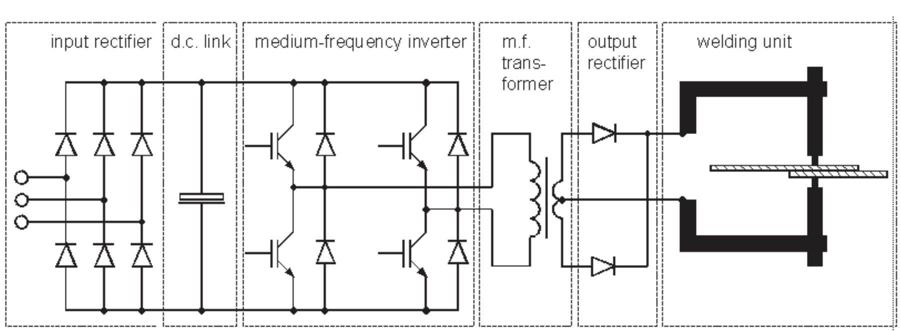


Figure 1: Block diagram of the power unit of a resistance welding inverter.

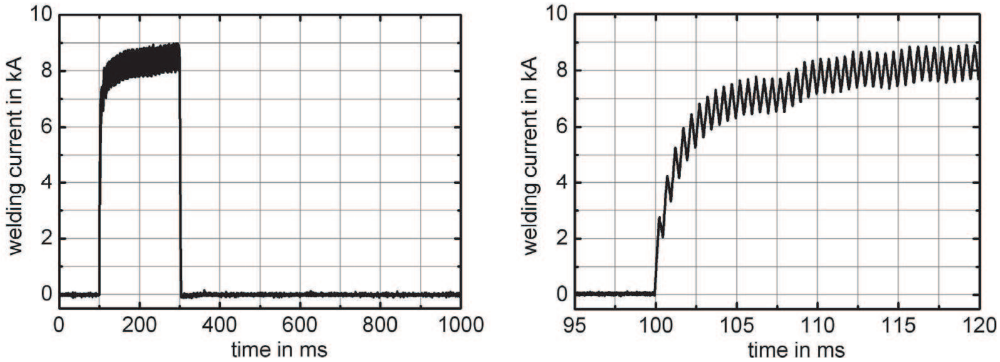


Figure 2: Example of a measured welding current of a resistance welding inverter, welding time 200 ms (right zoomed in).

range of increased exposure can be taken as a basis. With reference to the frequency of field alteration f_p , from diverse formulas attached to certain ranges of the f_p-value the relevant one has to be selected to determine the permissible change rate of magnetic flux density. For $f_p = 1,000$ to 48,000 Hz Equation (1) has to be applied.

$$\frac{dB}{dt}_{mean\ permiss} = 0.72\frac{\mathrm{T}}{\mathrm{s}} \cdot 10^{-3} \cdot f_p \cdot V \tag{1}$$

The magnetic field pulses occurring within a certain integration time T_I (in the case of resistance welding the given maximum of $T_I = 1\,\mathrm{s}$) have to be analyzed. V is a weighting factor to consider the short-time exposure, calculated according to (2),

$$V = \sqrt{\frac{T_I}{\text{exposure time}}} \quad V_{max} = 8 \tag{2}$$

resulting in an increase of permissible values. The permissible peak values of magnetic flux density can be obtained by using Equation (3), in which $\tau_{p\,min}$ is the minimum of rising or falling time of the ripple current.

$$B_{peak\ permiss} = \frac{dB}{dt}_{mean\ permiss} \cdot \tau_{p\,min} \tag{3}$$

To decide whether the exposure situation is permissible or not, the determined limit values have to be compared with the flux density values occurring in the workplace area. In principle, these values can be determined using appropriate field calculation programs but mostly they might be determined by means of measurements. In the case of a fixed position of the operator, the main measuring points are in the regions of head, breast and pelvis of the operator [4]. To consider the possibility of changing positions of the operator to the welding circuit and a possible field exposure of sojourning persons, the distribution of the magnetic flux density in the vicinity of the resistance welding equipment should be analyzed. As an example, in Fig. 3 (a) results measured in a horizontal plane crossing the center of the welding circuit of an inverter-type resistance welding machine are displayed.

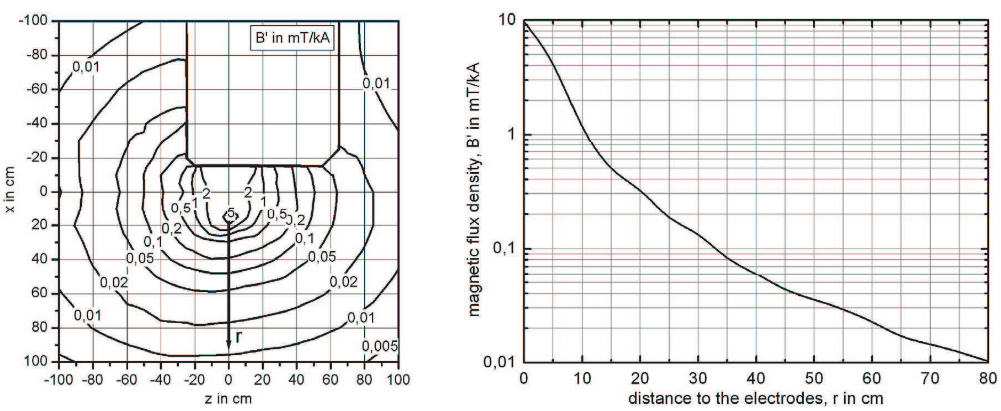
The current-related form of the magnetic flux density B' (in mT/kA) allows to calculate the respective value of the magnetic flux density B by multiplication with the concrete welding current i_w (its RMS value or certain waveform components, e.g., the 2 kHz ripple component).

To capture complete field distributions is very time-consuming. However, often it is adequate to measure only along characteristic axes, e.g., in the direction towards the operator (Fig. 3(b)). If such magnetic flux density-distance-dependence is available, it can be used for comparison with calculated limit values to determine minimum distances. To observe these minimum distances to the resistance welding equipment (particularly to its welding circuit) will ensure compliance with the field exposure limitations.

The concrete shape of the flux density-distance-dependence (advantageously presented in the here proposed form using the current-related magnetic flux density B') is mainly influenced by the size of the output current loop of the machine forming its welding circuit. This is illustrated by the example of a second resistance welding machine (Fig. 4(a)) whose welding unit can be varied enabling different sizes of the output current loop. The power unit of this machine represents a conventional arrangement with a single-phase a.c. power controller connected to the primary winding of the welding transformer. Due to the a.c. welding current, such machines are named a.c. resistance welding machines. Inductance of the output current loop depends on the size of the welding circuit. Therefore, this size influences the time duration parameters of the single current intervals occurring in the waveform of the welding current due to the applied phase-angle control. In the BGV B11 based procedure to determine the permissible magnetic flux density values of pulsed fields, the intervals in which an alteration of field parameters takes place, are considered. Thus, in the case of a.c. resistance welding machines the field exposure limit values are additionally influenced by the respective size of the welding circuit, besides the obvious impact of the output current adjustment (Fig. 4(b)). In the case of inverter-type resistance welding machines this influence can be neglected, because there is no change of the time parameters of the crucial ripple-current waveform since the current is flowing continuously.

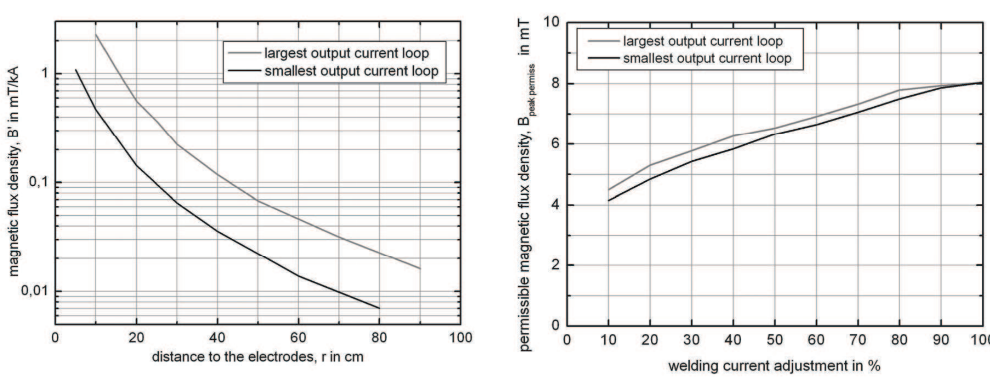
Regarding a.c. resistance welding machines the influence of the size of the welding circuit on the field exposure limit values has to be compared with its impact on the magnetic field distribution in the surrounding of the machine and especially on the magnetic flux density values in the workplace area of the operator. Changing the size of the welding circuit from the smallest to the largest output current loop results in an increase of the permissible magnetic flux density values by less than 10% in the considered example, whereas the occurring magnetic flux density values increase approximately fourfold.

In the examples considered in this paper for both resistance welding machine types a welding current of about 8 kA (RMS) and a welding time of 200 ms gives a permissible minimum distance of



(a) (b)

Figure 3: Current-related magnetic flux density B'. (a) In the surrounding of a selected resistance welding machine, measured in a horizontal plane crossing the center of welding circuit. (b) Dependence on distance along axis r. Size of the welding circuit: throat depth x throat height 345 mm × 445 mm.



(a) (b)

Figure 4: Characteristics of an a.c. resistance welding machine (largest output current loop 570 mm × 415 mm, smallest output current loop 370 mm × 240 mm). (a) Dependence of current-related magnetic flux density B' on distance to the electrodes. (b) Dependence of permissible magnetic flux density on welding current adjustment in the case of welding time 200 ms).

the body of the operator to the welding electrodes in an order of magnitude of about 20 cm. Under different conditions (higher welding currents, longer welding times) the required minimum distances might be remarkably larger resulting in an inadmissibility of occurring exposures of operators of resistance welding machines.

In general, the different characteristics of the waveforms (of magnetic flux density or welding current respectively) which have to be used as the basis for the determination of permissible magnetic flux density values lead to slightly lower permissible distances of the operator to the machine in the case of inverter-type resistance welding machines compared to a.c. machines provided that other relevant parameters (welding current, welding time, size of output current loop) coincide.

4. EXPOSURE ASSESSMENT CONSIDERING BASIC RESTRICTIONS

Due to the conservative derivation of the reference levels from the basic restrictions, there might be numerous cases, especially in resistance welding applications, in which the reference levels are violated but the basic restrictions are met. However, in such a case the current density values in relevant tissues of the body have to be determined and compared with the permissible current density values representing the basic restriction (it can be expected that in the future the current density will be replaced by the electric field strength as the quantity to be considered; this is already intended in [3]). The values of current density and electric field strength in the body can be determined by means of contemporary high-performance computers using numerical field calculation. For the complex calculations an appropriate model of the body including human anatomy and relevant characteristics of the different tissues has to be applied.

In the area of resistance welding such investigations have been carried out so far considering the magnetic field exposure caused by spot welding guns in the case of sinusoidal welding current (50 Hz) [6]. It will be an important issue in the future, to clarify in detail the connection between the magnetic fields caused by resistance welding equipment and the resulting values of the electric quantities in the relevant tissues of the human body considering the various influencing factors, e.g., waveform and amperage of the welding current, size of the welding circuit, distance and posture of the operator.

5. CONCLUSIONS

Concerning the assessment of field exposure at workplaces a complex situation exists considering the variety of national and international guidelines and standards. Evaluation of occupational field exposure shall be basically ruled by an upcoming EU directive. The limit values given in existing regulations are based on scientifically proven effects occurring in a human body resulting from field exposure. At present an evaluation of occupational exposure to pulsed fields can be performed based on the guideline BGV B11. Applying the given procedure with special emphasis on magnetic field alteration aspects to resistance welding machines results in certain influences of the electrical and constructional design of the machines on the required minimum distance of the operator.

Generally, it has to be stated that the field exposure limits in existing guidelines (especially in the case of the reference levels) have been derived in a conservative way. Future guidelines should guarantee health and safety of the employees based on the current state of knowledge regarding the connection between the characteristics of field exposure and resulting effects inside the human body considering the typical characteristics of contemporary field-exposure causing industrial equipment, e.g., pulsed fields and non-homogeneous field distributions.

ACKNOWLEDGMENT

The IGF research projects 12.147 B and 14.818 BG of the Research Association for Welding and Allied Processes of the German Welding Society "Forschungsvereinigung Schweißen und verwandte Verfahren des DVS, Aachener Straße 172, 40223 Düsseldorf" were promoted via the Federation of Industrial Research Associations AiF within the scope of the program for promotion of conjoint industrial research (IGF) based on a resolution of the German parliament.

REFERENCES

1. ICNIRP, "Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields," *Health Physics*, Vol. 74, No. 4, 494–522, 1998.
2. "Directive 2004/40/EC of the European Parliament and of the Council on the minimum health and safety requirements regarding the exposure of workers to the risks arising from physical

agents (electromagnetic fields)," *Official Journal of the European Union*, L 184/1, April 29, 2004.
3. ICNIRP Draft, "Guidelines for limiting exposure to time-varying electric and magnetic fields (1 Hz to 100 kHz)," July 29, 2009, http://www.icnirp.org/openELF/ICNIRPConsultationELF-0709.pdf.
4. BGV B11 (VBG 25), *Berufsgenossenschaftliche Vorschrift "Elektromagnetische Felder" (Regulations for the Prevention of Industrial Accidents)*, Edited by the German Professional Association of Precision Mechanics and Electrical Engineering (BGFE), Cologne, Germany, June 2001.
5. Mecke, H., R. Doebbelin, T. Winkler, T. K. Czarnecki, R. Winkler, and U. Gaertner, "EMVU-relevante Feldemissionen von Widerstandsschweißmaschinen (Field emissions of resistance welding machines relevant concerning environmental electromagnetic compatibility)," Final Report of Research Project AiF 12.147 B, Magdeburg, Duisburg, Germany, 2002.
6. Boerner, F., "Elektromagnetische Felder an handgeführten Punktschweißzangen (Electromagnetic fields of hand-operated spot welding guns)," BGIA Report 2/2009, Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung, Sankt Augustin, Germany, 2009.

Matrix Converter Commutation Time Reduction

J. Lettl, L. Linhart, and J. Bauer
Department of Electric Drives and Traction, Faculty of Electrical Engineering
Czech Technical University in Prague, Czech Republic

Abstract— The reduction of matrix converter commutation time can be obtained by means of number of steps decreasing. For this purpose the two-step voltage oriented method was developed. But some limitations of this method occurred, so the solution consisted in combining with four-step method that is naturally two times longer. The new developed two-step current oriented method can overtake the function of four-step one and fix the disadvantage of voltage oriented algorithm. Unfortunately both two-step methods are inapplicable separately.

1. INTRODUCTION

Commutation process means conventionally (in matrix converter terminology) switch-over of one output phase from one input phase to another input phase [1]. Commutation strategy resolves the process of transition between two steady switching states. This transition is performed through the sequence of some steps whose number is determined by the commutation strategy type. One commutation step involves either the switch-off or switch-on of the IGBT and its duration corresponds to the generated deadband period. It is possible to realize the commutation in 1, 2, 3 or 4 steps. The reduction of the number of steps yields decreasing of commutation time and increasing of commutation algorithm complexity [3, 4].

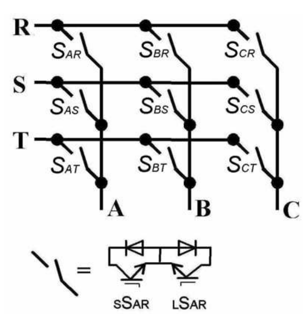


Figure 1: Matrix converter topology.

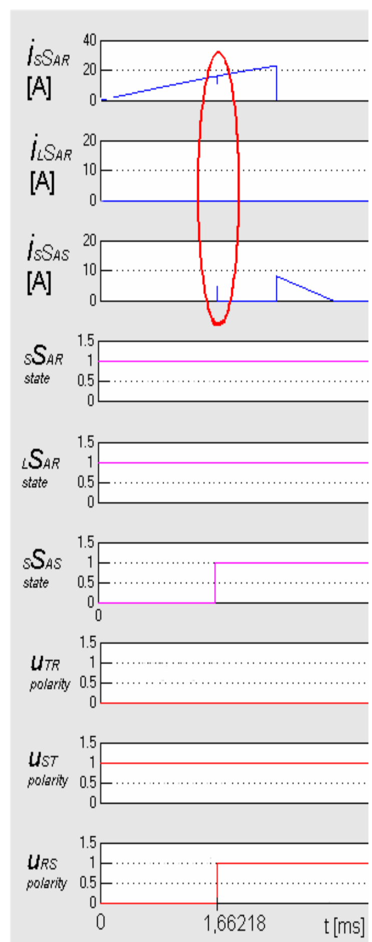


Figure 2: Parallel currents issue.

Figure 1 shows topology of 3×3 matrix converter. For simplification bi-directional switch (that is able to conduct currents and blocking voltage of both polarities) is replaced by imaginary contact.

In following passages IGBTs are marked in agreement with established convention. Prefixes "S" and "L" mean IGBT location — source or load side. Suffixes mean which phases are connected through this bi-directional switch. e.g., "${}_LS_{AR}$" is IGBT between phases A and R on the load side.

Best type of switching minimizes the number of steps to the lowest possible value — which is one [BRNO]. For this purpose it is necessary to control both the input phase-to-phase voltage and the output current. The prearranged state for next commutation should be changed when any of these control variables is changed. It is the most difficult method of commutation. It is based on the principle that the output phase is always connected with one, two or three input phases. However, only one transistor is switched-on in the frame of each connected input phase. The polarity of input phase-to-phase voltages determines the number of input phases (and the number of transistor as well) that are connected. This connection is always realized only through transistors on the same side — either source side, or load side. The direction of the current determines this side [3–6]. Because this method is the most difficult and requires precise measurement of current (which has not been implemented so far) we instead choose the two-step method [5].

2. TWO-STEP COMMUTATION METHOD

Two-step method was developed to reduce commutation time. Presently, two types of two-step commutation method were developed. The older release was oriented to phase-to-phase voltage as a control variable [7, 9]. Second one, two-step current oriented method, is the newest invention. Both commutation algorithms were developed in VHDL language for FPGA.

Two-step method is much complicated, but commutation time is half compared with four-step method. In case of four-step method only IGBTs of connected input phase are switched-on. It means that only IGBTs ${}_SS_{AR}$ and ${}_LS_{AR}$ are switched-on, if the input phase R is connected to output one. In steady state four-step method does nothing else. Commutation process between two steady states is in case of four-step method based on zigzag principle. IGBTs of connected phase and required phase take turns. IGBT of one phase will be switched-on, after deadband IGBT of another phase will be switched-off, and so on in four steps [3–5].

3. VOLTAGE ORIENTED TWO-STEP METHOD

This method is established on idea that the maximum number of IGBTs has to be switched on. Therefore some IGBTs of unplugged input phases have to be switched on and they create a prearranged state for next commutation. This prearranged state must not cause an interphase short-circuit and output current must not be interrupted [5].

The polarity of input phase-to-phase voltage is chosen as control variable for commutation. It means that the prearranged state is set based on polarity of phase-to-phase voltage between connected phase and other two phases. e.g., if phase R was connected to output, we would monitor polarity of voltage U_{RS} and U_{TR}. If phase R is connected to output and the modulator will require switching to phase S, the following sequence of steps will be performed:

First step: All IGBTs which are placed on the same side (load side or source side) like the switched-on-IGBT on the phase S will be switched off. This IGBT on phase S will be let switched on.

Second step: Second IGBT of phase S will be switched on. The new prearranged state will be set. Now we have to realize that the new prearranged state will be set based on polarity of voltage of another phase than in previous case. Now polarity of voltages U_{ST} and U_{RS} is relevant [5, 11].

4. CURRENT ORIENTED TWO-STEP METHOD

This method is unlike the previous one based on idea that the minimum number of IGBTs has to be switched on. That means only one IGBT. A problem appears when the current decreases under the value for which we are uniquely able to determine the current direction. In this case it is necessary to switch on the second IGBT of the plugged phase as well. After the unique determination of the flowing current direction we switch off the IGBT which is not conductive. As clearly visible, we have to provide additional switching also beyond the intrinsic commutation as well as in the case of previous method. Once again, if phase R is connected to output and the modulator will require switching to phase S, the following sequence of steps will be performed:

First step: IGBT which is placed on the same side (load side or source side) like the switched-on-IGBT on the phase R will be switched on. This IGBT on phase R will be let switched on.

Second step: IGBT on the phase R will be switched off. The phase S is now current-carrying.

It should be noted, the necessary additional switching is more complicated in case of current waviness. Thus the second IGBT has to be switched on earlier than in case of smoothed current [3].

5. TWO-STEP METHOD LIMITATIONS

Since the beginning of development it was clear that the critical region of both commutation methods (voltage and current controlled) will be an interval of control variable change. At first the voltage controlled method seemed to be more progressive, but enough limitations occurred.

Close to polarity change of relevant phase-to-phase voltage some currents peaks occur. One of these peaks is shown in red zoom in Figure 2. Three top rows show currents (blue colour) that flow through IGBTs. Magenta colour represents on/off state of these IGBTs as shown in the three middle rows. The three bottom rows show polarities of line-to-line voltages.

It is obvious that the current seems to be drained out from the linked phase because one peak is complementary to the second peak. It is actually not a short circuit but rather parallel conducting a current through two phases. This parallel current flow must be caused by preliminary diode of prearranged IGBT.

Before the change the phase R is connected and the current flows only through IGBT ${}_SS_{AR}$ and his preliminary diode. After the change the new prearranged state is set — ${}_LS_{AS}$ is switched off and ${}_SS_{AS}$ is switched on. At this moment the parallel current flows through ${}_SS_{AR}$ and his preliminary diode as well. It is evident that this diode is polarized forward. After the polarity change the voltage on this relevant diode is still positive and so the current flow has no limitations. Time to decrease of this voltage to the threshold voltage is about 5 microseconds. For this short time the preliminary diode is ready to conduct and that is hazardous state for us. Thus it is necessary to switch-off all IGBTs of prearrangement in this short interval. At the time of absence current oriented two-step method the commutation request in this critical interval should be bridged by means of the four-step algorithm. Now I can imagine that the two-step current method will be set to compensate this voltage method disadvantage.

Another trouble is the relevant phase-to-phase polarity changing during the second-step deadband time. In this case input phase R is connected to the output and commutation to phase S is required. Commutation process is performed. At the beginning of second step current is commutated to phase S very fast. But close to polarity change, current is recommutated back to phase R because of voltage situation on IGBTs. Actually two commutations were performed within the second step. So deadband of this step have to be doubled and it causes enlarging of whole two-step commutation.

6. CONTROL VARIABLE COMPARISON

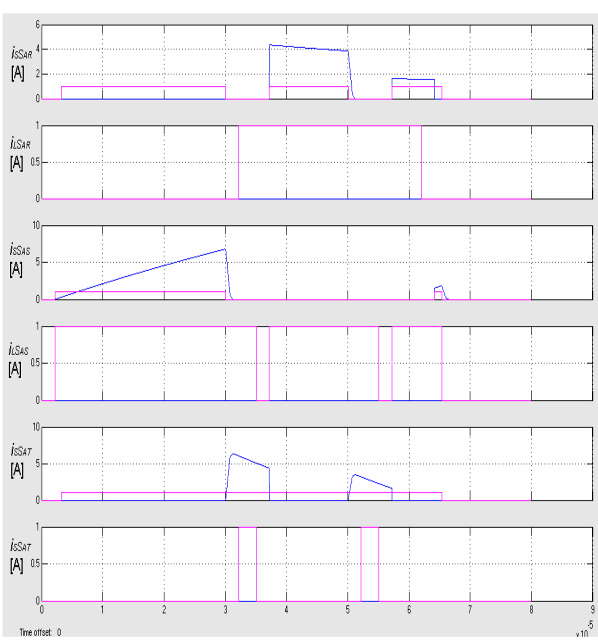


Figure 3: The behaviour of voltage oriented method.

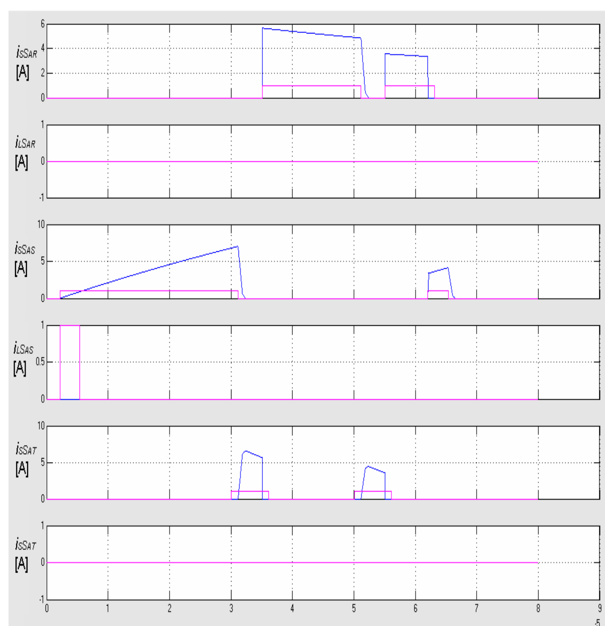


Figure 4: The behaviour of current oriented method.

The results from Matlab-Modelsim are shown in Figure 3 (voltage oriented method) and Figure 4 (current oriented method). At first sight the both curves embody the same situation. In case of voltage oriented method the prearrangement activity is visible. Each commutation leads to change IGBT prearrangement of unplugged input phases compared with current method where only conductive IGBT is switched on. In case of current method two IGBTs are switched on at the beginning of curve. The process of commutation is with absence of current in circuit leads to both plugged phase IGBTs switching on. After current threshold crossing and unique location of conductive IGBT, non-conductive IGBT can be switched off as shown in Figure 4. If the commutation occurs when both IGBTs are switched on we can not perform switching, because it is not known which IGBTs are able to cause an interphase short circuit. In case of small phase currents two-step current method is inapplicable thus the commutation has to be delayed.

From the Figure 3 and Figure 4, it is also obvious that the first commutation (in the time 30 µs) is in case of voltage method performed immediately — current is recommutated without delay. In case of current method the current is still increasing in phase S (by means of ${}_SS_{AS}$). As soon as the ${}_SS_{AS}$ switches off, the current is commutated to the phase T. It means that the phase T is plugged in the frame of second step. The last commutation (in the time 62 µs) is interesting as well. It is evident that last current peak in the phase S is smaller in case of voltage oriented method. It is caused not only by second step taking over of current but also by initial value of current before this commutation. In case of the voltage method this value is cca 2 A, but in second case it is almost 4 A. It is evident that switching over in the frame of the first or second step has effect on magnitudes of currents. The results of confrontation of both methods are not absolutely identical.

The two-step current oriented method has in case of 50 Hz frequency two critical regions of zero crossing in the frame of one output phase. In case of voltage method there are 6 zero crossing, but only 4 are critical for us. Always one phase-to-phase voltage is irrelevant for controlling. It makes the current method more attractive. Increasing of the matrix converter output frequency leads to increasing the current zero crossing at the same time interval and it compensate this advantage. Effectiveness of current method increases depending on increasing of matrix converter switching frequency and decreasing of matrix converter output frequency. So the main disadvantage of current method can be its changefulness (phase-to-phase voltage is more predictable).

7. CONCLUSIONS

Apparently, the best solution may be synergy of both two-step methods. There are two several ways depending on method subordination. The current method can be master and critical region can be bridged by means of voltage method, or it can be contrary. In the first case the number of critical regions is variable, but this number can be in certain circumstances reduced to minimum. Second solution (master method is voltage oriented two-step method) seems to be easier if the better prediction and delay compensation of phase-to-phase voltage are considered. Till now the four-step voltage oriented method was used to bridge the two-step voltage oriented method critical regions. So if the commutation was required in this critical region, the commutation time was two times longer. This idea can be easily replaced by switching off of all prearrangement IGBTs and non conductive IGBT as well. Herewith two-step current oriented method can be applicable to solve the commutation requirement. The knowledge about polarity of current is enough — exact measurement of current is not necessary.

REFERENCES

1. Alesina, A. and M. Venturini, “Analysis and design of optimum-amplitude nine-switch direct ACAC converters,” *IEEE Transactions on Power Electronics*, Vol. 4, No. 1, 101–112, 1989.
2. Casadei, D., G. Serra, A. Tani, and L. Zarri, “Matrix converter modulation strategies: A new general approach based on space-vector representation of the switch state,” *IEEE Transactions on Industrial Electronics*, Vol. 49, No. 2, 370–381, 2002.
3. Posta, P., “Stavový automat modulátoru pro kompaktní maticový menic,” Diploma Thesis, Czech Technical University in Prague, Prague, 2007.
4. Flígl, S., “Matrix converter in hybrid drives,” Ph.D. Thesis, Czech Technical University in Prague, Prague, 2006.
5. Linhart, L., “Modulacní strategie maticového menice,” Diploma Thesis, Czech Technical University in Prague, Prague, 2008.
6. Zarri, L., “Control of matrix converters,” Ph.D. Thesis, University of Bologna, Bologna, 2007.

7. Linhart, L., J. Lettl, J. Bauer, and M. Bednar, "Cosimulation of two-step commutation method of the matrix converter in simulink/modelsim," *TRANSCOM Proceedings*, 59–62, Zilina, 2009.
8. Lettl, J. and S. Fligl, "Simulation and realization of matrix converter drive system," *TRANSCOM Proceedings*, Vol. 8, 69–72, Zilina, 2005.
9. Linhart, L., J. Lettl, and P. Bauer, "Matrix converter two-step commutation method development," *EDPE*, Dubrovnik, Croatia, 2009.
10. Lettl, J. and S. Fligl, "Matrix converter output voltage control with overmodulation," *PIERS Proceedings*, 375–379, Moscow, Russia, August 18–21, 2009.
11. Linhart, L., P. Posta, and J. Lettl, "Comparison of different commutation strategies of matrix converter," *EDS Proceedings*, 479–484, Brno, 2009.

The Use of Prediction to Improve Direct Torque Control

D. Kuzmanovic and J. Lettl
Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

Abstract— In this paper, an improved Direct Torque Control scheme for induction motor drives, employing prediction, is presented. The DTC is an efficient induction machine control method. However, in order to keep a low torque ripple, it requires a high switching frequency ($\geq$ 20 kHz). If the switching frequency is not sufficiently high, the only way to improve the drive's performance is to increase the switching frequency by modifying the used hardware. In this article, the use of prediction is offered as an alternative, cost effective solution, which doesn't require expensive hardware changes.

1. INTRODUCTION

The principle of the DTC was first published by Depenbrock in 1985 [1] (as Direct Self Control — DSC), and Takahashi and Noguchi in 1986 [2] (also known as Switching Table Based DTC, ST-DTC for short). Since then, many DTC schemes have been developed by various researches: schemes employing space vector modulation (SVM) or discreet SVM, schemes using PI, dead-beat or artificial intelligence (artificial neural networks, fuzzy logic sets, fuzzy-neural networks, genetic algorithm based systems and generic algorithm assisted systems) based controllers, etc. [3–6]. However, the most widely used is the ST-DTC and its modifications. In fact, if not specified otherwise, the term direct torque control refers to the ST-DTC, and that is the convention which will be used throughout this article.

In comparison with the Field Oriented Control, where the torque and flux values are controlled by precisely controlling the current vector amplitude and angle relatively to the selected flux vector, involving complicated computations, the DTC approach is very simple — The torque and flux values are not precisely driven to their set values, but are kept within their tolerance bands by switching the most suitable voltage vectors (hysteresis control). The DTC is characterized by fast torque response and a simple control scheme without inner current control loops and coordinate transformations.

The block diagram of the DTC scheme is shown in Figure 1. The estimated torque and stator flux are compared with their reference values, and if they are out of their tolerance bands, an adequate voltage vector is applied to the motor stator windings. The flux and torque are controlled be two-level and three-level hysteresis controllers respectively. The outputs of these controllers (c_Ψ and c_T), together with the flux vector sector ($k_{S\text{-}\Psi}$), are used to select the appropriate voltage vector from *Voltage Vector Selection Table*.

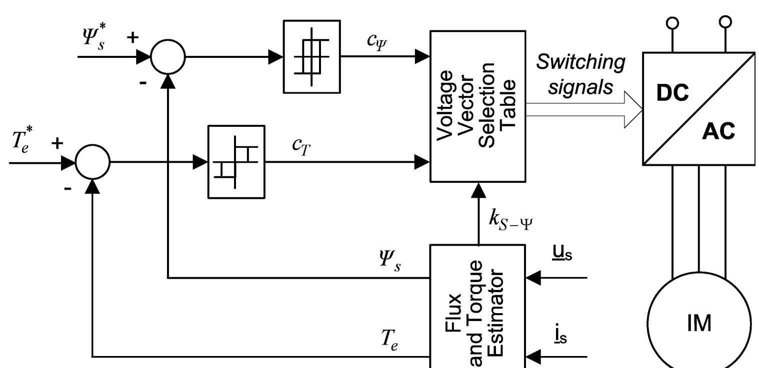


Figure 1: DTC block diagram.

2. DTC USING PREDICTION

To keep a low torque ripple, DTC requires high switching frequency ($\geq$ 20 kHz). If the switching frequency is low, the controller will not be able to keep the torque within the configured tolerance band. To demonstrate that, DTC with sampling period T_{SW} = 150 µs (switching frequency

$f_{SW} = 6.67\,\mathrm{kHz}$) was implemented on a test bed consisting of a 3 kW, 4-pole, 50 Hz cage induction motor with nominal torque 20.2 Nm. For simplicity, two-level hysteresis torque controller was used. The converter output voltage is estimated from the converter input voltage and switching signals. Experiment results are presented in Figure 2. Although the torque tolerance is ±1 Nm, the torque ripple is much higher, owing to the low switching frequency.

A zoomed detail of the graph from Figure 2 is shown in Figure 3. Each tick mark on the horizontal axis represents one control cycle. By analyzing the control cycle 2, it can be seen that the estimated torque value has exceeded its upper limit. That causes the torque controller to issue a command to decrease the torque, which as the result has switching of a zero voltage vector in the next cycle. That should cause the stator flux vector to stop rotating and the torque to decrease. However, the torque continues to increase for one more cycle and starts to decline in cycle 3. The same observation can be made for cycle 12. Similarly, in cycle 9 the estimated torque was below its lower limit, which causes the torque controller to issue a command to increase it. The result should be the switching of an active vector, which should lead to the torque increase. However, the torque continues to decrease for one more cycle.

This behavior can be explained by analyzing the control cycle structure. In Figure 4, the most important parts of the control cycle are depicted. At the beginning of control cycle N, A/D conversion of measured stator currents and converter input voltages takes place. The stator current, along with the estimated converter output voltage is used for the flux and torque estimation, which are then used in the DTC control algorithm for the selection of the reference voltage vector for cycle $N+1$. However, the stator current value will change during the control cycle N under the influence of the voltage vector set in the previous cycle. Consequently, the torque and flux values will change as well.

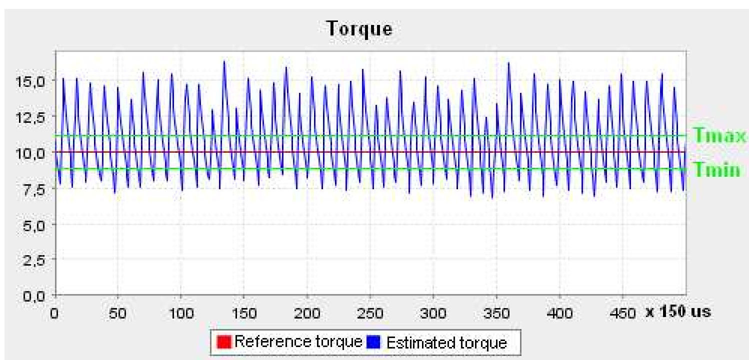


Figure 2: Torque control [Nm] (reference = 10 Nm, tolerance = ±1 Nm); $T_{\max}$ and $T_{\min}$ are tolerance band limits.

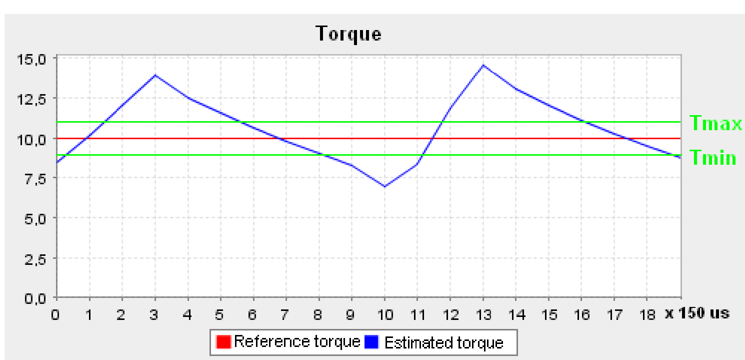


Figure 3: Torque control — zoomed detail.

Now, let's apply these observations to cycle 2 in Figure 3. The torque controller issues a command to decrease the torque, and sets a zero reference voltage vector for cycle 3. However, the torque will continue to increase during cycle 2 under the influenced of the active vector set in cycle 1 and generated over cycle 2, and will start to decline in cycle 3, i.e., one cycle later than expected.

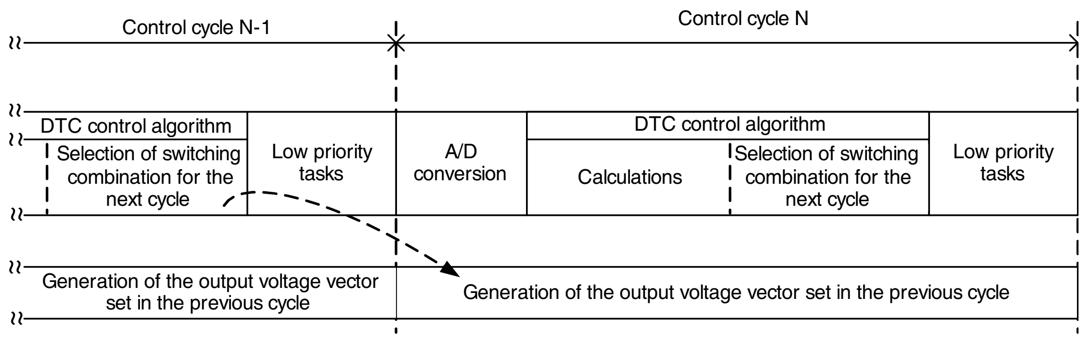


Figure 4: Control cycle structure.

From here it can be seen that the DTC could be significantly improved if the reference voltage for the next cycle could be selected based on the torque value from the end of the current cycle.

In order to improve the torque control, a prediction of the motor state at the end of the control cycle (or at the beginning of the next one), based on the measured current and motor speed at the beginning of the cycle and the reference voltage, was implemented. The prediction is carried out in the following way. From the induction motor flux equations, the stator and rotor current vectors can be expressed as

$$\underline{i}_s = \frac{\underline{\Psi}_s - \frac{L_h}{L_r}\cdot\underline{\Psi}_r}{L_s - \frac{L_h^2}{L_r}} \qquad \underline{i}_r = \frac{\underline{\Psi}_r - \frac{L_h}{L_s}\cdot\underline{\Psi}_s}{L_r - \frac{L_h^2}{L_s}} \tag{1}$$

Substituting (1) into the motor voltage equations (2)

$$\underline{u}_s^k = R_s\cdot\underline{i}_s^k + \frac{d\underline{\Psi}_s^k}{dt} + j\omega_k\cdot\underline{\Psi}_s^k \qquad 0 = \underline{u}_r^k = R_r\cdot\underline{i}_r^k + \frac{d\underline{\Psi}_r^k}{dt} + j\left(\omega_k - \omega\right)\cdot\underline{\Psi}_r^k \tag{2}$$

yields

$$\frac{d\underline{\Psi}_s}{dt} = \underline{u}_s - \frac{R_s}{L_s - \frac{L_h^2}{L_r}}\cdot\underline{\Psi}_s + \frac{R_s\cdot\frac{L_h}{L_r}}{L_s - \frac{L_h^2}{L_r}}\cdot\underline{\Psi}_r \tag{3}$$

$$\frac{d\underline{\Psi}_r}{dt} = \frac{R_r\cdot\frac{L_h}{L_s}}{L_r - \frac{L_h^2}{L_s}}\cdot\underline{\Psi}_s - \left(\frac{R_r}{L_r - \frac{L_h^2}{L_s}} - j\omega\right)\cdot\underline{\Psi}_r \tag{4}$$

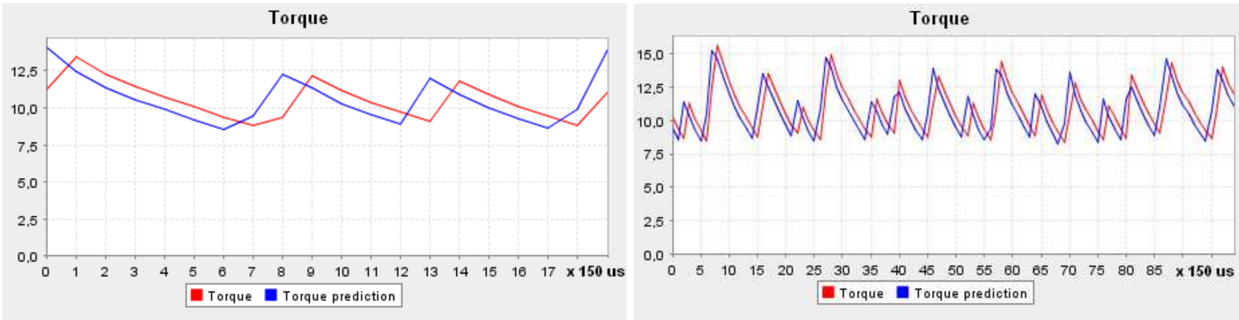
Equations (3) and (4) are rewritten in the discrete-time form

$$\frac{\underline{\Psi}_s^p(N+1) - \underline{\Psi}_s(N)}{T_c} = \underline{u}_s^*(N) - \frac{R_s}{L_s - \frac{L_h^2}{L_r}}\cdot\underline{\Psi}_s^p(N+1) + \frac{R_s\cdot\frac{L_h}{L_r}}{L_s - \frac{L_h^2}{L_r}}\cdot\underline{\Psi}_r^p(N+1) \tag{5}$$

$$\frac{\underline{\Psi}_r^p(N+1) - \underline{\Psi}_r(N)}{T_c} = \frac{R_r\cdot\frac{L_h}{L_s}}{L_r - \frac{L_h^2}{L_s}}\cdot\underline{\Psi}_s^p(N+1) - \left(\frac{R_r}{L_r - \frac{L_h^2}{L_s}} - j\omega(N)\right)\cdot\underline{\Psi}_r^p(N+1) \tag{6}$$

where $\underline{\Psi}_s^p(N+1)$ and $\underline{\Psi}_r^p(N+1)$ are the predicted stator and rotor flux vectors at the beginning of the next cycle and T_c is the control cycle duration.

At the beginning of the cycle N, the stator currents and the motor speed values are obtained by the A/D conversion. From these values and the voltage generated in the previous cycle, the stator



(a) (b)

Figure 5: Torque prediction [Nm].

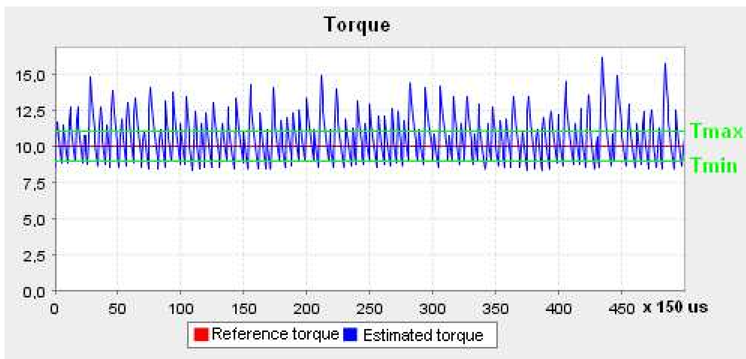


Figure 6: Torque control with prediction [Nm].

and rotor flux values at the beginning of the cycle ($\underline{\Psi}_s(N)$ and $\underline{\Psi}_r(N)$) are calculated using the selected flux estimator. By substituting these values, the speed and the reference voltage $\underline{u}_s^*(N)$ into (5) and (6), and solving these equations, the predicted stator and rotor flux values are obtained. The motor speed used in (6) is measured at the beginning of the cycle, but since it changes very slowly comparing to the voltage and current, it is considered constant during the control cycle. The predicted torque is calculated as

$$T_e^p(N+1) = \frac{2}{3} \cdot \frac{L_h}{\sigma \cdot L_s \cdot L_r} \cdot p \cdot \left(\Psi_{r\alpha}^p(N+1) \cdot \Psi_{s\beta}^p(N+1) - \Psi_{r\beta}^p(N+1) \cdot \Psi_{s\alpha}^p(N+1) \right) \tag{7}$$

3. EXPERIMENTAL RESULTS

The proposed algorithm was implemented on the same test best and under the same conditions as described at the beginning of the previous section. Examples of the torque prediction are presented in Figure 5. Red lines represent the torque values calculated at the beginning of the cycle, and blue lines represent the predicted torque values. From these examples it can be seen that the predicted values precede the estimated ones by one cycle, as expected. In Figure 6, the results of the torque control under the same conditions like in Figure 2 are shown, but now with activated torque prediction. The torque ripple decrease, when the prediction is used, is obvious.

4. CONCLUSIONS

In order to decrease the torque ripple in DTC schemes, when the switching frequency is not sufficiently high, the use of prediction was proposed in this paper. The torque and flux values are first estimated based on the stator current obtained by the A/D conversion at the beginning of the control cycle. By using that state as the basis, the torque and flux values at the end of the control cycle are estimated from the measured motor speed and the reference voltage. The effect of the use of the prediction is similar to the increase of the switching frequency. Drawbacks are more complicated calculations and higher sensitivity to motor electric parameter errors then in the standard DTC, and the fact that it requires a speed encoder. The last drawback disappears in schemes that already include a speed encoder as part of the speed control loop.

REFERENCES

1. Depenbrock, M., "Direkte selbstregelung (DSR) für hochdynamische srehfeldantriebe mit stromrichterspeisung," *ETZ Archive*, 211–218, 1985.
2. Takahashi, I. and T. Noguchi, "A new quick-response and high-efficiency control strategy of an induction motor," *IEEE Transaction on Industry Applications*, 1820–1827, September 1986.
3. Buja, G. S. and M. P. Kazmierkowski, "Direct torque control of PWM inverter-fed AC Motors-A survey," *IEEE Transaction on Industrial Electronics*, 744–757, August 2004.
4. Xue, Y., X. Xu, T. G. Habelter, and D. M. Divan, "A low cost stator flux oriented voltage source variable speed drive," *IEEE Industry Applications Society Annual Meeting*, 410–415, October 1990.
5. Habetler, T. G., F. Profumo, M. Pastorelli, and L. M. Tolbert, "Direct torque control of induction machines using space vectors modulations," *IEEE Transaction on Industry Applications*, 1045–1053, September/October 1992.
6. Toufouti, R., S. Meziane, and H. Benalla, "Direct torque control for induction motor using intelligent techniques," *Journal of Theoretical and Applied Information Technology*, 35–44, 2007.

Measurement and Signal Processing for Electric Drive Control System

P. Brandstetter, P. Bilek, J. Szotkowski, and P. Vaculik
Department of Electronics, VSB-Technical University of Ostrava, Czech Republic

Abstract— Paper is focused on presentation of possibilities which come with application of digital signal processing into area of electric drives control. In the paper, description of control system, structures for measurement and signal processing of analogue and digital signals are presented. Hardware and software realization of these devices are shown, and also description of control structure which is used for digital signal processing. In conclusion, experimental results of used control system are presented.

1. INTRODUCTION

Expansion and miniaturization of microprocessor control systems allows significant reduction of cost, dimensions and also simplify usage of microprocessor control systems into wide area of application. Digital signal processing allows to great improvement of control parameters and increases quality of technological processes. It allows simple remote control of the process and communication with superior control system. The application of control system with digital signal processing will be demonstrated on control system of electric drive.

2. DESCRIPTION OF CONTROL SYSTEM

Structure of control system may be divided to three levels. This division is realized according to type of signals which are present at particular level of control system. The basic structure of control system is shown at Figure 1. The highest level is the digital signal processing (DSP) unit itself. A software signal processing and A/D conversion is realized at this level.

To connect DSP unit to the drive, second level called interface is necessary. At this level, input signal level conversion is realized to enable connection of sensors to the DSP unit. Next function of interface level is power boosting of output signals which allows connection of DSP unit to the driver of electric drive.

The bottom level of control system is drive level. Acquisition of necessary signals from electric drive is provided by measuring elements.

3. MEASUREMENT

It is necessary to require measured signals in enough quality to provide good functionality of the whole control structure. To provide this, several demands must be adhered [3]: enough accuracy, high noise resistance, dynamic range, galvanic insulation between measured signal and sensor output.

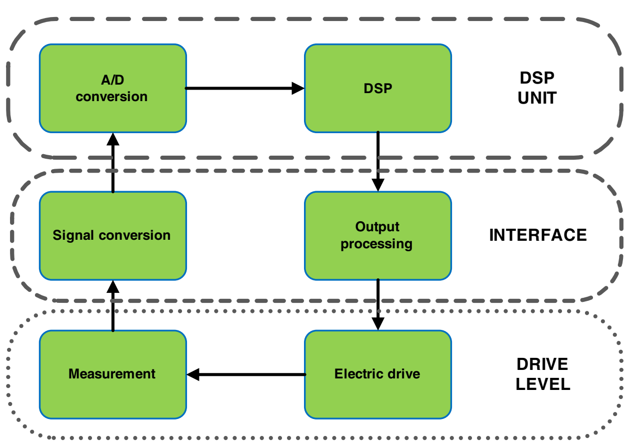


Figure 1: Structure of control system.

3.1. Measurement of Voltages and Currents

To measure voltages and currents, sensors based on Hall Effect were used. These sensors provide galvanic insulation and enough dynamic range of measurement. The output quantity of sensor is current which is converted to voltage signal on measurement resistor R_m. The measurement resistor is situated as close as possible to input of A/D converter. Schematic representation of sensor is shown at Figure 2.

3.2. Speed Measurement

Usually, the incremental encoder is used to measure speed of electric drive. There are two rectangular signals at the output of incremental sensor, shifted by 90° angle. Number of pulses per one round depends on mechanical construction of the incremental encoder. Rotor position change is evaluated by counter, direction of position change is evaluated from mutual shift of signals U_A a U_B. Output signals of incremental encoder are presented at Figure 3. For very low speed range, incremental encoder with sinusoidal output signals may be used. These signals cannot be directly connected with input of DSP, transformation to rectangular signal is necessary. Realization of this transformation is presented at next chapter.

4. SIGNAL LEVEL CONVERSION

Usually, signals from sensors cannot be directly attached to DSP, because level of voltage output differs with possibilities of input gate of DSP. Signal level conversion is necessary in many cases to allow connection between sensor and DSP. Next important function of interface level is over-voltage protection of DSP input gate.

4.1. Processing of Signals from Voltage and Current Sensors

Principle of signal conversion is shown at Figure 4. The signal level conversion may be realized by two ways. The first way is application of operational amplifier. This solution requires high quality low noise devices and many circuits are necessary to process larger quantity of signals. All these conditions increases cost of the interface level. Better solution of signal level conversion is application of reference voltage source. Reference voltage merged to measurement resistor provides the same voltage shift as provides operational amplifier, with less circuitry and cost. Schematics of both variants of signal level conversion block are shown at Figure 5. Input and output signals of signal level conversion block are presented at Figure 6.

4.2. Processing of Signals from Incremental Encoder

Application of incremental encoder with TTL signal output requires no signal conversion; interface level provides only protection of DSP.

The connection of incremental encoder with sinusoidal output to DSP requires signal conversion of output signals to voltage levels, which can be interpreted by counter of DSP (see Figure 7). Signal processing of signal from incremental encoder with sinusoidal output is shown at Figure 8.

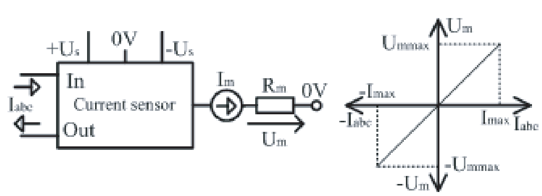


Figure 2: Schematic representation of sensor.

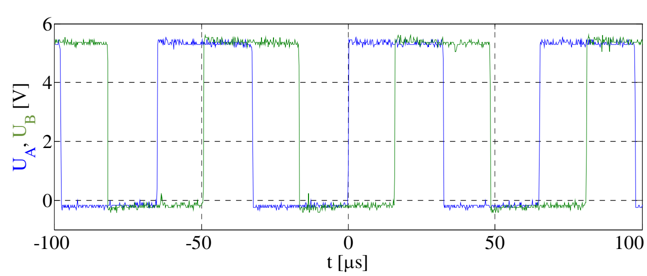


Figure 3: Signals of incremental encoder.

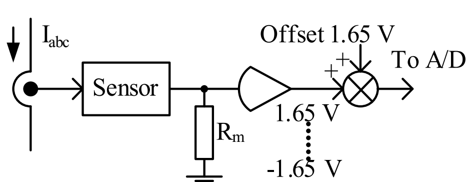


Figure 4: Principle of signal conversion.

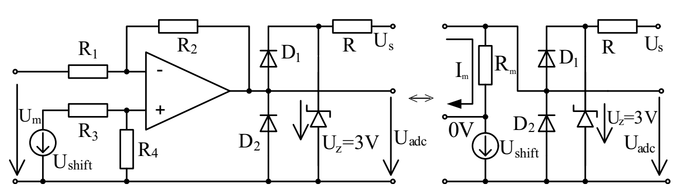


Figure 5: Schematics of signal level conversion blocks.

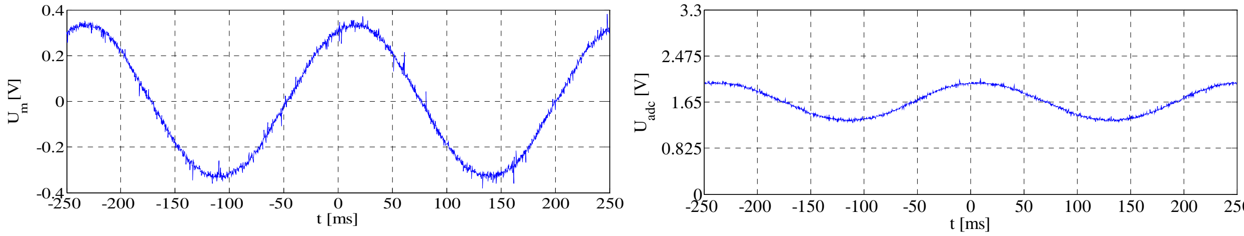


Figure 6: Input signal (left) and output signal (right) at signal level conversion block.

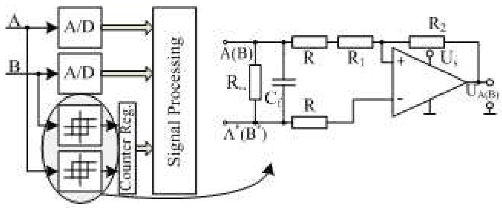


Figure 7: Incremental encoder signal processing block.

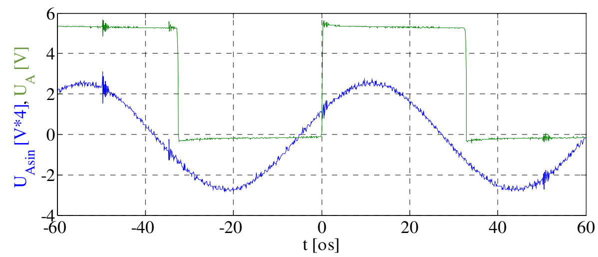


Figure 8: Signals from input and output of incremental encoder signal processing block.

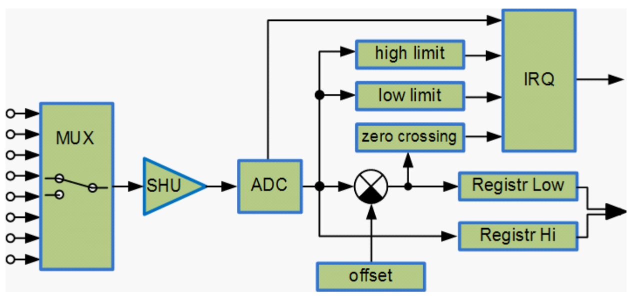


Figure 9: ADC block scheme.

5. DSP SIGNAL PROCESSING

5.1. Analogue Signal Conversion to Digital Form

The analog digital converter (ADC) consists of two converters working independently. Each of them include eight analog inputs and the own sample-hold unit (SHU). The basic parameters of ADC are: input analogue voltage from 0 to 3.3 V, the differential inputs are supported, 12 bit resolution (4096 levels), maximum speed is 1.7 M samples per second.

The signal processing of controlled AC machines it's necessary to convert at least two channels at same time. Two phase currents measuring of delta connected machines is the reason. The worst but possible variant is using two SHU and sequential samples conversion by single ADC. The conversion speed and number of distinguishable levels are the parameters influence the regulation quality. The benefit of the ADC is the minimum and maximum values monitoring, which is allowed to be use like software protection against over current and short-circuit. It is possible to say that used ADC is suitable for this application. The Figure 9 shows how the signal passes through ADC and its processing.

For channels connected with common signal ground there is easy relation between analogue and digital values. Arithmetic-logic unit 56F8037 allows to process two basic data types: integer 8, 16, 32 bits (signed or unsigned) and fraction 16 or 32 bits (only signed).

The both type can be converged for each other. The benefit of fraction type against integer is correct result at multiplying of high numbers. The overflow cannot occur because the product can be less than coefficient only. But on the other side the information loss is expected when the low number are multiplying. The main data type of our digital processing was chosen fraction type, especially owing to the software libraries, because the library's functions were created for fraction type mostly.

5.2. Digital Signal Processing

The software processing gives very interesting possibilities, how quickly, effectively and casual resolve the problems. We know that there is special software for resolving almost all kind of problems. In our case the facility is digital signal controller (DSC) 56F8037 and the software is development environment called Code Warrior.

5.3. Control Structure and the Main Principle

The signal processing was designed for the vector control of AC machines. The following Figure 10 shows control structure for permanent magnet synchronous motor (SMPM). After signal conversion to digital form expressed as integer type we have to convert it to fraction type. It is necessary to measure actual rotor position for creating the system of coordinates, which always has same direction like rotor. The position sensor is a source of two signals with 90 degrees phase shift. The signals are connected with timer through the differentiating amplifier. There is an actual rotor angle saved in timer register after its correct setting. The Clarke transformation (marked 3/2 on the figure) transfers measured phase currents into orthogonal coordination system. Using Park transformation (marked $\alpha\beta/dq$ on the figure) we obtain rotating orthogonal coordinate system oriented like rotor. The controllers were chosen PI type or I type with limitation their outputs. The last part is a modulation [1, 2].

5.4. Output Signals

The last block of signal processing is PWM modulator. It converts three digital values on also three two-level signal with changing their ratio. The obtained signals have voltage from 0 to 3.3 V and they will not damage by TTL logic if the open collector property is set. The driver need higher voltage (MOS levels) so there is a levels converter for readjusting differ voltages.

6. EXPERIMENTAL RESULTS

The control system with the DSC 56F8037 provides digital signal processing. The next Figures 11 and 12 show four stages of digital processing. The SMPM ran up, reversed and stopped on the end.

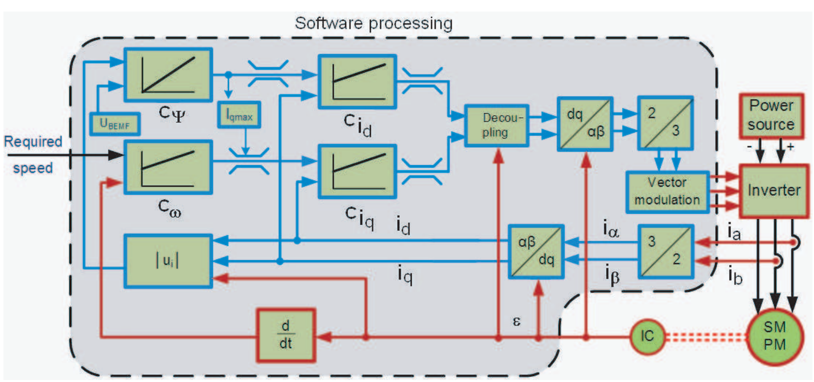


Figure 10: Control structure of AC drive with permanent magnet synchronous motor.

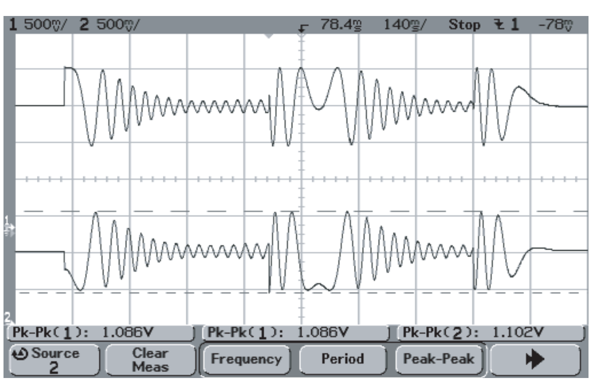


Figure 11: Analogue signals of phase currents i_a, i_b on the ADC inputs (1 V = 16.67 A).

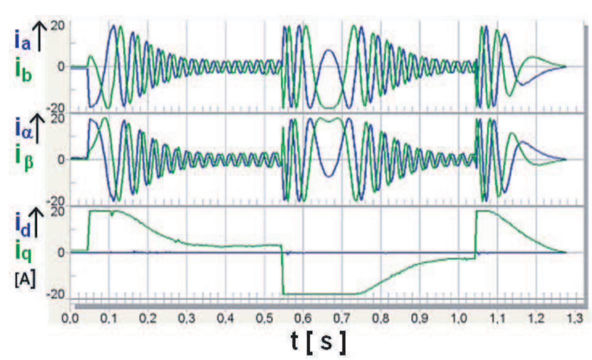


Figure 12: Measured stator currents in three different system of coordinates.

From the top there are measured phases currents at stator system of coordinates, two components and rotor coordinate system. At last one there are supply voltages in other words they are the output from regulators multiplied DC bus voltage.

7. CONCLUSION

The paper deals with digital signal processing for applications in electric drives control. It was shown the techniques with analog and digital signals which are verified on practice application of AC drive control.

ACKNOWLEDGMENT

In the paper there are the results of the project 102/08/0755 which was supported by The Czech Science Foundation (GA CR) and project SP/201098 which was supported by Student Grant Competition of VSB-TUO.

REFERENCES

1. Brandstetter, P., *A.C. Control Drives — Modern Control Methods*, VSB-Technical University of Ostrava, 1999, ISBN 80-7078-668-X.
2. Chlebis, P., P. Moravcik, and P. Simonik, "New method of direct torque control for three-level voltage inverter," *Proceedings of 13th European Conference on Power Electronics and Applications*, Barcelona, Spain, 2009, ISBN 978-1-4244-4432-8.
3. Osmancik, L., M. Polak, P. Simonik, L. Hrdina, P. Skotnica, and P. Palacky, "Digital signal processor TMS320F2812 and its application in electric drives," *Proceedings of International Conference on Applied Electronics*, 129–132, Pilsen, Czech Republic, 2006, ISBN 978-80-7043-442-0.

Sensorless Control of Asynchronous Motor Using Voltage Signal Injection

P. Brandstetter, P. Bilek, J. Szotkowski, and P. Vaculik
Department of Electronics, VSB-Technical University of Ostrava, Czech Republic

Abstract— Controlled induction motor drives without mechanical speed sensors at the motor shaft have the attractions of low cost and high reliability. To replace the sensor, the information on the rotor speed is extracted from measured stator voltages and currents at the motor terminals. Vector controlled drives require estimation of a magnitude and spatial orientation of the stator or rotor magnetic flux. Open loop estimators or closed loop observers are used for this purpose. They differ with respect to accuracy, robustness, and sensitivity against model parameter variations. Dynamic performance and steady-state speed accuracy in the low speed range can be achieved by signal injection, exploiting the anisotropic properties of the machine.

1. INTRODUCTION

Various concepts for controlled high-performance induction motor drives without speed sensor have been developed in the past few years. Eliminating the speed sensor on the motor shaft represents a cost advantage. This combines favorably with increased reliability due to the absence of this mechanical component and its sensor cable. Speed sensorless induction motor drives are well established in those industrial applications in which persistent operation at lower speed is not considered essential [1]. The injection techniques were to be developed for synchronous motor, but it is possible used to induction motor too. For induction motor, two injection methods are used. One of these is "high frequency" (HF) injection and the other is voltage testing impulses injection.

2. HIGH-FREQUENCY MODEL OF INDUCTION MOTOR

The injected signals may be periodic, creating either a high frequency revolving field, or an alternating field in a specific, predetermined spatial direction. Such signals can be referred to as carrier signals. The carrier signals, mostly created by additional components of the stator voltages, get modulated by the actual orientations in space of the machine anisotropies. The carrier frequency components are subsequently extracted from the machine currents. They are demodulated and processed to retrieve the desired information. Instead of injecting a periodic carrier, the high-frequency content of the switched waveforms in a PWM controlled drive system can be exploited for the same purpose. The switching of the inverter produces a perpetual excitation of the transient leakage fields [2]. Their distribution in space is governed by the anisotropies of the machine. Measuring and processing of adequate voltage or current signals permits identifying their spatial orientations.

The equations which describe a model of an induction machine in the stator reference frame $[\alpha, \beta]$ are defined as follows:

$$\mathbf{u}_S^S = R_S \mathbf{i}_S^S + \frac{d\boldsymbol{\psi}_S^S}{dt} \tag{1}$$

$$0 = R_R \mathbf{i}_R^S + \frac{d\boldsymbol{\psi}_R^S}{dt} - j\omega_m \boldsymbol{\psi}_R^S \tag{2}$$

$$\boldsymbol{\psi}_S^S = L_S \mathbf{i}_S^S + L_h \mathbf{i}_R^S \tag{3}$$

$$\boldsymbol{\psi}_R^S = L_h \mathbf{i}_S^S + L_R \mathbf{i}_R^S \tag{4}$$

The quantities $\mathbf{q}^S$ are vectors in stator reference frame:

Stator flux vector $\boldsymbol{\psi}_S^S = \psi_{S\alpha} + j\psi_{S\beta}$, rotor flux vector $\boldsymbol{\psi}_R^S = \psi_{R\alpha} + j\psi_{R\beta}$, stator voltage vector $\boldsymbol{u}_S^S = u_{S\alpha} + ju_{S\beta}$, stator current vector $\boldsymbol{i}_S^S = i_{S\alpha} + ji_{S\beta}$, R_S and R_R are resistances of stator and rotor winding, L_S and L_R are the stator and rotor inductances, L_h is the magnetizing inductance.

Total stator and rotor inductances L_S, L_R (L_h is magnetizing inductance, $L_{S\sigma}$ is stator leakage inductance, $L_{R\sigma}$ is rotor leakage inductance) are defined as follows:

$$L_S = (L_h + L_{S\sigma}) \quad L_R = (L_h + L_{R\sigma}) \quad L_{S\sigma} = L_S - \frac{L_h^2}{L_S L_R} \tag{5}$$

At high frequencies, a voltage drop on the stator and rotor resistances R_S and R_R can be neglected [2]:

$$\mathbf{u}_S^S \cong \frac{d\boldsymbol{\psi}_S^S}{dt} \cong L_S \frac{d\mathbf{i}_S^S}{dt} + L_h \frac{d\mathbf{i}_R^S}{dt} \tag{6}$$

$$0 \cong \frac{d\boldsymbol{\psi}_R^R}{dt} \cong L_h \frac{d\mathbf{i}_S^S}{dt} + L_R \frac{d\mathbf{i}_R^S}{dt} - j\omega_m \left(L_h \mathbf{i}_S^S + L_R \mathbf{i}_R^S\right) \tag{7}$$

For fixed-frequency excitation at an injected frequency ω_i, we can write [2]:

$$0 \cong j\left(\omega_i - \omega_m\right)\left(L_h \mathbf{i}_S^S + L_R \mathbf{i}_R^S\right) \tag{8}$$

$$\left(\omega_i - \omega_m\right) \neq 0 \Rightarrow \left(L_h \mathbf{i}_S^S + L_R \mathbf{i}_R^S\right) = 0 \tag{9}$$

$$\mathbf{i}_R^S \cong \frac{L_h}{L_R} \mathbf{i}_S^S \quad \mathbf{u}_S^S \cong j\omega_i L_{S\sigma} \mathbf{i}_S^S \tag{10}$$

When a rotor-position-dependent saliency is present in the machine, the stator leakage inductance is not constant and instead becomes a function of the rotor position.

In the stator reference system $[\alpha, \beta]$, the stator leakage inductance matrix can be represented as shown in (11).

$$\begin{bmatrix} u_{S\alpha} \\ u_{S\beta} \end{bmatrix} = j\omega_i \left\{ \begin{bmatrix} L_{S\sigma} + \Delta L_{S\sigma}\cos(h\varepsilon) & -\Delta L_{S\sigma}\sin(h\varepsilon) \\ -\Delta L_{S\sigma}\sin(h\varepsilon) & L_{S\sigma} + \Delta L_{S\sigma}\cos(h\varepsilon) \end{bmatrix} \begin{bmatrix} i_{S\alpha} \\ i_{S\beta} \end{bmatrix} \right\} \tag{11}$$

$L_{S\sigma}$ mean value of stator leakage inductance.

$\Delta L_{S\sigma}$ differential stator leakage inductance caused by the saliency.

h harmonic number of the saliency (equal to twice the ratio between the angular speed that the saliency rotates at and the speed of the rotor, and can be positive, negative, or zero).

ε rotor position in electrical degrees.

For a saliency which rotates at the same speed as the rotor, h= 2 [2].

3. VOLTAGE SIGNAL INJECTION

One form of the high-frequency voltage injection signal with constant amplitude U_i and angular frequency ω_i is described as follows [2–5]:

$$\begin{bmatrix} u_{S\alpha\_i} \\ u_{S\beta\_i} \end{bmatrix} = U_i \begin{bmatrix} \cos(\omega_i\, t) \\ -\sin(\omega_i\, t) \end{bmatrix} = U_i\, e^{j\omega_i t} \tag{12}$$

When a salient machine, with a stator transient inductance modeled by Eq. (11), is excited with the carrier-signal voltage excitation (Eq. (12)), carrier signal currents will be induced equal to

$$\mathbf{i}_{S\_i}^S = \begin{bmatrix} i_{S\alpha\_i} \\ i_{S\beta\_i} \end{bmatrix} = I_{i\_p} \begin{bmatrix} \sin(\omega_i t) \\ \cos(\omega_i t) \end{bmatrix} - I_{i\_n} \begin{bmatrix} \sin(h\varepsilon - \omega_i t) \\ \cos(h\varepsilon - \omega_i t) \end{bmatrix} \tag{13}$$

$$\mathbf{i}_{S\_i}^S = -jI_{i\_p} e^{j\omega_i t} + jI_{i\_n} e^{j(h\varepsilon - \omega_i t)} \tag{14}$$

$$I_{i\_p} = L_{S\sigma} \frac{U_i}{\omega_i\left(L_{S\sigma}^2 - \Delta L_{S\sigma}^2\right)} \qquad I_{i\_n} = \Delta L_{S\sigma} \frac{U_i}{\omega_i\left(L_{S\sigma}^2 - \Delta L_{S\sigma}^2\right)} \tag{15}$$

Current with high frequency, which will originate by injection of the additional voltage signal, includes two basic components [2–4]:

Positive component I_{i_p}, which does not contain information of rotor position ε.

Negative component I_{i_n} includes information of position rotor position ε.

The injection is performed in stator reference system $[\alpha, \beta]$ by additional symmetrical three phase sinusoidal voltage signal which is superposed to reference stator voltage (see Fig. 2). Frequency injected signal must be enough high, to was not problem with remove basic frequency in inference current with high frequency, however with regard to the value of switching frequency must be conserved sufficient accuracy.

4. SYNCHRONOUS FILTER

The synchronous filter (see Fig. 1) can be derived from Eq. (14). The band-pass filter removes the fundamental component and leaves HF component. By transforming the HF current to a reference frame synchronous with the injection (positive sequence), the carrier current component becomes DC and the negative sequence then rotating with $2\omega_i$. A high-pass filter completely eliminates the DC component of the current signal and the low cut-off frequency of this high-pass filter ensures that the component at $2\omega_i$ is largely unaltered. Finally, a rotation back attached to the frame synchronous with the negative sequence produced the position signal at baseband. The position angle ε can be obtained by using the arctangent function:

$$\varepsilon = \frac{1}{2}\,\mathrm{arctg}\left(\frac{i_{n\_i\beta}}{i_{n\_i\alpha}}\right) \tag{16}$$

5. SENSORLESS CONTROL OF INDUCTION MOTOR

The Fig. 2 shows sensorless control structure of induction motor drive (ω_{m_ref} is reference speed value). Vector controlled induction motor need two constants as input references: The torque component and the flux component. Two motor phase currents i_{Sa} and i_{Sb} are measured by current sensors and are transformed by Clarke transformation (T3/2) module. The outputs of this projection are designated $i_{S\alpha}$ and $i_{S\beta}$. These two components of the current are the inputs of the Park transformation that gives the current in the $[x, y]$ rotating reference frame. The outputs of the current controllers are components u_{Sx_ref} and u_{Sy_ref}, they are applied to the inverse Park transformation. The outputs of this projection are reference voltage components $u_{S\alpha\_ref}$, $u_{S\beta\_ref}$. The carrier voltages $u_{S\alpha\_i}$ and $u_{S\beta\_i}$ are added to the reference voltage components $u_{S\alpha\_ref}$, $u_{S\beta\_ref}$. An amplitude must be elect so, to injection does not affect motor activity, but at the same time current response on voltage injection must be enough high with the view of next processing. Usually amplitude is at interval 5–20 V. Frequency votes within the range 0.5–2 kHz.

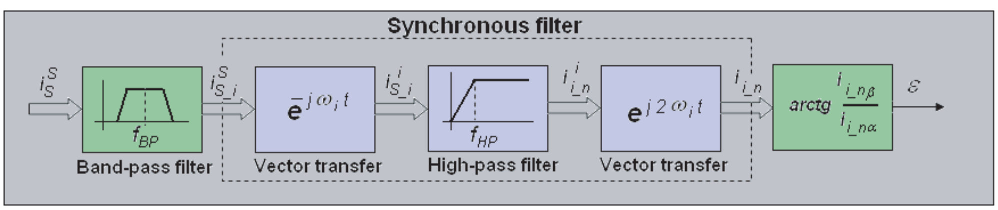


Figure 1: Block of synchronous filters.

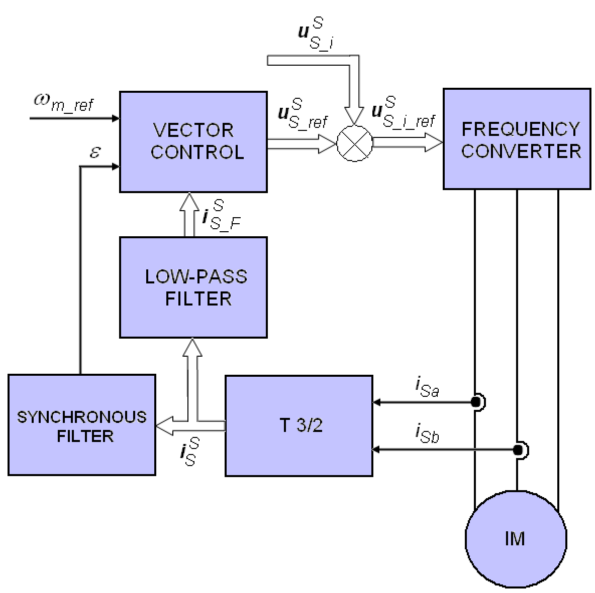


Figure 2: Sensorless control structure of IM.

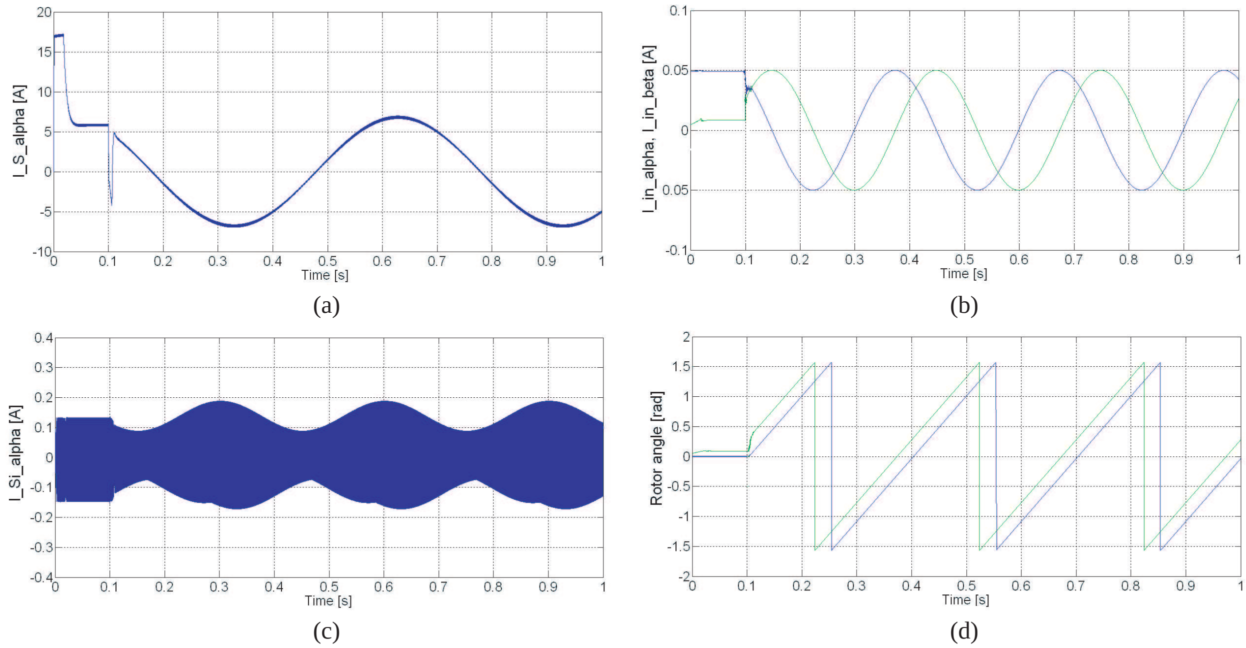


Figure 3: Signals processing in the synchronous filter (a) $i_{S\alpha} = f(t)$, (b) $i_{i\_n\alpha} = f(t)$, $i_{i\_n\beta} = f(t)$, (c) $i_{S\_i\_\alpha} = f(t)$, (d) $\varepsilon = f(t)$.

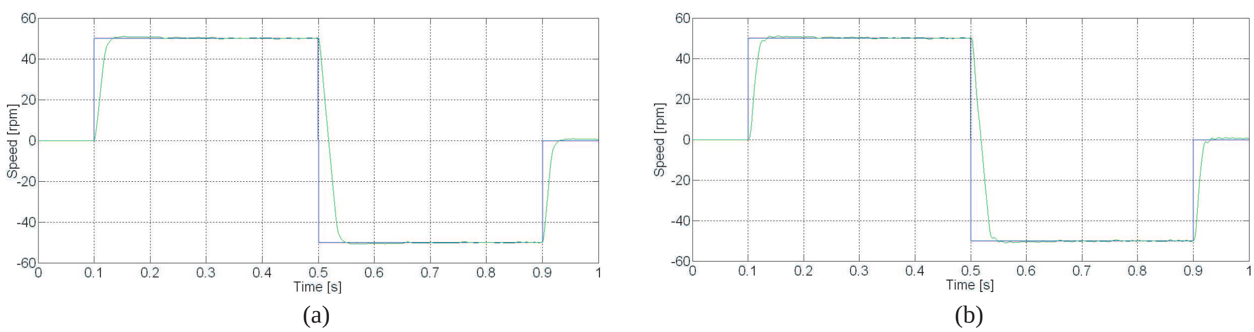


Figure 4: (a) Reference and real speed, (b) reference and estimated speed.

6. SIMULATION RESULTS

The described HF injection method is simulated by MATLAB-SIMULINK on a vector controlled drive which is explained in Fig. 2. All the simulations carried out at ±50 rpm and the estimation of quantities sensed in the open loop whereas drive ran in the close loop. The magnitude of the HF injection voltage is set up at 20 V and $f_i = 2$ kHz. The Fig. 3 shows some important quantities of the described demodulator (synchronous filter) and the resulting position signal from Fig. 1. The Fig. 4 shows the real and estimated speed during reversation.

7. CONCLUSION

The paper has basically presented one of the several non-model based (or self-sensing) estimation methods (concretely HF injection voltage — rotating α-β injection) for the sensorless vector control of the induction motor and main presumption of machine to their application in that methods. Next was described the demodulator which is necessary for this type of control. The signal injection method requires that the induction motor possesses saliency. The frequency converter must produce the additional HF carrier voltage which will be produce additional losses, both in the induction motor and frequency converter. This is naturally perception as a disadvantage. The carrier voltage is added to the fundamental voltage, this reduces the available voltage that can be used for control purposes. If the signal injection method is utilized in the whole speed region, this will reduce the maximum speed that can be reached. Therefore is very suitable use these methods only for the low speed range. For induction motor with small saliency, the voltage distortion due to the dead time of power switching devices can have a significant influence on the speed and position estimation. A suitable dead-time compensation method should be used. On the other hand these methods are interesting for the future research.

ACKNOWLEDGMENT

In the paper, there are the results of the project 102/08/0755 which was supported by The Czech Science Foundation (GA CR) and project SP/201098 which was supported by Student Grant Competition of VSB-Technical University of Ostrava.

REFERENCES

1. Holtz, J., "Sensorless control of induction motor drives," *Proceedings of the IEEE*, Vol. 90, No. 8, 1359–1394, 2002.
2. Degner, M. W. and R. D. Lorenz, "Position estimation in induction machines utilizing rotor bar slot harmonics and carrier-frequency signal injection," *IEEE Transactions on Industry Applications*, Vol. 36, No. 3, 736–742, 2000.
3. Gao, Q., G. Asher, and M. Sumner, "Sensorless position and speed control of induction motors using high-frequency injection and without offline precommissioning," *IEEE Transactions on Industry Applications*, Vol. 54, No. 5, 2474–2481, 2007.
4. Jansen, P. L. and R. D. Lorenz, "Transducerless position and velocity estimation in induction and salient AC machines," *IEEE Transactions on Industry Applications*, Vol. 31, No. 2, 240–247, 1995.
5. Ha, J. I. and S. K. Sul, "Sensorless field-orientation control of an induction machine by high-frequency signal injection," *IEEE Transactions on Industry Applications*, Vol. 35, No. 1, 45–51, 1999.

Comparison of Different Filter Types for Grid Connected Inverter

Jiri Lettl, **Jan Bauer**, and **Libor Linhart**
Department of Electric Drives and Traction, Faculty of Electrical Engineering
Czech Technical University in Prague, Technicka 2, 166 27 Prague 6, Czech Republic

Abstract— The acquisition of the electrical energy from renewable sources is very trendy in these days. That is also why the applications for renewable energy generation undergo rapid development. As the greatest weakness of renewable sources can be assumed the instability and dependence of energy amount, that they deliver. In order to stabilize the energy output and to give it some defined shape and value, the power converter must be connected to the output of the solar panel. For this purpose of application is the most suitable choice voltage source inverter (VSI). In order to suppress or reduce negative effects the filter is connected between the converter and the network. The filter must be designed precisely, because it must have sufficient attenuation at the inverter's switching frequency and it must not bring oscillations to the whole system.

This paper deals with design and simulation of such a filter type. Simulation models of the systems inverter-filter were made.

1. INTRODUCTION

The amount of the energy from renewable energy sources, as solar energy or wind turbine energy, that is delivered to supply network, significantly rises. That is why the grid friendly interface between the supply grid and the energy source is needed. The quality of generated energy is crucial, because non-sinusoidal currents delivered to the grid can cause additional non-sinusoidal voltage drop across the line impedances and therefore increase of the voltage distortions delivered to the load.

Next problem is that the energy supplied by these sources does not have constant value, but fluctuates according to the surrounding conditions (intensity of sun rays, water flow, etc.). These sources are supplemented by additional converters as voltage source inverters (VSI). This converter is then controlled by a modern sophisticated control algorithm such as PWM which ensures nearly sinusoidal current generation. This solution is illustrated in Fig. 1. The power from some renewable energy source is delivered to the DC-link and the output of the inverter is filtered in order to obtain low current distortion.

2. FILTER TOPOLOGIES

The output filter reduces the harmonics in generated current caused by semiconductor switching. There are several types of filters. The simplest variant is filter inductor connected to the inverter's output. But also combinations with capacitors like LC or LCL can be used. These possible topologies are shown in Fig. 2.

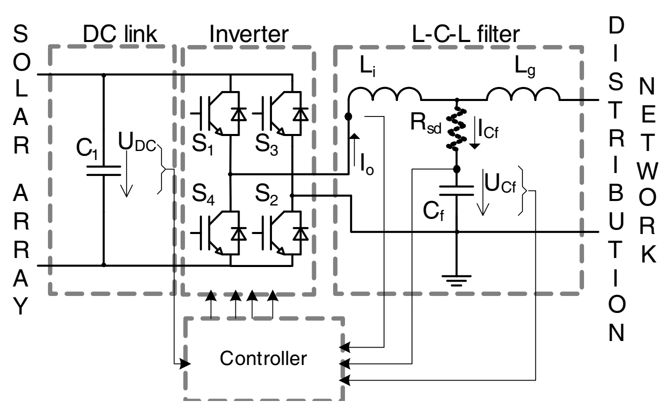


Figure 1: Block diagram of the VSI.

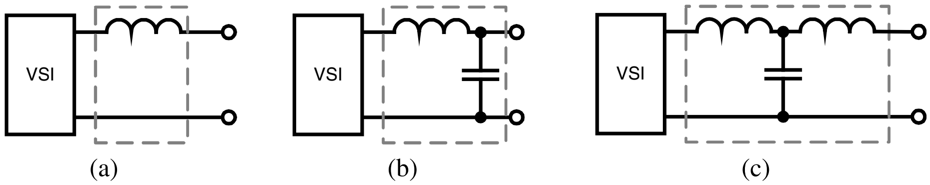


Figure 2: Basic filter topologies.

2.1. L-filter

The L-filter (Fig. 2(a)) is the first order filter with attenuation 20 dB/decade over the whole frequency range. Therefore the application of this filter type is suitable for converters with high switching frequency, where the attenuation is sufficient. On the other side inductance greatly decreases dynamics of the whole system converter-filter. Transfer function of the L-filter is depicted in Fig. 3 as a black dashed line.

2.2. LC-filter

The LC-filter is depicted in Fig. 2(b). It is second order filter and it has better damping behaviours than L-filter. This simple configuration is easy to design and it works mostly without problems. The second order filter provides 12 dB per octave of attenuation after the cut-off frequency f_0, it has no gain before f_0, but it presents a peaking at the resonant frequency f_0. Transfer function of the LC-filter is

$$F(s) = \frac{1}{1 + s \cdot L_F + s^2 \cdot L_F \cdot C_F} \tag{1}$$

It is depicted in Fig. 3 by red colour. In order to suppress the negative behaviours near cut-off frequency the damping circuit is added to the filter. The damping can be either series or parallel. The damping circuit selection influences the transfer function of the filter (Eq. (2) resp. Eq. (3)). The influence is depicted in Fig. 3.

$$F(s) = \frac{1 + s \cdot R_{PD} \cdot C_{PD}}{1 + s \cdot R_{PD} \cdot C_{PD} + s^2 \cdot L_F \cdot (C_F + C_{PD}) + s^3 \cdot L_F \cdot C_F \cdot R_{PD} \cdot C_{PD}} \tag{2}$$

$$F(s) = \frac{R_{SD} + s \cdot (L_F + L_{SD})}{R_{SD} + s \cdot (L_F + L_{SD}) + s^2 \cdot L_F \cdot C_F \cdot R_{SD} + s^3 \cdot L_F \cdot C_F \cdot L_{SD}} \tag{3}$$

The own design of the filter is a compromise between the value of the capacity and inductance. The high capacity has positive effects on the voltage quality. On the other hand higher inductance value is required to achieve demanded cut-off frequency of the filter. Connecting system with this kind of filter to the supply grid, the resonant frequency of the filter becomes dependent on the grid impedance and therefore this filter is not suitable, too.

2.3. LCL-filter

The attenuation of the LCL-filter is 60 dB/decade for frequencies above resonant frequency, therefore lower switching frequency for the converter can be used. It also provides better decoupling between the filter and the grid impedance and lower current ripple across the grid inductor. Therefore LCL-filter fits to our application. Transfer function of the LCL-filter is depicted in Fig. 3.

The LCL filter has good current ripple attenuation even with small inductance values. However it can bring also resonances and unstable states into the system. Therefore the filter must be designed precisely according to the parameters of the specific converter. In the technical literature we can find many articles on the design of the LCL filters [4, 5]. Important parameter of the filter is its cut-off frequency. The cut-off frequency of the filter must be minimally one half of the switching frequency of the converter, because the filter must have enough attenuation in the range of the converter's switching frequency. The cut-off frequency must have a sufficient distance from the grid frequency, too. The cut-off frequency of the LCL filter can be calculated as

$$f_{res} = \frac{1}{2\pi} \sqrt{\frac{L_i + L_g}{L_i L_g C_f}} \tag{4}$$

The LCL filter will be vulnerable to oscillations too and it will magnify frequencies around its cut-off frequency. Therefore the filter is added with damping. The simplest way is to add

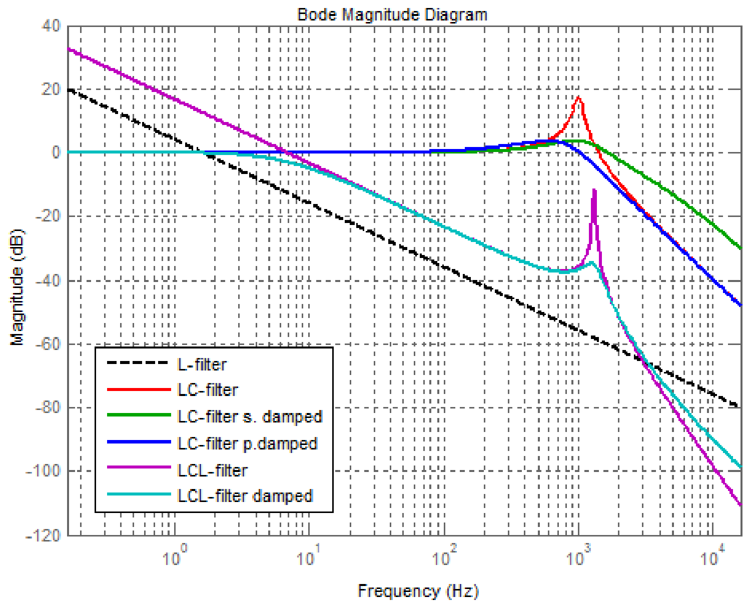


Figure 3: Filter transfer functions.

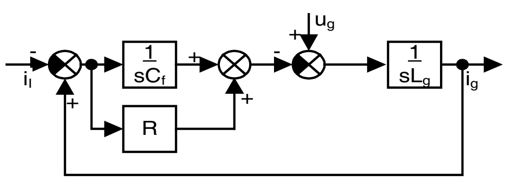


Figure 4: Filter model in operator area.

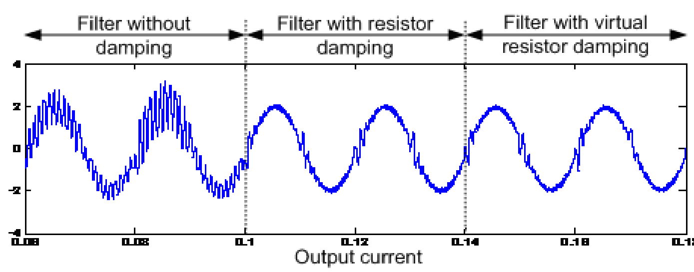


Figure 5: Effects of the damping circuit.

damping resistor. In general there are four possible places where the resistor can be placed — series/parallel to the inverter side inductor or series/parallel to filter capacitor. The variant with resistor connected in series with the filter capacitor has been chosen. The value of the damping resistor can be calculated as

$$R_{sd} = \frac{1}{3\omega_{res} C_f} \tag{5}$$

Transfer function of the filter with damping resistor is depicted in Fig. 3 by cyan colour. The peak near resonant frequency has nearly vanished. This is simple and reliable solution, but it increases the heat losses in the system and it greatly decreases the efficiency of the filter. This problem can be solved by active damping. The filter can be modelled as shown in Fig. 4.

The effect of the damping resistor is clear from Fig. 4. The resistor reduces the voltage across the capacitor by a voltage proportional to the current that flows through it. This can be also done in the control loop. The current through C_f is measured and differentiated by the term sC_fR_{sd}. A real resistor is not used and the calculated value is subtracted from the demanded current. In this way the filter is actively damped with a virtual resistor without losses. The disadvantage of this method is that an additional current sensor is required and the differentiator may bring noise problems because it amplifies high frequency signals.

3. FILTER DESIGN AND SIMULATION RESULTS

For this filter were taken these limitations into account:

1) the cut-off frequency of the filter must be minimally 10 times greater then grid frequency and simultaneously maximally one half of the converter switching frequency
2) The decrease of the power factor caused by the filter capacitance should be lower than 5%

Table 1: Parameters for calculating the filter components.

Grid Voltage (V)	230	Inverter Side Inductance Li (mH)	17.7
Output Power of the Inverter (kVA)	1.5	Grid Side Inductance Lg (mH)	5.7
DC link Voltage (V)	400	Filter Capcitor Cf (μF)	3.45
Grid Frequency (Hz)	50	Damping resistor R (Ω)	11.3
Switching Frequency (Hz)	3000	Cut-off Frequency (Hz)	1300

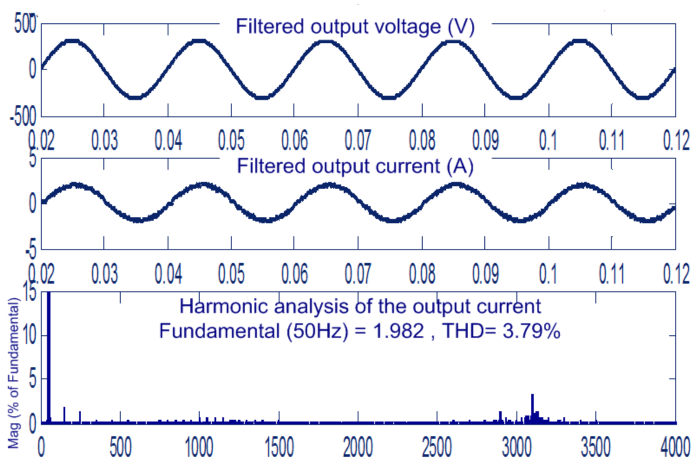


Figure 6: Output of the inverter with filter.

The parameters of the designed filter as well as parameters required for the filter design are summarized in Table 1.

The simulation results of the inverter with the designed filter with active damping connected to its output are in Fig. 6. Fig. 5 shows the effect of the damping resistor and active damping with the virtual resistor.

4. CONCLUSIONS

The output current filter has been designed and simulated. The obtained results seem to be promising. However, we will be able to evaluate the functionality of the filter after the whole system is realized and the filter will be connected to the output of the inverter.

ACKNOWLEDGMENT

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS 10 800630. The research described in the paper was supervised by Prof. J. Lettl, Csc.

REFERENCES

1. Hinz, H., P. Mutschler, and M. Calais,“ Control of a single phase three level voltage source inverter for grid connected photovoltaic systems,” *PCIM*, 1997.
2. Liserre, M., F. Blaabjerg, and S. Hansen, “Design and control of an LCL-filter based three-phase active rectifier,” *Industry Applications Conference, 2001, Thirty-Sixth IAS Annual Meeting, Conference Record of the 2001 IEEE*, Vol. 1, 2001.
3. Araújo, S. V., A. Engler, and B. Sahan, “LCL filter design for grid-connected NPC inverters in offshore wind turbines,” *The 7th International Conference on Power Electronics*, Daegu, Korea, 2007.
4. Dahono, P. A., “A method to damp oscillations on the input LC filter of current-type ac-dc pwm converters by using a virtual resistor,” *Telecommunications Energy Conference INTELEC’03*, 2003.
5. Raoufi, M. and M. T. Lamchich, “Average current mode control of a voltage source inverter connected to the grid: Application to different filter cells,” *Journal of Electrical Engineering*, 2004.

Soft-switched Converter for Ultracapacitors

P. Chlebis, Z. Pfof, A. Havel, and P. Vaculik
Department of Electronics, Faculty of Electrical Engineering and Computer Science
VŠB — Technical University of Ostrava, 17. Listopadu 15, Ostrava-Poruba 70833, Czech Republic

Abstract— The paper deals with design and optimization of a converter for ultracapacitors with zero voltage switching (ZVS) utilization. The converter for ultracapacitors consists of two converters — buck converter and bidirectional half-bridge converter; both of them with ZVS. The converters with ZVS require a different design approach in comparison with common hard-switched converters. Therefore, the paper describes influence of parasitic circuit parameters for the converters with ZVS utilization and their design optimization. The paper provides an analysis of the stages during switching cycle of before-mentioned converters with ZVS as well.

1. INTRODUCTION

The ultracapacitor is a capacitor with large capacitance (up to 5000 Farads) and high efficiency (up to 98%). It leads to the idea of using ultracapacitors as an alternative source to batteries, or for extracting higher efficiency from existing power sources, e.g., fuel cells. The ultracapacitor also has other advantages — It is capable of very fast charges and discharges, a millions of cycles without degradation, extremely low internal resistance or ESR, high output power, etc.. Ultracapacitors cannot yet be used as a primary power source in automotive applications. On the other hand, they can be a good choice as a secondary power source that works as an electric power storage system. They are able to deliver peak power for drive demands for acceleration or can be used for storing regenerative braking energy.

Zero voltage switching (ZVS) is soft-switching technique, which is suitable for high-frequency operation. ZVS can be realizable on the semiconductor switch VT complemented by resonant elements L_r and C_r. A switch with resonant elements creates a so-called resonant switch (Fig. 1). ZVS eliminates turn-on loss due to zero voltage on the capacitor or bypass diode conducting. Within the turn-off time, the capacitor creates a slowdown semiconductor switch voltage build-up, hence turn-off loss is reduced. The output of converters can be regulated by variable on-time or switching frequency control. MOS-FET transistor is the most suitable semiconductor device for ZVS [2].

The converter for ultracapacitors requires a high dynamics that can be achieved by a high switching frequency. But high switching frequency causes a high switching looses, which reduces the efficiency of whole converter. Therefore, the zero voltage switching keeps the high efficiency of a converter with preservation of high switching frequency. Furthermore, the ultracapacitors has a high efficiency (up to 98%), the converter must have a high efficiency too; otherwise the converter reduces the efficiency of a whole unit. The zero voltage switching is used for other advantages as well such as reduction of magnetic components sizes, increase of a power density of a converter and suppression of EMI of a converter [1]. On the other hand, the zero voltage switching has disadvantages as well. These troublesome disadvantages make more difficult design and construction of ZVS converters. Consequently, the design of ZVS converters needs an optimization, which is described in detail in the following chapter 3.

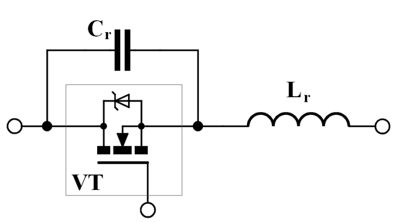


Figure 1: Schematic diagram of a zero-voltage resonant switch.

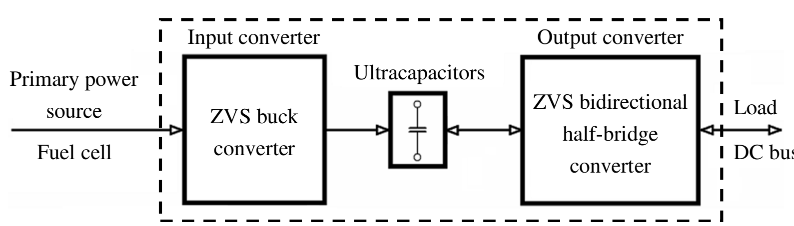


Figure 2: Block diagram of the soft-switched converter for ultracapacitors.

2. A CONVERTER FOR ULTRACAPACITORS

The soft switched converter for ultracapacitors is designed as a converter for delivering the high peak power in dynamic conditions; whereas the input power source will be operate in steady-state; that means no overloading in the course of delivering the high peak power to the load. Thus, as shown in Fig. 2, the converter for ultracapacitors is connected between the primary power source and the inverter, for example. From the figure can be observed, that the converter for ultracapacitors consists of ZVS buck converter (input converter) and ZVS bidirectional half-bridge converter (output converter). The input converter is designed for ultracapacitors charging, ultracapacitors steady-state voltage value maintaining and for delivering the steady-state power to the load as well. The output converter is designed for voltage step-up from ultracapacitors to the output and for voltage step-down from output to the ultracapacitors. A fuel cell is used as a primary power source for the converter for ultracapacitors; therefore the converter is designed in the middle-power range area with low-voltage level and high-current level input.

2.1. Input Converter — ZVS Buck Converter

In this chapter, the behavior of the buck converter with zero voltage switching will be explained. The principle diagram of ZVS buck convertor is shown in Fig. 3(a). Consequently, the time diagram for specification of ZVS buck convertor behavior is shown in Fig. 3(b). In following analysis, every circuit components will be considered as idealized components.

The switch VT is turned off at zero-voltage conditions at time t_0 and the current commutates into the charged capacitor C_r. The C_r is overcharged over the voltage U_1 to the maximum voltage U_{CrM} that can be defined as

$$U_{CrM} = I_O\sqrt{L_r/C_r} + U_1 \tag{1}$$

When the current i_{Lr} changes polarity (t_2), the capacitor C_r is discharged for achieving zero-voltage switching conditions. The off-time for switch VT is invariable and can be determined as

$$T_{off} = t_4 - t_0 = \pi\sqrt{L_rC_r} \tag{2}$$

When the voltage u_{Cr} reaches zero (t_4), bypass diode VD is conducting and the switch VT can be turned on with no losses. The current i_{Lr} changes polarity at t_5 and the switch VT is conducting. The current rises to current value I_O, which is given by load parameters. The current

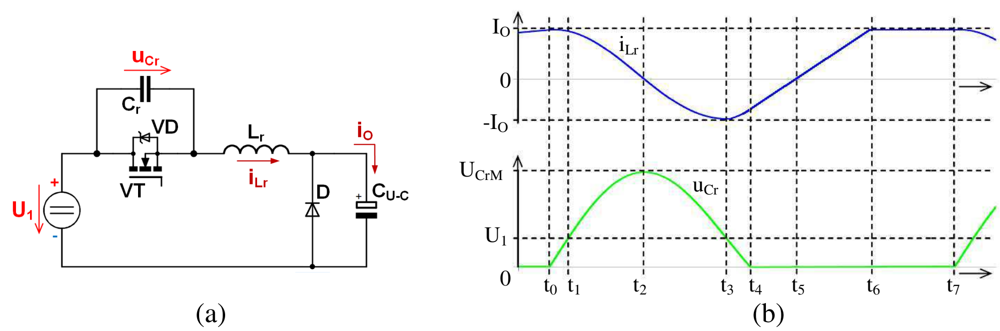


Figure 3: (a) Principle diagram of ZVS buck converter. (b) Time diagram of ZVS buck converter.

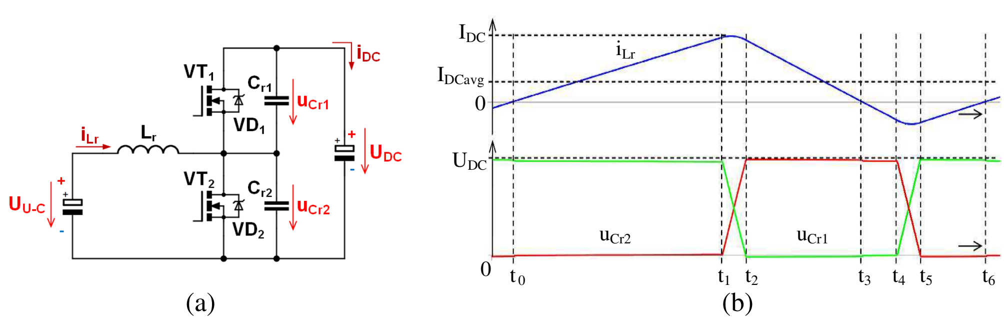


Figure 4: (a) Principle diagram of ZVS bidirectional half-bridge converter for voltage step-up. (b) Time diagram of ZVS bidirectional half-bridge converter for voltage step-up.

i_{Lr} reaches the constant value I_O at t_6 and after what the switch VT is tuned off at t_7, the switching period is repeated.

2.2. Output Converter — ZVS Bidirectional Half-bridge Converter

The behavior of the ZVS bidirectional half-bridge converter for voltage step-up will be explained in this chapter. The converter for voltage step-down is the same for current waveform, if the switches are reverse switched. The principle diagram of the convertor is shown in Fig. 4(a) and the time diagram is shown in Fig. 4(b). Just like in previous chapter, every circuit components will be considered as idealized components in analysis.

The switch VT_2 is turned on at time t_0, the current i_{Lr} rises to the current value I_{DC} according to

$$I_{DC} = U_{U-C} T_{on} / L_r \tag{3}$$

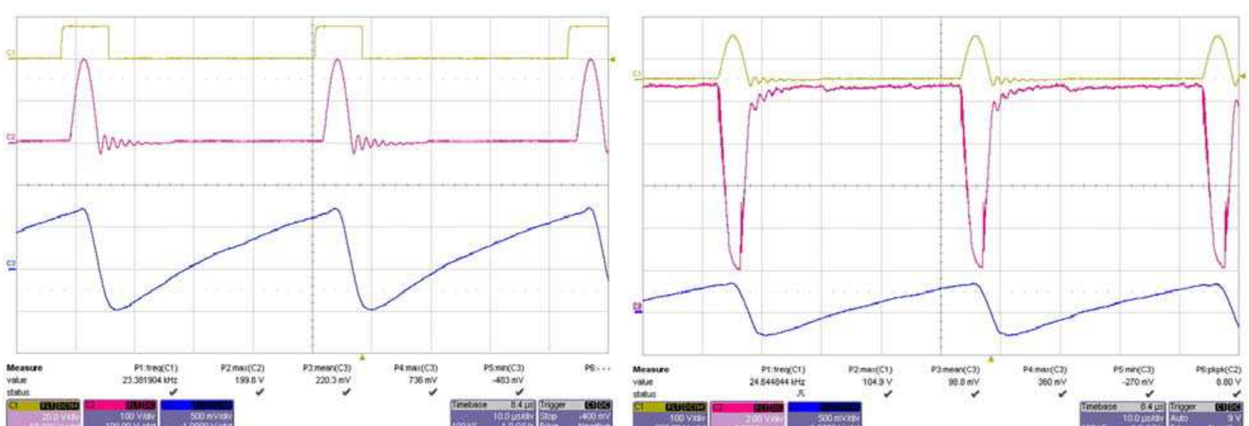
When the current reaches desired current I_{DC} (t_1), VT_2 is turned off at zero-voltage conditions caused by slowdown voltage build-up on capacitor C_{r2}. While C_{r2} is charged by the current i_{Lr}, capacitor C_{r1} is discharged for achieving ZVS conditions. When the voltage u_{Cr1} reaches zero (t_2), bypass diode VD_1 is conducting and VT_1 can be turned on with no losses. The passing current through VD_1 decreases and changes polarity at t_3. The switch VT_1 is now conducting for a short time, which is necessary for ZVS conditions maintenance. The switch VT_1 is turned off at t_4 and capacitor C_{r1} is charged to voltage value U_{DC}, whereas capacitor C_{r2} is discharged to zero voltage. The switch VT_2 can be turned on with no losses at t_5, because the bypass diode VD_2 is conducting. When the current changes polarity at t_6, the switch VT_2 is conducting and switching period starts repeatedly.

3. DESIGN OPTIMIZATION OF ZVS CONVERTERS

The design optimization of above mentioned converters will be described in this chapter. In general soft-switched converters are more exacting in comparison with hard-switched converters, especially for control and regulating microprocessor system. Furthermore, the resonant inductor as well as capacitor for soft-switched converters must have low losses; otherwise the efficiency of a converter is uselessly reduced and consequently can be even lower than efficiency of a hard-switched converter.

In the case of a high current passing through ZVS buck converter, the maximum current value can be a dangerous for a semiconductor switch. The high maximum current value is causes by the steep current rising, because of demand on high switching frequency achievement, thus short conducting time of the switch. From the current waveform of the ZVS buck converter (Fig. 5(a)) can be observed that the average current value of 22 A is quite different in comparison with the maximum current value of 73.6 A. This maximum current value is a value of a cut-off current of the switch.

During the resonant stage, the switch is stressed by a voltage value on the parallel resonant capacitor, which achieving the maximum value according to Eq. (1). The voltage stress is dangerous



(a) (b)

Figure 5: The voltage and current waveforms of a buck converter. (a) C1 — gate pulses, C2 — resonant capacitor voltage, C3 — resonant inductor current (10 mV/A). (b) C1 — resonant capacitor voltage, C2 — ultracapacitors voltage, C3 — resonant inductor current (10 mV/A).

too, because it can easily exceed the maximum drain-source voltage value of a semiconductor switch. The Fig. 5(a) shows that the maximum voltage value reaches to 200 V DC, whereas the input voltage value is 25 V DC.

The influence of parasitic inductances of soft-switched converter circuit can be observed during the resonant cycle, thus during the off-time of the switches. Especially, the parasitic inductance of ultracapacitors causes the change of ultracapacitors voltage value. The Fig. 5(b) shows the voltage waveform across the ultracapacitors. The used ultracapacitors are BOOSTCAP® ultracapacitors with a capacitance of 3 000 F and a voltage of 2.7 V DC; in total seventeen ultracapacitors are connected in series. The parasitic inductance of ultracapacitors was derived by the simulation of the measured waveform. It was found out that self-inductance of the one ultracapacitor cell is 16 nH. There is impact of the parasitic inductance of wires as well. The self-inductance of wires causes similar problem as a parasitic inductance of ultracapacitors. Thus, the change of voltage value across the wire can be observed during the resonant cycle.

The solution to influence of parasitic inductances consists in filter network utilization. Especially, the influence of parasitic inductance of ultracapacitors can be suppressed by appropriately designed LC smoothing filter network. It was found out that the LC filter can suppress the influence of parasitic inductances, but no eliminate. That means the influence of parasitic inductances can be neglected with filter network utilization.

The ZVS bidirectional half-bridge converter needs an optimization of control system for switching algorithm and of output DC bus design as well in comparison with hard-switched converter. The switching algorithm must be optimized, because both of the switches must be switched during one switching period (Fig. 4(b)). Furthermore, the converter needs a power source on the output, which can be realized by a capacitor with appropriately capacitance value. The power source on the converter output is important for achieving a zero voltage switching conditions. As was mentioned before (Fig. 4(b)), the auxiliary switch is switched within the time t_3 and t_4, thus the current direction is from output DC bus to ultracapacitors. Duration of conducting time of auxiliary switch is depending on the capacitance value of resonant capacitors.

4. CONCLUSION

The design and optimization of a converter for ultracapacitors with zero voltage switching was presented. Soft-switched converters have some advantages and also disadvantages, which make more difficult design and construction of these converters. Nowadays, increased demands on controlling microprocessor are insignificant, because of increased computing power of microprocessor [3]. Other demands need some design optimization. Consequently, the ZVS converter for ultracapacitors can works with high switching frequency with no restriction, which commonly make trouble hard-switched converters due to higher switching power losses.

ACKNOWLEDGMENT

The research described in this paper was supervised by Prof. Ing. Petr Chlebiš, CSc., Department of Electronics, VŠB — TU Ostrava and supported by GACR project 102/08/0775: New structures and control algorithms of mobile hybrid systems.

REFERENCES

1. Rashid, M. H., *Power Electronics Handbook*, Academic Press, California, USA, 2001, ISBN 0-12-571650-2.
2. Dudrik, J. and N.-D. Trip, "Soft-switching PS-PWM DC-DC converter for full-load range applications," *IEEE Transactions on Industrial Electronics*, Vol. 57, No. 8, 2010, ISSN 0278-0046.
3. Osmancik, L., M. Polak, P. Simonik, L. Hrdina, P. Skotnica, and P. Palacky, "Digital signal processor TMS320F2812 and its application in electric drives," *Proceedings of International Conference on Applied Electronics*, 129–132, Pilsen, Czech Republic, 2006, ISBN 978-80-7043-442-0.

Applications of Resonant and Soft Switching Converters

V. Sladecek, P. Palacky, T. Pavelek, and P. Hudecek
Department of Electronics, Faculty of Electrical Engineering and Computer Science
VŠB-Technical University of Ostrava, 17, listopadu 15, Ostrava-Poruba 70833, Czech Republic

Abstract— This paper covers the circuit modification of the power part of the inverter with auxiliary resonant poles utilising configuration of switches realised with routinely produced IGBT modules. Covered is also the control optimisation which goal is the minimisation of switching of the auxiliary resonant pole. Presented results were gained on a prototype of an inverter laboratory sample.

1. INTRODUCTION

High frequency operations of PWM converters allow reduction of the size and weight of their magnetic components. However, at high switching frequency, switching losses and EMI emissions become significant and must be reduced. Traditional high frequency switch — mode supplies, which rely on generating an AC waveform have used power transistors to "hard-switch" the unregulated input voltage at this rate. This means that a transistor turning on will have the whole raw input voltage, across it as it changes state. During the actual switching interval (less than 50 microseconds) there is a finite period as the transistor begins to conduct where the voltage begins to fall at the same time as current begins to flow. This simultaneous presence of voltage across the transistor and current through it means that, during this period, power is being dissipated within the device. A similar event occurs as the transistor turns off, with the full current flowing through it. More recently, new power conversion topologies have been developed which dramatically reduce the power dissipated by the main power transistors during the switching interval. The most common technique employed has been a constant frequency resonant switching scheme, which ensures that the actual energy being dissipated by the active device is reduced to nearly zero. This method, commonly called is "**Zero Voltage Switching**" (ZVS), "**Zero Current Switching**" or "**Soft Switching**". These converters use the LC resonance circuit. When using resonant circuits to reduce switching losses, by resonant inductance connected or disconnected from the resonant circuit at zero current passing through this inductance, or that connects or disconnects the resonance capacity at zero voltage. Leader in this area are quasiresonant converters, which use the properties of resonant LC circuit, only in moments of commutation switches. Reduction in the steepness of the starting and trailing edge of the output voltage when using resonant inverter also has a favorable influence on the electromagnetic interference.

2. POWER PART OF AN INVERTER WITH RESONANT POLE

In Fig. 1, there is shown a block scheme of the power part of used three-phase converter with soft switching. The converter with resonant poles consists of a conventional voltage source inverter with zero (reverse) diodes and resonant poles. Each branch of the inverter bridge contains switches VT_{11} and VT_{12} (VT_{21}, VT_{22} and VT_{31}, VT_{32}). The switches activate circuits Lr_1, Cr_1, respective Lr_2, Cr_2 and Lr_3, Cr_3 (see Fig. 1). These circuits are initialized in the instants, when the current of load is too low for fast overcharging, or wrong polarity for overcharging of resonant capacitors. If the current of load is high enough and right polarity it is possible to overcharge without resonant circuit utilization. The main switches of the converter use zero voltage switching and the auxiliary switches use zero current switching.

3. DESIGN OF METHODS FOR OPTIMIZATION OF RESONANT CIRCUIT CONTROL WITH RESPECT TO LOAD CURRENT

For design of optimalization of resonant circuit control it is used the scheme for one phase of resonant circuit (Fig. 2). It is clear that switching of IGBT transistors has to be realised in such way to no IGBTs use hard switching. It is possible to perform optimization of switching algorithm with respect to the value of load current in the following way — to minimal number of resonant circuit activation occurs and to short-term currents of resonant circuit have a necessary value only. The optimization can be divided into two basic problems.

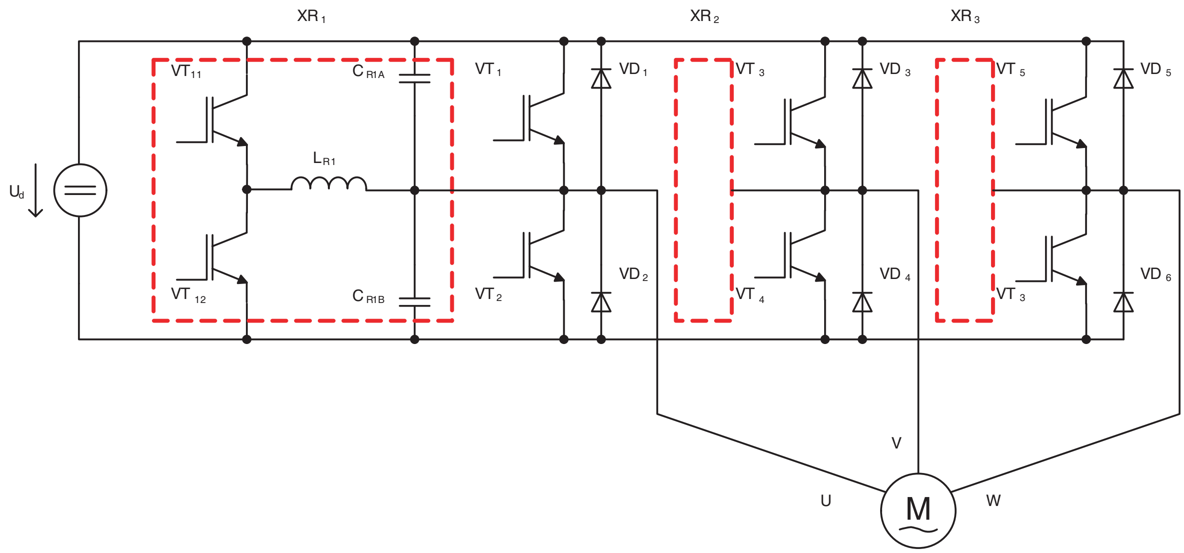


Figure 1: Inverter with resonant poles — power part.

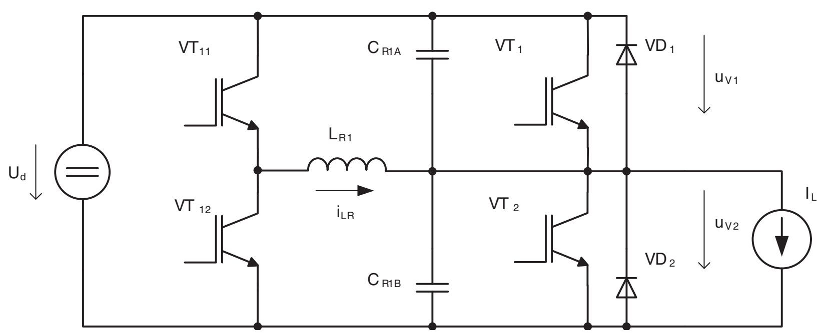


Figure 2: One phase of quasi-resonant inverter.

- Optimization of switching algorithm of the switches in the bridge with respect to the value of I_L.
- Optimization of current impulse size in the resonant circuit with respect to the value of I_L.

If the load current achieves sufficient value and right direction it is possible to overcharge the resonant capacitors by the current without resonant circuit activation. In the second case the situation is following, load current has right direction, but it is not high enough for fast overcharging the resonant capacitors. In this case the load current overcharges the resonant capacitors with the help of the resonant circuit without its activation. For right understanding it is presented detailed explanation. Each situation follows from waveforms of the load current. The first assumption — we take into account an ideal load current, we substitute it by a sinusoidal waveform. The waveform is divided into a few sectors according to size and direction of load current (see Fig. 3). Current directions, positive and negative, correspond to labelling in Fig. 2. The load current is also divided into sectors I, II, III with respect to its absolute value. Switching algorithm for control of one phase of quasi-resonant converter depends on the direction and sector, where load current I_L lies. From the fact follows that it exists five switching sequences of each IGBT in the quasi-resonant bridge.

4. BLOCK SCHEME OF THE CONTROL PART FOR ONE PHASE OF THE CONVERTER USING SOFT-SWITCHING

In Fig. 4, it is shown a block scheme of the control part for one phase of the converter using soft-switching. The input is a PWM signal connected to the analog part, where it is compared the load current and resonant current, it is evaluated position of load current with respect to the position in the given sector (see Fig. 4) and it is chosen a switching sequence for logical part of the control circuits.

5. EXPERIMENTAL RESULTS

In Figs. 5 to 9 there are depicted the time responses of output voltage and current together with the current of resonant coil during small and large values of the output current. On the pictures are also included the details of switching in various regimes.

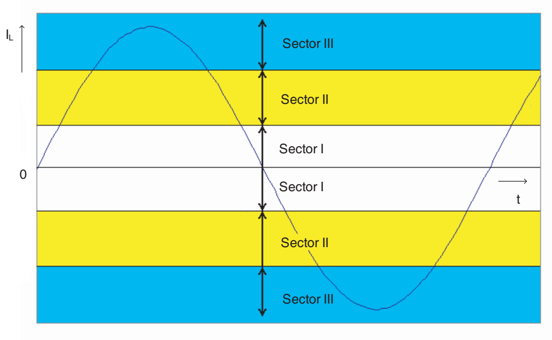


Figure 3: Waveform of load current and its division into particular sectors.

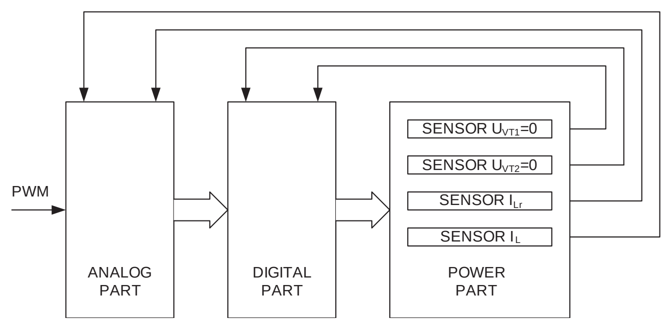


Figure 4: Block scheme of the control part for one phase of the converter using soft-switching.

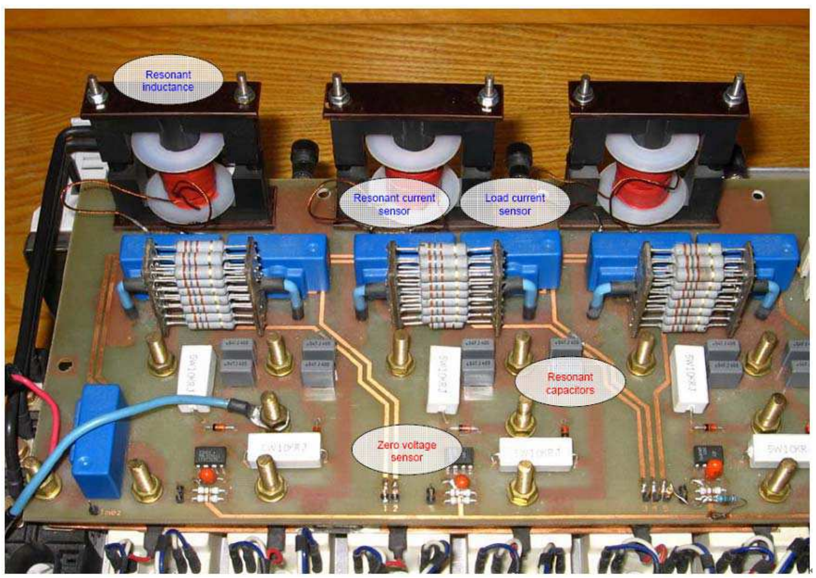


Figure 5: Inverter with resonant circuit.

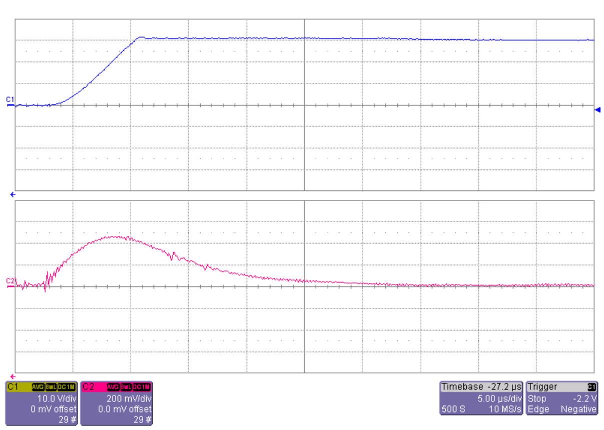


Figure 6: Voltage U_{VT_2} (blue) and resonant current i_{LR} (red) for small I_L.

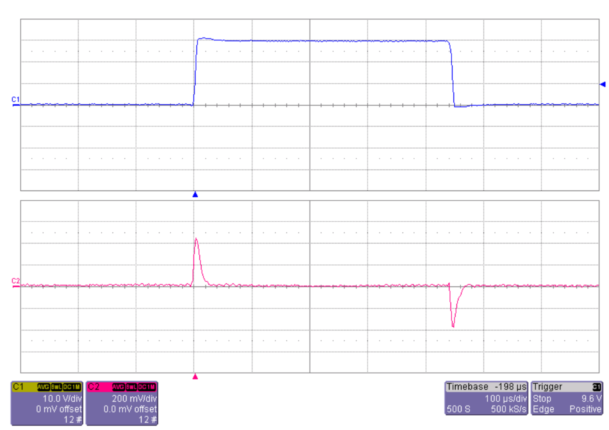


Figure 7: Voltage U_{VT_2} (blue) and resonant current i_{LR} (red) for one switching cycle VT_2 and small I_L.

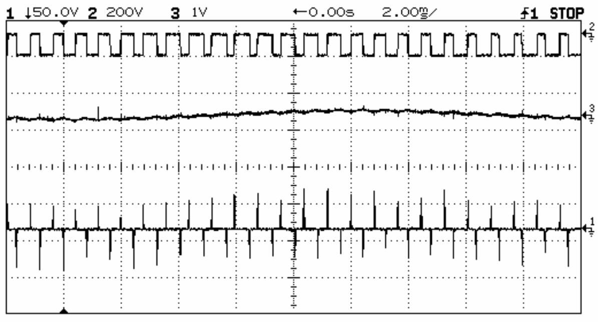


Figure 8: Output voltage, current and resonant current for $f = 50$ Hz and small output current.

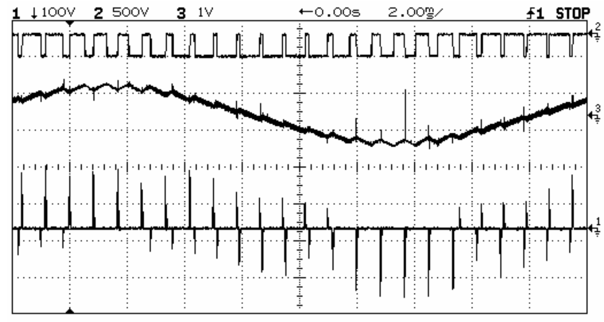


Figure 9: Output voltage, current and resonant current for $f = 50$ Hz and large output current.

ACKNOWLEDGMENT

The paper was created with the support of the projects MSM: 618910007, ENET: CZ.1.05/2.1.00/-03.0069.

REFERENCES

1. Brandstetter, P., T. Krecek, P. Rech, et al., "Facilities for a digital signals processing of AC drives," *Proceedings of International Conference on Applied Electronics*, 64–64, Pilsen, Czech Republic, September 2009.
2. Brandstetter, P. and P. Simonik, "Signal processing for vector control of AC drive," *20th International Conference Radioelektronika*, 93–96, Brno, Czech Republic, 2010.
3. Brandstetter, P., P. Rech, and P. Simonik, "Sensorless control of permanent magnet synchronous motor using luenberger observer," *PIERS Proceedings*, 424–428, Cambridge, USA, July 5–8, 2010.
4. Dudek, J. and P. Chlebis, "Comparison of power losses and quality of output voltage from different types of inverters," *Proceedings of 35th Annual IEEE Power Electronics Specialists Conference, PESC*, 3120–3126, Aachen, Germany, 2004.
5. Chlebis, P., P. Moravcik, and P. Simonik, "Method of direct torque control for three-level voltage inverter," *Proceedings of 13th European Conference on Power Electronics and Applications, EPE*, 4051–4056, Barcelona, Spain, September 2009.
6. Chlebis, P., P. Vaculik, P. Moravcik, et al., "Direct torque control methods for three-level voltage inverter," *Proceedings of 10th International Scientific Conference Electric Power Engineering*, 352–356, Kouty nad Desnou, Czech Republic, May 2009.

Space Vector Control for Quasi-resonant DC Link Inverter

Tomas Pavelek, Petr Palacky, Vaclav Sladecek, and Petr Hudecek
Department of Electronics, Faculty of Electrical Engineering and Computer Science
VŠB-Technical University of Ostrava, 17. Listopadu 15, Ostrava-Poruba 70833, Czech Republic

Abstract— There are summarized results of research of the quasi-resonant voltage type converter in this paper. Results show that the given type of the quasi-resonant converter is a good adept for vector pulse width modulation. The control part is based on the digital signal processor TMS320Cxx line by Texas Instruments, as for direct control of the voltage DC link as for vector pulse width modulation control algorithm too.

1. INTRODUCTION

High switching frequency of modern power converters leads to increased share of switching losses in total losses of converters. Reduction of the switching losses can be achieved by zero voltage switching or zero current switching. These techniques utilize resonant features of LC circuit, which is included in the power part of a converter. Converters, which allow soft switching, are called resonant converters. There exists a special group of converters, so-called quasi-resonant. Resonant circuit of a voltage source inverter with quasi-resonant DC-link is active in the instant of power switches commutation only. For realization of vector modulation it is needful to use a high performance microcontroller, which provides sufficient computing power. The vector modulation is implemented into digital signal processor TMS320LF2407 powered by Texas Instruments.

2. THE POWER PART OF INVERTER WITH QUASI-RESONANT DC-LINK WITH HIGH FREQUENCY TRANSFORMER

The power part consists of conventional three phase voltage source inverter with quasi-resonant DC-link, which contains the main switch SW1, auxiliary switches SA1, SA2, high frequency transformer and resonant capacitors Cr1 and Cr2 [1]. DC link is initialized only at the moment when the output space voltage vector changes. The power switches in the inverter are turned on/off under a principle of ZVS condition.

3. VECTOR MODULATION OF OUTPUT VOLTAGE OF QUASI-RESONANT INVERTER

For a voltage source inverter it is possible to realize output voltage modulation with the help of space vectors. The principle is shown in Figs. 2 and 3.

In each phase of three-phase load are generated voltages by means of switching combination. The size of the voltages is shown in table (U_d = DC link voltage, see Table 1.). From the table it is clear, that the voltage source inverter enables to realize eight switching combinations. Eight voltage vectors $\underline{u}_0$ to $\underline{u}_7$ correspond to the mentioned switching combinations.

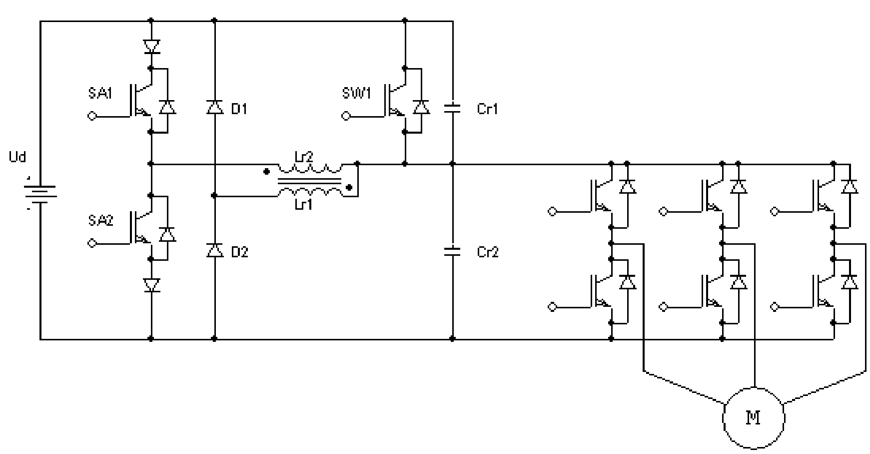


Figure 1: Basic scheme of the power part.

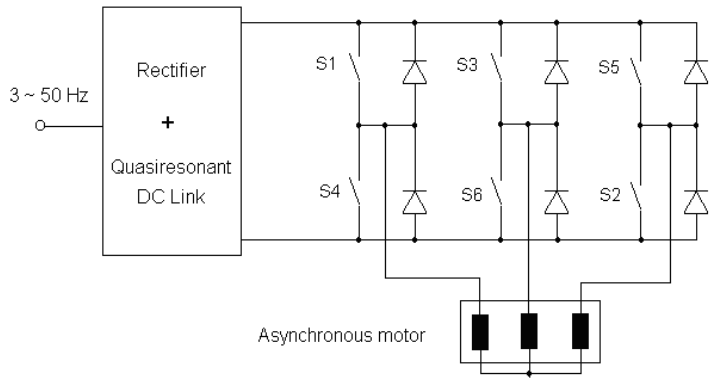


Figure 2: Scheme of a voltage source inverter.

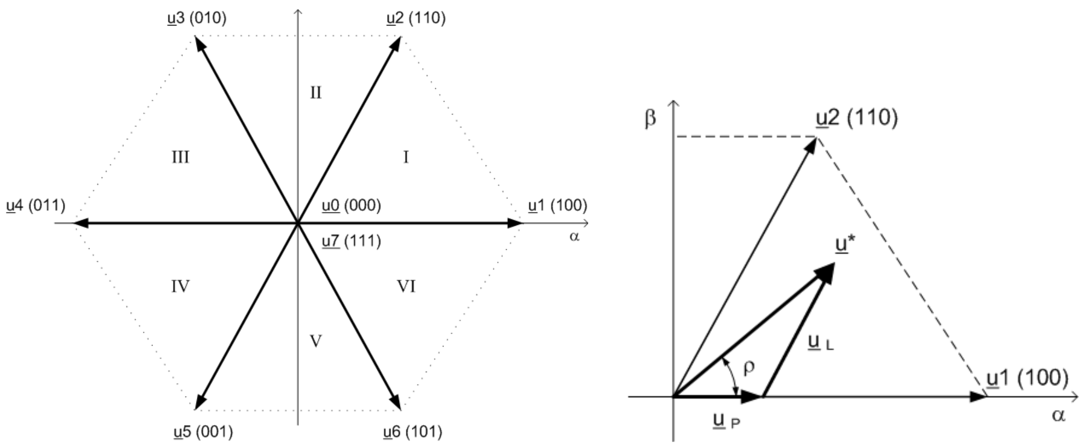


Figure 3: Voltage vectors.

Table 1: Phase voltages U_1, U_2, U_3 for the given switching combination.

Voltage vector	$\underline{u}_0$	$\underline{u}_1$	$\underline{u}_2$	$\underline{u}_3$
Switches	S_4, S_6, S_2 [000]	S_1, S_6, S_2 [100]	S_1, S_3, S_2 [110]	S_4, S_3, S_2 [010]
U_1	0	$2/3U_d$	$1/3U_d$	$-1/3U_d$
U_2	0	$-1/3U_d$	$1/3U_d$	$2/3U_d$
U_3	0	$-1/3U_d$	$-2/3U_d$	$-1/3U_d$
Voltage vector	$\underline{u}_4$	$\underline{u}_5$	$\underline{u}_6$	$\underline{u}_7$
Switches	S_4, S_3, S_5 [011]	S_4, S_6, S_5 [001]	S_1, S_6, S_5 [101]	S_1, S_3, S_5 [111]
U_1	$-2/3U_d$	$-1/3U_d$	$1/3U_d$	0
U_2	$1/3U_d$	$-1/3U_d$	$-2/3U_d$	0
U_3	$1/3U_d$	$2/3U_d$	$1/3U_d$	0

From the table it is possible to obtain the result that for constant DC-link voltage U_d size of a switched voltage vector can reach two values only:

$$|\underline{u}_0| \quad \text{or} \quad |\underline{u}_7| = 0 \cdot U_d \tag{1}$$

$$|\underline{u}_1| \quad \text{to} \quad |\underline{u}_6| = \frac{2}{3} \cdot U_d \tag{2}$$

Output voltage in the given instant, which is computed in higher level regulation system, is set to a control system in vector form, it means with the help of components u_α, u_β or in the form of voltage modulus U and angle ρ (polar coordinate — Fig. 3). Such defined vector of output voltage lies in some of six segments restricted by vectors $\underline{u}_1$ to $\underline{u}_6$ and it is determined by projection of adjacent vectors. For evaluation of pulse times T_P and T_L, the vector modulator utilizes the second method.

In the sampling interval of the converter, such defined interval can be realized as a mean value of vector sum of vector components, which lies on the right $\underline{u}_P$ and on the left $\underline{u}_L$ of reference vector, in this case vectors $\underline{u}_1$ and $\underline{u}_2$ and zero voltage vector. These vectors are active for the time corresponding to projection of reference vector $\underline{u}*$ to the directions of mentioned adjacent vectors. The pulse times T_1 and T_2 can be evaluated by means of the cosine theorem and substitution for vectors $\underline{u}_1$ and $\underline{u}_2$, or with the help of voltages obtained from higher level regulation system $\underline{u}_P$ and $\underline{u}_L$:

$$\underline{u}^* = \frac{\underline{u}_1 T_1}{T} + \frac{\underline{u}_2 T_2}{T} + \frac{\underline{u}_0 T_0}{T} = \underline{u}_P + \underline{u}_L + \underline{u}_0 \tag{3}$$

$$T = T_1 + T_2 + T_0 \tag{4}$$

$$\frac{\sin 120^\circ}{|\underline{u}^*|} = \frac{\sin(60^\circ - \rho)}{|\underline{u}_1| \frac{T_1}{T}} \tag{5}$$

$$T_1 = \sqrt{3} \cdot \frac{|\underline{u}^*|}{U_d} \cdot T \cdot \sin(60^\circ - \rho) \tag{6}$$

where T_1 is time for switching combination for the vector on the right from reference vector $\underline{u}*$, $\underline{u}_P$ is projection into direction of vector $\underline{u}_1$.

$$T_2 = \sqrt{3} \cdot \frac{|\underline{u}^*|}{U_d} \cdot T \cdot \sin\rho \tag{7}$$

where T_2 is time for switching combination for the vector on the left from reference vector $\underline{\mathbf{u}}*$, $\underline{u}_L$ is projection into direction of vector $\underline{u}_2$.

T is switching period and T_0 is time of switching combination for zero vectors $\underline{u}_0$, $\underline{u}_7$.

The relations inferred for the first sector are valid similarly in the next sectors. There is only shift of vectors by 60°.

In the case of quasi-resonant converters, zero switching vectors can be realized by activation of all switches in the converter bridge. This situation means short-circuit of power supply, but for quasi-resonant converters, the inverter is disconnected from DC-link by means of SW1.

For the reference vector realization it is not important the sequence of vector switching. To avoid hard switching it is necessary to insert an interval T_0 between intervals T_1 and T_2. Possible realization ways are shown in Fig. 4.

The program was realized in development software DMC Developer Pro. Developers board with digital signal processor TMS 320LF2407 contains 6 connectors, on which there are particular processor signals. The developers kit controls switching of transistors in DC-link SW1, Sa1, Sa2 and also switching of transistors in the converter bridge VT1-VT6. The inputs are signals from sensors of zero voltage on capacitors Cr1 and Cr2 and a signal from current sensor, which is important to indicate the end of resonant current.

4. EXPERIMENTAL RESULTS

To verify correctness of used algorithm it was realized simple measuring workplace, its each component are a quasi-resonant converter, asynchronous motor and control kit with digital signal

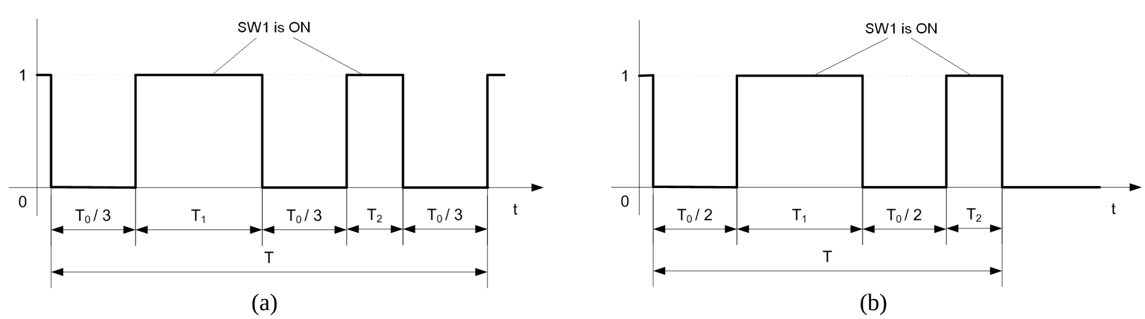
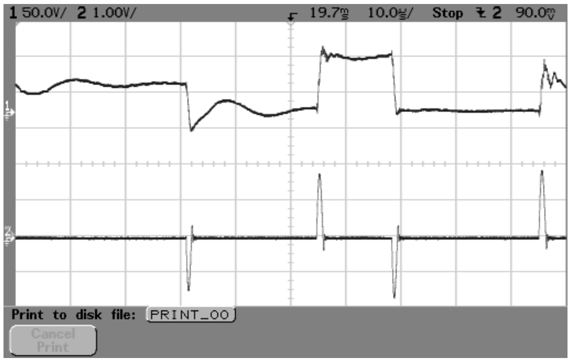
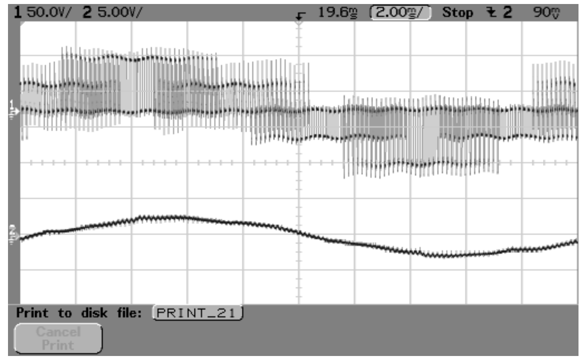


Figure 4: Possible division of a vector during the sampling period. (a) Division of zero vector into thirds. (b) Division of zero vector into half.



(a)



(b)

Figure 5: Experimental results. (a) Phase voltage and resonant current. (b) Voltage and current of one phase of induction motor.

processor.

It the paper it is presented two experimental results only, which show correct operation of the converter and vector modulator. In Fig. 5(a) it is shown the detail of phase voltage of the motor and resonant current of DC-link.

That is clear, that initialization of DC-link is performed in each change of switching combination.

In Fig. 5(b), there is shown voltage and current of one phase of the motor at switching frequency 6 kHz and output frequency 50 Hz. High performance DSP allows obtaining the switching frequency up to 10 kHz.

5. CONCLUSION

The work was focused to be the base for realization of vector control of voltage source inverter with quasi-resonant DC-link and the results to bring some improvement in the field of electrical drives and power electronics. In the future we will deal with measurement on the converter and its optimization.

ACKNOWLEDGMENT

The paper was created with the support of the project ENET: CZ.1.05/2.1.00/03.0069.

REFERENCES

1. Inaba, C. Y., E. Hiraki, S. Ohgaki, and M. Nakaoka, “Three phase soft commutation inverter with high frequency transformer assisted resonant snubbers,” *EPE — PEMC 2000 Košice*, Košice, Slovak Republic, 2000.
2. Brandstetter, P., L. Hrdina, and P. Simonik, “Properties of selected direct torque control methods of induction motor,” *IEEE International Symposium on Industrial Electronics — ISIE’10*, Bari, Italy, 2010.
3. Simonik, P., P. Brandstetter, and P. Chlebis, “DSP control unit for power electronics applications,” *International Conference Applied Electronics*, 311–314, Pilsen, Czech Republic, 2010, ISBN 978-80-7043-865-7.
4. Brandstetter, P., O. Francik, and P. Simonik, “AC drive with permanent magnet synchronous motor and sensorless control,” *14th International Power Electronics and Motion Control Conference*, Vol. 2, 66–71, Ohrid, Republic of Macedonia, 2010, ISBN 978-1-4244-7854-5.

Optimized Dual Randomized PWM Technique for Full Bridge DC-DC Converter

A. Boudouda[1], **N. Boudjerda**[1], **B. Nekhoul**[1], **K. El Khamlichi Drissi**[2], **and K. Kerroum**[2]

[1]LAMEL, University M. S. Ben Yahia Jijel, Algeria
[2]LASMEA, University Blaise Pascal, Clermont Ferrand, France

Abstract— This paper deals with the control of the full bridge dc-dc voltage converter by the Random Pulse Width Modulation (RPWM) technique. We propose a RPWM scheme based on a triangular carrier having two randomized parameters: the period T and the peak position βT (half period). First, we propose the modulating principle, then the voltage analysis based on Power Spectral Density (PSD) shows the advantage from the point of view of Electro-Magnetic Compatibility (EMC) of this scheme compared to conventional RPWM schemes having only one randomized parameter. Moreover, this analysis reveals the existence of an optimal value of the variation interval of parameter β for a maximum spread of the PSD. Once the optimization problem modeled, the resolution is performed using two powerful nonlinear methods: the trust region method and the simplex algorithm.

1. INTRODUCTION

Nowadays, much of the electrical energy is consumed through static converters, which allow adjusting this energy to the desired form. The use of PWM technique for the control of power converters allows adjusting the useful component of the voltage and eliminating some unwanted harmonics [1]. Thus, it is required for power converters to provide the desired electrical functionality and to meet international standards of Electro-Magnetic Compatibility (EMC) by reducing conducted and radiated emissions [2].

In order to better meet the EMC standards for conducted disturbances, we can use the filtering technique (passive and/or active power filters). Furthermore, RPWM technique is one of the most effective and least-cost solutions: it allows spreading the power spectrum over a wide frequency range while significantly reducing its amplitude and this constitutes a significant EMC advantage without any additional hardware.

Several works regarding this new control technique have been published recently, two conventional RPWM schemes are proposed: the scheme in which the switching period is randomized (Randomized Carrier Frequency Modulation: RCFM) and the scheme in which the period is kept constant and the pulse position is randomized (Randomized Pulse Position Modulation: RPPM) in both DC-DC and DC-AC [3–5]. It has been shown that RCFM scheme allows a better spreading of the spectrum then RPPM scheme [6]. However, to obtain a maximum spread, the combination of these two schemes has been proposed (RCFM-RPPM) that we call dual RPWM scheme [5, 6]. We propose in this paper an optimized scheme with two randomized parameters “optimized RCFM-RPPM”. The switching functions are generated by comparing a carrier having two random parameters with two deterministic reference signals: one switching function for each arm of the converter.

First, we propose the modulating principle of this technique. Then an analytical model of the Power Spectral Density (PSD) of the output voltage is developed and validated. This model is expressed directly in terms of the two random parameters of the carrier. Note that the simple schemes (RCFM and RPPM) are deduced directly from the proposed general model. The PSD analysis shows that the proposed dual RPWM scheme allows a better spread shape of PSD compared to the simple randomization schemes that is the desired EMC advantage. In addition, a parametric study shows that there is an optimum shape of the PSD based on statistical parameters, The treatment of this problem with two powerful nonlinear methods (Trust Region method and Simplex algorithm), gives the same results.

2. MODULATING PRINCIPLE

The structure of the converter is given in Fig. 1; it requires two switching functions q_a et q_b.

The modulating principle is illustrated in Fig. 2. The two switching functions q_a and q_b are obtained by comparing two deterministic reference signals r_a and r_b of magnitudes d_a and d_b

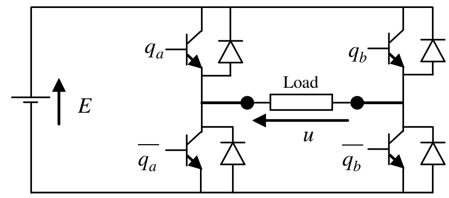


Figure 1: Full bridge DC-DC voltage converter.

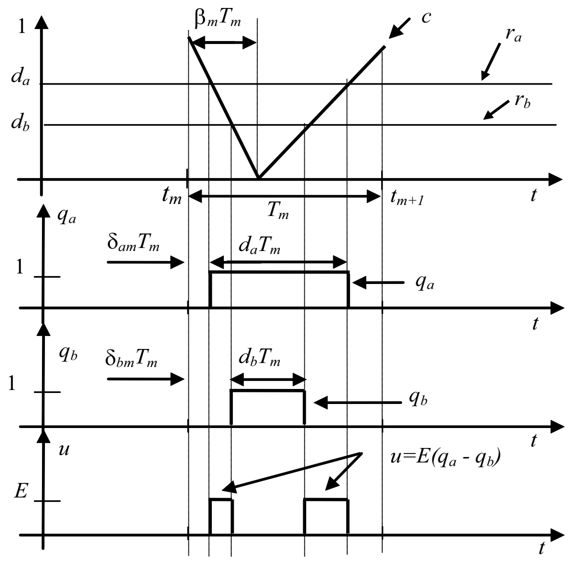


Figure 2: Modulating principle.

respectively, to a single randomized carrier c (Fig. 2). Generally, the magnitudes d_a and d_b are taken as follows:

$$\begin{array}{l} 0 < d_a < 1 \\ 0 < d_b < 1 \\ d_a + d_b = 1 \end{array} \tag{1}$$

Thus, the average value U_0 of the output voltage u is:

$$U_0 = (2d_a - 1)E \tag{2}$$

Note that a variation of d_a between 0 and 1 allows adjustment of the voltage U_0 between $-E$ and $+E$.

Each switching function is completely characterized by:

- The switching period T.
- The duty cycle d.
- The delay report δ.

Theoretically, the three parameters (T, δ and d) should be randomized in a separated or a combined way. However, in industrial applications, the duty cycle d is generally deduced from a reference signal and allows the control of output voltage. Therefore, only T and δ are really randomized, resulting in the four (4) following configurations (Table 1).

Table 1: RPWM schemes.

PWM scheme	β	T
DPWM	Fixed: $\beta = 0.5$	fixed
RPPM	randomized	fixed
RCFM	Fixed: $\beta = 0.5$	randomized
RCFM-RPPM	randomized	randomized

The output voltage u can be expressed using the switching functions q_a and q_b as follows:

$$u = (q_a - q_b)E \tag{3}$$

From Fig. 2, a random variation of parameter δ_a between 0 and $(1-d_a)$, "respectively δ_b between 0 and $(1 - d_b)$" is obtained by using a carrier with random parameter β between 0 and 1, which gives:

$$\begin{cases} \delta_a = \beta\left(1 - d_a\right) \\ \delta_b = \beta\left(1 - d_b\right) \end{cases} \tag{4}$$

The use of parameter β instead of parameters δ_a and δ_b allows defining the switching functions q_a and q_b and the output voltage u directly in term of the random parameters T and β of the carrier, therefore the voltage analysis can be performed directly from these two parameters. In addition, it provides only a single random variable "β" instead of the two parameters "δ_a" and "δ_b".

3. MATHEMATICAL MODEL OF PSD

A rigorous study of the spectral content of the output voltage is based on the Power Spectral Density (PSD) of random signals, defined as the Fourier transform of the auto-correlation of the considered signal:

$$W\left(f\right) = \int\limits_{-\infty}^{\infty} R\left(\tau\right) e^{-j2\pi f\tau} d\tau \tag{5}$$

and:

$$R\left(\tau\right) = \lim_{T_0\to\infty} \frac{1}{2T_0} \int\limits_{-T_0}^{T_0} u\left(t\right) u\left(t-\tau\right), dt \tag{6}$$

$R\left(\tau\right)$ and $W\left(f\right)$ are respectively the auto-correlation function and the power spectral density of the signal $u\left(t\right)$.

For a Wide Sense Stationary (WSS) pulse signal $u\left(t\right)$, the PSD can be expressed as follows [4–6]:

$$W\left(f\right) = \lim_{N\to\infty} \frac{1}{\overline{T}} E\left[\sum_{k=-N}^{N} U_m\left(f\right)\; U_{m+k}^*\left(f\right)\right] \tag{7}$$

$U_m\left(f\right)$ and $U_{m+k}^*\left(f\right)$ are respectively the Fourier transform of the signal $u\left(t\right)$ during the switching period T_m and its conjugate during the switching period T_{m+k}.

In relative magnitudes (per unit system), the expression (3) of the output voltage becomes:

$$u = q_a - q_b \tag{8}$$

and its Fourier transform over a switching period T_m is then:

$$U_m\left(f\right) = q_{am}\left(f\right) - q_{bm}\left(f\right) \tag{9}$$

$q_{am}\left(f\right)$ and $q_{bm}\left(f\right)$ are respectively the Fourier transforms of q_a and q_b during the switching period T_m.

After some mathematical transformations, the PSD of the output voltage for the scheme RCFM-RPPM can be written as follows [5]:

$$W\left(f\right) = \frac{1}{\overline{T}\left(\pi f\right)^2} \left\{ E_T\left[\left|U\left(f\right)\right|^2\right] + 2\mathrm{Real}\frac{E_{T,\beta}\left[U\left(f\right)e^{j2\pi fT}\right] E_{T,\beta}\left[U^*\left(f\right)\right]}{1 - E_T\left[e^{j2\pi fT}\right]} \right\} \tag{10}$$

where:

$$U\left(f\right) = \begin{pmatrix} e^{-j2\pi f\beta(1-da)T} e^{-j\pi f d_a T} \sin\left(\pi f d_a T\right) \\ -e^{-j2\pi f\beta(1-d_b)T} e^{-j\pi f d_b T} \sin\left(\pi f d_b T\right) \end{pmatrix} \tag{11}$$

$E_T\left[\;\right]$: Expected value related to the random variable T.

$E_{T,\beta}\left[\;\right]$: Expected value related to the random variables T and β.

From expressions (10) and (11), the simple schemes RCFM and RPPM can be obtained as particular cases: for RCFM scheme, the parameter β is constant ($\beta = 0.5$) and for RPPM, the period T is constant.

4. VALIDATION OF THE MATHEMATICAL MODEL OF PSD

The validation of the mathematical model of the PSD for the three RPWM schemes is performed by comparing the PSD calculated analytically with that estimated by the Welch method which can be obtained after simulation of the converter [7]. Moreover, for simple schemes (RCFM and RPPM), measurement results published in the literature are also used [1] for the comparison which is achieved with the following conditions:

- Input voltage: $E = 200\,\text{volts}$.
- Switching period T: randomized according to the uniform probability law between two values:

$$T_{\min} = 167\,\mu\text{s and } T_{\max} = 250\,\mu\text{s}.$$

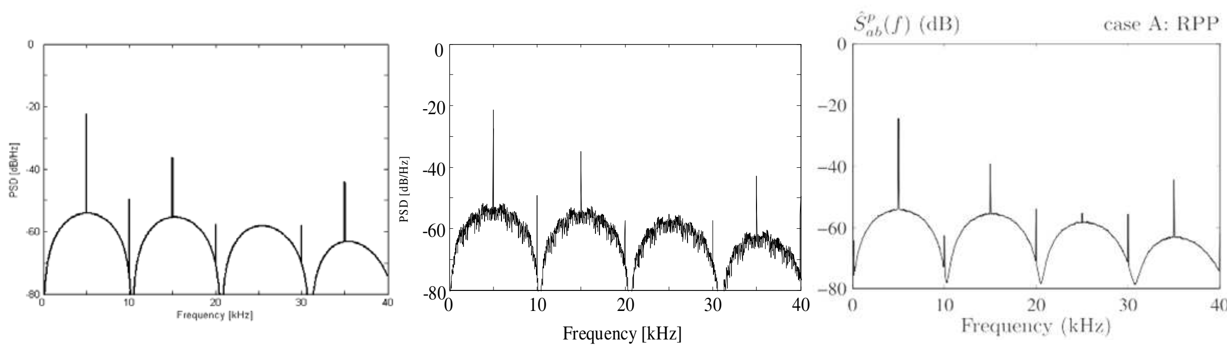
- β is randomized according to the uniform distribution in the interval:

$$\beta \in \left[\overline{\beta}\,(1 - R_\beta/2)\,,\ \overline{\beta}\,(1 + R_\beta/2)\right]$$

$\overline{\beta} = 0.5$: Statistical average.

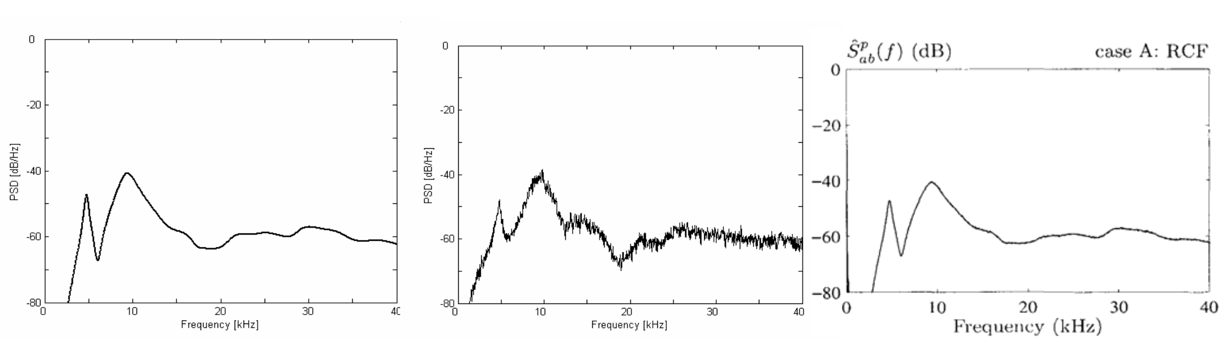
R_β: Randomness level, it allows defining the variation interval of parameter β, theoretically the maximum value is $(R_\beta)_{\max} = 2$.

Figures 3 and 4 show very good agreements between the estimate, the analytical calculation and the measurement for RPPM and RCFM schemes respectively while Fig. 5 shows a very good agreement between the estimation and the calculation for RCFM-RPPM scheme thereby validating our proposed model.



(a) (b) (c)

Figure 3: Validation for RPPM scheme. (a) Computed PSD. (b) Estimated PSD (Welch method). (c) Measured PSD [1].



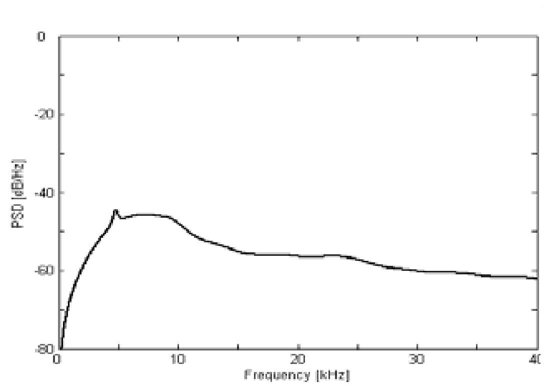
(a) (b) (c)

Figure 4: Validation for RCFM scheme. (a) Computed PSD. (b) Estimated PSD (Welch method). (c) Measured PSD [1].

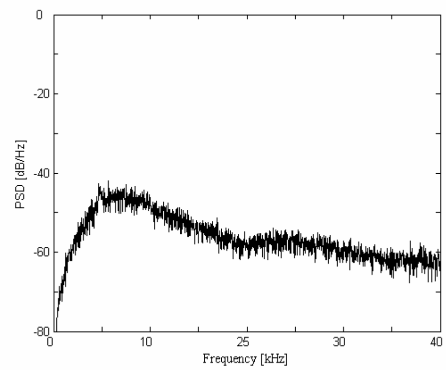
5. EMC ADVANTAGE OF THE RCFM-RPPM SCHEME

Figure 6 shows the PSDs of the output voltage obtained while varying R_β for three values of R_T ($R_T = 0.2$, $R_T = 0.3$ and $R_T = 0.4$) at one functioning point corresponding to the following duty ratio $d = d_a - d_b = 0.75 - 0.25 = 0.5$.

The effect of R_T appears clearly: important values of R_T give better spread PSD. In addition, the dual randomized scheme (RCFM-RPPM) adds a significant spread to the PSD compared to the simple scheme (RCFM). This spreading is also accompanied by a decrease in peaks. Nevertheless, a significant increase in R_β (Fig. 6) causes an increase in the peak of the PSD around Fs (F_s: Average frequency modulation) and a decrease in the peak of the PSD around $2Fs$. Knowing that our purpose is to spread best the PSD and to reduce its peaks (in order to meet EMC standards), a compromise between the two peaks of the PSD (at F_s and $2F_s$) can be achieved with an optimal value of R_β. The optimization problem is posed and solved in the following paragraph.

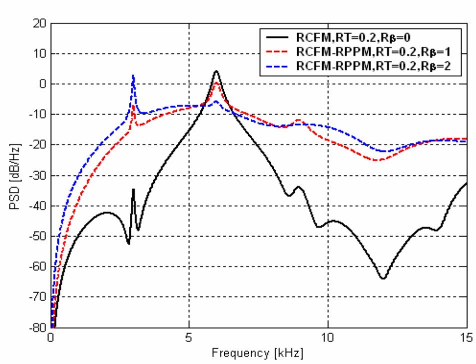


(a)

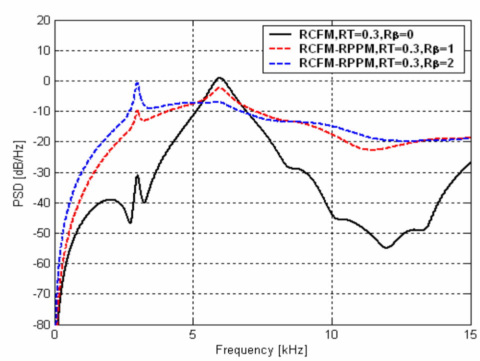


(b)

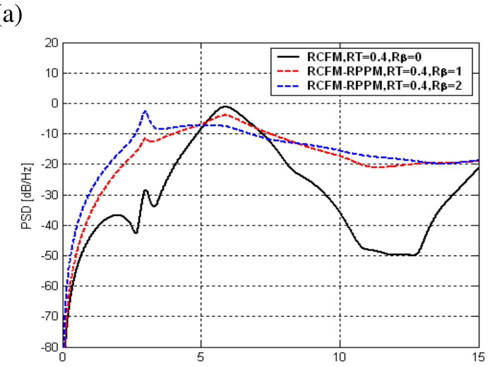
Figure 5: Validation for RCFM-RPPM scheme. (a) Computed PSD. (b) Estimated PSD (Welch method).



(a)



(b)



(b)

Figure 6: Effects of R_T and R_β on PSD shape. (a) PSD for $R_T = 0.2$ and ($R_\beta = 0$, 1 and 2). (b) PSD for $R_T = 0.3$ and ($R_\beta = 0$, 1 and 2). (c) PSD for $R_T = 0.4$ and ($R_\beta = 0$, 1 and 2).

6. OPTIMISATION OF PARAMETER R_β

Our purpose is to find an optimal value of the randomness level R_β (optimization parameter) in order to achieve the best spread of the PSD. R_β is subject to the following constraints of domain boundaries: $(0 \leq R_\beta \leq 2)$.

Generally, an optimization problem subject to the constraints of domain boundaries can be written as follows:

$$\begin{cases} \min\ F(x) \in \Re^n \\ l \leq x_k \leq u \quad k = 1, \ldots, n \end{cases} \tag{12}$$

where:

- x: n-dimensional vector $(x_k, k = 1, \ldots, n)$ representing the parameters of the problem to optimize.
- l and u: Lower and upper limits of the search area (domain boundaries).
- $\Re^n$: Search domain bounded by the limits$l$ and u.
- $F(x)$: Objective function or optimization criterion.

In our case, the criterion$F$ is chosen with the least square (quadratic). It is defined by the sum of m non-linear functions squares. Thus, the resolution of a least square optimization problem consists of searching the vector $x_* \in \Re^n$ which minimizes the objective function as follows [8]:

$$\min_{x \in \Re^n} F(x), \quad F(x) = \|f(x)\|_2^2 = \sum_{i=1}^{m} f_i^2(x) \quad m \geq n \tag{13}$$

where:
$f(x) = [f_1(x)\, f_2(x) \ldots f_m(x)]^T$: Vector containing m non-linear functions.
$\|f(x)\|_2$: Second order norm of vector $f(x)$.
$x \in \Re^n$: Vector of optimization parameters.

6.1. Mathematical Formulation

The parameter β varies randomly in the interval $[\beta_{\min}, \beta_{\max}]$, with $\beta_{\min} \geq 0$ and $\beta_{\max} \leq 1$. By introducing the statistical mean $\overline{\beta}$ and the randomness level R_b, we obtain:

$$\beta \in \overline{\beta}\left[1 - \frac{R_\beta}{2}, \quad 1 + \frac{R_\beta}{2}\right]$$

where: $\overline{\beta} = 0.5$ et $0 \leq R_\beta \leq 2$.

In expression (13), $f(x)$ is a vector of dimension m =2, it represents the PSD at frequencies F_s and $2F_s$ and x represents the searched value of R_β. We can then consider the probleme as an optimization problem in the sense of nonlinear least squares with constraints of the search area limits as follows:

- Unweighted problem:

$$\begin{cases} \min F(x) = W^2(R_\beta, F_S) + W^2(R_\beta, 2F_S) \\ 0 \leq R_\beta \leq 2 \end{cases} \tag{14}$$

- Weighted problem: Usually the low frequencies are more harmful for inductive loads, then we can introduce weighting coefficients in the objective function so as to give more important to the frequency Fs compared to $2F_s$ as follows:

$$\begin{cases} \min\ F(x) = (a * W(R_\beta, F_S))^2 + (b * W(R_\beta, 2F_S))^2 \\ 0 \leq R_\beta \leq 2 \end{cases} \tag{15}$$

where:

- R_β: Optimization parameter of the objective function.
- a and b: weighting coefficients, in practice the impedance of an inductive load is proportional to the frequency, thus its judicious to choose the ratio a/b inverse of the ratio between the frequencies F_s and $2F_s$, which gives $a = 2$ and $b = 1$.

- W: PSD of the voltage evaluated at frequencies F_s and $2F_s$ using the general expression:

$$W\left(R_\beta, f\right) = \frac{1}{\overline{T}\left(\pi f\right)^2}\left\{E_T\left[\left|U\left(R_\beta, f\right)\right|^2\right] + 2\text{Real}\left(\frac{E_{T,\beta}\left[U\left(R_\beta, f\right)e^{j2\pi fT}\right]E_{T,\beta}\left[U^*\left(R_\beta, f\right)\right]}{1 - E_T\left[e^{j2\pi fT}\right]}\right)\right\} \tag{16}$$

Note: In the expression (16) of PSD, the frequency f takes only two values (F_s and $2F_s$) and the variable R_β is introduced while replacing the expectations related to the random variable β, ($E_{T,\beta}[\ldots]$) by their expressions.

6.2. Optimization Methods

To solve this optimization problem, we propose two algorithms: the trust region method, available in MATLAB as the utility named **lsqnonlin** and to consolidate the results, we use the simplex algorithm (Nelder and Mead) available in MATLAB as the utility named **fminsearch**. These two algorithms are well suited to our problem, which is nonlinear with the constraints of the domain limits.

6.2.1. Trust Region Method

This method involves replacing the resolution of the problem posed by Equation (12) by solving a succession of simpler sub-problems where the optimization criterion F is replaced by a simpler function m called model function, which reflects the behavior of F in a neighborhood N around the point x, this neighborhood is called trust region and is generally spherical or ellipsoidal [9]. Originally, this algorithm is used in nonlinear optimization problems for both cases: without constraints and with constraints of the domain limits [10].

6.2.2. Simplex Algorithm

This algorithm is based on the concept of direct search: it attempts to solve the problem by using directly the objective function value without using its derivatives. It is suitable for strongly nonlinear optimization problems without constraints [11]. However, it may take into account the domain limits. A simplex is a geometric figure of n-dimension, created from $(n+1)$ points, each dimension corresponds to a parameter of the optimization problem. The minimum is sought by changing the simplex through standard operations: reflection, expansion and contraction [12].

6.2.3. Numerical Results of Optimization

For fixed R_T, we search the optimal value of R_β in the interval $[0, 2]$ with an arbitrary initial value ($R_{\beta_0} \in [0, 2]$). The obtained results are given in Tables 2, 3, 4 and 5. Note that these results remain unchanged for any value of R_{β_0}, demonstrating the robustness of these algorithms.

The two algorithms give almost the same results (difference less than 0.1%), which reinforce the obtained results.

- *Without frequency weighting* ($a = b = 1$)

Table 2: Obtained results (trust region).

R_T	Optimal R_β	$W\left(R_\beta, F_s\right)$	$W\left(R_\beta, 2F_s\right)$
0.2	1.1755	0.1649	0.2205
0.3	1.1791	0.0815	0.1150

Table 3: Obtained results (simplex).

R_T	Optimal R_β	$W\left(R_\beta, F_s\right)$	$W\left(R_\beta, 2F_s\right)$
0.2	1.1748	0.1645	0.2207
0.3	1.1782	0.0813	0.1151

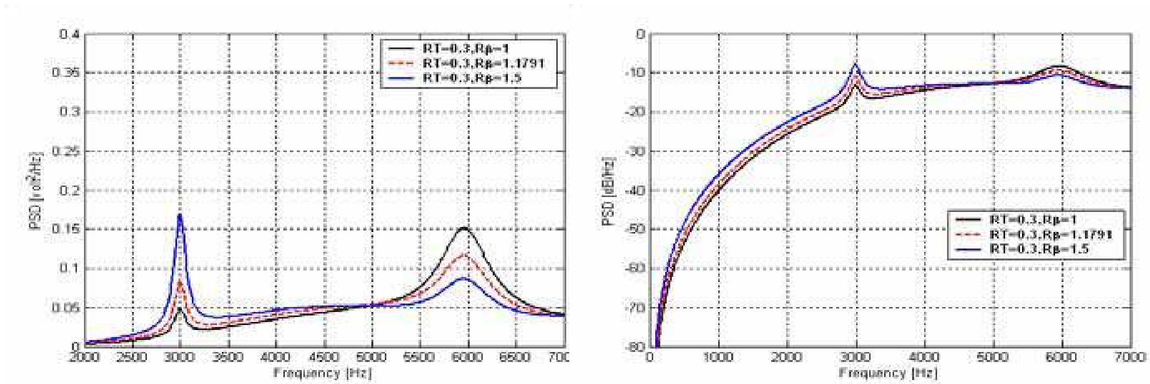
- *With frequency weighting* ($a = 2b = 2$)

Table 4: Obtained results (trust region).

R_T	Optimal R_β	$W\left(R_\beta, F_s\right)$	$W\left(R_\beta, 2F_s\right)$
0.2	1.0238	0.1036	0.2847
0.3	1.0167	0.0509	0.1457

Table 5: Obtained results (simplex).

R_T	Optimal R_β	$W(R_\beta, F_s)$	$W(R_\beta, 2F_s)$
0.2	1.0241	0.1037	0.2845
0.3	1.0177	0.0511	0.1455



(a) (b)

Figure 7: Comparison of PSDs. (a) PSD in volt^2/Hz. (b) PSD in dB/Hz.

To further support our results, we compare in Fig. 7, the variation of the PSD around the optimum value of R_β for R_T fixed. We see clearly (Fig. 7) the compromise reached between the two peaks at F_s and at $2F_s$.

7. CONCLUSION

In this paper, we have proposed an optimized RPWM technique for the full bridge DC-DC converter. This technique is based on a carrier with two randomized parameters. To make a rigorous analysis of the output tension, we have developed and validated a mathematical model of the power spectral density of this voltage. A parametric study reveals the existence of optimal statistical parameters. This problem has been treated with two powerful nonlinear methods.

REFERENCES

1. Bech, M. M., "Random pulse-width modulation techniques for power electronic converters," Ph.D. Thesis, Aalborg University, Denmark, 2000.
2. Redl, R., "Power electronics and electromagnetic compatibility," *IEEE Proceedings of PESC'96, Record, 27th Annual IEEE Conf.*, Vol. 1, 15–21, June 1996.
3. Tse, K. K., H. S.-H. Chung, S. Y. Hui, and H. C. So, "A comparative investigation on the use of random modulation schemes for DC/DC converters," *IEEE Trans. on Industrial Electronics*, Vol. 47, No. 2, 253–263, April 2000.
4. Kirlin, R. L., M. M. Bech, and A. M. Trzynadlowski, "Analysis of power and power spectral density in PWM inverters with randomized switching frequency," *IEEE Trans. on Industrial Electronics*, Vol. 49, No. 2, 486–499, April 2002.
5. Boudjerda, N., M. Melit, B. Nekhoul, K. El Khamlichi Drissi, and K. Kerroum, "Reduction of conducted perturbations in DC-DC voltage converters by a dual randomized PWM scheme," *Journal of Communications Software and Systems*, Vol. 5, No. 1, March 2009.
6. Boudjerda, N., "Réduction des perturbations conduites dans les convertisseurs de l'electronique de puissance par une commande en MLI aléatoire," Thèse de doctorat en sciences, Université Ferhat Abbas-Setif, Juillet 2007.
7. Welch, P. D., "The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. on Audio and Electroacoustic*, Vol. 15, 70–73, June 1967.
8. Dennis, Jr., J. E., "Nonlinear least-squares," *State of the Art in Numerical Analysis*, D. Jacobs, Ed., 269–312, Academic Press, 1977.

9. Rodriguez, J. F., J. E. Renaud, B. A. Wujek, and R. V. Tappeta, "Trust region model management in multidisciplinary design optimization," *Journal of Computational and Applied Mathematics*, Vol. 124, No. 1–2, 139–154, December 2000.
10. Coleman, T. F. and Y. Li, "An interior, trust region approach for nonlinear minimization subject to bounds," *SIAM Journal on Optimization*, Vol. 6, 418–445, 1996.
11. Lewis, R. M., V. Torczon, and M. W. Trosset, "Direct search methods: then and now," *Journal of Computational and Applied Mathematics*, Vol. 124, No. 1–2, 191–207, December 2000.
12. Nelder, J. A. and R. Mead, "A simplex method for function minimization," *Computer Journal*, Vol. 7, 308–312, 1965.

Skin Effect in Squirrel Cage Rotor Bars and Its Consideration in Simulation of Non-steady-state Operation of Induction Machines

Marcel Benecke[1], **Reinhard Doebbelin**[1], **Gerd Griepentrog**[2], **and Andreas Lindemann**[1]

[1]Institute of Electric Power Systems, Otto-von-Guericke-University, Magdeburg, Germany
[2]Siemens AG, Corporate Technology CT T DE, Germany

Abstract— This paper deals with squirrel cage induction machines and their modeling concerning the occurrence of current displacement in rotor bars caused by skin effect. Considering non-steady-state operation such as direct on-line start and especially in the case of sustained ramp-up, e.g., by using Y-D starter or soft starters (thyristor controlled AC voltage regulator), an accurate calculation of current and torque is required. An appropriate numerical calculation model is pointed out and an application example by simulating a double-cage induction machine start-up demonstrates the benefit of this extended machine model.

1. INTRODUCTION

In order to start-up three-phase induction machines very often three-phase AC voltage controllers based upon SCRs are being used to achieve a smooth transition to nominal speed and avoid high torque and current values. Especially widely used squirrel-cage motors are intentionally designed to develop a higher starting torque by utilizing the skin effect in the rotor bars. This, in turn, affects the stator currents as well.

To analyze the characteristics of these machines in non-steady-state operations with higher rotor slip, e.g., start with SCR based AC voltage controller or simple Y-D starter, the skin effect in rotor bars (also known as deep bar effect or current displacement) has to be taken into consideration in order to obtain valid results for torque and stator currents during ramp-up.

2. PRINCIPLE OF SKIN EFFECT IN ROTOR BARS

The rotor of current-displacement influenced induction machines is a squirrel cage consisting of a number of bars (N_2) arranged all-over the rotor perimeter and grouted with short circuit rings. The energized three-phase stator windings generate a rotary field to which the squirrel cage is exposed. This magnetic flux of the stator rotary field induces voltages which excite mesh currents in the adjacent rotor bars. Their values and phase angles are determined by the resistance and the leakage inductance of rotor bars and ring segments and also by the rotor slip s. The N_2 "windings" constitute an N_2-phase system which generates a rotor field acting like the field of a three-phase wound rotor [1]. So it is possible to model the rotor as a three-phase equivalent winding and to model the whole motor using the well-known T-equivalent circuit.

Under nominal operating conditions the current in the rotor bar cross-section is homogeneously distributed and the leakage flux lines are shaped like illustrated in Fig. 1(left). In non-steady-state operation, e.g., starting, the rotational speed does not correspond to the number of revolutions of the stator rotary field. Therefore, high slip values occur. In the case of a low rotational speed, which is connected with increased values of rotor current frequency, the current in rotor bars is cumulatively displaced in radial direction to the air gap. This effect is caused by the slot leakage field in the environment of the bars.

It is assumed that the rotor bar is divided into several elements (Fig. 1). Combined with the adjacent rotor bars they form partial coils connected by the corresponding short-circuit ring segments. Thus, the in radial direction internally located coil is exposed to a stronger leakage field and shows the highest leakage inductance value compared to the leakage inductance values of the upper coils close to the air gap [1]. At increased frequency f_2 of the rotor current i_2 (decreased rotational speed n) the leakage reactance predominates compared to the resistance and the current concentrates in the upper coils (Fig. 1(right)). Therefore, the effective conducting cross-section decreases and with it the resistance increases. As a consequence, leakage reactance and resistance values of the rotor depend on slip [2, 3].

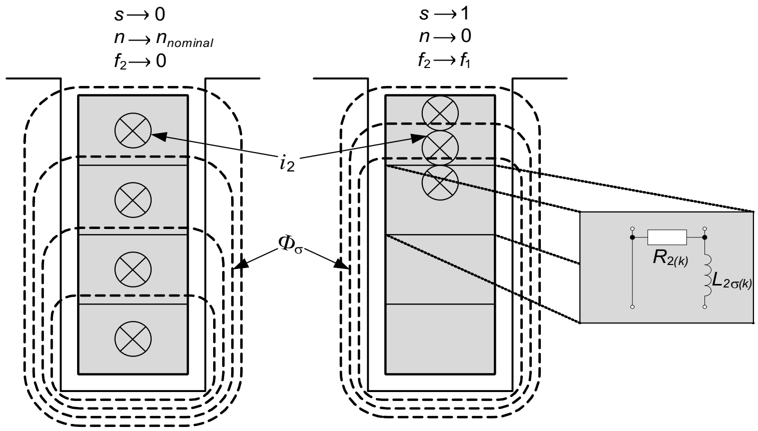


Figure 1: Principle of leakage field distribution in slots of a bar-wound rotor, with s — slip, n — rotational speed, f — electrical frequency (index: 1 — stator, 2 — rotor), i_2 rotor current, Φ_σ — leakage flux [1].

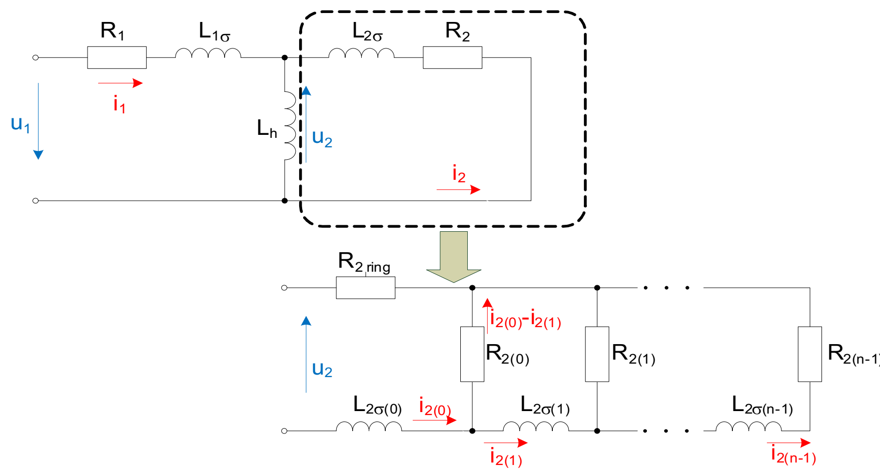


Figure 2: Common T-equivalent circuit diagram of an induction machine and its variation with the right side extended by RL-ladder network.

3. MODELING OF SKIN EFFECT

One method to model this slip-dependent effect is the determination of correction factors for sinusoidal supply and fundamental slot geometry characteristics [1, 4, 5]. This approach is sufficient in case of sinusoidal supply and for simple rotor bar geometries.

Apart from that, another option is modeling of the above defined partial coils by using lumped resistance and inductance elements in an RL-ladder network consisting of n stages, each with a resistance $R_{2(k)}$ and a leakage inductance $L_{2\sigma(k)}$ (see Fig. 1(right)) and replacing the right side of the T-equivalent circuit by this network [4, 6, 7], as illustrated in Fig. 2. This approach allows the numerical calculation of slip-dependent rotor parameters for more complex rotor bar geometries (e.g., double-cage induction machines) and non-sinusoidal feeding of the machine.

The concrete rotor bar geometry affects the resulting current displacement. Therefore the calculation of the bar-element's electrical parameters $R_{2(k)}$ and $L_{2\sigma(k)}$ is necessary. In the case of unknown rotor-bar geometry this can be done by an approximation assuming a homogenous rectangular rotor-bar divided into n elements. The resistance values $R_{2(k)}$ can be calculated as parallel connection and the inductance values can be calculated on closer examination of the current fragmentation in the RL-ladder network (see Fig. 2). In the first element flows the total current $i_{2(0)}$ and in every subsequent element (every k-th element) a current representing a portion of $(n-k)/n$ of the total current. Based on the law of energy conservation and the known equation for the partial sum $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$ the following calculation formula can be deduced:

$$L_{2\sigma(k)} = L_2 \cdot \frac{6n}{(n+1)(2n+1)} \tag{1}$$

In the case of known rotor bar geometry (esp. shape of the cross sectional area of the bars) approximating the real geometry is possible and more accurate. An example concerning modeling a double-bar geometry is given in Fig. 3. The parameters of the approximated rectangular elements can be calculated from geometry [8]:

$$\frac{R_{2(k)}}{l} = \frac{1}{\kappa \cdot \Delta x \cdot b_k} \qquad \frac{L_{2\sigma(k)}}{l} = \frac{\mu_0 \cdot \Delta x}{b_k} \tag{2}$$

with κ — electrical conductance, μ_0 — magnetic permeability, Δx — height of each element, b_k — width of one element, l — length of the rotor bars.

By extending the T-equivalent circuit as described, an implementation algorithm for computing the current $i_{2(0)}$ is convenient. This can be done using a system of state-space equations (Equation (3)) to express the mesh equations of the RL-ladder network:

$$\begin{aligned} \frac{di_{2(k)}}{dt} &= A \cdot i_{2(k)} + B \cdot i_{2(0)} \\ i_{2(1)} &= C \cdot i_{2(k)} + D \cdot i_{2(0)} \end{aligned} \tag{3}$$

with the state vector representing the partial currents of the ladder elements calculated by means of coefficient matrices (A, B) consisting of R/L terms.

With this state-space model the relevant partial current and with it the entire rotor current can be calculated for known rotor bar parameters. Using this state-space model the effect of current displacement is considered in a calculation algorithm realized in MATLAB- Simulink® by implementation of a specific motor model.

4. APPLICATION EXAMPLE: SIMULATION OF A DOUBLE-CAGE INDUCTION MACHINE

An industrial induction machine (Table 1) with double-cage rotor profile (Fig. 3) is the basis to parameterize the proposed model enabling the simulation of the non-steady-state behavior. The intention of using a double-cage rotor profile in induction motors is utilization of the skin effect by a special deep bar geometry (Fig. 3), to obtain a powerful starting torque.

To show the impact of modeling of current displacement, comparable simulations with consideration of skin effect (established motor model) and without consideration of skin effect (common motor model in Simulink) were performed. The simulation results of a direct start-up are illustrated in Fig. 4 by comparison of mechanical speed, electric torque and stator current.

As expected the ramp-up-time ($t = 0\,\mathrm{s} \ldots t = 0.1\,\mathrm{s}$) is shorter if the skin-effect is considered because of a higher starting torque. The oscillation of the torque is more damped in this case,

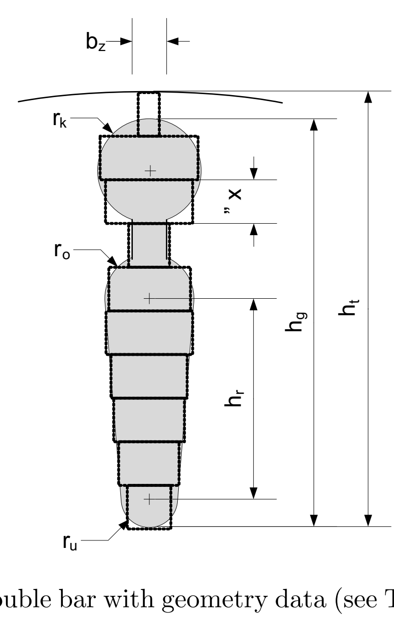


Figure 3: Profile section of a rotor double bar with geometry data (see Table 1) and illustrated approximation with rectangular elements ($n = 10$).

Table 1: Parameters of the used industrial induction motor.

rotor bar geometry:	Machine model 1LA5-207-4AA
$h_t = 30.85\,\mathrm{mm}$	(data for Y-connection)
$h_g = 29.60\,\mathrm{mm}$	$P_N = 30$ kW
$h_r = 16.45\,\mathrm{mm}$	$n_N = 1465$ U/min
$r_k = 3.05\,\mathrm{mm}$	$m_N = 196$ Nm
$b_z = 1.50\,\mathrm{mm}$	$p = 2$
$r_o = 2.75\,\mathrm{mm}$	$I_N = 55$ A
$r_u = 1.30\,\mathrm{mm}$	$J = 0.24\ \mathrm{Nm \cdot s^2}$
Number of ladder elements:	T-equivalent circuit parameters:
$n = 10$	$R_1 = 98\ \mathrm{m\Omega}$
	$R_2' = 62\ \mathrm{m\Omega}$
	$R_{Fe} = 50\ \mathrm{k\Omega}$
	$L_{1\sigma} = 1.235$ mH
	$L_{2\sigma}' = 1.146$ mH
	$L_{h3} = 35.3$ mH

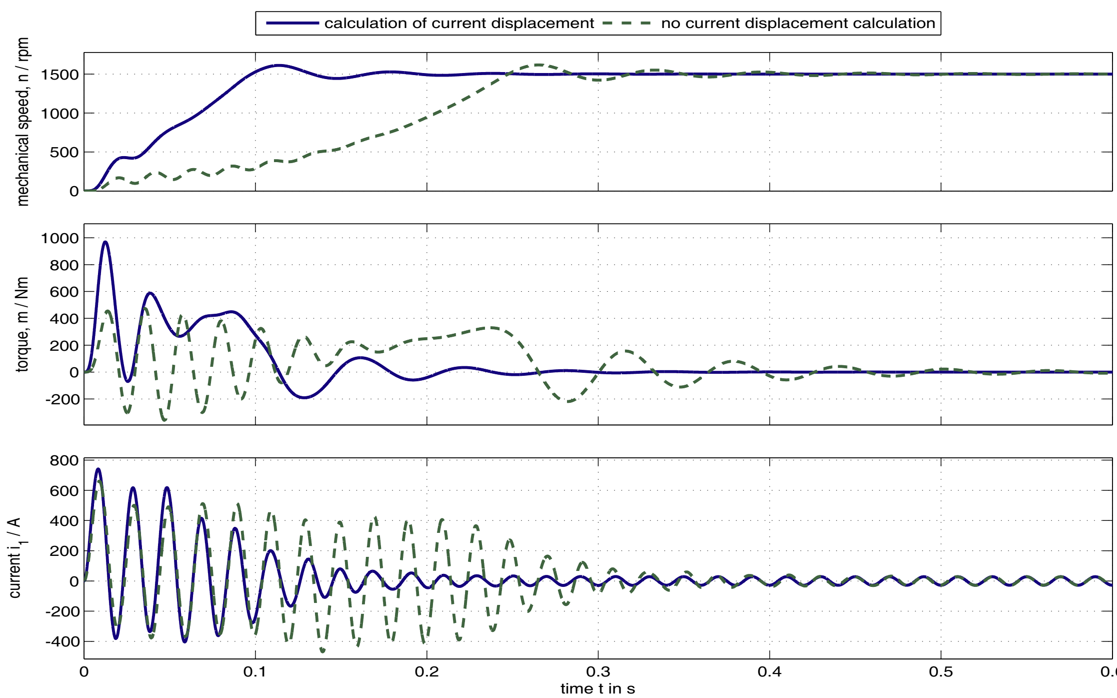


Figure 4: Comparison of machine models with (calculation of current displacement) and without consideration of skin effect (no current displacement calculation) by means of the example of a direct ramp-up, simulation of mechanical speed, electric torque and the stator current of one phase.

which yields a faster decline of the stator current amplitude. The simulation results of the skin effect model are in accordance with calculation results from a reference calculation tool and with data sheet specifications of inrush current ($i_{RMS} = 385\,\mathrm{A}$ corresponding to the RMS-value of the current i_1 between $t = 0\,\mathrm{s}$ and $t = 0.1\,\mathrm{s}$ in Fig. 4) and average starting torque ($M = 510\,\mathrm{Nm}$ corresponding to the average value between $t = 0\,\mathrm{s}$ and $t = 0.1\,\mathrm{s}$ in Fig. 4).

5. CONCLUSIONS

The paper presents a method to calculate the slip-dependent rotor current based on an extended T-equivalent circuit for non-sinusoidal power supply and arbitrary rotor-bar geometry. For complex rotor-bar geometries such as double-cage rotor profile these parameters can be obtained by approximation of the rotor bar cross-section using rectangle elements and deriving the $R_{2(k)}$ and $L_{2\sigma(k)}$ values for each element. The calculation algorithm has been realized in MATLAB-Simulink® by implementation of a specific motor model. The simulation results demonstrate the advantages of the proposed model in comparison to existing machine models.

REFERENCES

1. Mueller, G. and B. Ponick, *Grundlagen elektrischer Maschinen*, 9th Edition, Wiley-VCH Verlag GmbH und Co. KGaA, Weinheim, 2006.
2. Riefenstahl, U., *Elektrische Antriebssysteme*, 2nd Edition, Teubner Verlag, Wiesbaden, 2006.
3. Bottauscio, O., M. Chiampi, A. Manzin, and M. Zucca, "Additional losses in induction machines under synchronous no-load conditions," *IEEE Transactions on Industrial Electronics*, Vol. 40, No. 5, 3254–3261, 2004.
4. Oraee, H., "Numerical determination of skin-effect factors and their influence on starting performance of cage induction motors," *Electromotion*, Vol. 3, No. 3, 116–120, 1996.
5. Wuensch, D. and L. G. Del Castillo, "Erstellung eines EDV-Programmes zur Berechnung des dynamischen Verhaltens der Asynchronmaschine," *Technische Forschung Stahl*, EUR 8277 DE, 1984.
6. Babau, R., I. Boldea, T. J. E. Miller, and N. Muntean, "Complete parameter identification of large induction machines from no-load acceleration-deceleration tests," *IEEE Transactions on Industrial Electronics*, Vol. 54, No. 4, 1962–1972, 2007.
7. Seok, J.-K. and S.-K. Sul, "Pseudo-rotor flux oriented control for deep bar effect compensation," *IEEE Industry Applications Conference Record*, Vol. 1, 234–239, 1996.
8. Kuepfmueller, K. E. H., W. Mathis, and A. Reibiger, *Theoretische Elektrotechnik — Eine Einfuehrung*, Springer-Verlag Berlin, Heidelberg, 2008.

Direct Electromagnetic Torque Control of Induction Motors Powered by High Power PWM Inverters for Two Levels or Three Levels

M. R. Douiri[1], M. Cherkaoui[1], T. Nasser[2], and A. Essadki[3]

[1]Mohammadia School of Engineers (EMI), Morocco
[2]National School of Computer Science and System Analysis (ENSIAS), Morocco
[3]Superior School of Technical Education (ENSET), Morocco

Abstract— This study aims to develop a control strategy for high power induction motor capable of providing, in solicitations binding load torque, electromagnetic torque responses in wide dynamic. Direct torque control will achieve those goals. Indeed, by selecting from a table of switching vectors of the inverter output voltage, it imposes directly the states of power switches based on the electromagnetic state of the motor. Two applications are processed as part of this work. The first concerns the control of an induction motor fed by an inverter 2-voltage levels. The second develops new switching tables for direct torque control of induction motor fed by an inverter 3-voltage levels and structure of NPC. The characteristics of these tables justify the use of such a control strategy for systems implementing high power components such as GTOs. Particular attention is paid to maintaining the balance of the midpoint of the final structure of inverter.

1. INTRODUCTION

With the advancement of power electronics and digital technologies command, several control structures for the AC machines were proposed, in order get performance identical to those of the DC machine [1]. Among these structures, the direct torque control has been in recent years towards the most important research and best suited to industrial requirements [2, 3]. In addition, the development of variable speed control of induction machines has encouraged the use of three-level inverters. The increase in levels of the latter proves to be the best solution in high power drives. This structure of multilevel inverter was introduced by A. Naba and H. Akagi in 1981 [4–8], the aim was to reduce the amplitude of harmonics injected by the inverter. This term describes the connection point "O" through the diodes $S4'$, S^1 and $S1'$. The paper is organized as follows. The principle of classical DTC is presented in the second section. Section three and four describes the two level inverters and three level inverters fed DTC drive respectively. Section five presents the simulation results of the proposed DTC drive, compares them to those obtained with a classical DTC.

2. PRINCIPLE OF DIRECT TORQUE CONTROL

The principle of the command DTC is different. The objective is the direct regulation of the couple of the machine, by the application of the various vectors of tension of the inverter, which determines her state. The two controlled variables are: the flow statorique and the electromagnetic couple which who are usually commanded by regulators in hysteresis [1, 2]. It's about maintaining the greatnesses of statorique flux and the electromagnetic couple inside these bands of hystrsis. The output of this regulator determines the voltage vector of the optimal inverter to be applied to each switching instant. The use of this type of regulators supposes the existence of a frequency of switching in the variable converter requiring a step of very low calculation [2, 3].

Estimation of Stator Flux and Torque: The estimation of the flux and torque can be realized from the measures of the greatnesses stator current and motor voltage, we obtain the components α and β of the vector $\bar{\Psi}_s$ and Γ_e:

$$\begin{cases} \Psi_{s\alpha} = \int_0^t (V_{s\alpha} - R_s I_{s\alpha})dt \\ \Psi_{s\beta} = \int_0^t (V_{s\beta} - R_s I_{s\beta})dt \\ \Gamma_e = p_p(\Psi_{s\alpha} I_{s\beta} - \Psi_{s\beta} I_{s\alpha}) \end{cases} \tag{1}$$

Vector Control of Torque: The general expression of electromagnetic torque:

$$\Gamma_e = p_p \frac{L_m}{\sigma L_s L_r} \Psi_s \Psi_r \sin(\widehat{\bar{\Psi}_s \bar{\Psi}_s}) \tag{2}$$

The torque depends on the amplitude of both vectors $\bar{\Psi}_r$ and $\bar{\Psi}_s$ and their relative position. If we succeed in controlling perfectly the flux $\bar{\Psi}_s$ (from $\bar{V}_s$) in module and in position, we can thus control the amplitude and the relative position of $\bar{\Psi}_r$ and thus the torque.

Vector Control of Flux: The statorique flux of the asynchronous machine is obtained from following equation:

$$\bar{\Psi}_s = \int_0^t (\bar{V}_s - R_s \bar{I}_s) dt \tag{3}$$

3. DIRECT TORQUE CONTROL WITH TOW-LEVEL INVERTER

The switches of the inverter of voltage (see Figure 1) must be commanded so as to maintain the flux and the torque of the motor [5]. The vector of the voltage stator can be written under the shape:

$$V_s = \sqrt{\frac{2}{3}} E_0 \left(S_a + S_b e^{j\frac{2\pi}{3}} + S_c e^{j\frac{4\pi}{3}} \right) \tag{4}$$

where (S_a, S_b, S_c) represent the logical state of three switches: $S_i = 1$ mean that the high switch is closed and the low switch is opened ($V_i = +E_0$) and $S_i = 0$ mean that the high switch is opened and the low switch is closed ($V_i = -E_0$). We seek to control the flux and the torque via the choice of the vector of tension which will be made by a configuration of switches. As we have three switches, there are thus $2^3 = 8$ possibilities for the vector V_s. Two vectors (V_1 and V_8) correspond to the zero vector: $(S_a, S_b, S_c) = (0, 0, 0)$ and $(S_a, S_b, S_c) = (1, 1, 1)$.

Switching Table: The control table is built according to the state of variables $d\Psi$ and $d\Gamma_e$ and to the S_i zone and $\bar{\Psi}_s$ position, and so, it is shaped as presented in the Table 1.

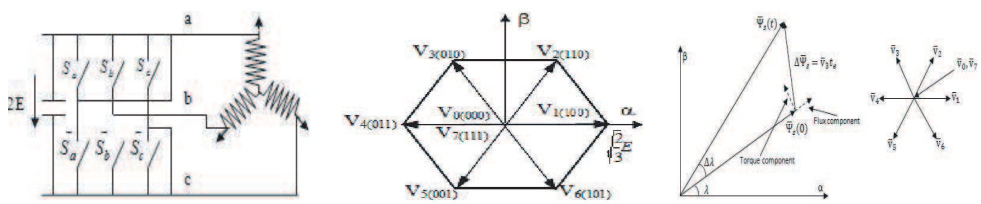


Figure 1: Partition of the $(\alpha; \beta)$ plane into six sectors and Schematic diagram of a two-level GTO inverter.

Table 1: Switching table for direct torque control (two levels inverter).

Sectors (S_i: i = 1 to 6)							
$d\Psi$	$d\Gamma$	S_1	S_2	S_3	S_4	S_5	S_6
1	1	V_2	V_3	V_4	V_5	V_6	V_1
	0	V_7	V_0	V_7	V_0	V_7	V_0
	−1	V_6	V_1	V_2	V_3	V_4	V_5
0	1	V_3	V_4	V_5	V_6	V_1	V_2
	0	V_0	V_7	V_0	V_7	V_0	V_7
	−1	V_5	V_6	V_1	V_2	V_3	V_4

4. DIRECT TORQUE CONTROL WITH TREE-LEVEL INVERTER

Figure 2 presents the general scheme of the inverter voltage three levels of structure called the neutral point "clamped" (NPC Neutral-Point-Clamped), it is one of the structures of inverter at three levels of tension [7, 8]. It has many advantages, such as the number of generated voltage is higher, less harmonic distortion and low frequency switching. Each arm of the inverter is comprised of four switches: $S_i, S_i^{'}, S_j, S_j^{'}$ Switches S_i and $S_i^{'}$ have a complementary function. The combination of four switches of the same arm ($S_i, S_i^{'}, S_j, S_j^{'}$), we can impose on the phase three levels of different tension: $(0, 0, 1, 1) \rightarrow -\frac{E}{2}$,$(0, 1, 1, 0) \rightarrow 0$, $(1, 1, 0, 0) \rightarrow \frac{E}{2}$ the combinations $(1, 1, 1, 0)$ and $(0, 1, 1, 1)$ realize a short circuit of the one both demies sources of continuous tension for it he are prohibited [8].

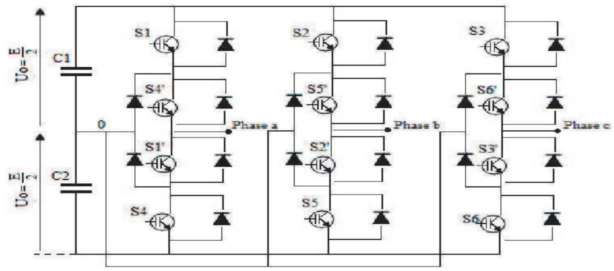


Figure 2: Structure of a three-level voltage inverter.

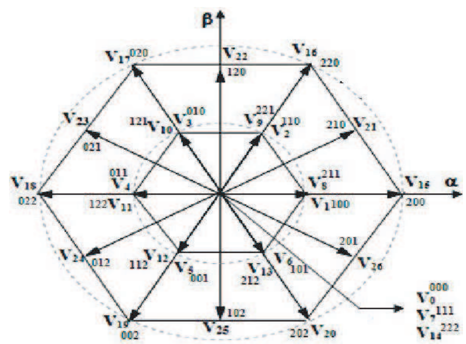


Figure 3: Hexagon of the voltages of a three-level inverter.

Vectors of output voltage of the inverter at three levels: The set of vectors voltages supplied by an inverter at three levels as well as sequences corresponding phase levels are shown in Figure 3.

The group of vectors "Zero voltage": they are obtained by three different combinations from states of three arms: $(1, 1, 1)$, $(-1, -1, -1)$ and $(0, 0, 0)$, and that we named respectively V_7, V_{14} and V_0. They have no influence on the voltage of the middle point of the inverter.

The group of vectors "voltage half": we can decompose this group into two other subgroups: The first one is constituted by vectors named V_1, V_2, V_3, V_4, V_5 and V_6. Other one is constituted by vectors $V_8, V_9, V_{10}, V_{11}, V_{12}$ and V_{13}. These vectors constitute the internal hexagon "voltage half". The application of a vector of the one or the other subgroup has an opposite effect on the evolution of the voltage of the middle point E, indeed, the application of a vector of the first subgroup (respectively of the second) will cause a discharge of input capacitor C_1 (respectively of capacitor C_2) [5, 7, 8].

The group of vectors "voltage full": this group contains vectors voltage named $V_{15}, V_{16}, V_{17}, V_{18}$, V_{19} and V_{20}. These vectors constitute the outside hexagon "voltage full". The voltage of the middle point middle E, is not affected by the application of these vectors, because the current which circulates in C_1 and in C_2 is the same.

The group of vectors "intermediate voltage": vectors voltage of this group are called V_{21}, V_{22}, V_{23}, V_{24}, V_{25} and V_{26}. During the application of these vectors, we cannot know if he is going to be to increase him or to decrease the tension of the middle point E, where the we are going to seek both capacitors, but currents which will cross them will not be equal. There will be an imbalance of E which depends on currents circulating in the phases during this functioning [4, 8].

The construction of switching tables (Tableau 2) is based on choice of the stator voltage vector applied to allow you to increase or decrease the modulus of the stator flux and electromagnetic torque value. A particular attention was dedicated to the synthesis of the table and to the comparators in hystrsis. In our case we use a hystrsis comparator in five level for the torque and at two levels for the regulation of flux in more we shall suppose that $U_{c1} = U_{c2} = \frac{E}{2}$.

5. SIMULATION RESULTS

Induction motor parameters: $P_n = 3\,\mathrm{Kw}$, $V_n = 220\,\mathrm{v}$, $R_s = 5.27\,\Omega$, $R_r = 5.07\,\Omega$, $L_s = 0.416\,\mathrm{H}$, $L_r = 0.423\,\mathrm{H}$, $L_m = 0.458\,\mathrm{H}$, $J = 0.2\,\mathrm{kg}\cdot\mathrm{m}^2$, $p = 2$.

Figures below represent the answer of the electromagnetic torque, flux statorique, and stator current for DTC 2-levels and DTC 3-levels. The reference torque Γ_e^* is a level of [7–20–8] and a reference flux of $\Psi_s^* = 1\,\mathrm{Wb}$. Figures 4(a) and 5(b), show that in the case of the inverter 3-levels, the good dynamics of the torque with fewer oscillations and overtaking of instruction, the torque follows

Table 2: Switching table for direct torque control (three levels inverter).

		Sectors (S_i: i = 1 to 12)											
$d\Gamma$	$d\Psi$	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}
-2	1	V_{20}	V_{26}	V_{15}	V_{21}	V_{16}	V_{22}	V_{17}	V_{23}	V_{18}	V_{24}	V_{19}	V_{25}
	0	V_{25}	V_{20}	V_{26}	V_{15}	V_{21}	V_{16}	V_{22}	V_{17}	V_{23}	V_{18}	V_{24}	V_{19}
-1	1	V_{13}	V_8	V_1	V_2	V_9	V_{10}	V_3	V_4	V_{11}	V_{12}	V_5	V_6
	0	V_5	V_6	V_{13}	V_8	V_1	V_2	V_9	V_{10}	V_3	V_4	V_{11}	V_{12}
0	1	V_0	V_7	V_{14}	V_0	V_7	V_{14}	V_0	V_7	V_{14}	V_0	V_7	V_{14}
	0	V_0	V_7	V_{14}	V_0	V_7	V_{14}	V_0	V_7	V_{14}	V_0	V_7	V_{14}
1	1	V_2	V_3	V_{10}	V_{11}	V_4	V_5	V_{12}	V_{13}	V_6	V_1	V_8	V_9
	0	V_3	V_4	V_{11}	V_{12}	V_5	V_6	V_{13}	V_8	V_1	V_2	V_9	V_{10}
2	1	V_{22}	V_{17}	V_{23}	V_{18}	V_{24}	V_{19}	V_{25}	V_{20}	V_{26}	V_{15}	V_{21}	V_{16}
	0	V_{17}	V_{23}	V_{18}	V_{24}	V_{19}	V_{25}	V_{20}	V_{26}	V_{15}	V_{21}	V_{16}	V_{22}

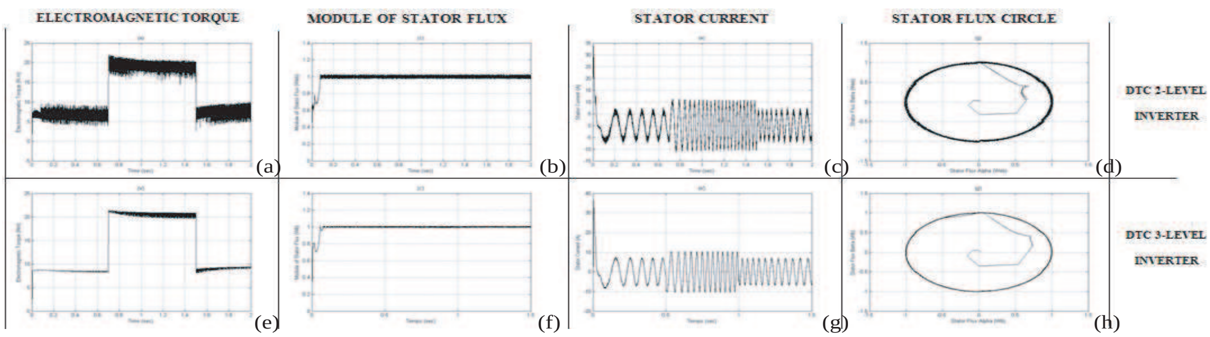


Figure 4: Comparison of the evolution electromagnetic torque,module of stator ux, stator current and stator Flux circle for DTC 2-level and DTC 3-level.

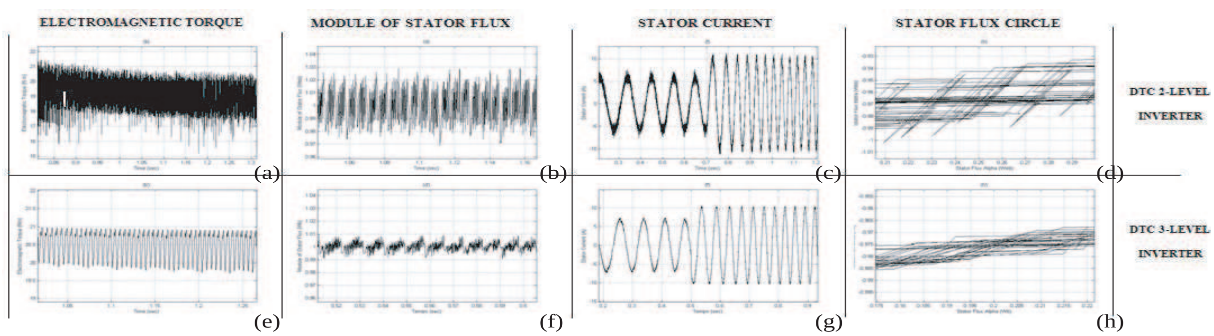


Figure 5: loupe: Comparison of the evolution electromagnetic torque,module of stator ux, stator current and stator flux circle for DTC 2-level and DTC 3-level.

perfectly its reference in regime establishes. According to Figures 4(c) and 5(d), we note that the establishment of the stator flux is a bit slower than the classical DTC Figures 4(c) and 5(d), but the plan continuous flux module presents a good response which is shown in Figure 4(c) and 5(d), where the evolution of the vector flux stator in the plan (α, β) is circular (4(g) and 5(h)). The Figures 4(e) and 5(f), show that the use of the inverter 3-levels entails a decrease of the undulations of the stator current, and the scheme of the current becomes purely sinusoidal.

6. CONCLUSION

A direct torque control of induction motor based on three levels and two levels inverter has been described. The system was analyzed, designed and performances were studied extensively by simulation to validate the theoretical concept. The main improvements by using a three levels inverter are:

- imitation of the current amplitude and low distortions for current and torque;
- No flux droppings caused by sector changes circular trajectory;
- Reduction in Flux, current and torque ripples,
- Stability of system.

REFERENCES

1. Vas, P., *Sensorless Vector and Direct Torque Control*, University Press, London, 1998.
2. Isao, T. and N. Toshihiko, "A new quick-response and high-efficiency control strategy of an induction mtor," *IEEE Transactions on Industry Applications*, Vol. 22, No. 5, 820–827, Sep./Oct. 1986.
3. Depenbrock, M., "Direct self-control (DSC) of inverter-fed induction machine," *IEEE Transactions on Power Electronics*, Vol. 3, No. 4, 581–588, Oct. 1998.
4. Noguchi, T., M. Yamamoto, S. Kondo, and I. Takahashi, "Enlarging switching frequency in direct torque — Controlled inverter by means of dithering," *IEEE Transactions on Power Electronics*, Vol. 35, No. 6, 1358–1366, Oct. 1998.
5. Lee,Y. H, B. S. Suh, and D. S. Hyan, "A novel PWM scheme for a three-level voltage source inverter with GTO thyristors," *IEEE Transactions on Power Electronics*, Vol. 33, No. 2, 260–268, 1996.
6. Trounce, J. C, S. D. Round, and R. M. Duke, "Comparison by simulation of three-level induction motor torque control schemes for electrical vehicle applications," *Proc. of International Power Engineering Conference*, Bejaia, Tunis, May 2001.
7. Wu, X. Q. and A. Steimel, "Direct self control of induction machines fed by a double three-level inverter," *IEEE Transactions on Industrial Electronics*, Vol. 44, No. 4, 519–527, Aug. 1997.
8. Lee, K. B, J. H Song, I. Choy, and J. Y. Yoo, "Improvement of low-speed operation performance of DTC for three-level inverter-fed induction motors," *IEEE Transactions on Industrial Electronics*, Vol. 48, No. 5, 1006–1014, Oct. 2001.
9. Escalante, M. F, J.-C. Vannier, and A. Arzand, "Flying capacitor multilevel inverters and DTC motor drive applications," *IEEE Transactions on Industrial Electronics*, Vol. 49, No. 4, 809–815, 2002.
10. Rodriquez, J., J.-S. Lai, and F. Z. Peng, "Multilevel inverters: A survey of topologies, controls, and applications," *IEEE Transactions on Industrial Electronics*, Vol. 49, No. 4, 724–738, Oct. 2002.

Broad Antireflection Grating by Apodization of One Dimensional Photonic Crystal

A. Mouldi, M. Kanzari, and B. Rezig
Laboratoire de Photovoltaïque et Matériaux Semi-conducteurs, ENIT
BP 37, Le Belvédère, Tunis 1002, Tunisie

Abstract— Dielectric multilayers can usually serve as perfect reflectors, so most researches are performed to broaden the optical band of the total reflexion. In this work, a new examination had been carried out to create a multilayer which acts as an antireflection system. It has been shown that applying an apodization to a one dimensional photonic crystal can significantly enhance the total optical transmission in the structure. The transmission characteristics of the apodized grating are analysed. So, in order to achieve a wideband of total transmission the number of layers, apodization profile and the reference wavelength should be optimized. The enhancement of optical transmission in this structure is very important since the performance of a number of optical devices could be improved.

1. INTRODUCTION

Currently, a wide range of modern integrated optoelectronics are based on photonic crystals (PCs) such us filters, waveguides, etc. Some applications like multiplexers, dispersion compensators, polarization filters, and image processors require a negligibly small reflection loss at the PC interface [1]. Therefore, applying antireflection interface structures, or antireflection coatings to the input and output interfaces of the PCs is important. This will also prevent unwanted reflections that can cause interference and cross talk between devices within PC-based compact integrated optical circuits [2]. Most of the studied structures are based on adding an antireflection coating to the photonic crystal, so the transmittance of the whole structure is increased [1–4]. Recent result shows that metallic-dielectric PCs themselves exhibit transparency in the visible and infrared spectra regions [5, 6]. So, it is interesting to realize structure presenting total transmission without adding antireflection coating and the structure itself operates similar to an antireflection coating. In this work, we propose a dielectric multilayer which is transparent in the visible light and near infrared region through a specific design based on the apodization. Apodization theory is frequently used for astronomy and for optics, for example in quantitative spectroscopy [7] or in fiber bragg grating in which the profile of gratings varies with the propagation distance [8–12]. In photonics, apodization is generally required to suppress the side lobes of the total reflection band and to widen it [8, 10–12]. Through the present work, apodization is used to create a band of total transmission and to enlarge it. A specific function applied to the refraction index of layers and having the form of a Gaussian apodization is exploited to enhance the transmission. So, each layer has a refractive index obeying to the Gaussian apodization according to its order through the grating. The transmission characteristics of the apodized grating are analysed. So, in order to achieve a wideband of total transmission, the number of layers and the apodization profile should be optimized. We define this band as the wavelength range when $R < 0.05\%$. The numerical method employed to obtain the reflection (or transmission) response of the structure is the transfer matrix method [13].

2. MODEL

The structure considered in this paper is formed by p layers and the refraction index of each layer is chosen to have the following form

$$n(j) = 1 + \exp[-(j - \mathrm{a})^2/b] \tag{1}$$

where j is the order of the layer in the grating, a and b are the apodization parameters. The optical thicknesses of the layers were taken quarter wavelengths. Therefore, geometrical thicknesses of the layers have the following form

$$d(j) = \frac{\lambda_0}{4 * n(j)} \tag{2}$$

In mathematics, Gaussian function takes the form:

$$f(x) = \alpha \exp\left[-\frac{(x-\beta)^2}{2\gamma^2}\right] \tag{3}$$

The graph of a Gaussian is a characteristic symmetric "bell curve" shape that quickly falls off towards plus/minus infinity. The parameter α is the height of the curve's peak, β is the position of the centre of the peak, and γ controls the width of the "bell". The form used in this work is

$$f(x) = 1 + \exp\left[-\frac{(x-\mathrm{a})^2}{b}\right] \tag{4}$$

If we compare this form to the general one of the Gaussian function, we can say that the function is composed of two parts, one is constant and equal to one, the second have a Gaussian profile with $\alpha = 1 \quad \beta = \mathrm{a} \; \gamma = \sqrt{\frac{b}{2}}$. So, it is a profile which tends to 1 at the edges of studied interval. Thus, refraction index are taken in the range 1-2. Dielectric materials with refractive indices involving in this range are for example SiO_2 (1.45), MgF_2 (1.38), MgO (1.78), Si_3N_4 (1.9) etc.

3. RESUTLS AND DISCUSSION

We purpose to minimize the number of layers with realising a broad high transmission band. So, we study the optical reflection by varying the layers number. We optimize the apodization parameters a and b for each number of layers. Table 1 presents the optimized couples (a, b) according to p where p varies between 6 and 21. These values were chosen to correspond to a reasonable optical reflection.

We note some regularity in a and b for the higher values of p. For $p \geq 11$, a generally takes the value $p - 2$ while b takes the values 17, 18 and 19. For $p \geq 14$, b keeps the value 18. Figure 1(a) shows some profiles of refractive indices for some p *values*. Table 1 was illustrated after numerical simulations of the optical reflection versus wavelength for some p, a and b values.

Figure 1(b) gives some optical reflection spectra obtained by apodization technique with λ_0 chosen to correspond to the middle of the visible spectrum (0.5 µm). It is worth noting that apodization enhanced the transmission through the grating. For $p = 6$, some narrow bandwidths appeared through the visible range and it is clearly that one band covers the reference wavelength 0.5 µm. For $p \geq 9$, by increasing the number of layers the curve flattens and the zero reflexion band becomes wider. Table 2 illustrates the band properties for different design (characterized by the triplet (p, a, b)), λ_{short} and λ_{long} are the wavelengths of respectively the lower and the upper band edges.

For $p \geq 10$, it has been found that p has no regular effect on the bandwidth but the curve becomes flatter by increasing p. From $p = 12$, we achieve a total transmission band covering almost

Table 1: Optimized apodization parameters for different numbers of layers.

p	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**
a	4	6	7	8	9	9	10	11	12	13	14	15	16	17	18	19
b	6	7	9	9	9	17	18	19	18	18	18	18	18	18	18	18

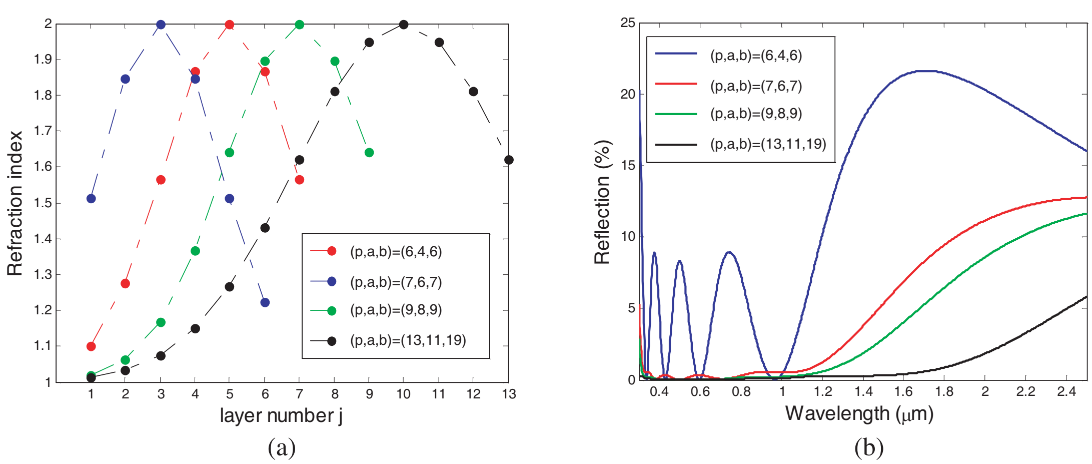


Figure 1: (a) Optimized refractive index profiles for different numbers of layers. (b) Reflection spectrum of the structure with reference wavelength 0.5 mm for different numbers of layers and refractive index profiles.

all the visible spectral range. The efficiency of apodization becomes maximum when we use 13 layers and the bandwidth attains 0.4314 μm. On the other hand for $p \geq 16$, no effect on the bandwidth was observed with regular 0.4148 μm value. So, by comparing these results with the classical Bragg dielectric mirror, the bandwidth increases by increasing the number of the layers. However, the bandwidth of the bragg mirror becomes stable for a critical p value [14]. It is clear, that the same result was found for the design simulated in the present work to have a total transmission band.

To have the maximum of the visible range covered by the total transmission band, we can adjust the reference wavelength. Figure 3 gives the response spectrum of the structure as a function of the central wavelength. We note that the band broaden and shifts towards the higher wavelengths when increasing the central wavelength (changing the central wavelength induces changing geometric

Table 2: The zero reflection band properties of apodized structure with reference wavelength 0.5 μm for different numbers of layers.

p	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
$\Delta\lambda$ (μm)	0.036	0.2462	0.3113	0.2723	0.0941	0.4252	0.4314	0.4252
λ_{short} (μm)	0.4825	0.4064	0.3894	0.3987	0.4575	0.366	0.3644	0.366
λ_{long} (μm)	0.5185	0.6527	0.7007	0.6707	0.5516	0.7912	0.7958	0.7912
p	**15**	**16**	**17**	**18**	**19**	**20**	**21**	
$\Delta\lambda$ (μm)	0.4094	0.4148	0.4148	0.4148	0.4148	0.4148	0.4148	
λ_{short} (μm)	0.3684	0.3676	0.3676	0.3676	0.3676	0.3676	0 .3676	
λ_{long} (μm)	0 .7778	0.7824	0.7824	0.7824	0.7824	0.7824	0.7824	

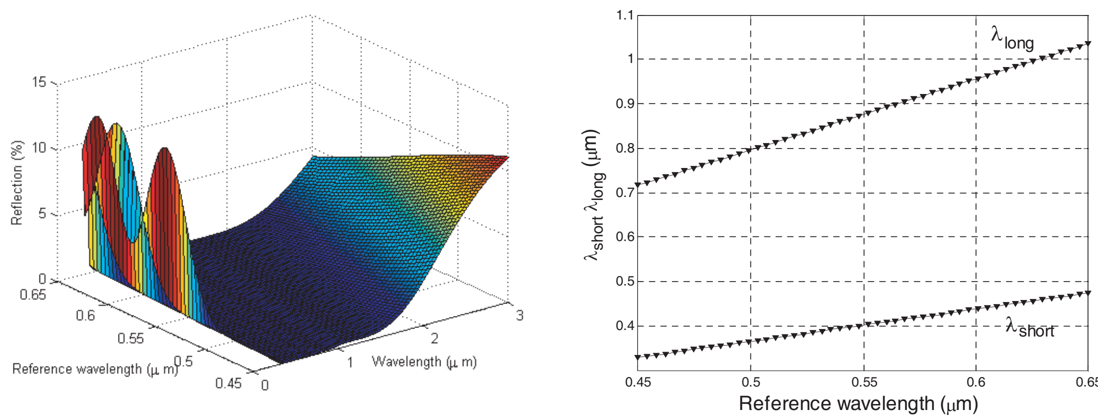


Figure 2: (a) Reflection spectrum of apodized structure with 13 layers as function of the reference wavelength. (b) The shift of the zero reflection band edges as function of the reference wavelength.

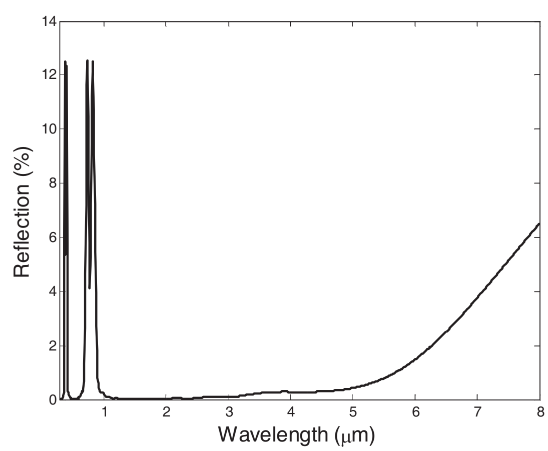


Figure 3: Reflection spectrum of apodized structure with 13 layers and reference wavelength 1.55 μm.

layers thicknesses since we keep the optical length of layers as quarter wavelength). So, we choose the central wavelength which permits to have the largest bandwidth in visible range. We can fix the reference wavelength at 0.55 μm.

We perform the same optimization process to realise a broad total transmission band in the near infrared range. For this, we take as reference wavelength the well known telecommunication wavelength 1.55 μm, the region in which almost all the world's optically transmitted information is passed. The optimum number of layers is also 13 for this case. With 13 layers, we can have a total transmission bandwidth of 1.3363 μm covering the range [1.1311, 2.4675] μm (Figure 3). This achievement is of military and commercial interest because the technique can be used to enhance or better transmit infrared images.

4. CONCLUSIONS

In this work, the reflectance properties of the proposed structure are different with the conventional multi-layered structure, because in this case dielectric multilayer gives a band structure with zero reflection under a specific apodized design. The reflectance properties of the grating with apodized structure are investigated in detail by numerical simulations. Apodization not only enhance transmission but also it permits the broadening of a total transmission band by optimization of the apodization parameters as well as the others parameters of the structure. The structure can be used as an antireflection coating which transmits a large range of wavelengths.

REFERENCES

1. Ushida, J, M. Tokushima, M. Shirane, A. Gomyo, and H. Yamada, "Systematic design of antireflection coating for semi-infinite one-dimensional photonic crystals using bloch wave expansion," *Appl. Phys. Lett.*, Vol. 82, No. 1, 2003.
2. Li, Z., E. Ozbay, H. Chen, J. Chen, F. Yang, and H. Zheng, "Resonant cavity based compact efficient antireflection structures for photonic crystals," *J. Phys. D: Appl. Phys.*, Vol. 40, 5873–5877, 2007.
3. Lee, S. G., J. S. Choi, J. E. Kim, and H. Y. Park, "Reflection minimization at two-dimensional photonic crystal interfaces," *Opt. Express*, Vol. 16, No. 6, 4270–4277, 2008.
4. Ohtera, Y., D. Kurniatan, and H. Yamada, "Antireflection coatings for multilayer-type photonic crystals," *Opt. Express*, Vol. 18, No. 12, 12249–12261, 2010.
5. Choi, Y. K., Y. K. Ha, J. E. Kim, H. Y. Park, and K. Kim, "Antireflection film in one-dimensional metallo-dielectric photonic crystals," *Optics Comm.*, Vol. 230, 239–243, 2003.
6. Chen, S., Y. Wang, D. Yao, and Z. Song, "Absorption enhancement in 1D Ag/SiO_2 metallic-dielectric photonic crystals," *Optica Applicata*, Vol. XXXIX, No. 3, 473–479, 2009.
7. Ceccherini, S., C. Belotti, B. Carli, and P. R. Chen, "Use of apodization in quantitative spectroscopy," *Opt. Lett.*, Vol. 32, No. 10, 1329–1331, 2007.
8. Qin, Z.-X., Q.-K. Zeng, D.-J. Feng, Y. Xiang, L. Ding, G.-Y. Kai, Z.-G. Liu, S.-Z. Yuan, X.-Y. Dong, N. Liu, C.-J. Liao, and S.-H. Liu, "Numerical study of the apodization profile functions, optimal profiles and lengths of a linearly chirped fibre bragg grating" *Chin. Phys. Lett.*, Vol. 18, No. 2, 239–241, 2001.
9. Liu, G. J., Q. Li, G. L. Jin, and B. M. Liang, "Transfer matrix method analysis of apodized grating couplers," *Optics Comm.*, Vol. 235, 319–324, 2004.
10. Tsai, T. Y., Y. C. Fang, Z. C. Lee, and J. R. Chen, "Design and simulation of apodized wavelength filters using Gaussian-distributed sidewall grating," *Opt. Quant. Electron.*, Vol. 39, 571–575, 2007.
11. Sun, N.-H., J.-J. Liau, Y.-W. Kiang, S.-C. Lin, R.-Y. Ro, J.-S. Chiang, and H.-W. Chang, "Numerical analysis of apodized fiber bragg gratings using coupled mode theory," *Progress In Electromagnetics Research*, Vol. 99, 289–306, 2009.
12. Sun, N. H. and J. J. Liau, "Design of gaussian apodized fiber bragg grating and its applications," *Int. J. Eng. Sci. Tech.*, Vol. 2, No. 5, 1419–1424, 2010.
13. Ben Ali, N. and M. Kanzari, "Designing of omni-directional high reflectors by using one-dimensional modified hybrid Fibonacci/Cantor band reflectors by using one-dimensional modified hybrid Fibonacci/Cantor band-gap structures," *J. Mod. Opt.*, Vol. 57, No. 4, 287–294, 2010.
14. Orfanidis, S. J., *Electromagnetic Waves and Antennas*, online book, 2004–2010.

High Frequency Back Scattering from a Real-scale Aircraft Using SBR and PTD-EEC Method

N. Altın[1] **and E. Yazgan**[2]

[1]Turkish Aerospace Industries, Inc.
Fethiye Mah., Havacılık Bulvarı No. 17, Kazan Ankara 06980, Türkiye
[2]Electrical & Electronics Engineering Department, Hacettepe University
Beytepe, Ankara 06532, Türkiye

Abstract— Monostatic Radar Cross Section (RCS) of a three dimensional target is calculated Shooting and Bouncing Ray (SBR) method and Physical Theory of Diffraction-Equivalent Edge Current (PTD-EEC) method. The aim of the presented algorithm is modeling of a real-scale aircraft, so the SBR ray tracing method is combined with PTD-EEC method, thus providing fast intersection routines as well as an accurate calculation of the resulting scattered field strengths. The size and complexity of the object is virtually unlimited. Calculated results with SBR and PTD-EEC methods are compared to Physical Optics Method.

1. INTRODUCTION

The electromagnetic scattering problem is an important subject in defense and aerospace applications. Realistic 3-D targets are modeled in computer aided design (CAD)-based computation environment, i.e., Unigraphics and/or Rhinoceros program. Models obtained from these programs have triangular mesh. First-order scattering, the shadowing and multiple bounce effects from 3-D real scale targets are investigated with SBR method which is one of the most powerful high frequency methods.

Recently several RCS calculation codes have been reported. RCS predicting algorithms are investigated and are developed software codes for large targets using method based on Physical Optics (PO) [15].

In the Shooting and Bouncing Ray (SBR) method [1–3], a set of parallel rays is launched from the incident direction toward the target. Each ray is traced as it bounces from one part of the target to another, until it exits the target. The field on the ray is calculated by the theory of geometrical optics, including the effects of ray tube divergence, polarization and material reflection coefficient. At the last reflection point on the target, a physical optics type integration of the induced surface currents that is performed to calculate the far field contribution from this ray tube. Contributions from all rays are summed up at a far field observation point to give rise to the final scattered field.

The far field contribution from a small ray tube is given formulation of the Huygens' Principle. The expression based on Huygens' Principle is useful in shooting and bouncing rays for solving complex scattering problems [2].

Effects of edge diffractions do not take into account in the SBR method. So, by including a routine that handles diffraction effects (i.e., a code based on Physical Theory of Diffraction (PTD)), the code can be used in stealth design applications. In the literature, quite a lot of work has been done and is the most widely used has done in 1984 by Michaeli' [1].

Complex targets as a F16, which requires a full three dimensional analysis, has not been studied in the literature. In this study, RCS of complex targets as a F16 is calculated SBR-PTD-EEC method for vertical and horizontal polarizations and the calculated results by this method have compared with physical optics results and only shooting and bouncing ray method. RCS of F16 has been investigated versus frequency, elevation and azimuth angle.

2. THEORY

2.1. Shooting and Bouncing Ray (SBR) Method Formulation

The SBR method is explained in detail in earlier publications [1–4]. In the present application, this method is applied over scattering surface and it is assumed that all of scattering surface is covered with Huygens' surface. The ray scattering formulation for validation is based on Huygens' Principle as in the two-dimensional case. A dense grid of geometrical optics (GO) rays representing the incident plane wave is shot toward the target. Whenever the rays cross the Huygens' surface S, the contributions of the rays to the scattered field are calculated by a ray-tube integration scheme.

Rays are traced according to the laws of geometrical optics as they bounce around the target. At the exit point of each ray, ray-tube integration is performed to sum up its contribution to the total scattered field [5]. While the basic idea behind the SBR methodology is simple, when combined with CAD tools for geometrical modeling and fast ray-tracing algorithms developed in computer graphics, this technique becomes a very general tool for characterizing the scattering from large, complex targets.

In this work in the RCS analysis with SBR method, ray window illuminated to target or scattering surface is located fixed xyz plane and the scattering surface is rotated.

At an observation point (r, θ, ϕ) in the far field, the contribution from this ray tube is expressed by

$$\vec{E}_s = \frac{e^{-jkr}}{r}[\hat{\theta}A_\theta + \hat{\phi}A_\phi] \tag{1}$$

When the SBR method is used, the contribution of the existing rays to the scattered field is given by:

$$\begin{bmatrix} A_\theta \\ A_\phi \end{bmatrix} = \frac{jk}{4\pi} \int\limits_{\substack{raytube\\projection\\on\ S}} e^{j\vec{k}_0 \cdot r} \left(\begin{bmatrix} -\hat{\phi} \\ \hat{\theta} \end{bmatrix} \times \vec{E}(r) + \eta \begin{bmatrix} \hat{\theta} \\ \hat{\phi} \end{bmatrix} \times \vec{H}(r) \right) \cdot \hat{n} ds' \tag{2}$$

$$\begin{bmatrix} A_\theta \\ A_\phi \end{bmatrix} = \frac{jk}{4\pi} \begin{bmatrix} A_\theta \\ A_\phi \end{bmatrix} = \sum_{i\, rays} \begin{bmatrix} B_\theta \\ B_\phi \end{bmatrix} \left(\frac{jk}{2\pi} \right) (\Delta A)_{exit} S(\theta, \phi) e^{jk \cdot r_A} \tag{3}$$

In the above expression, r_A is the position vector of point A where the ray-tube integration is carried out. Point A is usually chosen to be the last hit point on the target for the ray (see Fig. 2). $(\Delta A)_{exit}$ is the cross section of the exit ray tube at A and $S(\theta, \phi)$ is the shape function corresponding to the radiation pattern from the ray tube. $S(\theta, \phi)$ can usually be assumed to be unity if the ray tube area is sufficiently small, since the radiation from the ray tube will be nearly isotropic. B_θ, B_ϕ are explicitly related to the aperture fields at A as:

$$\begin{aligned} B_\theta = {} & 0.5[-s_1 \cos\phi E_3 - s_2 \sin\phi E_3 + s_3(\cos\phi E_1 + \sin\phi E_2)] \\ & +0.5Z_0[s_1(\cos\theta \sin\phi H_3 + \sin\theta H_2) + s_2(-\sin\theta H_1 - \cos\theta \cos\phi H_3) \\ & +s_3(\cos\theta \cos\phi H_2 - \cos\theta \sin\phi H_1)] \end{aligned} \tag{4}$$

$$\begin{aligned} B_\phi = {} & 0.5[s_1(\cos\theta \sin\phi E_3 + \sin\theta E_2) + s_2(-\sin\theta E_1 - \cos\theta \cos\phi E_3) \\ & +s_3(\cos\theta \cos\phi E_2 - \cos\theta \sin\phi E_1)] \\ & +0.5Z_0[s_2(\cos\phi H_3 + s_2 \sin\phi H_3) + s_3(-\cos\phi H_1 - \sin\phi H_2)] \end{aligned} \tag{5}$$

where $E(A) = E_1\hat{x} + E_2\hat{y} + E_3\hat{z}$ and $H(A) = H_1\hat{x} + H_2\hat{y} + H_3\hat{z}$ are respectively the electric and magnetic field associated with each ray at A, and $\hat{s} = s_1\hat{x} + s_2\hat{y} + s_3\hat{z}$ is the exit ray direction.

In summary, the far field contribution from a ray tube is given in Eq. (1) and Eq. (3). This result is obtained by applying the PO theory to the scattered field over the scatterer surface.

2.2. Physical Theory of Diffraction-Equivalent Edge Current (PTD-EEC) Method Formulation

3-D CAD model of complex target has edge and wedge. For scatterers with edge, it is important edge diffraction contribution is needed to add SBR method. In CAD based computation environment, the best method to implement edge diffraction is to use Michaeli's diffraction coefficient [1].

For a given scatterer geometry built by a CAD model, it is possible to identify edges automatically by checking the interior wedge angle between any two adjacent geometrical entries. In the CAD model, smooth surfaces are approximated by flat plates (i.e., half plane). If the angle between edges is different from zero, interior wedge angle is defined and it is considered as a "true wedge" and it will be used in subsequent edge calculation.

Once an edge is defined, the space curve that describes the edge is divided into linear segments, say, at a density of 10 segments per wavelength. For each segment, the following geometrical parameters are needed: (r_s, r_e) starting and ending position vectors of the segment, $(\hat{n}_1, \hat{n}_2)$ two unit outward normal of the wedge faces.

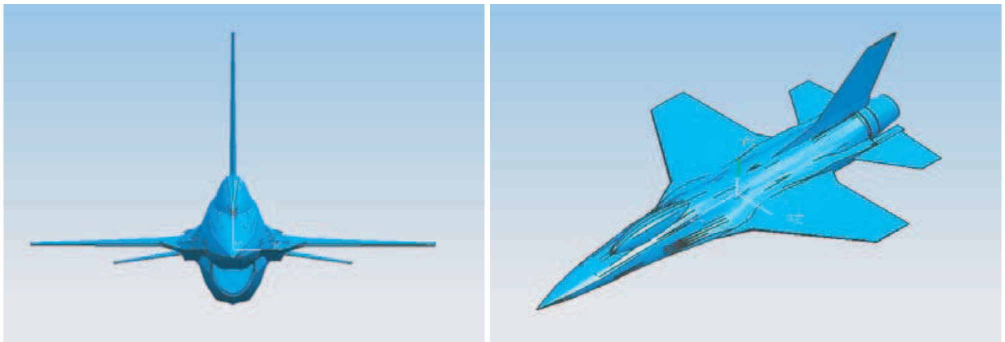
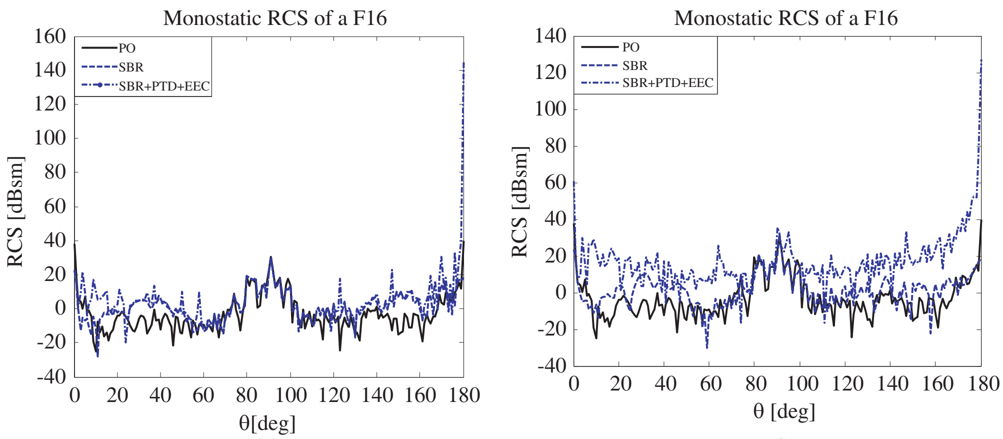


Figure 1: F16 solid model.



(a) (b)

Figure 2: RCS of F16 is calculated by PO, SBR and SBR-PTD-EEC (a) for vertical polarization, (b) for horizontal polarization.

The final edge diffracted field is given by,

$$E = \frac{e^{-jkr}}{r} \sum_{segments} \begin{bmatrix} f_{\theta\theta} & f_{\theta\phi} \\ f_{\phi\theta} & f_{\phi\phi} \end{bmatrix} \begin{bmatrix} E^i_\theta \\ E^i_\phi \end{bmatrix} \tag{6}$$

$$f_{uv} = \frac{dl}{4\pi} e^{jk(\hat{s}-\hat{l})\cdot d\hat{l}} [(\hat{u}\cdot\hat{t})(\hat{t}\cdot\hat{v})D_s + \hat{u}\cdot(\hat{t}\times\hat{s})(\hat{t}\times\hat{l})\cdot\hat{v}D_h + (\hat{u}\cdot\hat{t})(\hat{t}\times\hat{l})\cdot\hat{v}D_x] \tag{7}$$

Here $\hat{t}$ is the unit tangent vector along the direction $(r_s - r_e)$. Diffraction coefficient (D_h, D_s, D_x) are completely specified by the four geometrical parameters and the incident and scattered directions (i, s). Diffraction coefficients have been calculated by Michaeli' [1] formula. Explicit details are explained in [1]. The best available method for calculating the first order edhe diffraction is given by Eq. (4). In the sumamry, SBR results with the diffraction contribution are obtained the sum of Eq. (3) and Eq. (4).

3. NUMERICAL RESULTS

Monostatic RCS of complex target as an F16 is examined with SBR-PTD-EEC method and SBR method. Dimensions of F16 located in the xyz plane are $x = 7$ m, $y = 4$ m, $z = 10$ m, operating frequency is 10 GHz and diffraction angle is 90°. In Fig. 5, it shows 3-D F16 Unigraphics Model. The F16 is represented by 4092 triangle patches as shown in Fig. 4. In the SBR method, rays are shot towards the F16 with $\lambda/10$ density, where λ is wave length. The F16 was illuminated with 640.000 rays. For vertical and horizontal polarization, RCS according to θ elevation angle (0°–180°) is investigated and results obtained with SBR-PTD-EEC method have been compared to results obtained by Physical Optics (PO) and SBR method. Vertical polarization and horizontal polarization results are showed respectively in Fig. 5 and Fig. 6. In Fig. 7, both elevation (θ) and azimuth (ϕ) angles are 0°, RCS of F-16 is investigated versus frequency range (2–8 GHz) and results are compared to PO method.

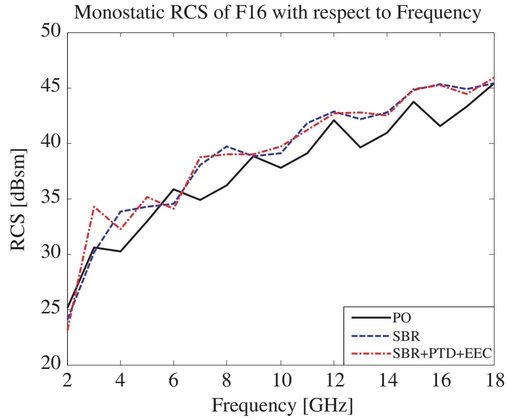
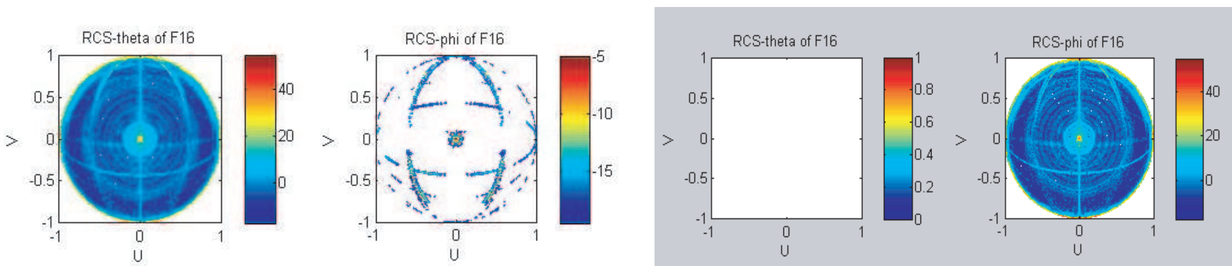


Figure 3: RCS of F16 is calculated by PO, SBR and SBR-PTD-EEC for vertical polarization while the azimuth angle is 90°.



(a) (b)

Figure 4: 3D result of RCS of F16. (a) For vertical polarization and (b) for horizontal polarization.

The results in Fig. 5(a) and Fig. 5(b) show the PO results, the SBR results without the diffraction contribution and the SBR results with the diffraction contribution for vertical and horizontal polarization, respectively. Ray tracing capability of the SBR code handles the multiple scattering and shadowing problems but SBR code is unable to model diffraction. Diffraction effects are calculated with PTD-EEC method.

In Fig. 5(a), PO, SBR and SBR-PTD-EEC results are perfectly matching between $\theta = 60°$ to 110°. For vertical polarization, in the elevation angles which diffraction is effective, the results are not matching well with the PO results. In Fig. 5(b), PO and SBR results are perfectly matching between $\theta = 60°$ to 110°. For horizontal polarization, results are not agreeing at angles where diffraction effects are dominant. Only PO, SBR and SBR-PTD-EEC results are perfectly matching between $\theta = 80°$ to 90° and diffraction effects are not important at these angles.

The F16 was illuminated from $\phi = 0°$ and $\theta = 0°$ and its monostatic RCS was computed with respect to frequency from 2 GHz to 18 GHz for vertical polarization. The results of the SBR and SBR-PTD-EEC are compared with the PO result and given in Fig. 7, When SBR result without diffraction is compared to SBR result with diffraction and diffraction is effective between 2–8 GHz as shown in Fig. 7.

Frequency is 10 GHz, elevation angle is (0°–180°), azimuth angle is (0°–360°), in Fig. 8 and in Fig. 9 show 3D RCS results for vertical and horizontal polarization respectively.

When 3D results are examined, it is shown that RCS value of F16 has large values at 0°, 90° and 180°.

4. CONCLUSION

As the conclusion of this work, we can say that equivalent edge current method which is suggested by Michaeli has given better result than other techniques used in the literature.

In conclusion, high frequency techniques are probably the optimum approach for analyzing very large and 3-D real targets. First-order scattering and multiple bounce effect from 3-D real scale targets are investigated with SBR method which is one of the most powerful high frequency methods. SBR method does not contain diffraction effects. Diffraction effects are examined PTD-EEC method and these effects are added to SBR results. Monostatic RCS of F16 with respect to frequency, angle and polarization are examined in our work. In addition to, three dimensional RCS of F16 are investigated for vertical and horizontal polarization.

REFERENCES

1. Ling, H., R. C. Chou, and S. W. Lee, "Shooting and bouncing rays: Calculating the RCS of an arbitrarily shaped cavity," *IEEE Trans. on Antennas and Propagat.*, Vol. 37, No. 2, 194–205, 1989.
2. Lee, S. W., H. Ling, and R. C. Chou, "Ray-tube integration in shooting and bouncing ray method," *Microwave and Optical Technology Letters*, Vol. 1, No. 8, 286–289, October 1988.
3. Ling, H., S. W. Lee, and R. C. Chou,, "High frequency RCS of open cavities with rectangular and circular cross section," *IEEE Trans. on Antennas and Propagat.*, Vol. 37, No. 5, 648–654, 1989.
4. Ling, H. and R. Bhalla, "Image-domain ray-tube integration formula for the shooting and bouncing ray technique," Air Force Wright Laboratory through NASA Grant NCC 3-273, July 1993.
5. Michaeli, A., "Equivalent edge currents for arbitrary aspects of observation," *IEEE Trans. on Antennas and Propagat.*, Vol. 32, No. 3, 252–258, 1984.
6. Michaeli, A., "Correction to equivalent edge currents for arbitrary aspects of observation," *IEEE Trans. on Antennas and Propagat.*, Vol. 33, No. 2, 227, 1985.
7. Umul, Y. Z., "Rigorous expressions for the equivalent edge currents," *Progress In Electromagnetics Research B*, Vol. 15, 77–94, 2009.

Thin Wires Structure for Decoupling of Multiple-antenna Terminals

I. B. Bonev, **O. Franek**, **and G. F. Pedersen**
Antennas, Propagation and Radio Networking Section
Department of Electronic Systems, Faculty of Engineering and Science
Aalborg University, Denmark

Abstract— In this paper, a thin wires (TW) structure has been used in order to reduce the coupling between two very close antennas on a common ground plane. Two dual-band planar inverted F antennas (PIFA) have been considered as the edge to edge distance at low band between both elements is $0.016\lambda_0$. After optimization of the location and the size of the TW structure, at low band, for the proposed two elements antenna we have achieved an isolation of 8.1 dB compared to the 3.1 dB for the initial antenna structure.

1. INTRODUCTION

When implementing Multiple Input Multiple Output (MIMO), the integration of several antennas operating in a diversity scheme at the base station side of the link is today a well-known and well-used solution. However, gathering together several antennas in a small device is much more complicated task due to the small allocated space for them and the highly complex fields within the antenna volume. Moreover, with the deployment of the Long Term Evolution (LTE) technology, the need to find efficient solution for isolating the antennas especially at low frequencies becomes extremely important. Consequently the academic research on multiple antennas operating at the same frequency, in a small communicating device is still of a great interest.

Identical planar inverted F antennas (PIFA) or monopoles-like antennas have been moved along the mobile phone printed circuit board (PCB) and different orientations have been investigated as well in [1, 2]. As could be expected, the best isolation values between the antennas were found for having the antennas spaced by the largest available distance, i.e., one at the top of the PCB and the other at the bottom. Several techniques to reduce the coupling between the antennas have been proposed [3–17]. Today future challenges concentrate on multi-antenna structures for the GSM 850/900 MHz [18, 19] and the long term evolution (LTE) 700 MHz bands [20]. At those low frequencies, indeed there is a huge challenge to isolate the ports of the antennas as the PCB is the unique radiator and the antennas only coupling elements. In this paper, we have proposed a thin wire (TW) structure to be placed in between two radiating elements of a PIFA type in order to reduce coupling between the antennas. An optimization of the best location of the TW structure and number of the elements within the TW structure has been performed using the finite-difference time-domain (FDTD) method [21]. The latter has been proved to be an efficient technique for solving complex electromagnetic problems [22].

2. ANTENNA CONFIGURATION

The investigated model with two PIFA antennas and the TW structure within the volume corresponding to a typical mobile phone handset are shown in Fig. 1.

The TW structure consists of periodically arranged wires with 1 mm distance between. In our study, the lateral number of the wires (M) has been fixed to 3, while the longitudinal number of wires (N) has been varied. The wires are 9 mm long and are connected to the ground plane at their bottom. An optical unit with size $40 \times 40 \times 10$ mm modeled by perfect electric conductor is attached to the ground plane. The aim of the optical unit is to avoid using conducting cables and, consequently, spoiling the antenna characteristics in measurement campaigns. An optical unit with that size has been designed and used during several measurement campaigns in Aalborg, Denmark as some more information and results can be found in [23]. In a real mobile phone, the place of the optical unit would be usually occupied by a battery of similar size, and the model would also be similar. The antennas resonate at 786 MHz and at 2.02 GHz which are very close to the LTE low and high bands. The edge to edge distance between both elements is $0.016\lambda_0$, where λ_0 is the wavelength in free space.

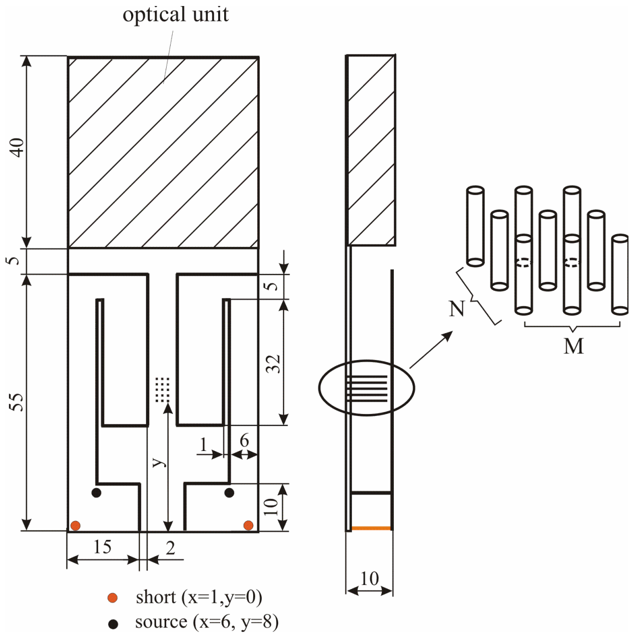


Figure 1: The handset model with two PIFA antennas and the decoupling TW structure.

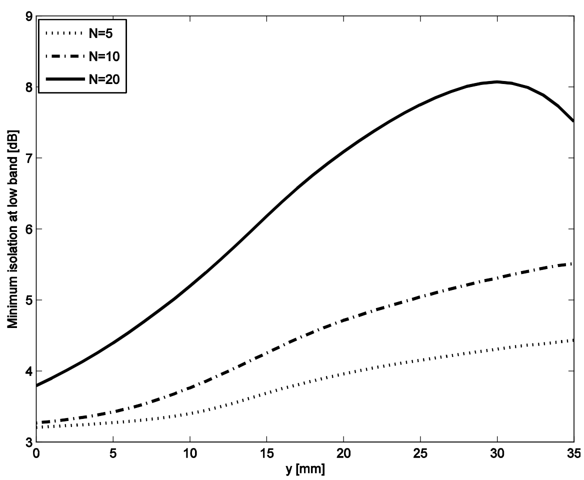


Figure 2: Minimum isolation of the proposed antenna at the low LTE band.

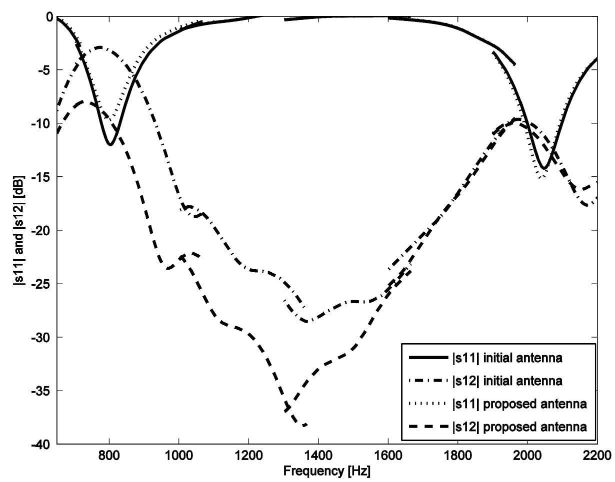


Figure 3: Simulated $|s_{11}|$ and $|s_{12}|$ of the initial and the proposed antenna (with the TW structure, y and N optimized).

Table 1: Minimum isolation of the initial antenna.

Low band	High band
3.1 dB	10.4 dB

3. NUMERICAL SIMULATIONS

In order to explain better the behavior of the $|s_{12}|$ we have defined a parameter called minimum isolation which is the maximum of the $|s_{12}|$ curve in the given frequency range. The minimum isolation of the initial antenna for both bands is shown in Table 1.

The results from the simulations for different distances y and sizes N of the TW structure are shown in Fig. 2.

The number of the wire elements N occurs to be an important parameter. The slope of the curves increases with increasing of y. Comparison of the simulated $|s_{11}|$ and $|s_{12}|$ between the initial antenna and the optimized one with the reduced coupling is shown in Fig. 3.

4. CONCLUSIONS

A thin wire structure for decoupling two PIFA antennas with edge to edge distance of $0.016\lambda_0$ at low LTE band has been proposed. The isolation has been improved from 3.1 dB to 8.1 dB. The future work includes investigation of this technique in the presence of a user's head and hand. Further, a planar version of the TW structure, supposed to be more easily manufactured, is under investigation. More optimizations of the TW structure dimensions will also be performed.

REFERENCES

1. Wong, K., J. Chou, S. Su, and C. M. Su, "Isolation between GSM/DCS and WLAN antennas in a PDA phone," *Microwave and Optical Technology Letters*, Vol. 45, No. 4, 347–352, May 20, 2005.
2. Su, C. M. and K. L. Wong, "Isolation between internal UMTS and WLAN antennas for a dual-mode wireless device," *Microwave and Optical Technology Letters*, Vol. 48, No. 10, 2001–2008, October 2006.
3. Ohishi, T., N. Oodachi, S. Sekine, and H. Shoki, "A method to improve the correlation and the mutual coupling for diversity antenna," *IEEE Antennas & Propagation Society Int. Symp.*, Washington, July 2005.
4. Shin, Y. S. and S. O. Park, "Spatial diversity antenna for WLAN application," *Microwave and Optical Technology Letters*, Vol. 49, No. 6, 1290–1294, June 2007.
5. Chiau, C. C., X. Chen, and C. G. Parini, "A compact four-element diversity-antenna array for PDA terminals in a MIMO system," *Microwave and Optical Technology Letters*, Vol. 44, No. 5, 408–412, March 5, 2005.
6. Yeap, S. B., X. Chen, J. A. Dupuy, C. C. Chiau, and C. G. Parini, "Integrated diversity antenna for laptop and PDA terminal in a MIMO system," *IEE Proc. Microwave Antennas Prop.*, Vol. 152, No. 6, 495–504, December 2005.
7. Yeap, S. B., X. Chen, J. A. Dupuy, C. C. Chiau, and C. G. Parini, "Low profile diversity antenna for MIMO applications," *IET Electronics Letters*, Vol. 42, No. 2, 69–71, January 19, 2006.
8. Kingsley, S. P., B. S. Collins, D. Iellici, J. W. Kingsley, S. G. O'Keefe, and T. J. Palmer, "Advancements in antennas: The development of dielectric antenna technology," *Antenna Systems 2003*, Denver, October 2003.
9. Villanen, J., J. Ollikainen, O. Kikevas, and P. Vainikainen, "Coupling element based mobile terminal antenna structures," *IEEE Transactions on Antennas and Propagation*, Vol. 54, No. 7, 2142–2153, July 2006.
10. Dossche, S., S. Blanch, and J. Romeu, "Optimum antenna matching to minimise signal correlation on a two-port antenna diversity system," *IET Electronics Letters*, Vol. 40, No. 19, 1164–1165, September 16, 2004.
11. Dossche, S., J. Rodriguez, L. Jofre, S. Blanch, and J. Romeu, "Decoupling of a two-element switched dual band patch antenna for optimum MIMO capacity," *Int. Symposium on Antennas and Propagations*, 325–328, Albuquerque, USA, July 2006.
12. Nilsson, A., P. Bodlund, A. Stjernman, M. Johansson, and A. Derneryd, "Compensation network for optimizing antenna system for MIMO application," *European Conference on Antennas and Propagations 2007, EuCAP07*, Edinburgh, UK, November 11–16, 2007.
13. Diallo, A., C. Luxey, P. Le Thuc, R. Staraj, and G. Kossiavas, "Study and reduction of the mutual coupling between two mobile phone PIFAs operating in the DCS1800 and UMTS bands," *IEEE Transactions on Antennas and Propagation*, Vol. 54, No. 11, Part I, 3063–3074, November 2006.
14. Diallo, A., C. Luxey, P. Le Thuc, R. Staraj, and G. Kossiavas, "Enhanced diversity antennas for UMTS handsets," *European Conference on Antennas and Propagations 2006, EuCAP2006*, Nice, France, November 6–10, 2006.
15. Diallo, A., C. Luxey, P. Le Thuc, R. Staraj, and G. Kossiavas, "Diversity performance of multi-antenna systems for UMTS cellular phones in different propagation environments," *International Journal on Antennas and Propagation (IJAP)*, Vol. 2008, Article ID 836050, 2008.
16. Hsu, C. C., K.-H. Lin, H.-L. Su, H.-H. Lin, and C.-Y. Wu, "Design of MIMO antennas with strong isolation for portable applications," *IEEE Antennas & Propagation Society Int. Symp.*, Charleston, June 2009.

17. Gummalla, A., M. Achour, G. Poilasne, and V. Pathak, “Compact dualband planar metamaterial antenna arrays for wireless LAN,” *IEEE Antennas & Propagation Society Int. Symp.*, San Diego, July 5–11, 2008.
18. Avendal, J., Z. Ying, and B. K. Lau, “Multiband diversity antenna performance study for mobile phones,” *IEEE Antennas & Propagation Society Int. Symp.*, Honolulu, USA, June 2007.
19. Plicanic, V., B. K. Lau, A. Derneryd, and Z. Ying, “Actual diversity performance of a multiband diversity antenna with hand and head effects,” *IEEE Transactions on Antennas and Propagation*, Vol. 57, No. 5, 1547–1556, May 2009.
20. Park, G., M. Kim, T. Yang, J. Byun, and A. S. Kim, “The compact quad-band mobile handset antenna for the LTE700 MIMO application,” *IEEE Antennas & Propagation Society Int. Symp.*, Charleston, June 2009.
21. Taflove, A., *Computational Electrodynamics: The Finite Difference Time Domain Method*, Artech House Publishers, Norwood, MA, 2005.
22. Wang, Y., I. B. Bonev, J. Nielsen, I. Kovacs, and G. F. Pedersen, “Characterization of the indoor multi-antenna body-to-body radio channel,” *IEEE Transactions on Antennas and Propagation*, Vol. 56, No. 12, April 2009.
23. Nielsen, J., B. Yanakiev, I. B. Bonev, M. Christensen, and G. F. Pedersen, “User influence on MIMO channel capacity for handsets in data mode operation,” accepted in *IEEE Transactions on Antennas and Propagation.*

Localized EBG Structure with DeCaps for Ultra-wide Suppression of Power Plane Noise

Jong-Hwa Kwon, Sang-Il Kwak, and Dong-Uk Sim
Radio Technology Research Group, Electronics and Telecommunication Research Institute (ETRI)
138 Gajeongno, Yuseong-gu, Daejeon 305-700, Korea

Abstract— The uni-planar electromagnetic bandgap (UC-EBG) structure is well known as a promising solution for suppressing power noise up to the GHz frequency range in high-speed digital systems. However, if a typical UC-EBG were adopted for the power and ground planes, the discontinuous reference plane would most likely degrade the high-speed signals passing over the EBG patterns because of the discontinuities of the etched reference plane. Also, EBG structures have a limitation in terms of the expansion of the lower bandgap frequency due to their physical size. In this paper, a localized EBG-patterned board with decoupling capacitors is proposed as a means of both suppressing noise propagation from extremely low frequencies and minimizing the effects of the discontinuous reference plane.

1. INTRODUCTION

In the multi-layer PCBs and packages of a high-speed digital system, a stable power distribution network (PDN) is needed to supply clean power to the core logic and I/O circuits. One of the major obstacles to building a stable PDN is known to be the simultaneous switching noise (SSN) generated in the power/ground (P/G) planes [1]. SSN can excite resonance modes between the parallel-plate waveguide-type P/G planes. As a result, the generated cavity resonance modes can lead to significant signal/power integrity (SI/PI) problems and electromagnetic interference (EMI) issues [2, 3]. Recently, a uni-planar compact electromagnetic bandgap (UC-EBG) structure was proposed as a promising means of effectively eliminating SSN [4–6]. However, if the UC-EBG is adopted for the P/G plane, problems of signal integrity may result from the high-speed signals crossing over the EBG patterns, because the perforated reference plane perturbs the return current flows. Also, a typical UC-EBG structure causes the resonance frequencies of the planes to move to a lower frequency range, which is outside the EBG's bandgap. So, adopting an EBG structure may make it worse at a lower frequency range below several hundred MHz [7–9].

This letter proposes a localized EBG board with decoupling capacitors ('DeCaps' hereafter) located only near the sources of noise as a means of both suppressing the power noise from extremely low frequencies and of minimizing the effects of the discontinuous reference plane. The SSN suppression performance of the proposed structure was validated and investigated by both simulation and measurement.

2. TEST BOARD DESCRIPTION

The main objective of this study is to investigate the effect of the localized EBG unit cell with DeCaps located near the noise source, on the assumption that the noise sources exist only in a specific area. Thus, a well-known UC-EBG structure was used as the unit cell [4–6]. Figures 1(a) and (b) show the unit cell structure and the side view of the proposed PCB board, which had the localized EBG unit cells and DeCaps on the power plane. The unit cells were etched onto the power plane with their corresponding geometrical parameters, $d = 30\,\mathrm{mm}$, $a = 28.5\,\mathrm{mm}$, $g = 1\,\mathrm{mm}$, $l_1 = 6.5\,\mathrm{mm}$, $l_2 = 7.5\,\mathrm{mm}$, $w_1 = 1.5\,\mathrm{mm}$, and $w_2 = 1.5\,\mathrm{mm}$, respectively. The dimensions of the two-layer PCB were $180\,\mathrm{mm} \times 180\,\mathrm{mm}$, with a 1.0 mm FR4 ($\varepsilon_r = 4.5$) substrate. The real decoupling capacitor behaves as a series RLC resonant circuit [10]. To suppress SSN at an extremely low frequency, 8 SMT-type DeCaps were positioned around the noise source, Port 1. The parameters' values of the used DeCap are $C = 100\,\mathrm{nF}$, $L_{ESL} = 0.55\,\mathrm{nH}$, and $R_{ESR} = 0.02\,\Omega$.

As shown in Figure 2, the three types of test PCBs used to verify the proposed structure's performance were designed and fabricated as follows: (a) reference plane, (b) localized EBG plane, and (c) localized EBG plane with 8 decoupling capacitors around port 1. The four ports were located as follows: P1 (45, 135, 0 mm); P2 (135, 45, 0 mm); P3 (45, 45, 0 mm); and P4 (135, 135, 0 mm). The point of origin was situated in the lower left-hand corner of the PCB, as shown in Figure 2. These ports were used to evaluate the insertion loss between the ports by both simulation and measurement.

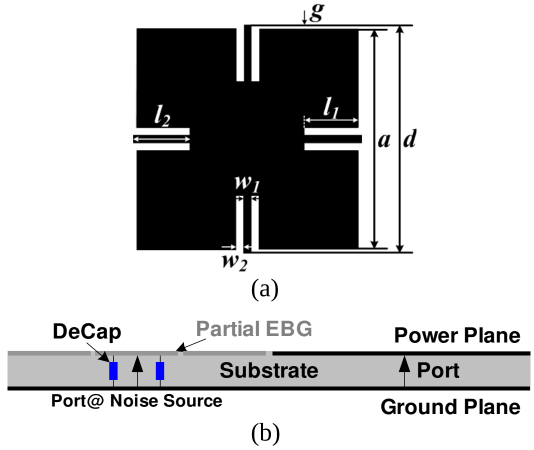


Figure 1: (a) EBG unit cell structure and parameters. (b) Side view of PCB with localized EBG unit cells with DeCaps.

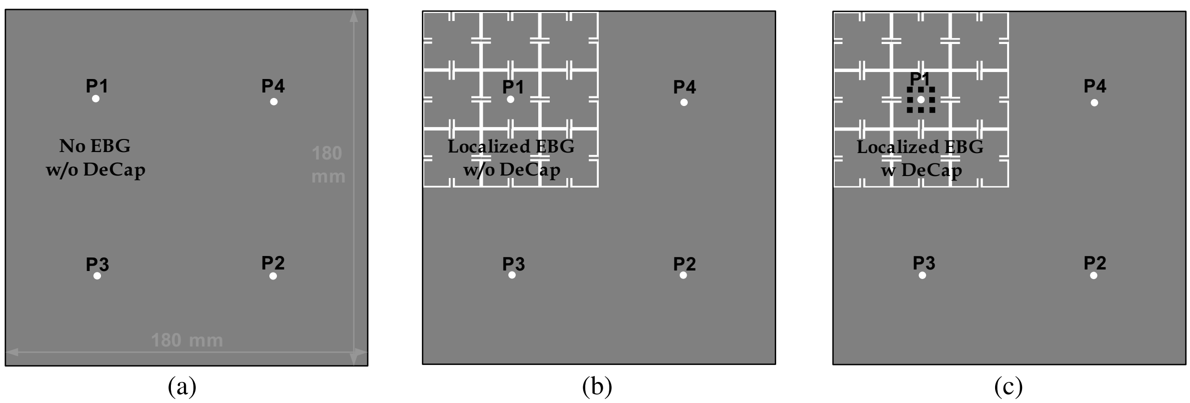


Figure 2: Test boards. (a) Reference board. (b) Localized EBG board. (c) Localized EBG with DeCaps on the power plane.

3. SIMULATED AND MEASURED RESULTS

The SSN suppression property can be confirmed with the insertion loss between the ports in the frequency domain. A 3D EM simulation tool (HFSS from Ansoft, USA) [11] and a vector network analyzer (Agilent 8236B) were used to obtain the simulated and measured insertion losses between the ports, respectively. Figures 3 and 4 show the simulated and measured insertion loss, i.e., S_{21} for the proposed localized EBG-patterned power planes with DeCaps located near the noise source, Port 1. The insertion losses of the reference board and the localized EBG board without DeCaps are also presented in the figures for comparison. As shown in the figures, reasonable agreement is obtained from DC to 5 GHz between the measurements and simulations.

The stop bandwidth was defined by 'S_{21} lower than -30 dB' for the purposes of this study. As shown in Figures 3 and 4, the suppression property of the proposed structure superposes the properties of the EBG planes and DeCaps. That is, the stop bandwidth of the proposed structure is the sum of the stop bands due to both the DeCaps at the lower frequency range and the EBG cells at the higher frequency range; and the average level of suppression within the forbidden bandgap is higher than that of the power plane with only the localized EBG unit cells.

In the case of the localized EBG board without DeCaps, the SSN is sufficiently reduced at the stop band ($0.95 \sim 4.4$ GHz) of the UC-EBG structure. But the resonances below the stop band in the EBG-patterned board become more severe than those of the reference board because a typical UC-EBG structure causes the resonance frequencies of the planes to move to a lower frequency range. To reduce the SSN at the lower frequency range, DeCaps are used in this study. It is clearly seen in the figures that the localized EBG board with DeCaps mitigates sufficiently the propagation of power noise across the entire frequency range from DC to 5 GHz. The stop bandwidth was defined by 'S_{21} lower than -30 dB' for the purposes of this study. As shown in

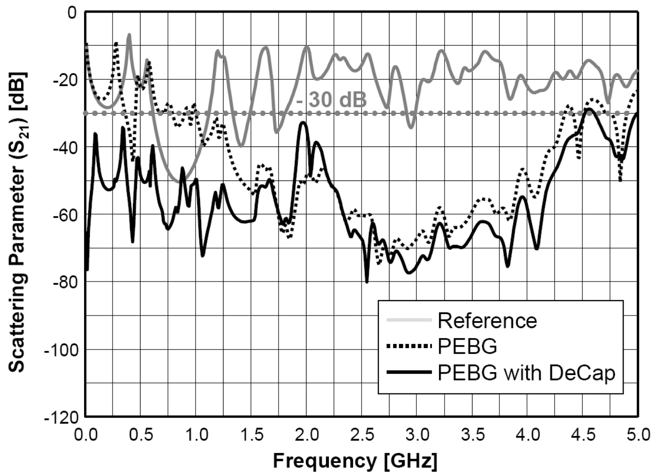


Figure 3: Simulated S_{21} between the fabricated EBG boards and the reference board.

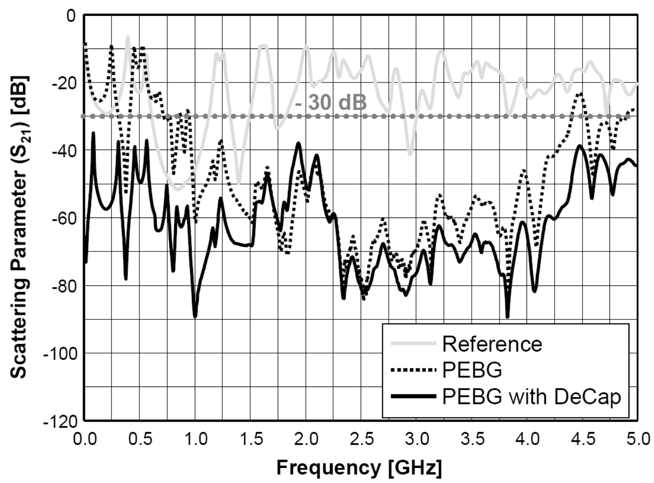


Figure 4: Measured S_{21} between the fabricated EBG boards and the reference board.

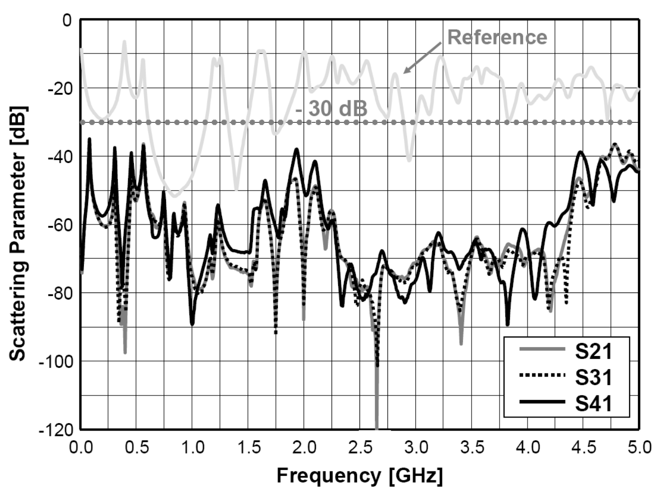


Figure 5: Measured SSN suppression behavior for the different receiving ports located on the test PCB with localized EBG unit cells with DeCaps.

Figures 3 and 4, the suppression property of the proposed structure superposes the properties of the EBG planes and DeCaps. That is, the stop bandwidth of the proposed structure is the sum of the stop bands due to both the DeCaps at the lower frequency range and the EBG cells at the higher frequency range; and the average level of suppression within the forbidden bandgap is higher than that of the power plane with only the localized EBG unit cells.

Figure 5 shows the measured SSN suppression behavior of the proposed localized EBG board with DeCaps for the receiving ports located in other different locations, namely ports 3 and 4. The noise excitation was at port 1. The insertion loss at port 2 is also presented in this figure for comparison. Although there are no EBG unit cells in the positions of ports 3 and 4, it was found that the SSN is still suppressed sufficiently in the ultra-wide frequency band from DC to 5 GHz. This behavior indicates that the proposed design is capable of suppressing the SSN on the PDN in an omni-directional manner.

In this study, only a portion (25%) of the power plane is used to adopt the EBG unit cells. So, if the solid part of the proposed power plane were used as the return current path for a high-speed signal, the SSN could be sufficiently suppressed and the signal quality improved to a greater extent than the conventional EBG-patterned PDN with fully located unit cells.

4. CONCLUSIONS

In this paper, the localized EBG patterned power plane with decoupling capacitors (DeCaps) located only near the noise source is proposed as a means of both suppressing the noise propagation from DC to several GHz and of minimizing the effect of the discontinuous reference plane. From the results obtained from simulation and measurement, it is seen clearly that the proposed localized EBG board with DeCaps suppresses sufficiently the propagation of power noise across the entire frequency range from DC to 5 GHz. Furthermore, the remaining solid part of the localized EBG PDN can be used as a stable return current path for high speed signals to minimize the effect of discontinuous EBG-patterned PDN.

REFERENCES

1. Swaminathan, M., J. Kim, I. Novak, and J.P. Libous, "Power distribution networks for system-on-package: Status and challenges," *IEEE Trans. Advanced Packaging*, Vol. 27, No. 2, 286–230, May 2004.
2. Berghe, S. V. D., F. Olyslager, D. D. Zutter, J. D. Moerloose, and W. Temmerman, "Study of the ground bounce caused by power plane resonances," *IEEE Trans. Electromagnetic Compatibility*, Vol. 40, 111–119, May 1998.
3. Lei, G.-T., R. W. Techentin, and B. K. Gilbert, "High frequency characterization of power/ground-plane structures," *IEEE Trans. MTT*, Vol. 47, No. 5, 562–569, May 1999.
4. Wu, T. L., Y. H. Lin, and S. T. Chen, "A novel power plane with super-wideband elimination of ground bounce noise on high speed circuits," *IEEE Microwave Wireless Compon. Lett.*, Vol. 15, No. 3, 174–176, Mar. 2005.
5. Wu, T. L., Y. H. Lin, T. K. Wang, C. C. Wang, and S. T. Chen, "Electromagnetic bandgap power/ground planes for the wideband suppression of ground bounce noise and radiated emission in high-speed circuits," *IEEE Trans. MTT*, Vol. 53, No. 9, 2935–2942, Sept. 2005.
6. Kim, B. and D.-W. Kim, "Bandwidth enhancement for SSN suppression using a spiral-shaped power island and a modified EBG structure for a $\lambda/4$ open stub," *ETRI Journal*, Vol. 21, No. 2, 201–208, Apr. 2009.
7. Kwon, J. H., D. U. Sim, S. I. Kwak, and J. G. Yook, "Partial placement of electromagnetic bandgap unit cells to effectively mitigate simultaneous switching noise," *Electronics Letters*, Vol. 44, No. 22, 1302–1303, Oct. 2008.
8. Kwon, J. H. and J. G. Yook, "Partial placement of EBG on both power and ground planes for broadband suppression of simultaneous switching noise," *IEICE Trasn. Commun.*, Vol. E92-B, No. 7, 2550–2553, Jul. 2009.
9. Kwon, J. H., S. I. Kwak, D. U. Sim, and J. G. Yook, "Partial EBG structure with decap for ultra-wideband suppression of simultaneous switching noise in a high speed system," *ETRI Journal*, Vol. 32, No. 2, 265–272, Apr. 2010.
10. Xu, M., T. H. Hubing, J. Chen, T. P. Van Doren, J. L. Drewniak, and R. E. DuBroff, "Power-bus decoupling with embedded capacitance in printed circuit board design," *IEEE Trans. EMC*, Vol. 45, No. 1, 22–30, Feb. 2003.
11. HFSS V11.0, www.ansoft.com, Ansoft, USA.