Basic Experiments with Model of Inductive Flowmeter

P. Fiala[1], **V. Sadek**[1], **P. Dohnal**[2], **and T. Bachorec**[1]

[1]Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejni 2906/4, 612 00 Brno, Czech Republic
[2]Department of Languages, Brno University of Technology
Udolni 53, 602 00 Brno, Czech Republic

Abstract— This article deals the numerical modelling and experimental verification of physical and chemical processes during measurement with inductive flowmeter. There are presented theoretical model and example numerical solution with the comparison of magnetic field measurement. The numerical models are based on combined finite element method (FEM) and finite volume method (FVM) of flowmeter electrodes. The model joins magnetic, electric and current field, flow field and chemical nonlinear ion model. Results were compared with accuracy magnetic field measurement inside of body flowmeter.

1. INTRODUCTION

The full magneto-hydro-dynamical (MHD) model of inductive flowmeter is coupled problem. There are coupled magnetic, fluid flow field and electric circuit and chemical (ions) models, Fig. 1. These models are solved with different view. The one of them is the analytical solution of flowmeter model with concentrated parameters application, solved in [1] and in further to use solution with electronics [2]. Complete MHD numerical model formulated with partial differential equations was published in research report [3] and paper [4]. Flowmeter calibration to zero drift before known was described in journal [5].

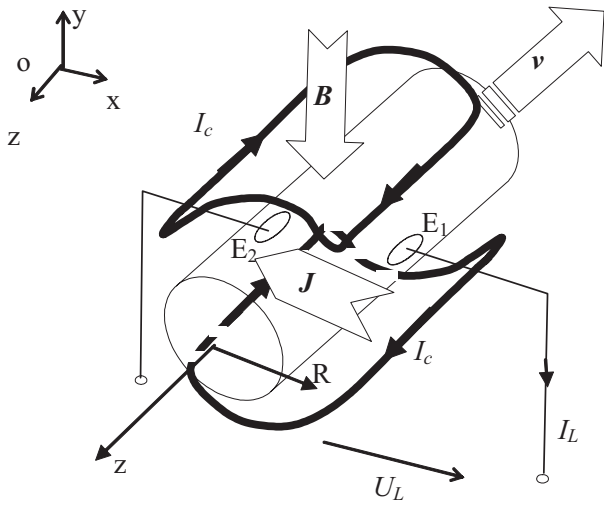


Figure 1: Principle of the induction flowmeter.

Our model was solved with combined finite element methods (FEM) and finite volume methods (FVM) like was presented in [4]. Results from numerical model and experiments were compared and the numerical equality was very good.

2. NUMERICAL MODEL

According to research report [3] electromagnetic part of the flowmeter is derived from reduced Maxwell equations

$$rot\boldsymbol{H} = \boldsymbol{J}, \tag{1}$$

$$div\boldsymbol{B} = 0, \tag{2}$$

where $\boldsymbol{H}$ is vector of magnetic field intensity, $\boldsymbol{B}$ is vector of magnetic field induction, $\boldsymbol{J}$ is vector of current density.

$$rot\boldsymbol{E} = 0, \tag{3}$$
$$div\boldsymbol{J} = 0, \tag{4}$$

where $\boldsymbol{E}$ is vector of electric field intensity. Material properties are represented by equation

$$\boldsymbol{E} = \boldsymbol{H}\mu_0, \tag{5}$$
$$\boldsymbol{J} = \boldsymbol{E}\gamma, \tag{6}$$

where μ_0 is permeability of vacuum, is specific conductance of measured liquid. Vector functions of electric and magnetic field are expressed by means of scalar electric ϕ_e and magnetic potentials ϕ_m

$$\boldsymbol{E} = -\mathrm{grad}\,\phi_\mathrm{e}, \tag{7}$$
$$\boldsymbol{H} = -\mathrm{grad}\,\phi_\mathrm{m}. \tag{8}$$

Final current density from (4) $\boldsymbol{J}$ is influenced by velocity $\boldsymbol{v}$ of the flowing ions solution and outer magnetic field

$$\boldsymbol{J} = \gamma(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}). \tag{9}$$

If electrodes E_1 and E_2 have different electrical potentials (Fig. 2) then current density $\boldsymbol{J}$ is created in the Ω area according to (9) and current I_L flows in the ion solution

$$I_L = \iint\limits_{\boldsymbol{S}_e} \boldsymbol{J} \cdot d\boldsymbol{S} = \iint\limits_{\boldsymbol{S}_e} \gamma(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \cdot d\boldsymbol{S} \tag{10}$$

where $\boldsymbol{S}_e$ is a directed area of electrodes E_1 and E_2 into space Ω. We obtain voltage between flowmeter electrodes E_1, E_2 from

$$U_L = \int\limits_{E_1}^{E_2} \boldsymbol{E} \cdot d\ell, \tag{11}$$

where electric field intensity is derived from the force $\boldsymbol{F}$ which affects a charge q. Current density $\boldsymbol{J}(\boldsymbol{v})$ depends on immediate ion velocity between E_1 and E_2. After modification voltage on flowmeter electrodes is

$$U_L = \iiint\limits_{\Omega} \left(\frac{\boldsymbol{J}(\boldsymbol{v})}{I_L} \times \boldsymbol{B} \right) \cdot (\boldsymbol{v}_{io} + \boldsymbol{v}) dV. \tag{12}$$

The model of fluid velocity distribution is derived for incompressible fluid and was able to formulate from balance of forces the Navier-Stokes equation for the fluid element

$$\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \mathrm{grad}\,\boldsymbol{v} = \boldsymbol{A} - \frac{1}{\rho}\mathrm{grad}\,p + \boldsymbol{v} \cdot \Delta\boldsymbol{v}, \tag{13}$$

where $\boldsymbol{A}$ is an external acceleration and $\boldsymbol{v}$ kinematic viscosity. Next step was derivation of FEM and FVM model [3]. The final term for output voltage on flowmeter electrodes which was evaluated is

$$U_L \cong \left(\frac{1}{|ik^+|} + \frac{1}{|ik^-|} \right) \frac{1}{2F_c^2 I_L} \sum_{e=1}^{N_\Omega} \frac{|\boldsymbol{J}_2|}{\Delta V_e^2 \sqrt{\left(|\boldsymbol{v}_{m,\,e}|^2 + \left(\frac{|\boldsymbol{v}_{ok,\,e}^+| + |\boldsymbol{v}_{ok,\,e}^-|}{2} \right)^2 \right)}} (\boldsymbol{J}_e \times B_e)$$
$$\cdot \left(\boldsymbol{J}_e \left(\frac{1}{|ik^+|} + \frac{1}{|ik^-|} \right) + \Delta V_e F_c \boldsymbol{v}_{m,\,e} \right) \tag{14}$$

where

$$\boldsymbol{v}_{ok,\,e}^{+} = \frac{\boldsymbol{J}_e}{F_c \Delta V_e i k^{+}}, \qquad \boldsymbol{v}_{ok,\,e}^{-} = \frac{\boldsymbol{J}_e}{F_c \Delta V_e i k^{-}},$$
$$ik^{+} = \sum_{k=1}^{N_{ion+}} c_k^{+} N_k^{+ion} = 1,2902 \cdot 10^{-5}\,\mathrm{mol/m^3},$$
$$ik^{-} = \sum_{k=1}^{N_{ion-}} c_k^{-} N_k^{-ion} = -1,3175 \cdot 10^{-5}\,\mathrm{mol/m^3}, \tag{15}$$

and where F_c is Faraday constant, $F_c = 96484\,\mathrm{C.mol^{-1}}$, $\boldsymbol{E}_e$ electric field intensity in direction of ions motion in an element of mesh, c^{+} positive ions concentration, c^{-} negative ions concentration, ΔV_e is element volume, N_k^{+ion} is integer multiple of electron charge for specific positive ion, N_k^{-ion} is integer multiple of electron charge for specific negative ion, q_e^{-} is whole charge of negative ions in one element, q_e^{+} is whole charge of positive ions in one element, N_{ion}^{+} is number of different positive charge carriers (elements, compounds), N_{ion-} is number of different negative charge carriers (elements, compounds). Potable water has for instance this composition of ions with volume density m_{io}:

Positive ions	**Negative ions**		**Neutral substances:**
Na ... 32,71 mg/dm^3	F ... 1,58 mg/dm^3	HCO_3 762,4 mg/dm^3	CO_2 ... 4063 mg/dm^3
K ... 1,525 mg/dm^3	Cl ... 5,350 mg/dm^3		H_2O ... 1000000 mg/dm^3
Mg ... 43,81 mg/dm^3	SO_4 ... 13,08 mg/dm^3		
Ca ... 157,7 mg/dm^3	NO_3 ... 0,540 mg/dm^3		

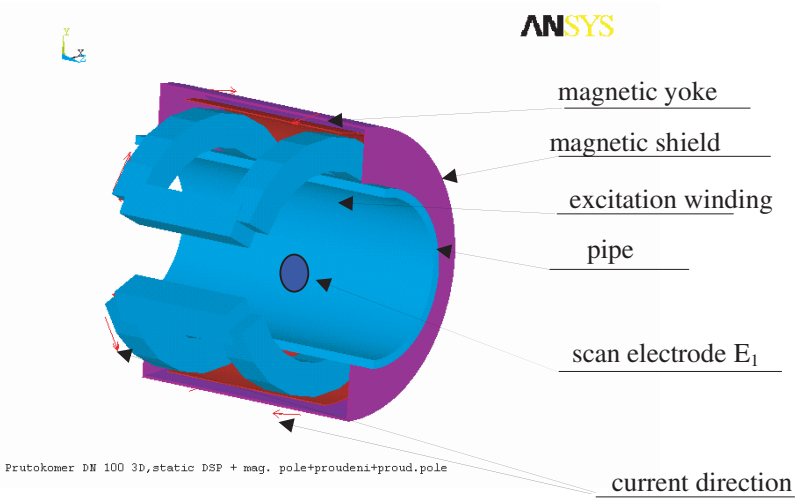
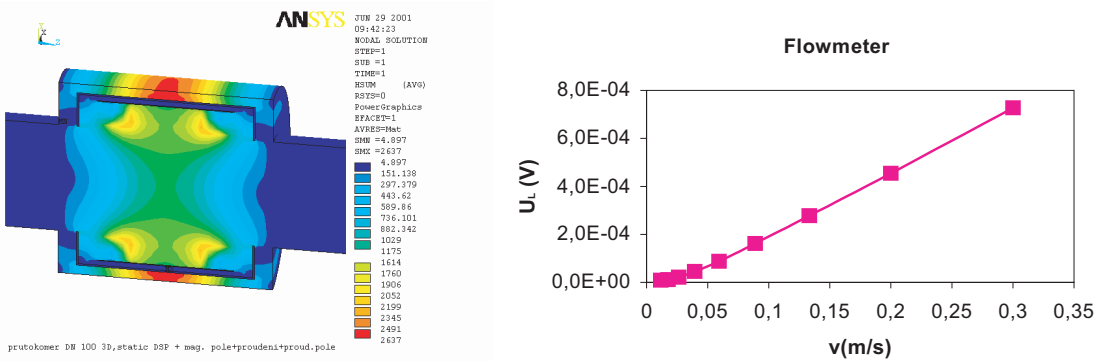


Figure 2: Geometrical model of flowmeter body.

3. NUMERICAL MODEL AND EXPERIMENTAL MEASUREMENT

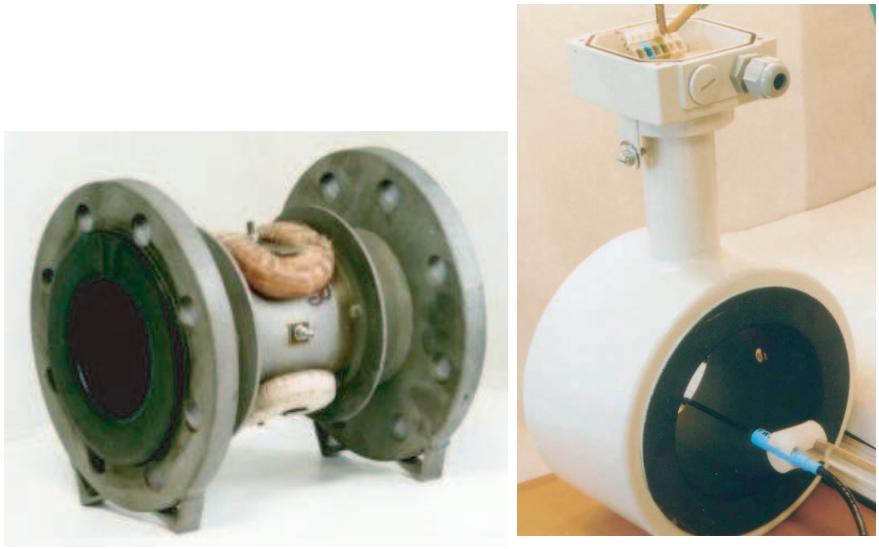
The partial results of MHD model are showed in Fig. 3(a). Dependence of output voltage on fluid velocity is showed in Fig. 3(b). This result is practically identical with experimental measurement on DN-100 ELIS s. r. o Brno Fig. 4(b), where the part of magnetic flux density $\boldsymbol{B}$ measurement as a function of radius R and coordinate z according Fig. 1 is shown. The results of the experiments on Fig. 5 correspond with the numerical model described above in Equations (1) to (15). The deviation between measured and modeled flowmeter never exceeds 3.4%.

The voltage U_L dependence on flow velocity v was tested on flowmeter Fig. 4(b). The deviation of numerical result and experimental was less than measurement uncertainty.



(a) (b)

Figure 3: Numerical modelling results, (a) Distribution of a magnetic field intensity H in the flowmeter body, (b) Dependence of voltage U_L on flow velocity v.



(a) (b)

Figure 4: Experimental flowmeter body, (a) Open body, (b) Magnetic field verification.

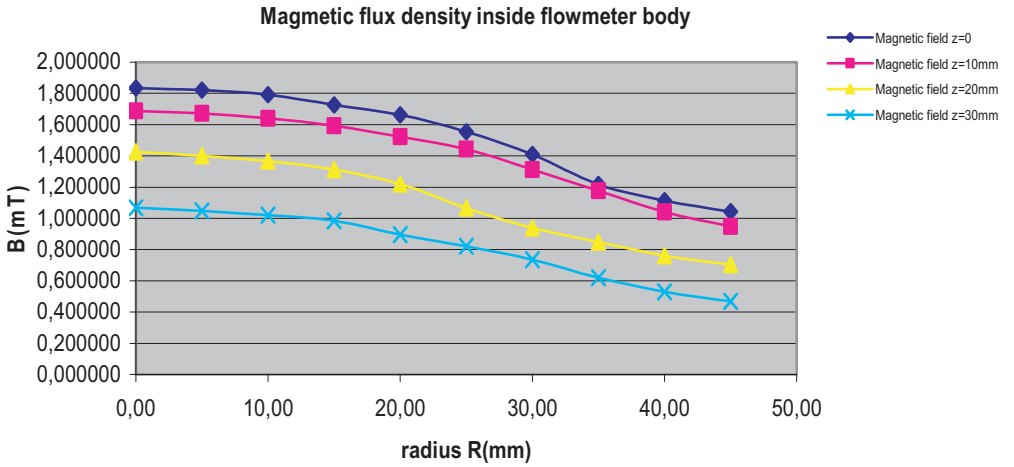


Figure 5: Experimental results, magnetic flux density depend on position of active part flowmeter.

ACKNOWLEDGMENT

The research described in the paper were financially supported by FRVŠ by research plan No. MSM 0021630513 ELCOM, No. MSM 0021630516 and grant GAAV No. B208130603.

REFERENCES

1. Maalouf, I. A., "The derivation and validation of the practical operating equation for electromagnetic flowmeters: Case of having an electrolytic conductor flowing through," *IEEE Sensors Journal*, Vol. 6, No. 1, February 2006.
2. Maalouf, I. A., "A valid model for the electromagnetic flowmeter's measuring: Case of having an electrolytic conductor flowing through," *IEEE Sensors Journal*, Vol. 6, No. 3, June 2006.
3. Fala, P., "Model of inductive flowmeter DN-100," Research report No. 2/01, 1–23, Laboratory of modelling and optimisation of electromechanical systems BUT FECT, Brno, Czech Republic, June 6, 2001.
4. Fiala, P., T. Jirku, and I. Behunek, "Numerical model of inductive flowmeter," *Progress In Electromagnetics Research Symposium Proceedings*, 971–975, Beijing, China, March 26–30, 2007.
5. Maalouf, I. A., "A valideted model for the zero drift due to transformer signals in electromagnetic flowmwter operating with electrolytic conductors," *IEEE Sensors Journal*, Vol. 6, No. 6, December 2006.

Experiments with Accuracy of Air Ion Field Measurement

M. Steinbauer[1], **P. Fiala**[1], **K. Bartušek**[2], **and Z. Szabo**[1]

[1]Department of Theoretical and Experimental Electrical Engineering
University of Technology Brno, Kolejni 4, 612 00 Brno, Czech Republic
[2]Institute of Scientific Instruments, Academy of Sciences of the Czech Republic
Královopolská 147, 612 64 Brno, Czech Republic

Abstract— An analysis of the electric state of air shows the presence of various ion sorts. The therapeutic effect of negative high-mobility ions of proper concentration is known. This positive effect was observed in caves that are used for speleotherapy. This article presents the capability of methods for measuring ion concentration and for ion spectral analysis.

1. INTRODUCTION

Electrical phenomena occurring in ionized gases involve physical-chemical-biologically reactive ions, radicals, and molecular species. These phenomena are encountered and overlap among diverse fields of chemistry, physics, engineering, meteorology, climatology, medicine, microbiology, physiology and industrial hygiene. The physical and chemical aspects of small air ions and radicals have been under investigation almost from the discovery of electricity. Plasma chemistry and discharge physics are inexorably intertwined. The health implications of air ionization have been reviewed elsewhere [1–5]. Improved diagnostics and mechanistic understandings of electrical discharges in gases [6, 7] have led to the development of engineered devices with highly controllable processes for the generation of non-thermal plasmas in the treatment of chemical [8, 27, 28] and biological contaminants [9–12]. Coupled with the increased interest in controlling the potpourri of airborne contaminants, there has been an awakening kindled in applying this technology for improving the air quality of enclosed indoor environments [13–17].

It is known the effect of negative ions to human healthy. The presence of negative air ions in an inhaled air is necessary for the normal vital activity of animals and humans [18]. Clean rural air contains some hundreds to some thousands negative air ions in a cubic centimeter. It has been established that during a long stay in closed rooms, people experience a deficit of the negative air ions, the so-called oxygen "starvation." It causes a worsening of a health condition and predisposes to the development of diseases [19, 20]. In the 1930s, A. L. Tchijevsky suggested compensating for the negative air ions deficiency by an electric ionization of air [19]. The problem of an artificial ionization of air remains urgent also at present. The indoor air in hospitals and offices where many people are present, as well as in air-conditioned rooms, contain, in the best case, as little as tens or hundreds of charges in a cubic centimeter, and the level of the negative air ions can fall even to zero [21]. It was shown by A. L. Tchijevsky [19, 22] and later by N. I. Goldshtein [23] that the complete absence of the negative air ions in the inhaled air is fatal to animals. Modern devices generate the negative air ions in different quantities, from a few hundreds to thousands and millions ions per cubic centimeter. These devices are used for both cleaning the air from dust, microbes, and volatile admixtures by high doses of the negative air ions [24] and enriching the air by the negative air ions for prophylactic and therapeutic purposes.

The Department of the Theoretical and Experimental Electrical Engineering of Brno University of Technology and the Institute of Scientific Instruments of the Academy of Sciences of the Czech Republic are involved in the research of ion field in office and living spaces [25, 26]. The objective is to increase the concentration of light air ions in these spaces. Another task is to set up a simulated therapy room, with conditions similar to speleotherapy caves. It sets the requirements for accurate measurement of ion field with good repeatability. The article deals with the design of gerdien condenser and peripheral measuring devices. An optimal design is important for eliminating the inaccuracy of ion concentration measurement.

2. MEASURING METHOD

Several methods are currently used to measure air ion fields: the dispersion method, the ionspectrometer method, the Faraday cage method, and the gerdien condenser method, whose principle is shown in Figure 1. Here is $d1$ — inner electrode diameter, $d2$ — outer electrode diameter, l

— length of gerdien condenser, M — air flow volume rate, v — air flow velocity, e — elementary charge of electron, $\oplus$ positive air particle (ion), $\ominus$ negative air particle (ion). The gerdien condenser consists of two electrodes. There is an electric field between the inner electrode (the collector) and the outer electrode. The field is imposed by voltage source U. Air ions flow from the fan through the gerdien condenser. Negative ions in the electric field impact the collector, and the current produced is measured by a pA-meter [25]. The current measured is proportional to air ion concentration. The model of the measuring system was presented in [26].

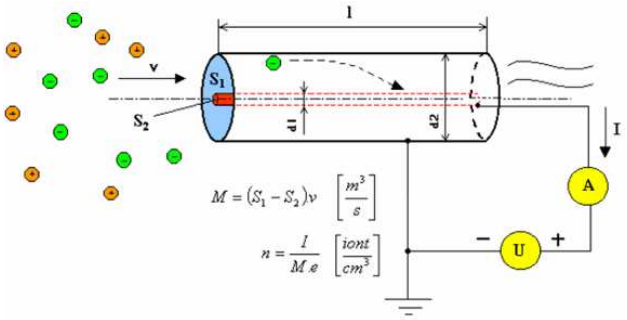


Figure 1: Principle of gerdien condenser method.

Gerdien condenser shape ensures that the flow of air is laminar. Air flow turbulence can distort the accuracy of measurement. The surface of the electrodes is required to be as smooth as possible. The design of gerdien condenser is shown in Figure 2.

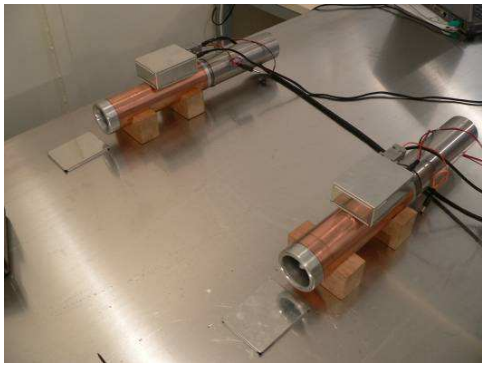


Figure 2: Gerdien condenser design.

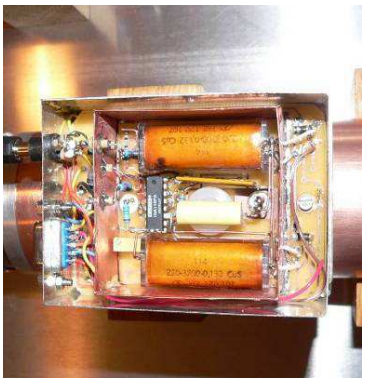


Figure 3: Design of low-level amplifier.

The current flowing through the gerdien condenser is due to the ion concentration. Current intensity depends on polarization voltage, on the dimension and parameters of gerdien condenser, and on ion concentration. The specific current range for the designed gerdien condenser is 10^{-10} A–10^{-13} A. The low-level amplifier is realized with INA 116 — Figure 3. The design of the amplifier is shown in Figure 3.

3. NUMERICAL MODELING OF DESIGN, VERIFICATION

It is possible to carry out analysis of a gerdien condenser model as a numerical solution by help of Finite element method (FEM). The electromagnetic part of the model is based on the solution of full Maxwell's equations. It was solved like simply electrostatic field. This results showed to new facts in gerdien condenser design [26]. New design of gerdien condenser was made with filter for the specific particles, Figure 2.

4. TESTS OF GERDIEN CONDENSERS

The gerdien condenser design was tested according to the special measuring methodology [26]. The experiments are shown in Figure 4 and Figure 5.

Measuring type	Mode	Fan	Voltage	Ion generator	Air flow Regulation	Comment
		F	V	I	R	
Absolute	Testing A	0	0	0	0	A
		0	1	0	0	A
		1	0	0	0	AF
		1	1	0	0	AFV
Diferential	Testing D	0	0	0	0	D
		0	1	0	0	DV
		1	0	0	0	DF
		1	1	0	0	DV
	Measuring D	0	0	1	0	DI
		0	1	1	0	DVI
		1	1	1	1	**DVIF**

Table 1. Table of gerdien tube measurement method.

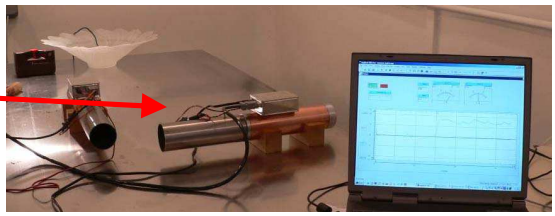


Figure 4: Gerdien condenser measuring method- test by electronic ion generator.

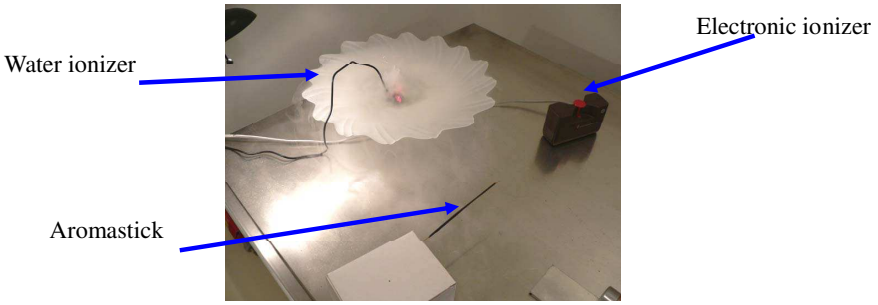


Figure 5: Gerdien condenser measuring-test by water, electronic generator and aromastic.

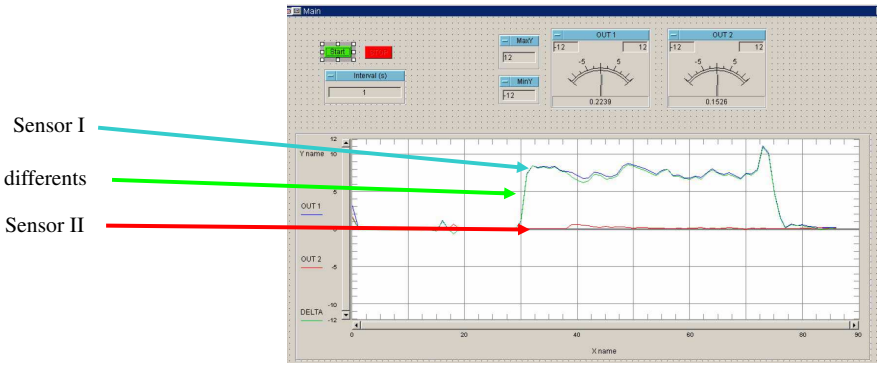


Figure 6: Results of differential measurement of electronic negative ion generator.

5. CONCLUSION

The new design of gerdien condenser and the optimization of peripheral measuring devices have minimized the systematic error of measurement. The system allows measuring air ion concentration with a sensitivity $> 50\,\mathrm{ions/cm^3}$. The ion mobility is in the interval 0.3–100 $\mathrm{cm^2V^{-1}s^{-1}}$. The system is used to measure ion field distribution in living and office spaces.

ACKNOWLEDGMENT

The paper were prepared within the framework of VUT Brno of the Czech Republic and with the support of the research plan MSM 0021630516 and grant GAAV No. B208130603.

REFERENCES

1. Kinne, S. M., "A public health approach to evaluating the significance of air ions," Health Science Center at Houston, School of Public Health, University Texas, Houston, TX, May 1997.

2. Kondrashova, M. N., E. V. Grigigorreko, A. N. Tikhonov, T. V. Sirota, A. V. Temnov, I. G. Stavrovskaya, N. I. Kosyakova, N. V. Lange, and V. P. Tikonov, "The primary physicochemical mechanism for the beneficial biological/medical effects of negative air ions," *IEEE Trans. Plasma Sci.*, Vol. 28, 230–237, Feb. 2000.
3. Krueger, A. P. and E. J. Reed, "Biological impact of small ions," *Science*, Vol. 193, No. 4259, 1209–1213, Sept. 24, 1976.
4. Soyka, F. and A. Edmonds, *The Ion Effect — How Air Electricity Rules Your Life and Health*, 1–181, E. P. Dutton, New York, 1977.
5. Kavet, R. I., Eds., *Boca Raton*, Ch. 9, 1–538, CRC, FL, 1987.
6. Roth, J. R., "Industrial plasma engineering," *Principles*, Vol. 1, 1–538, IOP, Philadelphia, PA, 1995.
7. Schutze, A., J. Y. Jeong, S. E. Babayan, J. Park, G. S. Selwyn, and R. F. Hicks, "The atmospheric-pressure plasma jet: A review and comparison to other plasma sources," *IEEE Trans. Plasma Sci.*, Vol. 26, 1685–1694, Dec. 1998.
8. Eliasson, B. and U. Kogelschatz, "Nonequilibrium volume plasma chemical processing," *IEEE Trans. Plasma Sci.*, Vol. 19, 1063–1077, Dec. 1991.
9. Laroussi, M., "Biological decontamination by nonthermal plasmas," *IEEE Trans. Plasma Sci.*, Vol. 28, 184–188, Feb. 2000.
10. Seo, K. H., B. W. Mitchell, P. S. Holt, and R. K. Gast, "Bactericidal effects of negative air ions on airborne and surface salmonella enteritidis from an artificially generated aerosol," *J. Food Protection*, Vol. 64, No. 1, 113–116, 2001. [Online] Available: http://www.seprl.ars.usda.gov/bmitchell/status.htm.
11. Zhang, R., T. Yamamoto, and D. S. Bundy, "Control of ammonia and odors in animal houses by a ferroelectric plasma reactor," *IEEE Trans. Ind. Appl.*, Vol. 32, 113–117, Feb. 1996.
12. Herrman, H. W., I. Henins, J. Park, and G. S. Selwyn, "Decontamination of chemical and biological warfare (CBW) agents using an atmospheric pressure plasma jet (APPJ)," *Phys. Plasmas*, Vol. 6, 2284–2289, May 1999.
13. Daniels, S. L., "Engineered solutions for mitigation of IAQ problems," *Proc. 2nd NSF Int. Conf. Indoor Air Health*, 243–249, Miami Beach, FL, Jan. 29–31, 2001.
14. "Control of VOCs and PM in indoor environments by air ionization, poster presentation, electromed 2001," *Proc. 2nd Int. Symp. Nonthermal Medical/Biological Treatments Using Electromagnetic Fields and Ionized Gases*, Portsmouth, VA, May 20–23, 2001.
15. "Applications of air ionization for control of VOCs and PM, paper #918," *Proc. 94th Annu. Conf. Exhibition Air and Waste Management Assoc.*, Orlando, FL, June 24–28, 2001.
16. Daniels, S. L. and M. T. Fox, "Engineering design for indoor air environments: Perception vs. reality, Paper E-5," *Proc. 1st NSF Int. Conf. Indoor Air Health: Impacts, Issues, and Solutions*, 65–74, Denver, CO, May 3, 1999.
17. Daniels, S. L. and C. Frost, "Applications of negative air ionization for removal of volatile organic compounds (VOCs) and particulate matter (PM)," *Proc. 2nd NSF Int. Conf. Indoor Air Health*, 346–52, Miami Beach, FL, Jan. 29–31, 2001.
18. Charry, J. M. and R. Kvet, *Air Ions: Physical and Biological Aspects*, CRC Press, Boca Ration, FL, 1987.
19. Tchijevsky, A. L., *Aeroionization: Its Role in the National Economy*, Office Naval Intell., Washington, DC, 1960.
20. Kondrashova, M. N., E. V. Grigorenko, A. V. Tikhonov, T. V. Sirota, A. V. Temnov, I. G. G. Stavrovskaya, N. I. Kosyakova, N. V. Lange, and V. P. Tikhonov, "The primary physicochemical mechanism for beneficial biological/medicinal effects of negative air ions," *IEEE Trans. Plasma Sci.*, Vol. 28, No. 7, 230–237, Feb. 2000.
21. Gerasimova, L. I., *Lasers in Surgery and Therapy of Thermal Burns. Manual for Physicians (in Russian)*, Medicine, Moscow, Russia, 2000.
22. Tchijevsky, A. L., *A Manual for The Application of Ionized Air in Industry, Agriculture, and Medicine (in Russian)*, Gosplanizdat, Moscow, Russia, 1959.
23. Goldstein, N. I. and T. V. Arshavskaya, "Is atmospheric superoxide vitally necessary?," *Z. Naturforsch*, Vol. 52, No. 5–6, 396–404, May–Jun. 1997.
24. Daniels, S. L., "On the ionization of air for removal of noxious effluvials," *IEEE Trans. Plasma Sci.*, Vol. 30, No. 4, 1471–1481, Aug. 2002.

25. Fiala, P., K. Bartusek, T. Vojtek, and T. Skoupil, "Accuracy of air ion field measurement," *Progress in Electromagnetics Research Symposium*, Boston, USA, 2006.
26. Fiala, P., K. Bartusek, T. Vojtek, and T. Skoupil, "Accuracy of air ion field measurement," *Progress in Electromagnetics Research Symposium*, Beijing, China, 2007.
27. Kadlecova, E., I. Behunek, and P. Fiala, "Numerical model of optimization lead-acid accumulator grids," *Progress in Electromagnetics Research Symposium*, Beijing, China, 2007.
28. Fiala, P., T. Jirku, and I. Behunek, "Numerical model of inductive flowmeter," *Progress in Electromagnetics Research Symposium*, Beijing, China, 2007.

Optical Methods Identifying of the Special Purpose Generator Pulses

P. Fiala, P. Drexler, and M. Steinbauer
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology
Kolejní 2906/4, 612 00 Brno, Czech Republic

Abstract— The article provides a brief description of measurement methods which are suitable for measurement of parameters of the high power electromagnetic pulses. Special requirements for measurement methods have to be considered because of the specific properties of EMPs and the parameters of output quantities of pulsed power sources. For measurement of current pulse which is generated in pulsed power source the magnetooptical method was proposed. This method was experimentally realized in collinear configuration. The properties of the EMP which is generated by the high power microwave source can be quantified by means of calorimetric method. On the basis of this method, the calorimetric sensor was constructed and utilized for experimental measurement.

1. INTRODUCTION

In the last few years we can observe the growing interest in non-lethal electromagnetic weapons development [1]. The purpose of this sort of weapons is to affect the functionality of electronic systems of intruders. In connection with this problematic, the EMP (electromagnetic pulse) is frequently cited term in the area of military and security applications. High power microwave sources, which are able to radiate the EMP requires a powerful source of energy. One of the promising way is the utilization of flux compression generators (FCG). FCG serves as a converter of the kinetic energy of detonation to the electromagnetic energy. A high current and high voltage pulse is produced by the FCG [2]. The value of the output pulse can reach up to hundreds of MA and hundreds of kV. After the pulse shaping the power pulse is fed to the microwave source, which emits the EMP. The arrangement of the high power microwave pulse generator is shown in Figure 1.

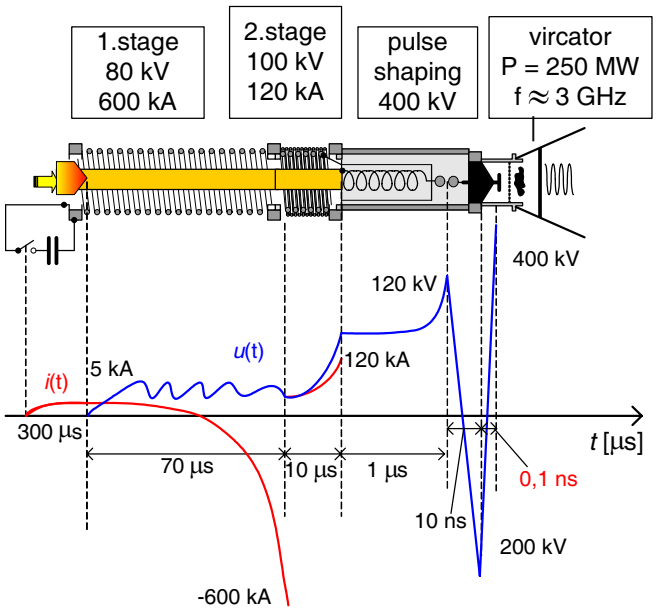


Figure 1: The arrangement of high power microwave pulse generator.

For optimization purposes of design and construction of the pulsed power generator and microwave source it is necessary to measure the electromagnetic quantities on the realized prototypes. In the case of the pulsed power generator the peak value of current pulse and its waveform is

the object of interest. The properties of the current measurement method are very demanding. They are limited by the time relations of the waveform in the sub-microseconds range and by the peak value of the current. The classical current measurement method (current shunt) are failing. The current shunts have to large parasitic inductance. For the MA current values it is necessary to achieve very low resistance to. The promising is the utilization of magneto-optical effect as a principle of measurement methods.

The basic idea about the optimal EMP generation process in the microwave pulse source can be obtained by the measurement of the power and the energy of the output pulse. EMP is emitted into the free-space obviously. The special antenna can be used for pulse measurement. But, this approach is inconvenient, because of the wide frequency spectrum of the pulse, its single-shot character and mostly unknown radiation pattern of the EMP emitter. It is not necessary to get the waveform of the EMP. The idea about the optimal function of the emitter gives the power level and the energy of the pulse. For this reason, the calorimetric method is suitable for application in EMP emitter measurement.

2. THE MAGNETO-OPTICAL MEASUREMENT METHOD

The magneto-optical measurement method. is based on the Faraday's magnetooptic effect. The connection between the sensor and the measuring device is implemented in the optical wavelength.

There are three basic types of the possible active sensors. The first type is a garnet with high Verdet constant, the second one is an optic fiber and the third one is based on magnetooptic properties of ferromagnetic mono/multi thin film. Other types of sensors are based on the magnetooptic Kerr's effects (MOKE), or surface MOKE (SMOKE) effect. By an available measuring devices application we can measure pulses with the limit length $T_{\max} = 0.1\,\text{ns}$ [3, 4].

The polarization rotation of light passing the MO sensor is affected by the magnetic part of EM pulse. The rotation is due to the magnetic field and properties of the sensor material (Verdet constant). The differential measurement method utilizing the MO glass element was experimentally realized. Laser beam with linear polarization passes the MO glass placed in Helmholtz coil. The laser beam is subsequently fed through an analyzer and the polarization rotation is converted to intensity modulation. The intensities orthogonal beams are sensed by a photodiodes. The magnetooptic glass FR-5 by Hoyoa Optics was used in this experiment.

The sensitivity of the sensor and its bandwidth are contradictory properties. Because of the high power level of EM pulse the sensitivity of the sensor is not a critical requirement. The bandwidth of the sensor has to be appropriate to the measured signal characteristic. When it has to be able to detect transients with duration 1 ns, the corresponding bandwidth reaches the value of 350 MHz. The experimental setup is depicted in Figure 2.

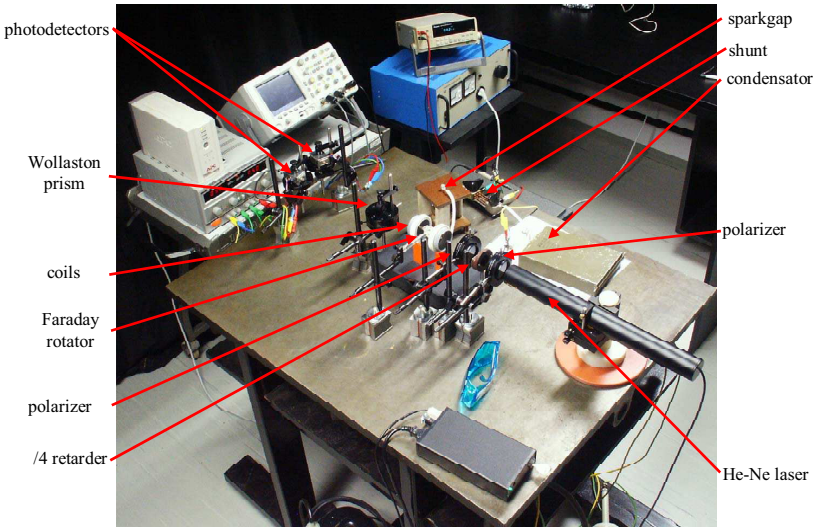


Figure 2: Experimental setup for magneto-optic method verification.

Measured current is fed through the coil system with 4 turns. The inductance was $L = 2\,\mu\text{H}$. The coil system is connected into the discharge circuit with the high voltage capacitor($C = 8\,\mu\text{F}$, $U = 2,2\,\text{kV}$) and spark-gap. For comparative current measurement a low resistance ($R_s = 5\,\text{m}\Omega$)

was used as a current shunt. The frequency of the periodic current transient was $f = 37\,\mathrm{kHz}$. A strong electromagnetic interference was generated by the spark-gap discharge. The interference was suppressed by means of differential measurement method. The example of the waveform obtained by the discharge are shown in Figure 3. On the basis of calculated field distribution in the coils and the Verdet constant of the Faraday rotator, the waveform of the current pulse was obtained. The peak value of the current pulse achieves $I_p = 2\,\mathrm{kA}$. The current waveform distortion by the parasitic inductance of the shunt is obvious in the Figure 3.

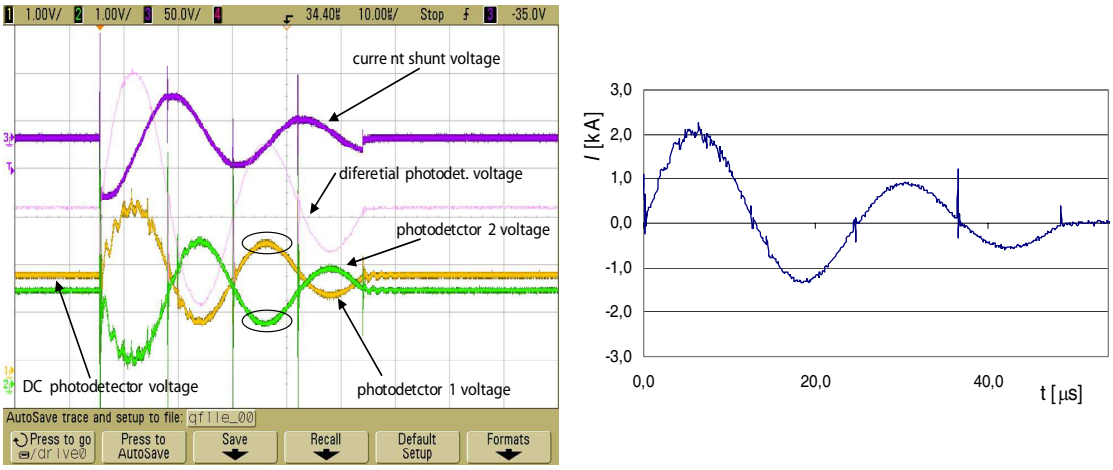


Figure 3: The example of the waveforms obtained by the magneto-optical method verification (left) and the evaluated current waveform (right).

3. THE CALORIMETRIC METHOD

For measurement of parameters of EMP which is generated in the last stage of the pulsed microwave source the calorimetric method was chosen. The group of calorimetric methods represents another type of converter to be introduced. We can measure power supplied by pulse. Figure 4 depicts four versions of the method utilizing calorimetric measurement.

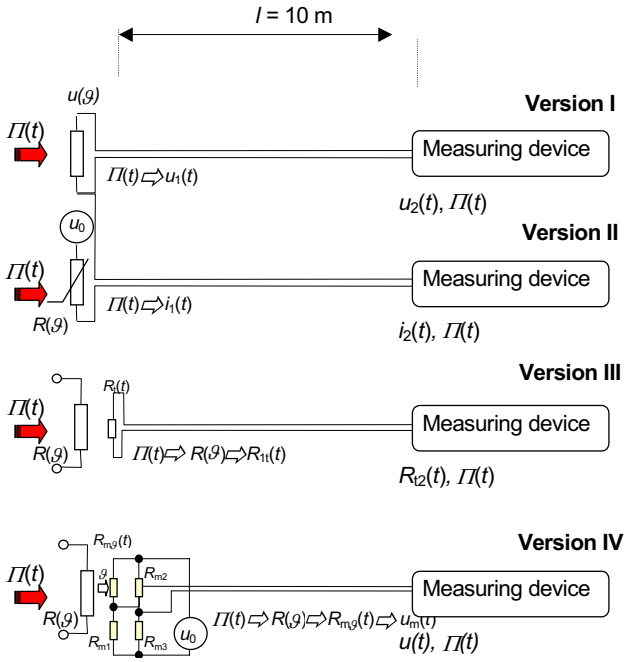


Figure 4: Calorimetric method for the single-shot pulses measurement.

Version I discussed in [5, 6] has a sensor in the form of an ideal resistor and enables measurement of the maximum value of microwave power $P_{\max}$. The analyzed peak voltage corresponds to peak

value of power $P_{\max}$. For available measuring devices we can measure pulses with the limit length $T_{\max} = 50\,\text{ps}$. Version II scans the change of resistance of the sensor, created by an evaporated thin layer, in dependence on the pulse energy. For available measuring devices we can reach the accuracy of 30% up to impulse limit length $T_{\max} = 0.1\,\text{ns}$. Version III is based on the measurement of the temperature change of the thermistor placed in contact with the layer. Under the same conditions as for the previous version we can reach the accuracy improvement of an order of magnitude. Version IV is the bridge connection of Version III. Several thermistors are attached in series to the evaporated layer; then, three resistors create a DC bridge of Weston type with the thermistors. The change of resistance in the thermistor arm is evaluated. The voltage in measuring the bridge diagonal is consequently integrated. Thus, the value of the pulse energy is obtained (and recorded by the measuring device). For available measuring devices we can measure pulses with the limit length $T_{\max} = 0.03\,\text{ns}$ with accuracy to 10

For the measurement of free-space EMP, the combined calorimetric sensor was built [7]. The sensor operation is based on Version I and Version IV of the calorimetric method in Figure 4. The first part (Version I) serves as the sensor of instantaneous power and the second part (Version IV) serves as the sensor of pulse energy. The realization of the combined sensor is shown in Figure 5. Both parts are equipped with Horn antennas to ensure the matching of the free-space EMG wave to the sensor input. The sensor was calibrated with an RF generator in an absorption room. The calibration was performed for microwave pulses with defined duration and power level. Due to safety requirements, the connection between the sensor and the measuring device was ensured by means of coaxial cable of the minimum length $l_{\min} = 10\,\text{m}$.

Figure 5: Realization of combined calorimetric sensor.

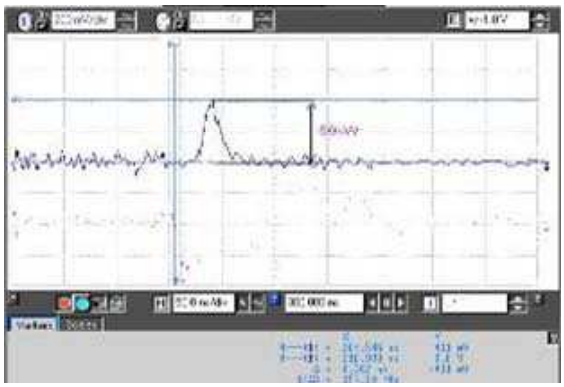


Figure 6: Waveform of microwave power, $P_{\max} = 50\,\text{kW}$.

The combined calorimetric sensor was used for the measurement of vircator-emitted EMP. The supply of the vircator was provided by pulse high-voltage source powered by Marx bank. When the vircator is in the operational mode, hard RTG emission is generated in addition to the microwave emission [8]. The energy of the electron beam is $W_b = 1\,\text{MeV}$. Therefore, safety requirements equal to those mentioned above have to be considered.

The example of the waveform of small microwave power $P_{\max} = 50\,\text{kW}$ is in Figure 6. However, vircator is able to emit EMP with peak value of hundreds of MW when supplied with pulsed power generators.

4. CONCLUSION

The overview of methods suitable for the measurement of short single-shot EMPs with high power level and high-value pulsed current waveform was given. The methods were experimentally realized and verified. A combined calorimetric sensor for free-space measurement was built and the functionality of the calorimetric sensor was proved by real measurement of vircator-emitted EMP. The magneto-optic method was found as suitable for application by the measurement on the pulsed power generator. Its advantages are high bandwidth, galvanic isolation and the possibility of common-induced interference rejection. For practical application an integral fiberoptic magneto-optical sensor is proposed [9].

ACKNOWLEDGMENT

The paper was prepared within the framework of the research plan No. MSM 0021630513 of the Ministry of Education, Youth and Sports of the Czech Republic.

REFERENCES

1. Kopp, C., "An introduction to the technical and operational aspects of the electromagnetic bomb," Technical Report, Air Power Studies Centre, Fairbairn, ISBN 0-642-264155, 1996.
2. Fiala, P., *Design and Modeling of Pulsed Power Source*, Brno University of Technology, Brno, 2004.
3. Riordan, J. A., F. G. Sun, Z. G. Lu, and X.-C. Zhang, "Free-space transient magneto-optic sampling," *Applied Physics Letters*, Vol. 71, 1452–1456, 1997.
4. Hwang, E. U. and B. Y. Kim, "Pulsed high magnetic field sensor using polymethylmethacrylate," *Measurement Science and Technology*, Vol. 17, 2006.
5. Kasal, M., *Multinuclear Measuring Channel of NMR Spectrometer*, FE BUT, Brno, 1984.
6. Pfeffer, M. and M. Kasal, "Automatic impulse reflectometer," *Sdělovací technika*, Vol. 10/11, Czech Republic, 1986.
7. Drexler, P. and P. Fiala, "FEM model of calorimetric sensor," *Proceedings of the International Conference Electronic Devices and Systems EDS '05*, Brno University of Technology, Brno, Czech Republic, 2005.
8. Jiang, W., K. Woolverton, J. Dickens, and M. Kristiansen, "Optimization of microwave generation by coaxial vircator," *ICOPS'99, 1999 IEEE International Conference on Plasma Science*, 1999.
9. Drexler, P., "Methods for measurement of ultra-short single-shot electromagnetic pulses," PhDThesis, Brno University of Technology, Brno, 2007.

A Passive Optical Location with Limited Range

P. Fiala[1], **T. Jirku**[2], **R. Kubasek**[3], **Z. Szabo**[4], **and P. Konas**[5]

[1,2,3,4]Department of Theoretical and Experimental Electrical Engineering, Brno University of Technology
Kolejni 2906/4, 612 00 Brno, Czech Republic
[5]Department Wood Science, Mendel University of Agriculture and Forestry
Brno Zemedelska 3, 613 00 Brno, Czech Republic

Abstract— We know active and passive methods of a location. This article deals only with a passive location of dynamic targets. The passive optics location is suitable just for tracking of targets with mean velocity which is limited by the hardware basis. The aim of this work is to recognize plasma, particles etc. It is possible to propose such kind of evaluation methods which improve the capture probability markedly. Suggested method is dealing with the short-distance evaluation of targets. We suppose the application of three independent principles how to recognize an object in a scanned picture. These principles use similar stochastic functions in order to evaluate an object location by means of simple mathematical operations. Methods are based on direct evaluation of picture sequence by the help of the histogram and frequency spectrum. We find out the probability of unidentified moving object in pictures. If the probability reaches a setting value we will get a signal. The processing of dynamic pictures and their filtration are a significant part of work. Static objects, background (trees, buildings) must be filtered off before. This filtration is being also done by means of the probability function. Probability distribution of an object position is gained from a sequence of more pictures.

1. INTRODUCTION

The present state of the image processing with the target detection aim is orientated mostly on the security [6], robotics [7], medicine [8], astronomy [10, 11, 14] or for example on ground target detection and recognition in the army area [9, 12, 15]. Also there are other application areas [13]. Mostly the main problem to detect relatively large and slow objects on the static background and to recognize objects, faces etc. and to identify them. For example in robotics is image processing used for the robot orientation and navigation, eventually for some simple environment analysis, in the security area could be image processing used for the face detection and identification. In the medical applications it should be used in combination with CT or NMR methods. Army uses image processing for example in the radar technology where image is often restored from some data obtained with help of some active radar system and it is restored in the cylinder or in the spherical coordinates. The optical location is by army used for the target detection, identification and tracking etc. Our system solves different problem. We are trying to detect and track relatively very fast objects like plasma particles etc. In the first approximation we can test the device on the birds flocks. We suppose usage of the several basic methods bounded by the probability function for the movement detection and tracking in the image. The theoretical solution was published [2]. There were done the first tests of the proposed methods. These methods were tested in the MATLAB environment. Some of the solution aspects are described in the following text. The proposed methods were tested by the set of the testing images and the object location methodology was tested on the real video sequence. These methods are recapitulated and resumed from the view of the method efficiency, speed and the other parameters during the image processing. There is also summarized hardware used for the algorithm implementation, tests and on-line detection experiments in the text.

2. THE COMBINED DIFFERENTIAL IMAGES METHOD

The differential images method described in the [2] was combined with the image segmentation. The good properties of the both methods are used with an advantage and so we obtained powerful tool, which meets the high computational requirements and the low resolution ability of the recorded image. The MATLAB environment supports two basic principles of the image segmentation — The boundary detection method and the edge detection method, both described in [3–5]. There was with an advantage used a *bwboundaries* function, implemented in the Imagetools package. This function is based on the boundary detection segmentation method. It can work on the binary image only. The searching process is following: from some automatically selected initial pixel its

vicinity pixels are probed for their value and it is decided if they belong to the object area. This step is repeated for each pixel in the object vicinity. There was tested another function from the MATLAB Imagetools package — *bwtraceboundary*. This function uses the same principle as the *bwboundaries* function described above but it has in comparison to the *bwboundaries* function one significant disadvantage — user must specify the initial pixel for the image segmentation. Another tested segmentation method was based on the edge detection principle. This method had bad results during the testing. We weren't able to prevent the removing of some target objects or their significant parts. The combined differential images method uses several-step algorithm. In the first step are the important areas selected in the image by the image segmentation. Then in these selected areas the detailed tracking is performed with the time domain differences assistance.

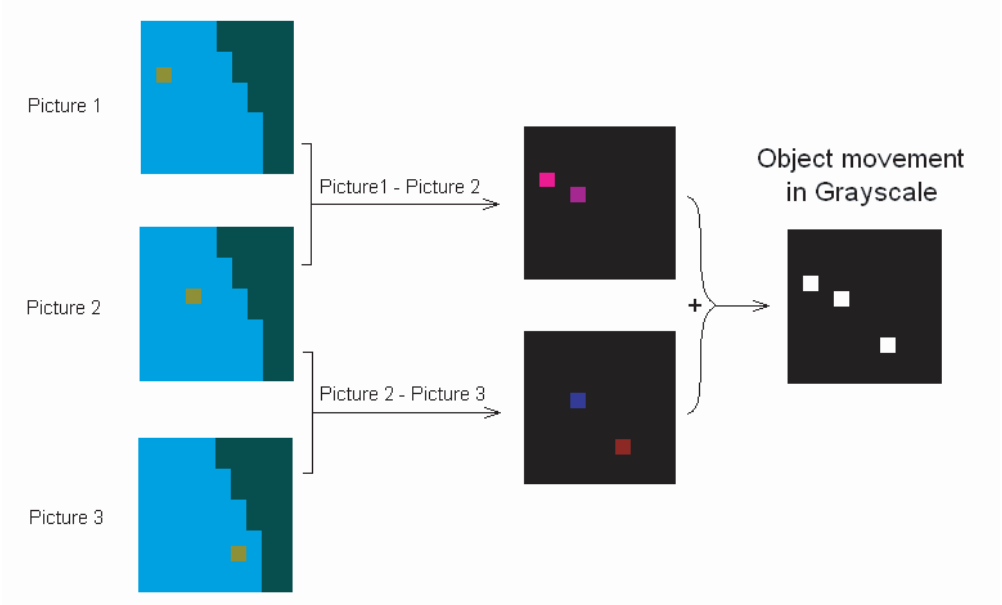


Figure 1: Differential images method principle.

3. THE HISTOGRAM METHOD

The histogram is the statistical information. It shows the intensity distribution in the image. It can be said in the first moment this method is not capable of the target movement detection in the image. The reason for this is: If the object is moving above the quasi-constant background, it is very difficult to recognize target movement or its position.

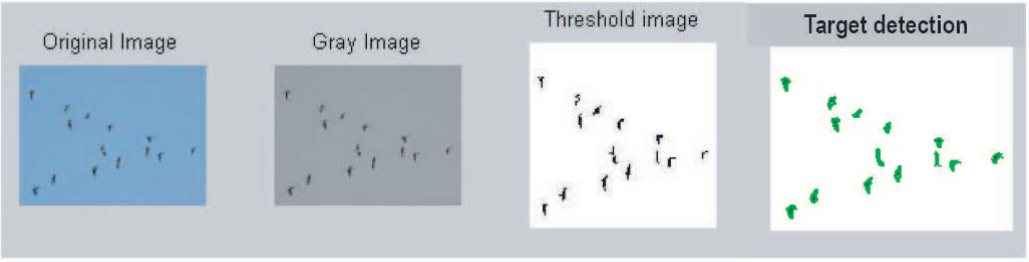


Figure 2: Image segmentation and target detection.

4. THE SPECTRUM ANALYSIS METHOD

The next suggested method is the spectrum's analysis method. In the figure bellow is shown the spectrum's difference for the diagonal and the flat object move above the simple background.

Particular spectrums are obtained by the 2D Fast Fourier transform algorithm — in MATLAB implemented in the *fft*2 function. The lines direction in the amplitude spectrum difference is perpendicular to the movement direction. If there are changed background colors, is in addition similar pattern visible in the phase spectrum difference. In the similar test images sequences is solution simple. More complicated method application is in the real recorded sequences. Step order was: At first were selected potential target areas by the RGB space analysis (or the combined differential images method) — to the feasible pixels was associated white color and to the others was associated the black color. In the next step the spectrums were calculated. During the difficult testing was discovered that method of the spectrum's analysis method doe's not give usable results according to the entered parameters of the object velocity, its size and quality and the image sequence density. The tested objects were detectable only with big difficulty. More detailed problem description is in [3–5].

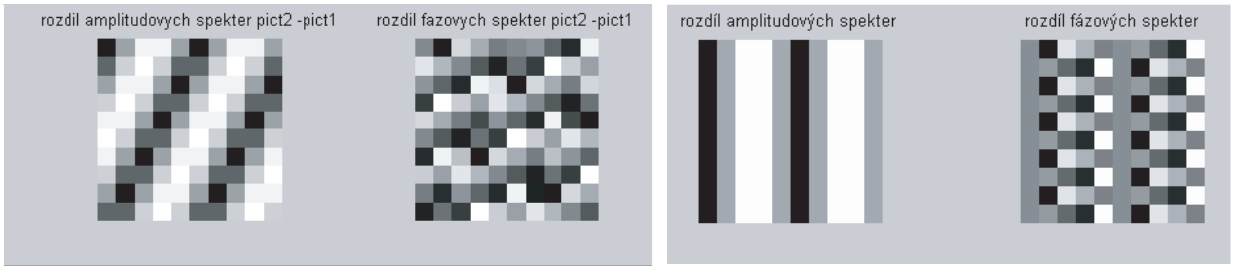


Figure 3: The difference of the amplitude and phase spectrum's during the diagonal and flat object move.

5. THE METHODS SUMMARY

From the comparison of the suggested methods (The combined difference images method, the histogram method and the spectrum's analysis method) follows that the methods are applicable on the images with much higher intensity level according to the background. It was determined that the image segmentation method based on the edge detection is not directly usable for the hardware solution implementation. The MATLAB environment during the edge detection applies some threshold on the image and this can filter out the target or his significant part. In the opposite the boundary detection method (the *bwboundaries* function) looks like very powerful tool. This function works above the binary image and it is very fast. Disadvantage for this method is that we are loosing some information during the transformation from the RGB space to the binary space. This can result loss for the small targets. Still this disadvantage has lower dependencies than the corresponding effect for the edge detection method. Very efficient is threshold segmentation method but this method can give bad results according to the threshold selection. In addition, if we know the color distribution in the RGB spectrum we can very simply filter the image before we try to detect the object and remove some redundant information. Then we can concentrate only to the reduced image area and detection progress can be faster. If we speak about the sensitivity and the precision of the elementary methods, big problem part lies in the fact, that the real image matrix is relatively very big according to the traced object. This disadvantage can be particularly removed by using the Subimaging. On these subimages is applied segmentation by the thresholding in the RGB space. Probable targets are processed by the suggested methods. (The combined difference images method, the histogram method and the spectrum's analysis method).

6. THE FALSE TARGET (BACKGROUND) FILTERING METHOD

Suppose that we have selected possible targets in the image. We can confirm or decline the correctness of the detected target using the object presence monitoring. Monitoring is done by the Kalman filter application [3–5]. If the target movement shows some hint of periodicity, we can reject it as the false target (ground targets, false targets, etc.). The second possibility of the false target detection is finding of its connection to the image boundary. According to this information we can reject this target like false target or select it like the probable target.

7. HARDWARE AND THE DSP PROGRAMMING METHOD SELECTED FOR EXPERIMENTS

For the experiments are selected following components-Camera: ST VS6724 with the ×24 evaluation board, CMOS chip, Resolution of 2Mpix, Possibility of training with small and cheap camera chip and then chip exchange to some more reliable for application, Processor: TI TMS320C6416 with the TMS320C6416T DSK evaluation board. Processor is programmed with help of the Real-time Workshop and Simulink system using the Real-time Workshop target. After the principles verifying in MATLAB follows simple compilation of MATLAB (SIMULINK) code to the C language using special libraries for selected DSP and import to the Code Composer Studio. There can be the C code optimized, compiled and uploaded to the DSP.

8. IMAGE PROCESSING PREVIEW [5]

(a) (b) (c)

Figure 4: (a) Original image — selected targets, (b) Segmented image — selected potential targets, (c) Final detected targets selected by method.

9. CONCLUSIONS

Basic theoretical methods were tested and owing to their suitable combination the method of dynamic object detection in the image was found. In the Section 8 you can see the images sequence in the particular method phases. Figure 9 shows the detected targets. More testing must be performed for the algorithm speed and accuracy tuning.

ACKNOWLEDGMENT

The research described in the paper were financially supported by FRVŠ by research plan No. MSM 0021630513 ELCOM, No. MSM 0021630516 and grant GAAV No. B208130603 of the Ministry of Education, Youth and Sports of the Czech Republic.

REFERENCES

1. Gonzales, C. G., R. E. Woods, and S. L. Eddins, *Digital Image Processing Using Matlab*, Pearson Prentice Hall, New Jersey, 2004.
2. Fiala, P., T. Jirku, R. Kubasek, P. Drexler, and P. Konas, "A passive optical location with limited range," *Progress In Electromagnetics Research Symposium*, Beijing, 2007.
3. Dostál, V., "Dynamic image processing," *The Bachelor's thesis*, DTEEE FEEC BUT, 2007.
4. Daniel, M., "A passive optical location," *The Bachelor's thesis*, DTEEE FEEC BUT, 2007.
5. Suchardal, J., "A dynamic object image processing," *The Bachelor's thesis*, DTEEE FEEC BUT, 2007.
6. Gang, L., Z. Ruili, and L. Ling, "Moving target detection in video monitoring system," *The Sixth World Congress on Intelligent Control and Automation*, Vol. 2, 9778–9781, Jun. 21–23, 2006.
7. Wang, J. and M. Sugisaka, "Study on robust target detection technique," *SICE 2004 Annual Conference*, Vol. 1, 772–775, Aug. 4–6, 2004.
8. Nunes, P. R. R. L., M. Morf, S. L. Wood, and J. M. Turner, "Image reconstruction techniques and target detection," *20th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes*, Vol. 20, Part 1, 351–352, Dec. 1981.

9. Rong, S. and B. Bhanu, "Modeling clutter and context for target detection in infrared images," *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 106–113, Jun. 18–20, 1996.
10. Sridhar, S. and G. Healey, "Point target detection in spatially varying clutter," *Proceedings of IEEE Workshop on Applications of Computer Vision*, 232–239, Nov. 30–Dec. 2, 1992.
11. Chang, C.-I. and D. C. Heinz, "Constrained subpixel target detection for remotely sensed imagery," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38, No. 3, 1144–1159, May 2000.
12. Li, G., X. Cai, and Y. Liu, "Efficient target detection from infrared image sequences using the sequential monte carlo method," *Proceedings of the 2006 IEEE International Conference on Mechatronics and Automation*, 549–554, Jun. 25–28, 2006.
13. Wang, X., J. Liu, and H. Liu, "Small target detection in sea clutter based on doppler spectrum features," *International Conference on CIE Radar*, 1–4, Oct. 2006.
14. Demirci, S., B. Yazgan, and O. Ersoy, "Multispectral target detection by statistical methods," *Proceedings of 2nd International Conference on Recent Advances in Space Technologies*, 653–659, Jun. 9–11, 2005.
15. Karacali, B. and W. Snyder, "Automatic target detection using multispectral imaging," *Proceedings of Applied Imagery Pattern Recognition Workshop*, 55–59, Oct. 16–17, 2002.

Numerical Method of Simulation of Material Influences in MR Tomography

M. Steinbauer[1], **R. Kubasek**[1], **and K. Bartusek**[2]

[1]Faculty of Electrical Engineering and Communications, Brno University of Technology, Czech Republic
[2]Institute of Scientific Instruments, Academy of Sciences of the Czech Republic, Czech Republic

Abstract— Generally all Magnetic Resonnce Imaging (MRI) techniques are affected by magnetic and electric properties of measured materials, resulting in errors in MR image. Using numerical simulation we can solve the effect of changes in homogeneity of static and RF magnetic fields caused by specimen made from conductive and/or magnetic material in MR tomograph. This paper deals with numerical simulation of material susceptibility influence to magnetic field.

1. INTRODUCTION

In MR tomography, strong magnetic field is used (above 1 T). Because MR is very sensitive to inhomogeneity of this field, even such weak induced magnetic field as from para- or diamagnetic material is significant. Second mechanism, which affects field homogeneity, is magnetic field off eddy current induced by RF impulse in conductive material. Aim of this paper is in simulation of weakly magnetic material influence to static magnetic field.

For numerical simulation, two approaches can be used: normal Finite Element Method (FEM) and calculation of reaction field considering the induced polarity in the region. These approaches have been discussed in [6] and [7], where the accuracy consideration and possible reduction of numerical method errors can be found. Using FEM for calculation of magnetic field in MR tomography requires double precision arithmetic and sufficiently fine mesh, because change of the magnetic field in vicinity of slightly magnetic materials is weak in comparison to basic static magnetic field. Second approach may use single precision arithmetic, because only reactional field (induced own field of magnetic material) is computed.

2. 2D ANALYTIC SOLUTION

Let's have specimen with susceptibility χ_{m1} surrounded by reference medium with known susceptibility χ_{m2} and placed into static primal magnetic field with magnetic intensity vector $\mathbf{H}_0$ oriented in $\mathbf{u}_z$ direction — see Fig. 1 left. We have to determine magnetic intensity $\mathbf{H}$ of incurred field, which is superposition of primal and reaction field $\mathbf{H}_{\mathrm{r}}$ (effect of specimen magnetization).

Because there are not variable currents in whole area, magnetic field is irrotational (rot H = 0) and we can use scalar magnetic potential

$$\mathbf{H} = -\mathrm{grad}\varphi_{\mathrm{m}}. \tag{1}$$

Magnetic potential of primal field of intensity $\mathbf{H}_0$ is

$$\varphi_{\mathrm{m0}} = -\int \mathbf{H}_0 \cdot \mathbf{u}_z \mathrm{d}z = -\mathbf{H}_0 z. \tag{2}$$

Incidence of magnetized specimen from Fig. 1 left we can replace with effect of field of surface magnetic charge with density σ_{m} on boundary of areas Ω_1 a Ω_2 — see Fig. 1 right, whereas susceptibility of areas is now zero. First we have to compute magnetic charge density distribution on bound Γ and consequently the intensity of reaction field $\Delta\mathbf{H} = \mathbf{H} - \mathbf{H}_0$

$$\Delta\mathbf{H}(\mathbf{r}) = \frac{1}{2\pi} \oint_{\Gamma} \sigma_{\mathrm{m}}(\mathbf{r}') \frac{\mathbf{u}_{\mathrm{r}}}{R(\mathbf{r},\, \mathbf{r}')} \mathrm{d}\Gamma. \tag{3}$$

Surface magnetic charge invokes scalar magnetic potential [5]

$$\varphi_{\mathrm{mr}}(\mathbf{r}) = -\frac{1}{2\pi} \oint_{\Gamma} \sigma_{\mathrm{m}}(\mathbf{r}') \ln R(\mathbf{r},\, \mathbf{r}') \mathrm{d}\Gamma. \tag{4}$$

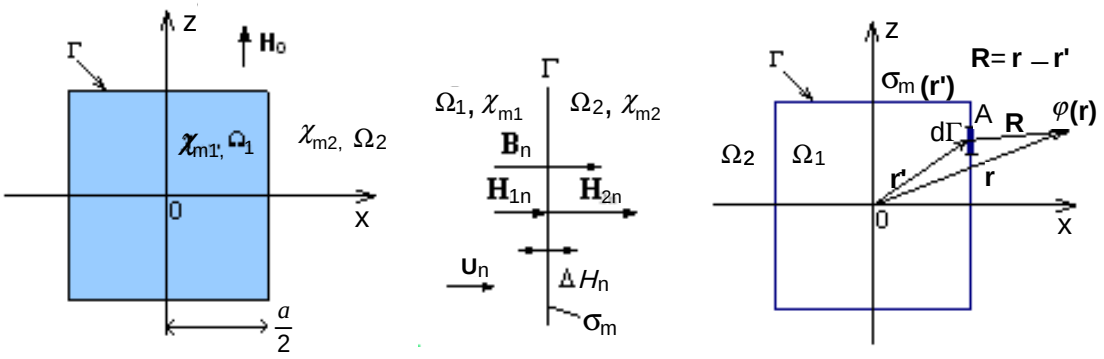


Figure 1: 2D analytic model for rectangular specimen (left), replacement of specimen magnetization effect by surface magnetic charge at the area boundaries (right) and boundary detail (middle).

Total scalar magnetic potential at point $\mathbf{r}$ is superposition of static primal field intensity (2) and contribution from charged bound (4)

$$\varphi_{\mathrm{m}}(\mathbf{r}) = -Hz - \frac{1}{2\pi}\oint_{\Gamma} \sigma_{\mathrm{m}}(\mathbf{r}')\ln R(\mathbf{r},\,\mathbf{r}')\mathrm{d}\Gamma. \tag{5}$$

Using condition of magnetic flux $\mathbf{B}_{\mathrm{n}} = B_{\mathrm{n}}\mathbf{u}_{\mathrm{n}}$ we obtain an integral formula for surface magnetic charge density normal component conjunction on bound Γ (see Fig. 1 middle)

$$B_{\mathrm{n}} = \mu_0(1+\chi_{\mathrm{m1}})H_{1\mathrm{n}} = \mu_0(1+\chi_{\mathrm{m2}})H_{2\mathrm{n}} \tag{6}$$

Analogically to the Gauss theorem causes magnetic charge of density σ_{m} at point A magnetic field of intensity

$$\Delta H_{\mathrm{n}} = \pm\frac{\sigma_{\mathrm{m}}(\mathrm{A})}{2}. \tag{7}$$

Using (5) and (1) we have the normal components of magnetic field intensity at point A (Fig. 1 middle)

$$H_{1\mathrm{n}} = H_0\mathbf{u}_z\mathbf{u}_{\mathrm{n}} + \frac{1}{2\pi}\mathrm{grad}\oint_{\Gamma,\, r\in\Omega_1} \sigma_{\mathrm{m}}\left(\mathbf{r}'\right)\ln R(\mathbf{r},\,\mathbf{r}')\mathrm{d}\Gamma\mathbf{u}_{\mathrm{n}}, \tag{8}$$

$$H_{2\mathrm{n}} = H_0\mathbf{u}_z\mathbf{u}_{\mathrm{n}} + \frac{1}{2\pi}\mathrm{grad}\oint_{\Gamma,\, r\in\Omega_1} \sigma_{\mathrm{m}}\left(\mathbf{r}'\right)\ln R(\mathbf{r},\,\mathbf{r}')\mathrm{d}\Gamma\mathbf{u}_{\mathrm{n}}. \tag{9}$$

Whenever $\mathrm{A}\in\Gamma$ and thus $\mathbf{r}\in\Gamma$, has integral in formulas (8) a (9) singularity at point A (where $\mathbf{r}=\mathbf{r}'$). We can remove this singularity omitting point $\mathbf{r}=\mathbf{r}'$ from integration and taking field contribution of this point using (7) instead. So we can write

$$H_{1\mathrm{n}} = H_0\mathbf{u}_z\cdot\mathbf{u}_{\mathrm{n}} + \frac{1}{2\pi}\oint_{\substack{\Gamma\\ \mathrm{r}\neq\mathrm{r}'}} \sigma_{\mathrm{m}}\left(\mathbf{r}'\right)\frac{1}{R(\mathbf{r},\,\mathbf{r}')}\mathrm{d}\Gamma\mathbf{u}_R\cdot\mathbf{u}_{\mathrm{n}} - \frac{\sigma_{\mathrm{m}}(A)}{2}, \tag{10}$$

$$H_{2\mathrm{n}} = H_0\mathbf{u}_z\cdot\mathbf{u}_{\mathrm{n}} + \frac{1}{2\pi}\oint_{\substack{\Gamma\\ \mathrm{r}\neq\mathrm{r}'}} \sigma_{\mathrm{m}}\left(\mathbf{r}'\right)\frac{1}{R(\mathbf{r},\,\mathbf{r}')}\mathrm{d}\Gamma\mathbf{u}_R\cdot\mathbf{u}_{\mathrm{n}} + \frac{\sigma_{\mathrm{m}}(A)}{2}, \tag{11}$$

where was used

$$\mathrm{grad}\ln R(\mathbf{r},\,\mathbf{r}') = \frac{1}{R(\mathbf{r},\mathbf{r}')}\,\mathbf{u}_R. \tag{12}$$

Substituting from (10) and (11) into (6) we have after some rearrangement

$$\frac{\chi_\Delta}{2\pi}\oint_{\substack{\Gamma\\ \mathrm{r}\in\Gamma,\,\mathrm{r}\neq\mathrm{r}'}} \frac{\sigma_{\mathrm{m}}\left(\mathbf{r}'\right)}{R\left(\mathbf{r},\,\mathbf{r}'\right)}\mathrm{d}\Gamma\mathbf{u}_R\cdot\mathbf{u}_{\mathrm{n}} + \frac{\sigma_{\mathrm{m}}(\mathbf{r})}{2} = -\chi_\Delta H_0\mathbf{u}_z\cdot\mathbf{u}_{\mathrm{n}}, \tag{13}$$

where differential susceptibility was introduced

$$\chi_\Delta = \frac{\chi_{m1} - \chi_{m2}}{\chi_{m1} + \chi_{m2} + 2}. \tag{14}$$

Formula (13) is not analytically solvable, thus we solve it numerically by mean of boundary element method. After solution of (13) using collocation method described in [5] we have obtained results, shown in next figure. Shape of magnetic flux density is in the Fig. 2. In this simulation the aluminium specimen (Ω_1) was considered with $\chi_{m1} = 22 \cdot 10^{-6}$, length of specimen $z = 20$ mm, thickness $a = (3,\ 5$ and $7)$ mm. Specimen was immersed into the water with $\chi_{m2} = -9 \cdot 10^{-6} - (\Omega_2)$.

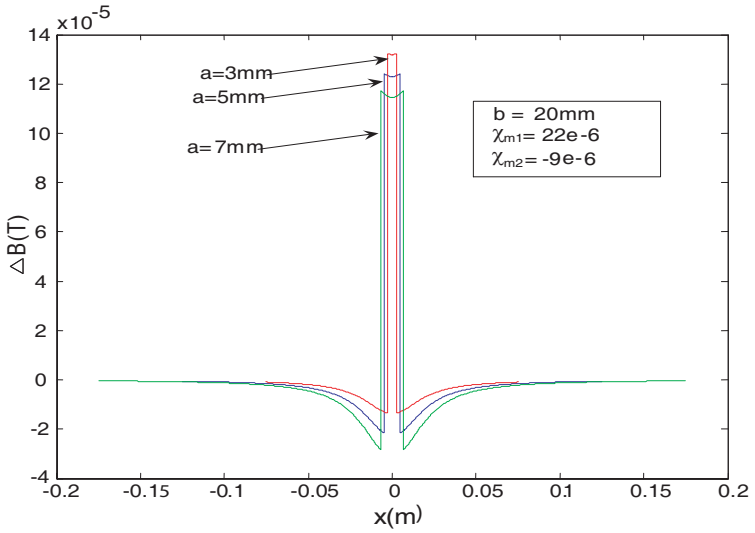


Figure 2: Result of numerical solution - magnetic flux density obtained by collocation method.

3. 3D NUMERICAL SOLUTION

Three dimensional numerical modeling was provided using FEM and Ansys software. The scalar magnetic potential was computed by solving of Laplace's equation

$$\Delta\varphi_m = \operatorname{div}\mu\,(-\operatorname{grad}\varphi_m) = 0. \tag{15}$$

One of used model is in Fig. 3. Here weakly paramagnetic specimen is surrounded by diamagnetic reference substance. The model was meshed with Solid96 element type. Boundary conditions were

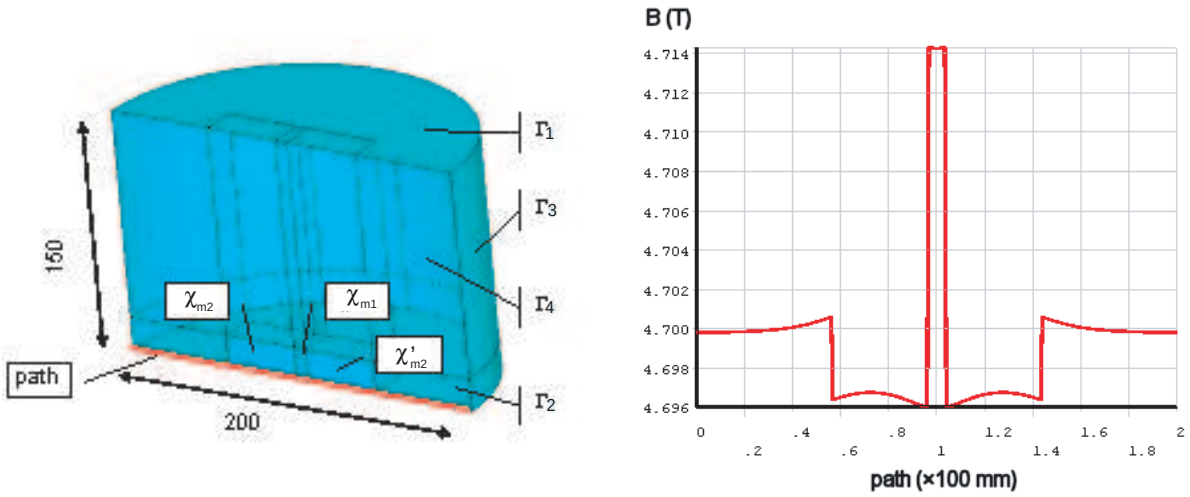


Figure 3: FEM model (left) and compute result (right) for magnetic flux density.

set up to achieve induction $B_0 = 4{,}700\,\mathrm{T}$ in z-axes direction: $\varphi_\mathrm{m} = const.$ on the surfaces Γ_1, Γ_2, $\frac{\partial \varphi_\mathrm{m}}{\partial \mathrm{n}} = 0$ on the shell surface Γ_3.

FEM modeling and solution is described in [8]. One of obtained results — the module of magnetic induction B along the "path" marked in Fig. 3 — is shown in the figure in the right.

4. CONCLUSIONS

Both method of numerical modeling of magnetic field deformation in MRI, caused by weakly magnetic specimen, which were described here, was compared with experimental results [3, 8]. Proximity of measured and numerically modeled data was good — see [5]. To enable comparison, simulation was adjusted to the same conditions as experiment: size of sample, susceptibilities and magnetic field of $\mathrm{B} = 4.7\,\mathrm{T}$.

Based on this simulation, MR measuring technique was founded [8], which is suitable for substances with no signal in MR tomography. The method uses Gradient Echo (GE) method and benefits from magnetic induction field shape in specimen vicinity, which is immersed in reference medium with measurable MR signal. After an optimization this method can be used for investigation of the materials used in MR tomography as well as of biological tissues affecting quality of MR images.

ACKNOWLEDGMENT

This work is supported by the grants GAAV B208130603 and GAAV B208130604.

REFERENCES

1. Ernst, R. R., G. Bodenhausen, and A. Wokaun, *Principles of NMR in One and Two Dimensions*, Oxford Science Publishing, 1987.
2. Sepúlveda, N. G., I. M. Thomas, and J. P. Wikswo Jr., "Magnetic susceptibility tomography for three-dimensional imaging of diamagnetic and paramagnetic objects," *IEEE Transaction on Magnetics*, Vol. 30, No. 6, 5062–5069, 1994.
3. Kubasek, R., M. Steinbauer, and K. Bartusek, "Material influences in MR tomography, measurement and simulation," *Journal of Electrical Engineering*, Vol. 8, 58–61, Zilina, 2006.
4. Wang, Z. J., S. LI, and J. C. Hasselgrove, "Magnetic resonance imaging measurement of volume magnetic susceptibility using a boundary condition," *Journal of Magnetic Resonance*, Vol. 142, 477–481, 1999.
5. Hwang, S. N. and F. W. Wehrli, "Experimental evaluation of surface charge method for computing the induced magnetic field in trabecular bone," *Journal of Magnetic Resonance*, Vol. 139, 35–45, 1999.
6. Eastham, J. F., R. J. Hill-Cottingham, I. R. Young, and J. V. Hajnal, "A method of inverse calculation for regions of small susceptibility variations," *IEEE Transaction on Magnetics*, Vol. 33, No. 2, 1212–1215, 1997.
7. Eastham, J. F., R. J. Hill-Cottingham, I. R. Young, J. V. Hajnal, P. J. Leonard, and W. Lin, "Use of finite element analysis to calculate field changes in low susceptibility materials," *Proceedings of Compumag 95 International Conference*, 546–547, Berlin, Germany, 1995.
8. Steinbauer, M., "Magnetic susceptibility measurement using MRI methods," PhD thesis, Brno University of Technology, 2005.

Experiments with the Effect of Non-homogenous Parts into Materials

Pavel Fiala[1], **Eva Kroutilova**[1], **Miloslav Steinbauer**[1]
Premysl Dohnal[1], **Michal Hadinec**[1], **and Karel Bartušek**[2]

[1]Brno University of Technology, Czech Republic
[2]Academy of Scince of the Czech Republic, Czech Republic

Abstract— The article describes possibility of the deformity magnetic field numerical modeling into surroundings of the measured dia-paramagnetic specimens. This deformation study is given for the purposes MR measuring, where MR images are deformed from the point of view of susceptibility heterogeneous materials. For the numerical simulation was chosen elementary configuration. The results was verified by MR experimental measurement. As the MR measurement method was used techniques of the gradient echo.

1. INTRODUCTION

Measurement method for measuring susceptibility of the homogenous materials by magnetic resonance (MR) techniques were detailed described in texts [4–7] and the results were verified by finite element methods (FEM) numerical simulations. For measuring susceptibility of the heterogeneous materials with the complicated shapes MR techniques (it was described in [8, 9]) hadn't always unique outgoing values. It is necessary find the method, how it is possible reliable measured results verify. One of the possibility of verification is the use of FEM numerical analysis with ANSYS system. The main advantages of the FEM ANSYS system is the ability to solve the multi-physical models, anisotropic materials and possibility to use of APDL programming language. This all advantages haven't the other programs together.

The aim of our work is to find simply model for numerical modeling and NMR experiment-verification. There can be changed heterogeneous material properties and next experimentally measured. The idea of increase of MRI is in the hybrid experimental and numerical inverse method. The numerical results are used in the MRI experimentally obtained data. The application of numerical results to the NMR post-processing can much more increase the final images.

2. GEOMETRICAL MODEL

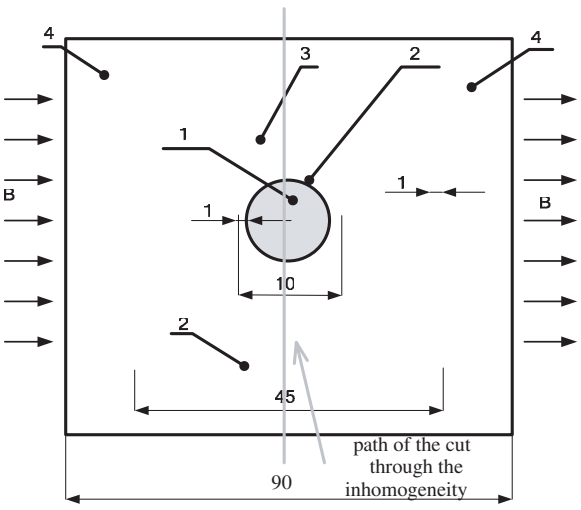


Figure 1: The sample geometry for numerical modeling, path of magnetic flux density evaluation.

The Fig. 1 describes the sample geometry for the numerical modeling. On both sides, the sample is surrounded by the referential medium (clay). As shown in Fig. 1, in the model there are defined five volumes with different susceptibilities. The materials are defined by their permeabilites: material No. 4 – the medium outside the cube (air, $\chi = 0$), material No. 3 – referential medium

(clay, $\chi = -9,92.10^{-6}$), material No. 2 – the cube and the inhomogeneity covering (sodium glass, $\chi = -11,7.10^{-6}$), material No. 1 – the material inside the inhomogeneity (water, $\chi = -12,44.10^{-6}$). The permeability rate was set with the help of the relation $\mu = 1 + \chi$.

For the sample geometry according to Fig. 1, the geometrical model was built in the system Ansys by FEM. In the model there was applied the mesh of elements, type Solid96 (ANSYS). The boundary conditions (4) were selected for the induction value of the static primary magnetic flux density to be $B_0 = 4,7000\,\mathrm{T}$ in the direction of the z coordinate (the cube axis) – corresponds with the real experiment carried out using the MR tomograph at the Institute of Scientific Instruments, ASCR Brno.

3. NUMERICAL ANALYSIS

The numerical modeling was realized using the FEM with the ANSYS system and it was detailed described into reference [4]. The numerical modeling results are represented in Fig. 2, Fig. 3. The numerical modeling results were then used for the representation of the module of magnetic flux density $\boldsymbol{B}$ along the defined path. For the model meshing, the element sizes selected as optimum were $0,25.10^{-3}\,\mathrm{m}$ and $0,125.10^{-3}\,\mathrm{m}$. The boundary conditions $\pm\varphi_m/2$ were set to the model edges, to the external left and right boundaries of the air medium, as represented in Fig. 1. The excitation value $\pm\varphi_m/2$ was set by using again the relation (4). This is derived for the assumption that, in the entire area, there are no exciting currents, therefore there holds for the $\mathbf{rot}\ \boldsymbol{H} = \mathbf{0}$ and the field is irrotational. Consequently, for the scalar magnetic potential φ_m holds

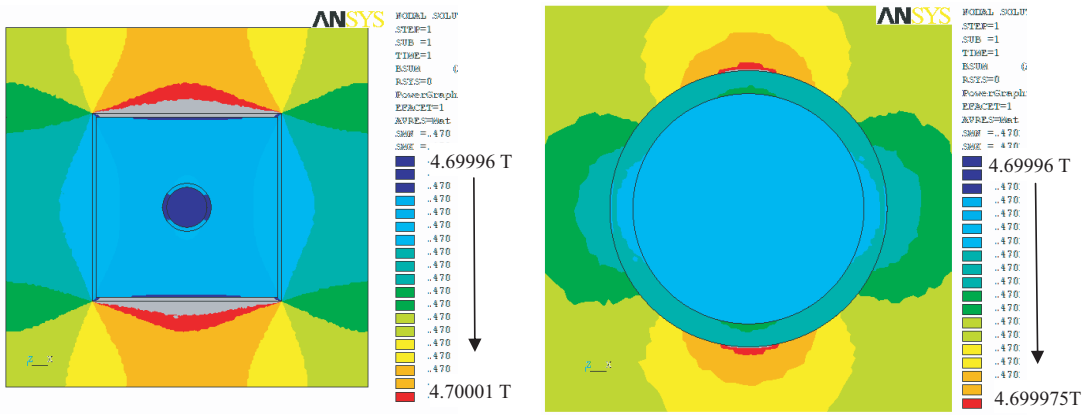


Figure 2: The numerical solution of the FEM model in the ANSYS system, distribution of module magnetic flux density B.

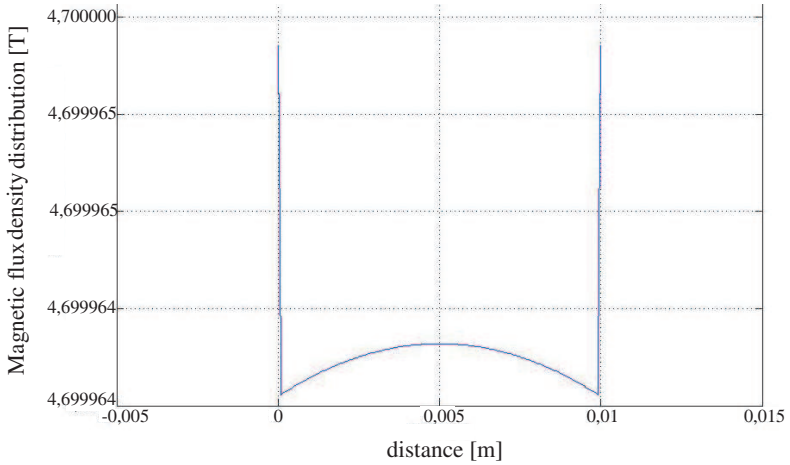


Figure 3: The magnetic flux density distribution change into water inhomogeneity $\Delta B = 36\,\mu\mathrm{T}$, ANSYS numerical analysis results.

$$\boldsymbol{H} = -grad\varphi_m \tag{1}$$

The potential of the exciting static field with intensity H_0 is by applying (2)

$$\varphi_m = \int \boldsymbol{H}_0 \cdot \boldsymbol{u}_z dz = H_0 \cdot z \tag{2}$$

where $\boldsymbol{u}_z$ is the normal vector on z direction. The

$$H_0 = \frac{B}{\mu_0 \cdot \mu_r} \tag{3}$$

Then

$$\pm \frac{\varphi_m}{2} = \frac{B \cdot z}{2\mu_0} = \frac{4,7000 \cdot 90 \cdot 10^{-3}}{2\mu_0} \tag{4}$$

where z is the total length of the model edge.

4. EXPERIMENTAL MEASUREMENT

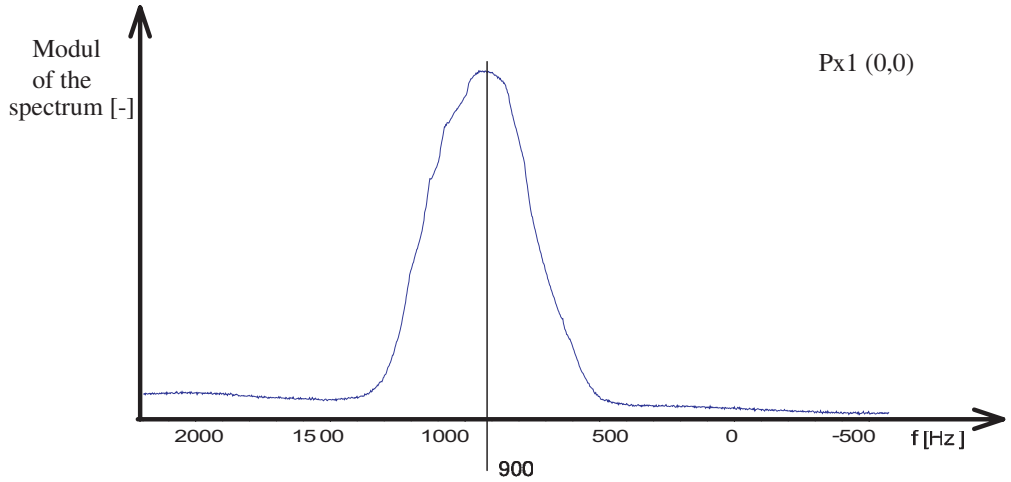


Figure 4: The experimentally measured spectral characteristic into water inhomogeneity.

The experimental measuring was realized using the MR tomograph at the Institute of Scientific Instruments, ASCR Brno. The tomograph elementary field $B_0 = 4,7000\,\mathrm{T}$ is generated by the superconductive solenoidal horizontal magnet produced by the Magnex Scientific company. The corresponding resonance frequency for the 1 H cores is 200 MHz. The numerical modeling and analysis of the task have verified the experimental results and, owing to the modifiability of the numerical model, we have managed to advance further in the experimental qualitative NMR image processing realized at the ISI ASCR. As the MR measurement method was used technics of the gradient echo. The results spectral characteristic into water inhomogeneity are given into Fig. 4. The relation for computation frequency to the magnetic flux density distribution change ΔB is

$$\Delta B = \frac{2\pi\Delta f}{2,67e^8} \tag{5}$$

where f (Hz) is measured frequency and ΔB (T) is magnetic flux density distribution change. For water inhomogeneity according to Fig. 1 is value magnetic flux density distribution change from graph Fig. 4 equal. $\Delta B = 4,7\,\mu\mathrm{T}$. The frequency $f_{max} = 900\,\mathrm{Hz}$ with maximum diagram value in Fig. 4 on histogram according to relation (5) correspond to $\Delta B = 21.15\,\mu\mathrm{T}$. This is the right results, because the difference of magnetic flux density from numerical results , Fig. 2 is approximately $\Delta B = 15.0\,\mu\mathrm{T}$. The accuracy result was obtained from path solution. The numerical results are showed in Fig. 3. There is the difference of magnetic flux density $\Delta B = 36.0\,\mu\mathrm{T}$. The better result can be done with the numerical result histogram solution.

5. CONCLUSIONS

The result showed, that it is possible use FEM modelling to the heterogeneous material analysis and use for the magnetic field deformation study into MR experiments. The mathematical experiments showed good results in surroundings of the heterogeneous objects with the complicated shapes.

ACKNOWLEDGMENT

The research described in the paper was financially supported by research plans GAAV B208130603, MSM 0021630516 and GA102/07/0389.

REFERENCES

1. *Ansys User's Manual*, Svanson Analysys System, Inc., Huston, USA, 1994–2006.
2. Fiala, P., E. Kroutilová, and T. Bachorec, Modelování elektromagnetických polí, počítačová cvičení, vyd. Brno: FEKT, VUT v Brně, Údolní 53, 602 00, s. 1–69, Brno, 2005.
3. Steinbauer, M., Měření magnetické susceptibility technikami tomografie magnetické rezonance. vyd. Brno: FEKT, VUT v Brně, Údolní 53, 602 00, Brno, 2006.
4. Fiala, P., E. Kroutilová, M. Steinbauer, M. Hadinec, K. Batrusek, "Inversion reconstruction of signals measured by the NMR techniques," *Progress In Electromagnetics Research Symposium Proceedings*, 430–433, Prague, Czech Republic, August 27–30, 2007.
5. Steinbauer, M. and E. Kroutilová, "Measurement of weakly magnetic material using new MR technique," *Applied Electronics 2006*, University of West BohemiaDepartment of Applied Electronics and TelecommunicationsUniverzitn 8306 14 PilsenCzech Republic, Pilsen, Czech Republic, ISBN: 80-7043-442-2, s. 87–91, 2006.
6. Steinbauer, E. and M. Batrusek, "Measurement and simulation of material influence in MR tomography," *Proceedings of the International Workshop ISEP-UTEE 3. -6. 9. 2006 Paris*, L'Institut Supérieur d'Électronique de Paris, Paris, France, ISBN: 80-214-3250-0, s. 87–92, 2006.
7. Hadinec, M., P. Fiala, E. Kroutilová, M. Steinbauer, K. Batrusek, "Magnetic field approximation in MR tomography," *PIERS Online*, Vol. 3, No. 8, 1250–1253, 2007.
8. Batrusek, K., Speciální metody měření difúzních koeficientů metodami nukleární magnetické resonance. n. n. Brno: VUTIUM, ISBN: 978-80-214-3379-3, 21 s, 2007.
9. Batrusek, K. and Z. Dokoupil, "Magnetic resonance diffusion measurement method in heterogenous systems," *Nodern Development of Magnetic Resonance*, 1. Kazan, Russian Federation, Zavoisky Physicel Technical Institute Kazan, s. 142–143, 2007.

Optimization Method of EMI Power Filters and Its Measurement

Z. Szabó, J. Sedláček, and M. Hadinec
BUT FEKT Brno, Czech Republic

Abstract— One of the most important problems solved nowadays is the improvement of the electronic systems immunity. This paper deals with the modelling and synthesis of EMC power filters. This is very important in the field of the electromagnetic compatibility and EMC filter design and optimisation. Various types of EMC filters are discussed. Idea of the synthesis and optimisation of EMC filters is illustrated on example. Results of our synthesis of EMC filter are shown at the conclusion.

1. INTRODUCTION

As the complexity of high-speed electronic system packages increase, engineers and designers are required to take control of more and more aspects of electrical and mechanical engineering early in the design cycle. In order to achieve the objective of faster time-to-market and to be cost effective one needs to be able to predict the electromagnetic radiated emission noises of the system design [1].

Sensitivity of devices to electromagnetic disturbances is increasing in many industrial fields. For example in EMC modeling and simulation on chip level [2], automotive equipments EMC modeling [3] etc.. Due to the increasing amount of devices sensitive to electromagnetic disturbances solution of problems coupled with electromagnetic compatibility is very important. Problems with electromagnetic compatibility can be suppressed by using special circuit elements. In area of EMC (electromagnetic compatibility) are very often used EMI (electromagnetic interference) filters. Modeling and synthesis of EMI filters is described in this paper. The created model enables to investigate influence of mismatched condition very quickly without measurement of filter. The great advantage of optimization method is that enable to optimize resulting filter model parameters by usage of usually accessible software for network analysis without requirement of special numerical programs what brings new possibility for many designers in area of EMC filter design and optimization. The great advantage of optimization method is that enable to optimize resulting filter model parameters by usage of usually accessible software for network analysis without requirement of special numerical programs what brings new possibility for many designers in area of EMC filter design and optimization.

2. EMI FILTERS

Electromagnetic interference (EMI) can be reduced to acceptable level using filter circuits usually referred as EMI or RFI filters. EMI filters are usually low-pass filter circuits with serial choke coils and parallel capacitors. These filters can be generally divided to two different groups. First group are named as data filters — are used namely in telecommunication systems. EMI data filters are performed as well known low-pass filter configurations (LC ladder circuits). Because these filters are constructed for constant load and generator impedances, design and optimization of filters can be realized according known design and optimization procedures.

The second group of EMI filters are filters used in power electronic. In comparison to EMI data communications filters EMI power filters operate typically under mismatched impedance conditions. This major problem of EMI filter design for power electronic equipment is caused by the arbitrary generator and load impedances. These impedances are really arbitrary because neither their value can be known, filters are installed in different equipments and supply network. The design of power EMI filters is different then well known procedures of classical filter design and requires some special view and procedures.

EMI filters are generally two-ports characterized by insertion loss (IL) rather then voltage attenuation. An insertion loss definition and measurement method is clear from Fig. 1. The difference between the measured voltage appearing beyond the insertion point before (switch position 1) and after the filter insertion (switch position 2) can be expressed as:

$$IL = 20 \log \left(\frac{U_{L1}}{U_{L2}} \right). \tag{1}$$

The voltage U_{L1} can be expressed using resistances of load and generator, and then insertion loss is given:

$$IL = 20 \log \left(\frac{U_g}{U_{L2}} \frac{R_L}{R_g + R_L} \right). \tag{2}$$

The requirement of insertion loss value must be fulfill in wide frequency range from DC to frequencies about hundred MHz. Thus analysis and measurement of the insertion loss must be made by filter design process in wide frequency range for many frequencies. Such a measurement procedure is not highly desirable in practical engineering. The chart in Fig. 2 presents typical frequency characteristic of insertion loss of EMI filters.

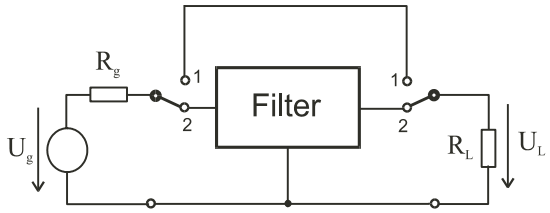


Figure 1: Insertion loss definition and measurement.

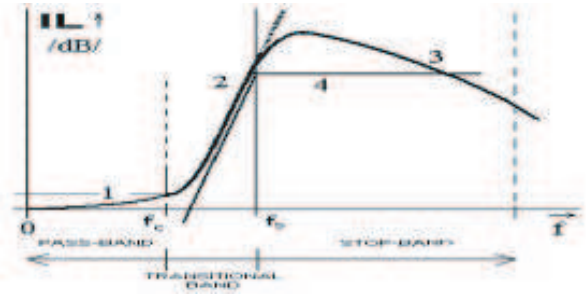


Figure 2: Typical frequency characteristic of EMI filter.

In the pass band insertion loss must be negligible from cut-off frequency f_c it monotonically increases. At the stop frequency f_s reaches insertion loss required value, up the stop frequency f_s due to parasitic effects exhibit curve some imperfections and usually decreases. After determining the required insertion loss in the stop band-pass, the next step of filter design is to choose a circuit configuration. Important factors may include a limitation on capacitive current for grounded equipments or the acceptable voltage drop across power line filters. For stringent suppression requirements must be also consider the mismatched impedance conditions. In area of power electronic EMC filter most often are used low-pass LC ladder filters in L, PI or T configurations. For high-performance applications are used also multistage LC circuits with higher number of basic sections. In power engineering practice, multistage filters having more then four stages are not very common. To suppress EMI on all wires, filter prototypes must be inserted in every wire of power lines. Thus power filter network becomes more complex with an increase in the number of wires to be filtered. The two-wire EMI filter should be studied as a six-terminal network. EMI power filters are inserted most often in three phase main supply lines and then must be filtered each wires including neutral. The complexity of EMI filters then significantly increases. The measurement of insertion loss in this case must be realized separately for all terminal pairs. According of used measurement system (symmetric, asymmetric or non-symmetric) the unused terminal pairs must be connected together to obtain the lowest insertion loss value. These specifications require the unused terminals to be grounded, ungrounded, or linked to ground through specific impedance [7].

3. MODELING OF EMI POWER FILTERS

The synthesis of proper filter models (equivalent circuits) including function elements as well as parasitic elements is one from important parts of successful EMI filter design and optimization. Using modeling techniques can be analyzed the effects of parasitic phenomena and impedance mismatch.

In present time PC technique enables to apply direct calculation method very easy. The direct calculation method is also the simplest approach for generating a complete EMC filter model. This modeling method is based on equivalent filter circuit synthesis by means of built-in filter elements. The models can be synthesized from the limited data available from manufacturers but also with measured data. To express filter performance in required wide frequency range, the basic filter elements must be assumed not ideal. Basic electrical element must be replaced by equivalent circuit including their parasitic elements (Fig. 3). The approximate values of parasitic elements of most often used EMI filter elements (inductors and capacitors) are summarized in Table 1 [4].

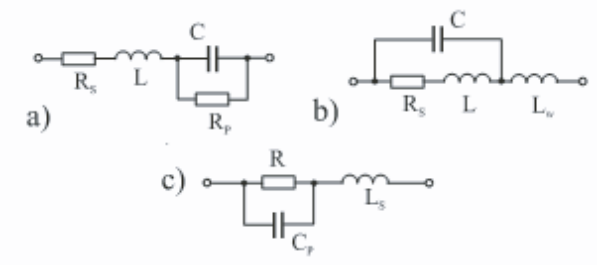


Figure 3: Equivalent circuits: (a) of capacitor, (b) of inductor, (c) of resistor.

C	$L_{parasitic}$	Remark
< 10 nF 10 nF - 1 μF >1 μF	10 - 20 nH 40 nH 30 - 100 nH	feedthrough type capacitors in orders lower then 1/10 values
L	R_S	$C_{parasitic}$
< 10 μH 50 μH - 200 μH >200 μH	1,5 mΩ 10 mΩ 0,5 Ω	2 pF 5 pF 10 - 30 pF

Table 1: Typical element values of real filter elements.

4. SYNTHESIS OF EMI POWER FILTER MODEL

The model of the filter for three phase power FN 256-64-52 is here presented as an example of EMC filter model synthesis and optimization. The first step of filter model synthesis was grown from known basic (from manufacturer's data sheet) filter topology (Fig. 4).

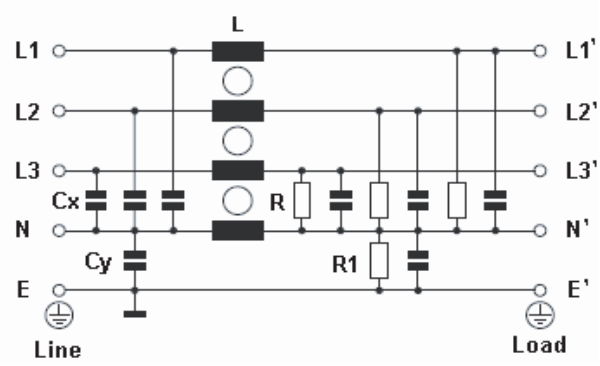


Figure 4: Typical electrical PI topology of power four-wire EMI filter.

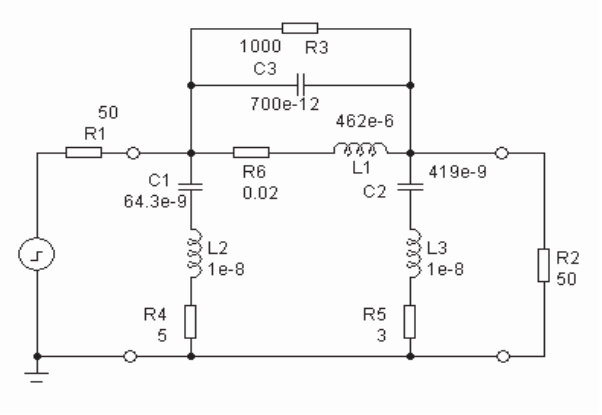


Figure 5: The resulting circuit diagram of filter model (filter FN 256-64-52).

In the second step the ideal basic elements (R, L, C) were replaced by real models for each of the tree lines. The initial filter value parameters were approximated. Using commercially available analyzers TINA and P-SPICE 9 the filter with equal load and generator resistors (50 Ω) was analyzed. Frequency curve of insertion loss was obtained from circuit analysis. It was compared with the frequency curve presented by the same measuring conditions in manufacturer's data sheet.

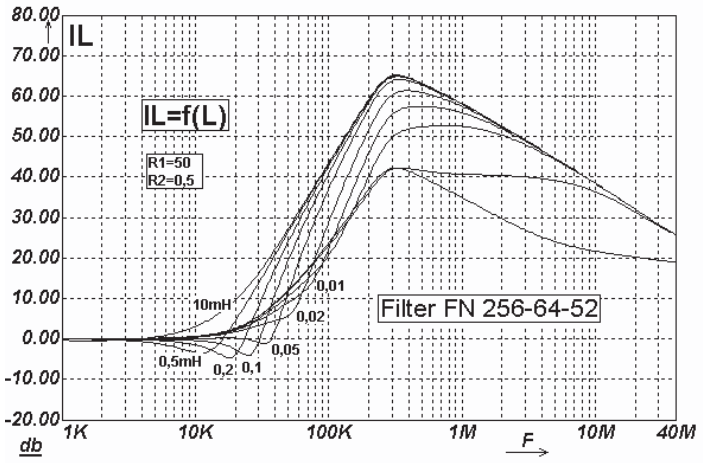


Figure 6: Insertion loss characteristics as function of load inductance.

Then the frequency curve the model of the filter was optimized using optimizer routines from analyzers. As the result of the optimization procedures the values of each element of the model of the filter were obtained. The resulting circuit diagram of the model of the filter with the values of its parameters is shown in Fig. 5.

Using created filter model an influence of resistance of generator R_1 and resistance load R_2 on insertion loss of the filter was investigated to determine worst case of operation. How is seen from curves (Figs. 9, 10), the effect of the mismatch conditions in the worst case can decrease initial insertion loss of about 20 dB in the entire working frequency range what must be by assumed.

We can see from Fig. 6, that in real conditions must be taken into account not only resistances, but also inductances of loads and generators.

Using measuring chamber is shown on figures (Fig. 7, Fig. 8) the insertion lost characteristics of the EMC four wire power filter was measured. Results of the measurement were compared with the insertion lost characteristics obtained from the filter modeling.

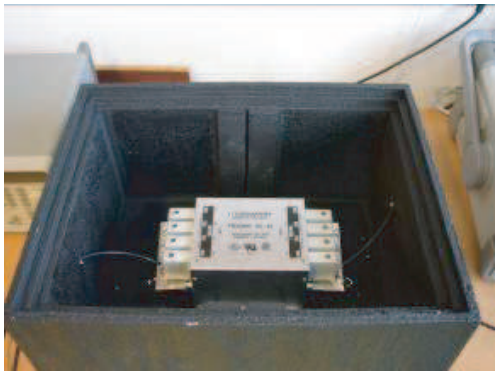


Figure 7: Measuring chamber with EMC four wire power filter.

Figure 8: Measuring place with measuring chamber and devices.

5. CONCLUSIONS

The paper deals with the problems of modeling of optimization of EMC filters. The method of synthesis and optimization was described on example of EMC filter. The resulting circuit diagram of the model of the filter with its parameters and the insertion lost characteristics were shown. Influence of mismatched conditions of the insertion lost was investigated. Resulted characteristic from synthesis were compared with experimentally obtained characteristics. In the future the method decrypted in this paper will be improved and the method will be tested on various types of EMC filters.

ACKNOWLEDGMENT

This work has been supported by the Research Project: MSM 0021630513 and Grant Agency of the Czech Republic under Grant 102/03/1181, 102/04/0442.

REFERENCES

1. Archambeault, B., S. Pratapneni, L. Zhang, D. C. Wittwer, and J. Chen, *2001 IEEE International Symposium on Electromagnetic Compatibility, EMC*, USA, 2001.
2. Steinecke, T., W. John, H. Koehne, M. Schmidt, "EMC modeling and simulation on chiplevel," *2001 IEEE International Symposium on Electromagnetic Compatibility, EMC*, 2001.
3. Poinsignon, J. M., P. Matossian, B. Mazari, and F. Duval, *IEEE International Symposium on Electromagnetic Compatibility, EMC'03*, 2003.
4. Thayne, L., *Electromagnetic Compatibility in Power Electronics*, IEEE Press, J. K. Eckert, Sarasota, USA, 1995, 2004.
5. Wiliams, T. and K. Armstrong, *EMC for Systems and Installations*, Newnes, Butterworth-Heineman, Oxford, Great Britain, 2000.
6. Rybak, T. and M. Steffka, *Automotive Electromagnetic Compatibility (EMC)*, Kluwer Academic Publishers, Norwell, Massachusetts, USA, 2004.
7. Dřínovský, J. and J. Svačina, "Estimation of EMI filter performance for the 'Worst-Case' system," *Radioengineering*, Vol. 15, No. 4, ISSN 04-510-79, Brno, 2006.

Estimation Method of Quasi-wavefronts for UWB Radar Imaging with LMS Filter and Fractional Boundary Scattering Transform

T. Sakamoto, K. Teshima, and T. Sato
Graduate School of Informatics, Kyoto University
Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501, Japan

Abstract— UWB (Ultra Wide-Band) radar has a variety of applications including security surveillance systems. The SEABED algorithm is a fast imaging method for UWB radar, that uses a reversible transform between the real and data spaces [1]. We introduce an intermediate space between the real and data spaces [2]. Curves in the intermediate space can be smooth, and can be used to extract quasi-wavefronts (the equi-phase surface). In this paper, we use LMS (Least-Mean-Square) filters in the intermediate space for imaging arbitrary target shapes.

1. INTRODUCTION

UWB (ultra-wideband) pulse radar is a promising candidate as an environment measurement, or sensing, method for robots. Radar imaging for a nearby target is known as an ill-posed inverse problem: a problem that has been extensively studied [3, 4]. However, conventional algorithms require long computational time, that makes it difficult to apply UWB to real-time operations for robots. We have proposed a fast radar imaging algorithm, the SEABED algorithm, for UWB pulse radar [5, 6]. This algorithm is based on a reversible transform, IBST (Inverse Boundary Scattering Transform), between the target shape and observed data. This transform enables us to estimate target shapes quickly and accurately in a noiseless environment. The SEABED algorithm extracts equi-phase surfaces (also called quasi-wavefronts) first, and then applies an IBST to obtain the estimated image. However, in a noisy environment the image estimated by the SEABED algorithm is degraded because the quasiwavefronts cannot be accurately estimated. In this paper, we introduce an FIBST (Fractional IBST) [2] to the quasi-wavefront extraction process. This transform is obtained by expanding the conventional IBST, which enables us to deal with the intermediate space between real and data spaces, and propose a stable quasi-wavefronts extraction algorithm . We show some application examples with numerical simulations.

2. SYSTEM MODEL

We assume a mono-static radar system. An omni-directional antenna is scanned along a straight line. UWB pulses are transmitted at fixed intervals and received by the antenna. The received data is A/D converted and stored in memory. We estimate target shapes using the obtained data. We deal with a 2-dimensional problem. We define a real space in which targets and antenna are located. We express the real space with the parameters (x, y). Both x and y are normalized by λ, which is the center wavelength of the transmitted pulse in air. We assume $y > 0$ for simplicity. The antenna is scanned along the x-axis in r-space. We define $s(X, Y)$ as the received waveform after applying a matched filter at the antenna-location $(x, y) = (X, 0)$. Here, we define Y with time t and the speed of the radiowave c as $Y = ct/(2\lambda)$. We define a data space expressed by (X, Y).

3. SEABED ALGORITHM

In previous work we developed a fast radar imaging algorithm, 'SEABED', based on a BST (Boundary Scattering Transform) [5–8]. The algorithm uses a reversible transform, BST, between target shapes and pulse delays. The BST is expressed as

$$X = x + y\frac{\mathrm{d}y}{\mathrm{d}x}, \tag{1}$$

$$Y = y\sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}, \tag{2}$$

where (X, Y) is a point on a quasi wavefront, and (x, y) is a point on the target boundary [1]. We have clarified that the inverse transform of the BST is given by

$$x = X - Y\mathrm{d}Y/\mathrm{d}X, \tag{3}$$

$$y = Y\sqrt{1 - (\mathrm{d}Y/\mathrm{d}X)^2}, \tag{4}$$

where we assume $|\mathrm{d}Y/\mathrm{d}X| \leq 1$. This condition is required because y should be a real number. First, quasi wavefronts are extracted from the received signals $s(X, Y)$ in the SEABED algorithm. Quasi wavefronts are extracted to satisfy the conditions $\mathrm{d}s(X, Y)/\mathrm{d}Y = 0$ and $|\mathrm{d}Y/\mathrm{d}X| \leq 1$. Finally, we apply the IBST to the quasi wavefronts, and obtain the final image. The extraction of quasi-wavefronts is critical to obtaining high-quality images with this algorithm. However, the quasi-wavefront cannot be accurately estimated for a noisy case. It is to solve this is the problem that we propose a new algorithm in this study.

4. FRACTIONAL BOUNDARY SCATTERING TRANSFORM

Here, we explain a fractional boundary scattering transform obtained by expanding the conventional boundary scattering transform [2]. We define the fractional boundary scattering transform, FBST (α) as

$$x^{(\alpha)} = x + \alpha y \frac{\mathrm{d}y}{\mathrm{d}x}, \tag{5}$$

$$y^{(\alpha)} = y\sqrt{1 + \alpha\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}. \tag{6}$$

These equations contain a parameter $\alpha\,(0 \leq \alpha \leq 1)$, which is not included in the conventional boundary scattering transform. We call $(x^{(\alpha)}, y^{(\alpha)})$ a 'fractional transform quasi wavefront'. We call the space expressed by $(x^{(\alpha)}, y^{(\alpha)})$ a 'fractional transform space'. The fractional transform quasi-wavefront is equivalent to the conventional quasi-wavefront for $\alpha = 1$ and the fractional transform quasi-wavefront is equivalent to the target shape for $\alpha = 0$. The fractional inverse boundary scattering transform, FIBST (α) is defined in relation to the FBST similarly to the relationship between BST and IBST; by changing the sign of α.

Our study [2] clarified that data in the fractional transform space can be smooth regardless of the shape of targets. This characteristic can be used in the extraction of quasi-wavefronts. Additionally, data in any space can be transformed to arbitrary space as in Fig. 1, a fact that can be effectively used to develop our new algorithm. An example of data in 3 spaces (real, data, and fractional transform) is shown in Fig. 2. The data has an edge around $X = 1.5$, but the data in the fractional transform space is smooth, where we set $\alpha = 0.5$. In this way, we can avoid edges in the real and data space by applying an FBST to transform the data to the fractional transform space.

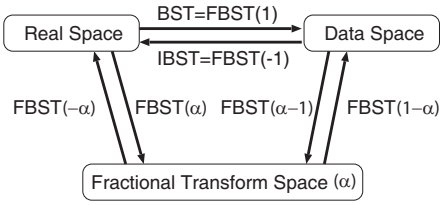


Figure 1: Relationships between spaces.

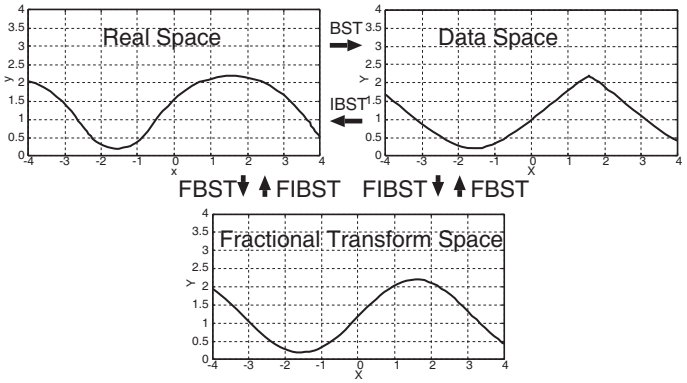


Figure 2: An example of data in real, data, and fractional transform spaces.

5. FRACTIONAL BOUNDARY SCATTERING TRANSFORM

A new extraction method for quasi-wavefronts is proposed here. 100 undesired interference points are assumed with the true quasi-wavefront points for each antenna position X. The first 10 true points are assumed and used as the initial value. We apply the FBST to the estimated quasi-wavefront to obtain the curve in the intermediate space, and apply a 5th-order LMS filter to estimate the entire curve. Then we apply the inverse FBST and obtain the predicted points. We

adopt the nearest point to the prediction as the estimation in the next step. This procedure is repeated. This process is shown in Fig. 3 and contrasted with the conventional method as used in the original SEABED algorithm [1].

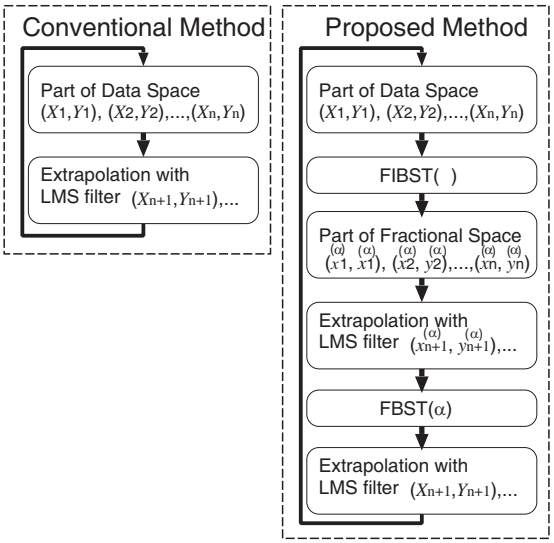


Figure 3: Procedures of the proposed and conventional quasi-wavefront extraction methods.

Figure 4 shows an example of the application of the proposed algorithm to the data shown in Fig. 2. The estimation of the quasi-wavefront until the 1st step in Fig. 4 is quite easy because of the smoothness in the data that means it does not depend on the method used. However, there is an edge around $X = 1.5$ in the data space. Simple LMS filtering fails to track the true quasi-wavefront here. The proposed algorithm applies FIBST to the data to obtain the fractional transform space data as black squares in step 2. Next, LMS filter prediction is applied to extrapolate the fractional transform space data in step 3. Finally, the FBST is applied to obtain the estimated quasi-wavefront in step 4. We adopt the nearest point to the prediction as the estimation in the next step. We repeat these procedures until the final point is estimated.

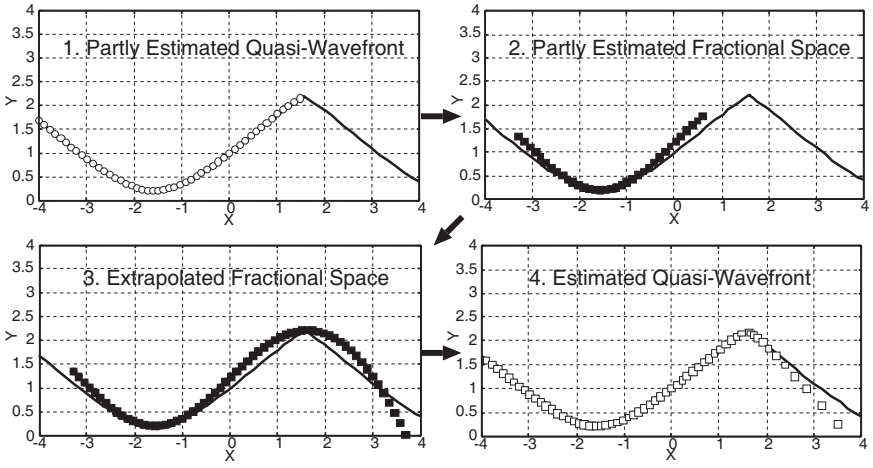


Figure 4: Application example of the proposed algorithm.

Figure 5 shows a comparison between the proposed method and the conventional method that applies the LMS filter in the data space rather than the intermediate space. The results show that the proposed method works while the conventional method produces a poor estimation. Fig. 6 shows the estimated image with the conventional method and the proposed method. For the conventional method, the shape for $x > 1.5$ is not estimated while our proposed method can correctly estimate the entire target shape.

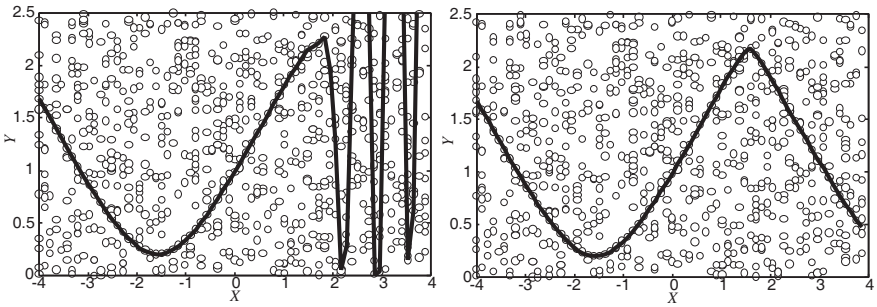


Figure 5: Quasi-wavefronts estimated with the conventional methods (left) and proposed methods (right).

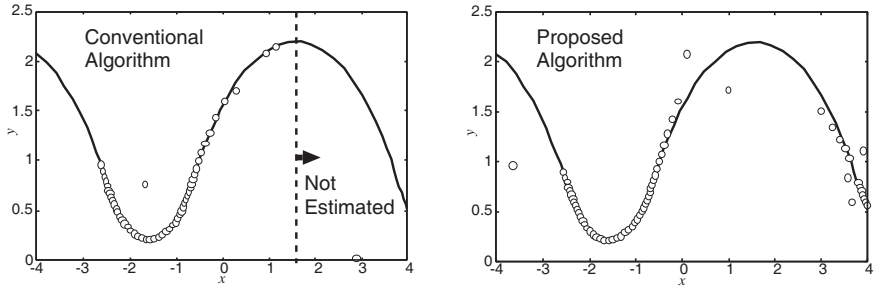


Figure 6: Images estimated with the conventional methods (left) and the proposed methods (right).

6. CONCLUSION

In this paper, we introduced a FIBST (Fractional IBST) to the quasi-wavefront extraction process, for the SEABED algorithm for UWB pulse radar imaging. This enables us to deal with the intermediate space between real and data spaces, and propose a stable quasi-wavefront extraction algorithm. The results of experimental application show that the proposed method maintains tracking data even in noisy environments. Additionally, the proposed method can estimate the entire target image while the conventional one cannot as it fails when multiple undesired points are caused by noise and interference.

REFERENCES

1. Sakamoto, T. and T. Sato, "A target shape estimation algorithm for pulse radar systems based on boundary scattering transform," *IEICE Transaction on Communications*, Vol. E87-B, No. 5, 1357–1365, May 2004.
2. Sakamoto, T., *IEICE Trans. on Commun.*, Vol. E90-B, No.1, 131–139, 2007.
3. Ferraye, R., J.-Y. Dauvignac, and C. Pichot, "An inverse scattering method based on contour deformations by means of a level set method using frequency hoping technique," *IEEE Trans. Antennas Propagat.*, Vol. 51, No. 5, 1100–1113, 2003.
4. Cui, T. J., W. C. Chew, X. X. Yin, and W. Hong, "Study of resolution and super resolution in electromagnetic imaging for half-space problems," *IEEE Trans. Antennas Propagat.*, Vol. 52, No. 6, 1398–1411, 2004.
5. Sakamoto, T. and T. Sato, "Fast imaging of a target in inhomogeneous media for pulse radar systems," *Proc. 2004 IEEE International Geoscience and Remote Sensing Symposium*, Vol. 3, 2070–2073, Sep. 2004.
6. Sakamoto, T. and T. Sato, "A fast algorithm of 3-dimensional imaging for pulse radar systems," *Proc. 2004 IEEE AP-S International Symposium and USNC/URSI National Radio Science Meeting*, Vol. 2, 2099–2102, June 2004.
7. Sakamoto, T. and T. Sato, "A phase compensation algorithm for high-resolution pulse radar systems," *IEICE Trans. Communications*, Vol. E87-B, No. 11, 3314–3321, Nov. 2004.
8. Sakamoto, T. and T. Sato, "A phase compensation algorithm for high-resolution pulse radar systems," *Proc. 2004 International Symposium on Antennas and Propagation*, 585–588, Aug. 2004.

Structural and Multiferroic Properties of $BiFe_{0.5}Co_{0.5}O_3$ Ceramics

Hai-Xia Lu, **Xiang-Yu Mao**, **Wei Wang**, **and Xiao-Bing Chen**
College of Physics Science and Technology, Yangzhou University, Yangzhou 225002, China

Abstract— The crystal and magnetic structure of polycrystalline $BiFe_{0.5}Co_{0.5}O_3$ prepared by a solid-state reaction method with multiple calcination. The sample are characterized by using various techniques: X-ray diffraction (XRD) study is carried out for phase determination and lattice parameter calculations ($a = 10.1812(4)$). The ferroelectricity of the samples has been confirmed by the hysteresis loops measurement but the leakage current is quite large, implying that Co substitution has no effect of the improvement on ferroelectricity in $BiFe_{0.5}Co_{0.5}O_3$ samples. The magnetization measurement performed at room temperature showed a perfect hysteresis loop with large remnant magnetization ($M_r \approx 0.55\,\mathrm{em\mu/g}$) and low coercive field ($H_c \approx 197\,\mathrm{Oe}$). The frequency dependence of admittance ($Y^{'}$) at room temperature turn out a resonance curve centered at $f \sim 12.6\,\mathrm{MHz}$ and a negative overshoot as a typical of resonancelike behavior. The dielectric property of the $BiFe_{0.5}Co_{0.5}O_3$ ceramics were investigated at intermediate frequencies ($10^2 \sim 10^6\,\mathrm{Hz}$) in the temperature rage of $100 \sim 560\,^{\circ}\mathrm{C}$.

1. INTRODUCTION

Multiferroic materials have renewed interest in recent years, in which both ferromagnetic and ferroelectric properties exist in the same phase [1]. As a result they have spontaneous magnetization which can be switched by an applied field, spontaneous polarization which can be reoriented by an electric field, and often some coupling between the two [2–4]. Special device application which have been suggested for such materials include multiple state memory elements, electric field controlled ferromagnetic resonance devices, and transducers with magnetically modulated piezoelectricity [5–8]. Many efforts have been devoted to find new materials with and to find multiferroic properties in known compounds. However, almost all those gigantic magnetoelectric effects occur essentially below liquid-nitrogen temperature.

$BiFeO_3$ (BFO) is an interesting candidate as a magnetoelectric materials because the ferroelectricity and antiferrimagnetic order present simultaneously at room temperature. G-type antiferromagnetic ordering takes place at 640 K, while ferroelectric order appears at a higher temperature of 1100 K [9]. One problem for BFO as a room-temperature multiferroics is its intrinsic antiferromagnetic ordering. In order to improve the properties of BFO ceramics, some attempts have been made including doping rare earth (RE) or Mn [10], respectively, on the Bi sites or Fe sites, and fabricating strained films [11]. Hhowever, little improvement in the magnetic properties of BFO has been achieved by element substitution, and the role of strain in magnetization also requires further investigation [12, 13].

Bucci had earlier prepared $BiCo_{1-x}Fe_xO_3$ and indicated a limiting value of 0.64 for x to obtain single phase material [14]. Vasudevan had also synthesized $BiCo_{1-x}Fe_xO_3$, Band confirmed that sample of $BiCo_{1-x}Fe_xO_3$ ($x \leq 0.7$) show complex magnetic behavior [14]. The magnetic behavior of sample with $x < 0.1$ is similar to that of $BiCoO_3$. When $0.1 \leq x \leq 0.2$, the sample show a slight field dependence of susceptibility. The samples with $0.5 \leq x \leq 0.7$ are ferromagnetic at room temperature. Magnetic hysteresis measurements behaved a value of 79.6 Oe for saturation magnetization in the $x = 0.7$ sample. When $x > 0.7$ were found to be biphasic. Therefore, it is worthwhile to experimentally investigate the effect of the substitution of Co for Fe on the physical properties of BFO.

$BiCo_{1-x}Fe_xO_3$ have, however, not been adequately reported about magnetic and other properties when x is a fix value. In this paper, the preparation of the Co (with $x = 0.5$) doped BFO sample is reported and the effects of Co substitution on structural, magnetic, and dielectric properties of $BiFeO_3$ are obtained. The results indicate that there exist a structural transition from rhombohedral phase to cubic in the $BiFe_{0.5}Co_{0.5}O_3$ and the addition of Co can improve significantly the ferromagnetic properties of $BiFeO_3$ bulk.

2. EXPERIMENTAL

Polycrystalline samples of $BiFe_{0.5}Co_{0.5}O_3$ were prepared by a solid state reaction method with multiple calcinations. The starting materials were Bi_2O_3 (99.99%), Co_2O_3 (99%), and Fe_2O_3

(99%), mixed together at the stoichiometric ratios designed for $BiFe_{0.5}Co_{0.5}O_3$ compositions with 5.0% excess Bi_2O_3 to compensate for the likely loss of Bi_2O_3 at higher temperatures. The oxide mixture were ball milled together for 24 h in ethanol, and were subsequently dried and sieved. The homogeneously mixed powders were calcined at 650∼700 °C for 24 h for several times. The fine calcined powders of BFCO were pressed into circular pellets of 12 mm in diameter and 2 mm in thickness. All the pellets were sintered at 830 °C for 1h for densification at atmospheric pressure, and then cooled down at a rate of 1 °C/min in air.

The XRD analysis was carried out to study the phase determination and lattice parameter calculations using $M03XHF_{22}$ diffractometry with Cu K_α radiation ($\lambda = 0.154056$ nm) (2θ range of 10∼70°). The magnetic moment of the samples were measured using an EV7–VSM in the range of fields -5 kOe $\sim$ 5 kOe at room temperature. Temprature dependence of the magnetic moments of the samples were measured from 25 to 500 °C at 2000 Oe by VSM (HH-15). The pellets were filed and polished to the thicknees of about 0.2 mm and 0.6 mm, coated with silver electrode for ferroelectric and dielectric measurement, respectively. The hysteresis loops of polarization verse electric field ($P - E$) were measured by a Precision LC (Radiant Technologies) standardized ferroelectric tester. The dielectric constant (ε') of the BFCO as a function of temperature ranging from 100 to 560 °C were performed in a Broadband Dielectric Spectrometer at a heating rate of 1 °C/min.

3. RESULTS AND DISCUSSION

Figure 1 shows the XRD pattern of the $BiFe_{0.5}Co_{0.5}O_3$ samples at room temperature and the inset is that of the $BiFeO_3$ sample. All the diffraction peaks of $BiFe_{0.5}Co_{0.5}O_3$ were indexed according to Ref. [14] and lattice parameters, derived by using A.C. Larson and R.B. Von Dreele, General Structure Analysis System (GSAS), are $a = b = c = 1.01812(4)$ nm and $\beta = 90°$C. Comparing with the XRD of $BiFeO_3$, the pattern indicates that Co-doped $BiFeO_3$ have a structural transformation from rhombohedral (space group $R3c$) [15] to cubic structure (space group $I23$).

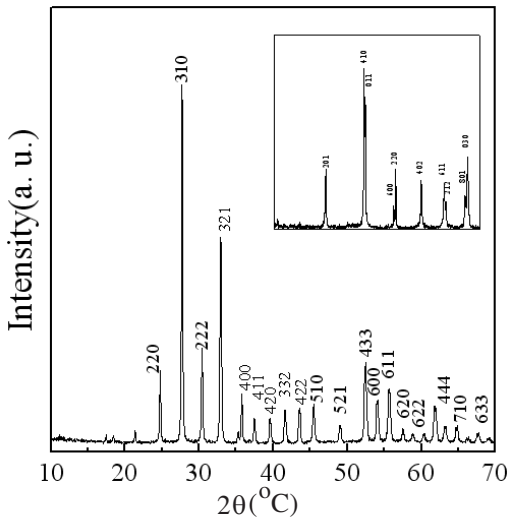


Figure 1: X-ray diffraction patterns of BFCO ceramics at room temperature. The inset shows the XRD of BFO ceramics.

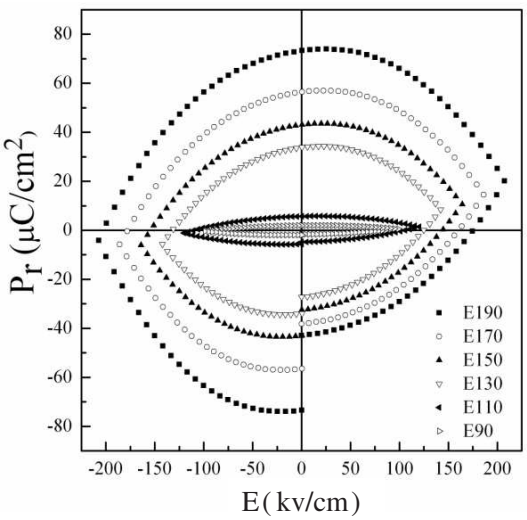


Figure 2: Ferroelectric hysteresis of BFCO cermics at different applied fields.

Figure 2 shows the ferroelectric hysteresis loops measurement on poled samples of BFCO at room temperature and different applied fields. BFCO did not give a perfect ferroelectric loop. However, on Co substitution, the leakage current is still quite large, implying that Co substitution has no effect of the improvement on ferroelectricity in BFCO samples. The reason may be that charged defects governed by Fe^{2+}, Co^{2+} irons, oxygen vacancies ($V_{O^{2+}}$), and /or bismuth vacancies ($V_{Bi^{3+}}$) may appear in the deoxygenated BFCO phases. These $V_{O^{2+}}$ and /or $V_{Bi^{3+}}$ vacancies will reduce the electrical resistivity of the sample, giving rise to high leakage currents in the samples [16]:

Figure 3 shows the magnetization M versus applied field H of BFCO at room temperature. Relatively symmetric hysteresis loop was obtained from which we measured the coercivity $H_c \approx$

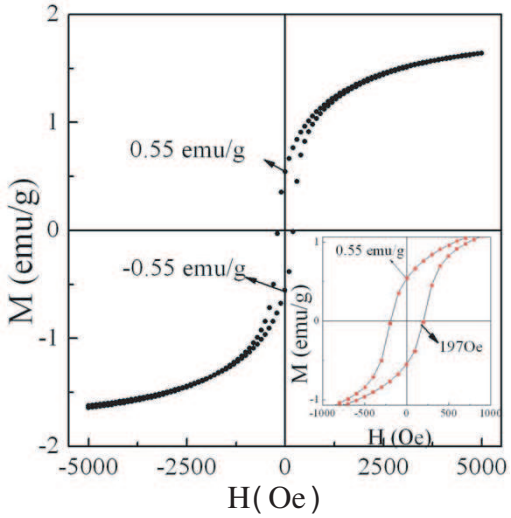


Figure 3: Magnetic hystersis loop of the BFCO ceramics. The inset shows the detail of the loop in the low field region.

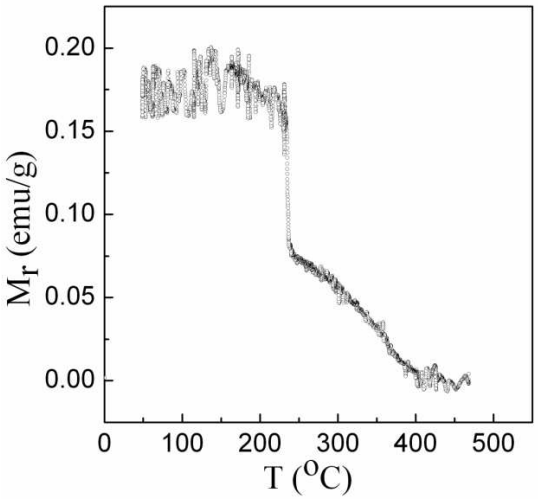


Figure 4: Temperature dependenc of the magnetic moments of BFCO ceramics, measured at 2000 Oe.

197 Oe, the average remanence magnetization $M_r \approx 0.55\,\mathrm{em\mu/g}$ and the magnetization at 5 kOe (defined as M_H) was 1.66 emμ/g. As shown in Fig. 4, the temperature dependence of magnetization indicates that the ferromagnetic phase can exist up to about 420 K, above 670 K the magnetic phase is paramagnetic, and between them an intermediate magnetic state is observed. The ferromagnetism is considerably improved by Co-doping. The enhancement of the magnetic parameters might be related to the increased magnetoelectric coupling inside Co-doped BFO. Presumably, the Fe^{2+}, Co^{2+} would coexist, giving rise to a relative large moment, in which the strong interaction would accordingly influence the magnetic nature of the compound.

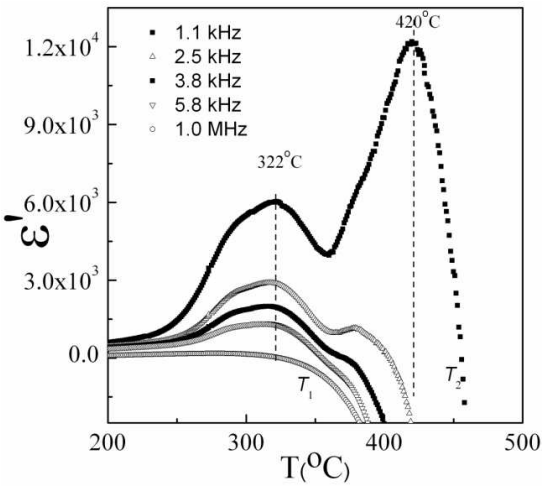


Figure 5: Dielectric constant vs temperature of BFCO ceramics at different frequencies.

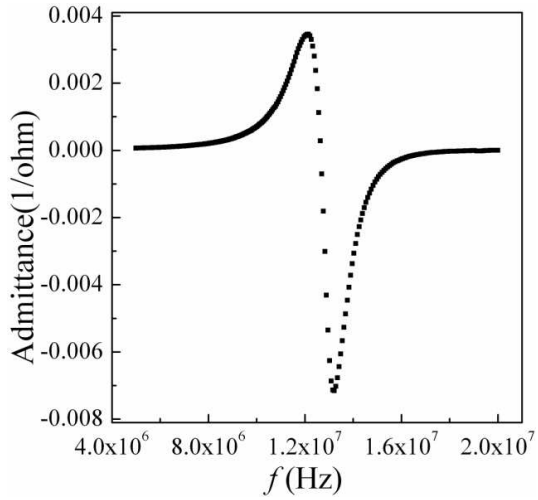


Figure 6: Admittance vs frequency of BFCO at room temperature.

Figure 5 shows the temperature variation of dielectric constant at different frequencies. Dielectric constant (ε') have two peaks with their heights decreasing as frequency increases. At low frequencies, the temperature of the first peak, $T_1 \sim 322\,^\circ$C, almost does not vary with frequency, while the second peak, $T_2 \sim 400\,^\circ$C, shifts to lower temperature side with increasing frequency. Above T_2 the dielectric constant (ε') reduces rapidly and changes to a negative value. At high frequencies when $f > 5.8$ kHz, the second peak was completely suppressed, meanwhile the first peak, T_1 also shifts to lower temperature as frequency increases. The T_1 is in the intermediate magnetic phase and may be related to the ferro-intermediate-magnetic-phase-transition. The negative dielectric permittivity is intriguing and its interpretation might be taken into account as the peculiar phase-separation formed by space-charge or interfacial polarization [17].

Figure 6 shows the plot of admittance versus frequency at room temperature for the BFCO. Prior to the measurement, the sample is electrically poled. It turns out a resonance curve centered at $f \sim 12.6\,\mathrm{MHz}$ and a negative overshoot as a typical resonance-like character in piezoelectric material. Such an anomalous resonance behavior confirms the relatively strong magnetoelectric interaction.

4. CONCLUSION

The $BiFe_{0.5}Co_{0.5}O_3$ ceramic sample was synthesized using solid-state reaction method with multiple calcinations. Co substitution at Fe site causes a structural transformation from rhombohedral (space group $R3c$) to cubic structure (space group $I23$). Though the electric properties are not enhanced by Co substitution, the improvement on both ferromagnetic and dielectric properties at room temperature are achieved. The magnetic phase can be determined as three regimes. The negative dielectric constants are observed in this ferromagnetic-ferroelectric system, requiring theoretical interpretation. It is hoped that this study would stimulate further investigation on this system.

REFERENCES

1. Schmid, H., *Ferroelectrics*, 162, 317, 1994.
2. Lottermoser, T. and M. Fiebig, *Phys. Rev. B*, 70, 220407, 2005.
3. Wang, J., J. B. Neaton, H. Zheng, et al., *Science*, 299, 1719, 2003.
4. Kimura, T., T. Goto, H. Shintani, K. Ishizaka, T. Arima, and Y. Tokura, *Nature*, 426, 55, London, 2003.
5. Dho, J. H., X. D. Qi, H. Kim, J. L. MacManus-Driscol, and M. G. Blamire, *Adv. Mater.*, Weinheim, Ger., 18, 1445, 2006.
6. Lorenz, B., A. P. Litvinchuk, M. M. Gospodinov, and C. W. Chu, *Phys. Rev. Lett.*, 92, 087204, 2004.
7. Lampis, N., C. Franchini, G. Satta, A. Geddo-Lehmann, and S. Massidda, *Phys. Rev. B*, 69, 064412, 2004.
8. Qi, X. D., M. Wei, Y. Lin, Q. X. Jia, D. Zhi, J. Dho, M. G. Blamire, and J. L. MacManus-Driscoll, *Appl. Phys. Lett.*, 86, 071913, 2005.
9. Kubel, F. and H. Schmid, *Acta Crystallogr., Sect. B: Struct. Sci.*, 46, 698, 1990.
10. Yang, C. H., T. Y. Koo, and Y. H. Jeong, *Solid State Commun.*, 134, 299, 2005.
11. Wang, J., J. B. Neaton, H. Zheng, V. Nagarajan, S. B. Ogake, B. Liu, D. Viehland, V. Vaithyanathan, D. G. Schlom, U. V. Waghmare, N. A. Spal-din, K. M. Rabe, M. Wuttig, and R. Ramesh, *Science*, 299, 1719, 2003.
12. Eerenstein, W., F. D. Morrison, J. Dho, M. G. Blamire, J. F. Scott, and N. D. Mathur, *Science*, 307, 1203a, 2005.
13. Ederer, C. and N. A. Spaldin, *Phys. Rev. B*, 71, 224103, 2005.
14. Vasudevan, S., C. N. R. Rao, et al., *Mat. Res. Bull.*, Vol. 14, 451-4-54, 1979.
15. Li, J. B., G. H. Rao, J. K. Liang, et al., *Appl. Phys. Lett.*, 90, 162513, 2007.
16. Yuan, G. L., S. W. Or, Y. P. Wang, et al., *Solid State Communications*, 138, 76–81, 2006.
17. Rivas, J., J. Mira, B. Rivas-Murias, A. Fondado, et al., *Appl. Phys. Lett.*, 88, 242906, 2006.

A New Broadband L-shaped Bend Based on Photonic Crystal Ring Resonators

M. Djavid, **F. Monifi**, **A. Ghaffari**, **and M. S. Abrishamian**

Department of Electrical Engineering, K.N.Toosi University of Technology, Tehran, Iran

Abstract— In this paper, we propose a new type of 2D photonic crystal L-shaped bent waveguide based on ring resonator with broadband acceptable bandwidth; FDTD method concludes output efficiency over 90%.

1. INTRODUCTION

One of the key elements of integrated optical devices is a bent waveguide. Many papers have investigated different geometries of bent waveguide in order to improve its transmission characteristics. Ring resonators are another useful element especially in add-drop filters. Because of high quality factor of the ring resonator and the ring's intrinsic single mode nature, very high spectral selectivity can be achieved [1, 2]. Therefore we used this element to achieve a new type of L-shaped Bent waveguide with acceptable bandwidth about 1537 nm to 1600 nm in third communication window.

2. PHOTONIC CRYSTAL ADD-DROP FILTER BASED ON RING RESONATOR COUPLING

A typical ring resonator in rectangular lattice photonic crystal of dielectric column in air host is shown in Figure 1. As shown in this figure, by adding the four scatterer rods at the corners of the ring resonator with the same properties as other rods, the performance of the ring resonator is optimized.

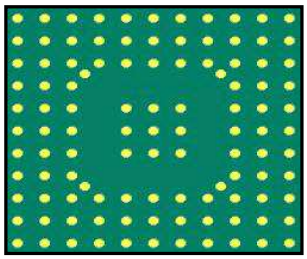
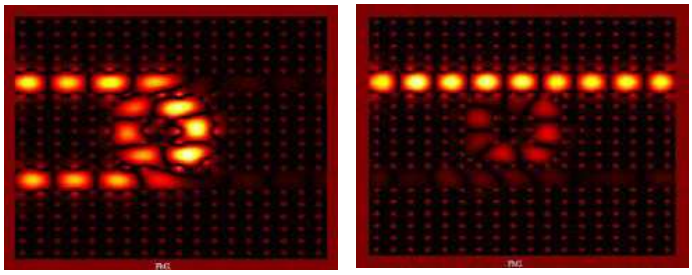


Figure 1: Single-ring photonic crystal ring resonator (PCRR).

By putting a ring resonator next to the waveguide, it can be coupled to the waveguide at its resonant frequency to trap the electromagnetic energy propagating in the waveguide and localized its energy [3]. In other word, the ring resonator drops a light from the top waveguide and sends it to the bottom waveguide, as shown in Figure 2.



(a) (b)

Figure 2: The electric field patterns for (a) drop (on-resonance $\lambda = 1577$ nm) and (b) through (off-resonance $\lambda = 1500$ nm) channel.

3. NEW L-SHAPED BENT WAVEGUIDE

A typical right-angle bent waveguide is shown in Figure 3(a). In many papers worked on this kind of bend and optimized it for high transmission by adding scatterer to its corner. But, bandwidth of their structure is narrow [3]. In this paper we design a new L-shaped bent waveguide with excellent power transmission as shown in Figure 3(a).

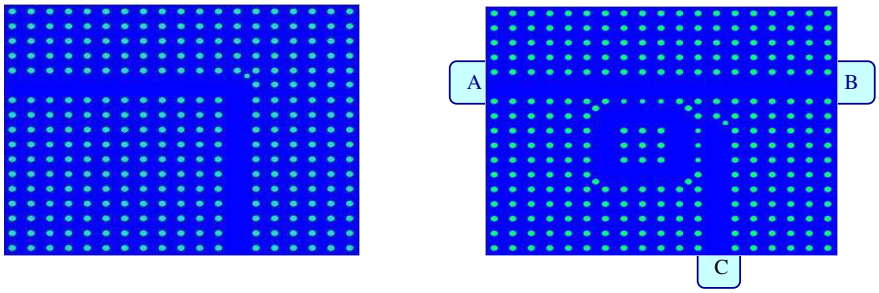


Figure 3: (a) A typical L-shaped bent waveguide, (b) a new L-shaped bent waveguide with ring resonator (PCRR).

An optimized ring resonator in rectangular lattice photonic crystal of dielectric column in air host is used in our structure as shown in Figure 3(b). This phenomenon occurs because of coupling between the waveguide and ring resonator at resonant frequency [4–7].

By closing top of the vertical waveguide and adding a scatterer to top of it, the performance of the bend is optimized, coupling rods are optimized separately. For better contrast, end of the horizontal waveguide is opened to show that whole signal through out port C. The structure and its three ports are labeled as A, B, and C, shown in Figure 3(b). As it will show later, the ring resonator drops a light from the horizontal waveguide and sends it to the vertical waveguide.

4. SIMULATIONS

The structure used in this paper is 2D rectangular lattice photonic crystal of Si rods in the air host. Refractive index of Si is 3.46, the radii of the rods is 18.5% of the lattice constant. The polarization of the signal in our simulation is TM. The spectrum of the power transmission is obtained with finite difference time domain method (FDTD) in *MATLAB* software. A perfect matched layer (PML) is used around the bent structure. The computer used in this simulation is P4 3.00 GHz and has 4 GB of RAM. The power transmission spectra are computed by taking the FFT of the fields that calculated by FDTD incorporating with integrating the pointing vector over the cells of the output ports (during 30000 time step, 45 min running time).

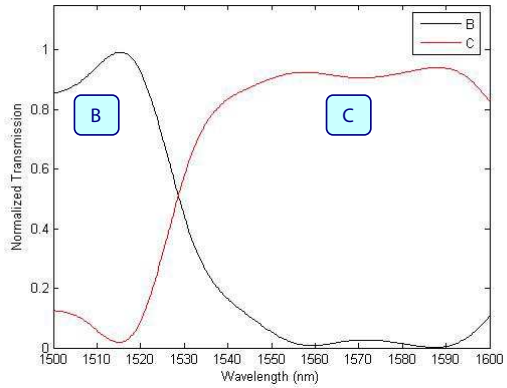


Figure 4: Optical power transmission characteristics of L-shaped bent waveguide.

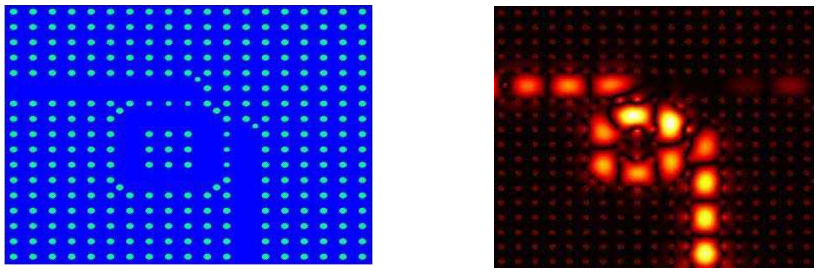
The result of the FDTD processing is shown in Figure 4. That shows the normalized transmissions of the structure over the third communication window. The normalized transmissions of the acceptable frequency range are presented in Table 1, as shown below. As shown in Table 1, normal-

ized transmission power in range 1537 nm–1600 nm is above 80% and in range 1549 nm–1595 nm is above 90%, therefore the acceptable frequency range is achieved.

Table 1: Normalized transmissions of the acceptable frequency range.

Wavelength range (nm)	Normalized transmission
1537-1549	80-90%
1595-1600	80-90%
1549-1595	Above 90%

Finally, the results of time domain simulation are depicted. The figure shows the electric field intensity of L-shaped bent waveguide at one of the ring resonator resonant wavelengths between 1537 nm to 1600 nm.



(a) (b)

Figure 5: (a) Final pattern of our L-shaped bend with ring resonator, (b) electric field intensity of L-shaped bent waveguide at one of the ring resonator resonant wavelengths in 1537–1600 nm.

5. CONCLUSIONS

In this paper, we investigated a two dimensional photonic crystal L-shaped bent waveguide based on ring resonator and we showed that we got the desired acceptable frequency range. Also we showed that the normalized transmission of this frequency range was above 90%.

ACKNOWLEDGMENT

This paper was supported by Iran Telecommunication Research Center.

REFERENCES

1. Qiang, Z., W. Zhou, and R. A. Soref, "Optical add-drop filters based on photonic crystal ring resonators," *Opt. Express*, 1823, 2007.
2. Manolatou, C., M. J. Khan, S. Fan, P. R. Villeneuve, H. A. Haus, and J. D. Joannopoulos, "Coupling of modes analysis of resonant channel add-drop filters," *IEEE J. Quantum Electron*, Vol. 35, 1322–1331, 1999.
3. Naka, Y. and H. Ikuno, "Analysis of 2-D photonic crystal L-shaped optical waveguide and its application to optical devices," *Proceedings of 2001 URSI International Symposium on Electromagnetic Theory*, 529–531, Victoria, Canada, 2001.
4. Kumar, V. D., T. Srinivas, and A. Selvarajan, "Investigation of ring resonators in photonic crystal circuits," *Photonics and Nanostructures*, Vol. 2, 199–206, 2004.
5. Notomi, M., A. Shinya, S. Mitsugi, E. Kuramochi, and H. Y. Ryu, "Waveguides, resonators and their coupled elements in photonic crystal slabs," *Opt. Express*, Vol. 12, 1551–1561, 2004.
6. Barwicz, T., M. Popovic, P. Rakich, M. Watts, H. Haus, E. Ippen, and H. Smith, "Microring-resonatorbased add-drop filters in SiN: fabrication and analysis," *Opt. Express*, Vol. 12, 1437–1442, 2004.
7. Romero-Vivas, J., D. N. Chigrin, A. V. Lavrinenko, and C. M. S. Torres, "Resonant add-drop filter based on a photonic quasicrystal," *Opt. Express*, Vol. 13, 826–835, 2005.

Photonic Lattice of Coupled Microcavities in Nonpermanent Gravitational Field Produced by Rotation

D. L. Boiko

Ecole Polytechnique Fédérale de Lausanne, Quantum Architecture group, 1015 Lausanne, Switzerland

Abstract— Rotation-induced splitting of the otherwise degenerate photonic bands is predicted for a two-dimensional photonic crystal made of evanescently coupled microcavities. The symmetry-broken energy splitting is similar to the Zeeman splitting of atomic levels or electron's (hole's) magnetic moment sublevels in an external magnetic field. The orbital motion of photons in periodic photonic lattice of microcavities is shown to enhance significantly such Coriolis-Zeeman splitting as compared to a solitary microcavity [11]. The equation of motion suggests that nonstationary rotation induces quantum transitions between photonic states and, furthermore, that such transitions might be used as a source for high-frequency gravitational waves.

1. INTRODUCTION

The Sagnac effect in a rotating ring cavity, known also as the Coriolis-Zeeman effect for photons, emerges as a frequency splitting of counterpropagating waves [1–4]. Thus, for a ring cavity with optical path length of M wavelengths, the modal shift is $M\Omega$ with Ω being a rotation rate. Nowadays, the effect is used in commercial He-Ne ring laser gyros of large ($M{\sim}10^6$) cavity size. Significant efforts were made towards designing miniature-sized solid-state devices based on a high optical gain medium [5–7]. Theoretical investigations have been carried out about the impact of rotation on the whispering gallery modes of a microdisc microstructure [8]. Here, the frequency splitting scales with the closed optical path length, while the field polarization is either not relevant or assumed to be parallel to the rotation axis $\mathbf{\Omega}$.

On the other hand, the Coriolis-Zeeman effect in a cylindrical microwave resonator rotating along the symmetry axis [9, 10] or in an optical Fabry-Pérot cavity rotating in the mirror plane is, at first sight, independent of the cavity size [11]. The frequency shift $(S{+}M)\Omega$ of polarization and transverse modes is set by the spin (±1) and the azimuth mode index M. However, the higher the mode index M, the larger the size of the cavity needed to support such a mode. By virtue of the complexity of the mode discrimination at high M and because of the small frequency splitting of the polarization modes, this effect has not yet been verified in experimental measurements.

Here, the Coriolis-Zeeman effect is considered in coupled microcavities arranged in a periodic two-dimensional (2D) lattice in the plane of rotation [Fig. 1(a)]. On example of a square-symmetry lattice, the possibility of enhanced Coriolis-Zeeman splitting, corresponding to $M{\sim}1000$, is predicted for the *low-order* photonic modes [12]. It is caused by the photon's orbital motion extended over the large number of lattice cells. The equation of motion, which is similar to the Hamiltonian for electrons and holes in magnetic field, suggests that nonstationary rotation induces quantum transitions between photonic states and, vice verse, that such transitions will generate a nonstationary gravitational field.

2. PHOTONIC CRYSTAL OF COUPLED MICROCAVITIES

Arrays of evanescently coupled microcavities belong to a particular sub-class of 2D photonic crystal (PhC) structures encompassing photonic crystal fibers and arrays of microcavities. Matrices of vertical cavity surface emitting lasers (VCSELs) are an example of such 2D photonic crystals [13, 14]. In such structures, only a small transversal component $\mathbf{k}_\perp$ of the propagation vector $\mathbf{k}$ undergoes Bragg reflections in the plane of periodic lattice [Fig. 1(a)]. The structures employ lattices of periods significantly exceeding the optical wavelength. They are typically realized by the mirror reflectivity patterning in a broad-area microcavity. As a model system for such PhC structures, a Fabry-Pérot cavity with patterned mirror reflectivity is considered here. The cavity length is one wavelength, $l_z{=}\lambda/n$ with n being the refractive index in the cavity. The reflectivity $R_1(x, y)$ of the one cavity mirror (e.g., of the top mirror) is modulated in two directions parallel to its plane. The reflectivity pattern $R_1{=}\exp i\varphi(x, y)$ consists of pixels with the relative phase shift $\varphi(x, y){=}\Delta\varphi$

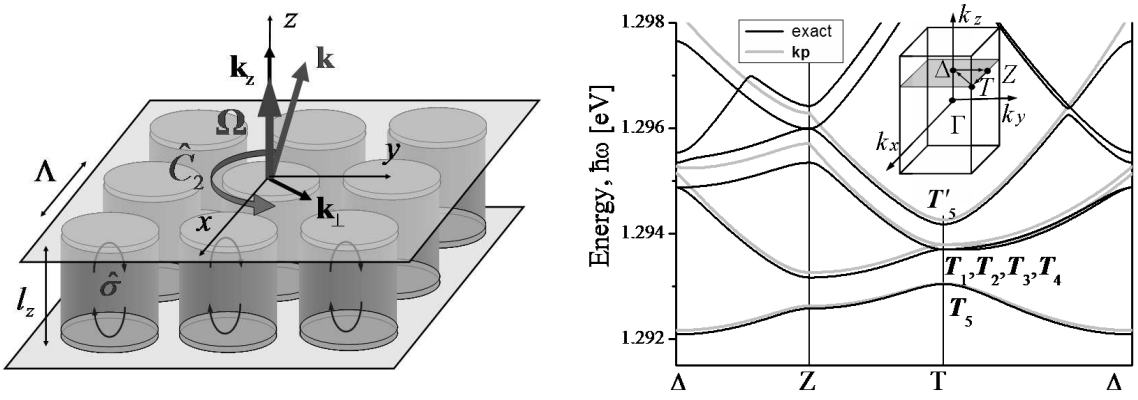


Figure 1: (a) Schematic of the array of coupled microcavities. (b) Band structure of square-lattice PhC calculated from (7) [black curves] and using 8×8 $\mathbf{k}\cdot\mathbf{p}$ approximation (10) [gray curves]. The inset shows the BZ and location of the high-symmetry triangle ΔZT. The parameters are λ=960 nm, n=3.53, Λ=4 μm, FF=0.65 and $\Delta\varphi$=0.02.

separated by a grid of $\varphi(x,y)$=0. The mirror is thus perfectly reflecting ($|R_1|$=1) and the phase modulation pattern $\varphi(x,y)$ defines the structure of the cavity modes. The period of the reflectivity pattern Λ is of a few micron pitch ($\Lambda\gg l_z$). The pattern is characterized by a fill factor FF that is the ratio between the area of the pixel and that of the unit cell. Like in typical VCSEL arrays, the phase contrast of the reflectivity pattern is small ($|\Delta\varphi|\sim10^{-2}$). The second mirror of uniform reflectivity (R_2=1) has no impact on the cavity modes.

3. MODEL HAMILTONIAN

The analysis is carried out here using an equivalent, unfolded cavity representation [15]. Multiple reflections at the cavity mirrors effectively translate the cavity into a structure that is periodic along the cavity axis (the z-axis). The unfolded PhC is thus three-dimensional and it can be analyzed in terms of a modal expansion on orthogonal plane waves (OPWs). The Coriolis-Zeeman effect in such photonic crystal is considered here using a frame of reference, which rotates together with the crystal. Such noninertial rest frame is characterized by a metric tensor g_{ik}, the off-diagonal space-time components $g_{0\alpha}$ of which are dependent on the angular rotation speed $\mathbf{\Omega}$ [16]. However, in the unfolded-cavity representation, g_{ik} differs from the diagonal Minkowski tensor, even in the absence of rotation. Thus $g_{0\alpha}\neq0$ at the subsequent mirror reflections since the reflection operator $\hat{\sigma}=\hat{I}\hat{C}_2$ includes rotation by π (about the z axis) followed by coordinate inversion [Fig. 1(a)]. Due to the patterned mirror reflectivity, the equivalent unfolded PhC is of periodically varying "noninertiality" in the xy plane [15]. In the approximation of the first-order terms $\Omega r/c$ and $\Delta\varphi$, the metric tensor g_{ik} has the following nonzero components:

$$\begin{aligned} g_{00} &= -g_{11}=-g_{22}=-g_{33}=1, \\ g_{0\alpha} &= -\tfrac{1}{c}e_{\alpha\beta\gamma}\Omega^{\beta}x^{\gamma}-\delta_{\alpha3}\tfrac{c}{\omega}\varphi(x^1,x^2)\textstyle\sum_j\delta(x^3-2jl_z), \end{aligned} \tag{1}$$

where $g_{0\alpha}$=$g_{\alpha0}$ ($\alpha=1,2,3$), the space-time coordinates are indexed according to the intervals dx^0=cdt, dx^1=dx, dx^2=dy and dx^3=dz; twice repeated Greek indexes indicate summation. The first term in $g_{0\alpha}$ accounts for rotation of the coordinate system [1–4]. The second term accounts for the multiple cavity roundtrips along the z-axis and reflections at the cavity mirrors. The z-period of the unfolded crystal is thus $2l_z$. The metric tensor (1) is validated by inspecting the system Hamiltonian [Eq. (7)] for the case of φ=0 (rotating FP cavity [11]) or Ω=0 (PhC in an inertial frame [15]).

In the approximation (1), the coordinate space is Euclidean, with the metric tensor $\gamma_{\alpha\beta}$=$-g_{\alpha\beta}+g_{0\alpha}g_{0\beta}/g_{00}$ being the Kronecker delta $\delta_{\alpha\beta}$. Proceeding in a standard manner [16], the covariant Maxwell's equations with metric (1) are converted to the usual form in terms of noncovariant field

vectors **B**, **H**, **D** and **E** that assume the constitutive equations [17]

$$\begin{aligned} \mathbf{D} &= \varepsilon\mathbf{E} + \mathbf{H}\times\mathbf{g}, \quad \mathbf{B}{=}\mu\mathbf{H} + \mathbf{g}\times\mathbf{E}, \\ \mathbf{g} &= \frac{\mathbf{\Omega}\times\mathbf{r}}{c} + \hat{\mathbf{z}}\frac{c}{\omega}\varphi(\mathbf{r}_\perp)\sum_j \delta(z-2jl_z), \end{aligned} \tag{2}$$

where $\hat{\mathbf{z}}$ is the unit vector along z-axis direction and the components of the vector $\mathbf{g}$ are $g_\alpha = -g_{0\alpha}/g_{00}$.

Maxwell' equations in photonic crystal (2) are solved here by separating fast oscillations in the z-axis direction and slow lateral field oscillations in the xy plane [17]:

$$\begin{bmatrix} E_\alpha \\ H_\gamma \end{bmatrix} = e^{ik_z z - i\omega t}\frac{1+\eta(z)}{\sqrt{2\pi}}\begin{bmatrix} Z^{\frac{1}{2}}\hat{\mathcal{E}}_{\alpha\beta} \\ Z^{-\frac{1}{2}}e_{3\beta\alpha}\hat{\mathcal{E}}_{\gamma\alpha} \end{bmatrix}\boldsymbol{\psi}_\beta(\mathbf{r}_\perp), \tag{3}$$

where $n{=}\sqrt{\varepsilon\mu}$ and $Z{=}\sqrt{\mu/\varepsilon}$ are the refractive index and impedance in the cavity. The gauge transformation is introduced here through the operator

$$\hat{\mathcal{E}}_{\alpha\beta} = \delta_{\alpha\beta}\left(1 - \frac{1}{4k_z^2}\frac{\partial^2}{\partial x_\gamma \partial x_\gamma}\right) + \frac{1}{2k_z^2}\frac{\partial^2}{\partial x_\alpha \partial x_\beta} + i\frac{\delta_{\alpha 3}}{k_z}\frac{\partial}{\partial x_\beta} + \frac{\Omega}{nc}e_{3\gamma\beta}\left(\delta_{\alpha 3}x_\gamma - i\frac{x_\alpha}{k_z}\frac{\partial}{\partial x_\gamma}\right), \tag{4}$$

where the terms $\sim k_\perp^2/k_z^2$ are taken into account. Such separation of variables is valid in conditions of the paraxial approximation ($\frac{1}{k_z^2|\boldsymbol{\psi}|}\left|\frac{\partial^2\psi_\alpha}{\partial x_\beta \partial x_\gamma}\right| \ll \frac{1}{k_z|\boldsymbol{\psi}|}\left|\frac{\partial\psi_\alpha}{\partial x_\beta}\right| \ll 1$) and of the low contrast of reflectivity pattern ($|\Delta\varphi| \ll 1$). The two-component vector $\boldsymbol{\psi}(\mathbf{r}_\perp){=}\begin{pmatrix}\psi_x \\ \psi_y\end{pmatrix}$ in the x-y plane is a slowly-varying function of coordinates. It defines the spatial patterns of the six electromagnetic field components (3) and it is considered here as the photonic state wave function. Its squared modulus $|\boldsymbol{\psi}(\mathbf{r}_\perp)|^2$ yields the intensity pattern of the main polarization component in (3). For $\Omega = 0$, Eqs. (3)–(4) are in agreement with the results obtained for the Gaussian beam [18].

In photonic crystals, the wave function $\boldsymbol{\psi}(\mathbf{r}_\perp)$ is a Bloch wave propagating in the xy plane [15],

$$\boldsymbol{\psi}_{q\mathbf{k}_\perp} = e^{i\mathbf{k}_\perp\mathbf{r}_\perp}\mathbf{u}_{q\mathbf{k}_\perp}(\mathbf{r}_\perp), \tag{5}$$

where $\frac{4\pi^2}{\Lambda^2}\int_{cell}\mathbf{u}^*_{q'\mathbf{k}_\perp}\mathbf{u}_{q\mathbf{k}_\perp}d^2\mathbf{r}_\perp{=}\delta_{q'q}$ [20, 21, 17]. The longitudinal part in (3) [the term $e^{ik_z z}\frac{1+\eta_{q\mathbf{k}}(z)}{\sqrt{2\pi}}$] is also a Bloch function. Within the z-period of the lattice, it has a small modulation depth $\langle|\eta_{q\mathbf{k}}|\rangle_{2l_z} = \frac{1}{2l_z}\int_{-l_z}^{l_z}|\eta_{q\mathbf{k}}|\,dz \sim \Delta\varphi$, which is set by an effective phase shift $\alpha_{q\mathbf{k}}$ at each reflection of the patterned mirror. The general form of such periodic function $\eta_{q\mathbf{k}}$ is

$$1{+}\eta_{q\mathbf{k}} = \exp\left\{i\alpha_{q\mathbf{k}}\sum_j\left[\theta(z{-}2jl_z){-}\frac{1}{2}\right] - \frac{iz\alpha_{q\mathbf{k}}}{2l_z}\right\} \tag{6}$$

where $\theta(z){=}\int_{-\infty}^{z}\delta(\zeta)d\zeta$ is the unit step function. Note that $\eta_{q\mathbf{k}}(z)$ is the odd function and $\langle\partial\eta_{q\mathbf{k}}/\partial z\rangle_{2l_z} \simeq 0$ by virtue of the small contrast of the reflectivity pattern.

By operating with $e_{3\alpha\beta}\hat{\mathcal{E}}^{-1}_{\beta\gamma}$ and $\hat{\mathcal{E}}^{-1}_{\alpha\beta}$ [from (4)] on Maxwell' equations for the curl of **E** and **H**, substituting the gauge (3) and averaging over the z-period of the lattice, the Maxwell' equations are converted into the same form of a Hamiltonian eigenproblem with respect to the photonic state wave function $\boldsymbol{\psi}_{q\mathbf{k}}(\mathbf{r}_\perp)$

$$\left[\frac{m_0c^2}{n^2} + \frac{\hat{\mathbf{p}}_\perp^2}{2m_0} - \frac{c\hbar}{2nl_z}\varphi(\mathbf{r}_\perp) - \frac{\Omega}{n^2}\left(\mathbf{r}_\perp\times\hat{\mathbf{p}}_\perp + \hbar\hat{S}_z\right)\right]\boldsymbol{\psi}_{q\mathbf{k}} = \hbar\omega_{q\mathbf{k}}\boldsymbol{\psi}_{q\mathbf{k}}, \tag{7}$$

where $m_0{=}n\hbar k_z/c$ and $\hat{S}_z{=}i\hat{\mathbf{z}}\times$ is the spin operator that reads $(\hat{S}_z)_{\alpha\beta}{=}ie_{\alpha 3\beta}{=}\left(\begin{smallmatrix}0 & -i \\ i & 0\end{smallmatrix}\right)$ in the basis of the two-component vector functions $\boldsymbol{\psi}{=}\begin{pmatrix}\psi_x \\ \psi_y\end{pmatrix}$. Solutions of Eq. (7) define the slowly-varying components of photonic modes (3). The difference between the exact equation for $(1{+}\eta_{q\mathbf{k}})\boldsymbol{\psi}_{q\mathbf{k}}$

and its z-period average [Eq. (7)] yields the equation for the periodic part of the fast longitudinal component:

$$\frac{\partial \eta_{q\mathbf{k}}}{\partial z} = i\alpha_{q\mathbf{k}}\left(1+\eta_{q\mathbf{k}}\right)\left[\sum_j \delta(z-2jl_z) - \frac{1}{2l_z}\right] \tag{8}$$

where $\alpha_{q\mathbf{k}}=\langle\psi_{q\mathbf{k}}|\,\varphi\,|\psi_{q\mathbf{k}}\rangle$. The solution of (8) is given by (6) provided that the eigen functions $\boldsymbol{\psi}_{q\mathbf{k}}$ of the Hamiltonian (7) are known.

In (7), the first term is due to the paraxial propagation along the z axis. The second and third terms are the in-plane kinetic energy and the periodic crystal potential, respectively. The last term in (7) is a perturbation induced by the Coriolis force. For φ=0 and Ω=0, within the accuracy of the time variable, Eq. (7) is just the paraxial wave equation. For φ=0, Eq. (7) yields the effective refractive index $\frac{c}{\omega}(k_z+\frac{k_\perp^2}{2k_z})$ of a circularly-polarized paraxial wave

$$n_{\text{eff}} = n + \boldsymbol{g\tau} \pm \frac{\boldsymbol{\Omega\tau}}{\omega n}, \tag{9}$$

where $\boldsymbol{\tau}=\frac{\mathbf{k}}{k}$ defines the propagation direction, $+/-$ sign is for the left/right handed polarization. Eq. (9) agrees with previously reported expressions for the axial nonreciprocity [1] (second term) and circular birefringence [11] (third term) induced by the Coriolis force for photons. Finally, the unperturbed (Ω=0) Hamiltonian has been verified by experimental measurements in VCSEL array PhC heterostructures [19]. These justify the approximation (7) of theHamiltonian.

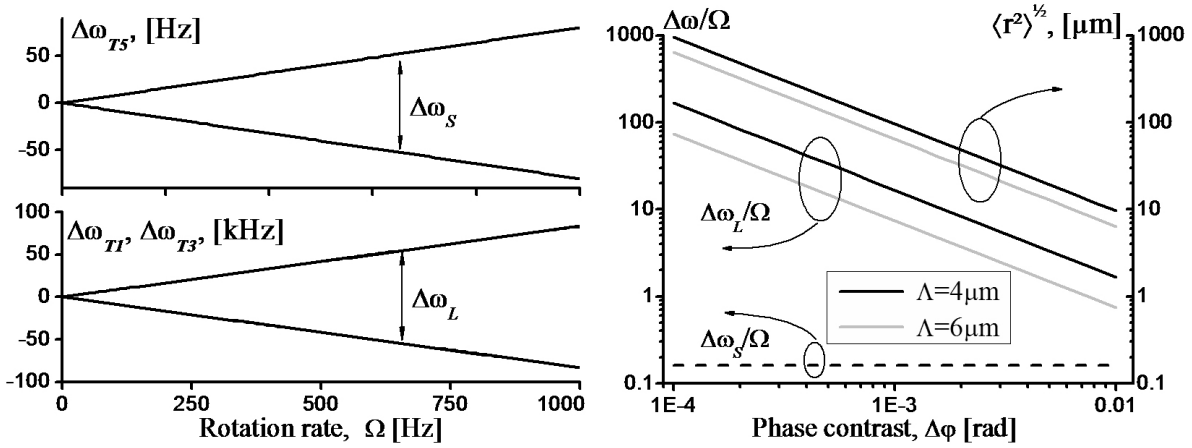
4. RESULTS AND DISCUSSIONS

Analytical similarities between the effective single-electron Hamiltonian in a semiconductor subjected to an external magnetic field and Eq. (7) allow the correspondence between the periodic crystal potential V and phase pattern φ ($V\rightarrow-\frac{c\hbar}{2nL}\varphi$), and the vector potentials $\mathbf{A}=\frac{1}{2}\mathbf{H}\times\mathbf{r}$ and $\mathbf{g}_\perp=\frac{1}{c}\boldsymbol{\Omega}\times\mathbf{r}$ ($\frac{e}{c}\mathbf{A}\rightarrow\frac{m_0c}{n^2}\mathbf{g}_\perp$). Photons in photonic crystal subjected to a nonpermanent gravitational field thus exhibit a behaviour similar to electrons (holes) in a magnetic field. Accordingly, the impact of rotation on the envelope function and periodic part of the photonic Bloch wave (5) is different. As in the case of electrons, [20, 21] the components of velocity operator $\hat{\mathbf{v}}_\perp=\frac{1}{i\hbar}[\mathbf{r}_\perp,\hat{H}]$ do not commute ($[\hat{v}_x,\hat{v}_y]=-\frac{2i\hbar}{m_0n^2}\Omega_z$, where $\hat{\mathbf{v}}_\perp=\frac{\hat{\mathbf{p}}_\perp}{m_0}-\frac{\boldsymbol{\Omega}\times\mathbf{r}_\perp}{n^2}$). However, the second-order terms $\sim\frac{\Omega^2r^2}{c^2}$ in the Hamiltonian (7) [and, respectively, in (1)] are needed to define whether the Landau-like quantization is possible for photonic envelope wave functions. In the rest of the Letter, the impact of rotation on the periodic part of Bloch functions is examined in detail using the example of a square-lattice PhC.

Figure 1(b) shows the typical band structure of a square-lattice PhC, which is calculated along the high symmetry lines Δ-Z-T in the Brillouin zone (BZ), using the OPW expansion in unperturbed (Ω=0) Hamiltonian. By virtue of the square lattice symmetry, all states are degenerate by the photon's spin (e.g., the doubly degenerate states $\mathbf{T}_5$ or $\mathbf{T}_5'$). Angular rotation removes the degeneracy of such states and splits their energies on $2\frac{\hbar\Omega}{n^2}$ [Fig. 2(a), top panel]. However, there are states, like the degenerate states $\mathbf{T}_1$, $\mathbf{T}_2$, $\mathbf{T}_3$ and $\mathbf{T}_4$, of the four-fold degeneracy, which is caused

Table 1: Basis functions (scalars) and photonic harmonics (vectors) of irreducible representations of the group of $\mathbf{k}$ at the T point of the BZ (C_{4v} point group)

T_i		$T_i\times T_5$	
T_1	S	$\mathbf{T}_5$	$S\hat{\mathbf{x}}, S\hat{\mathbf{y}}$
T_2	$XY(X^2-Y^2)$	$\mathbf{T}_5''$	$XY(X^2-Y^2)\hat{\mathbf{x}}, XY(X^2-Y^2)\hat{\mathbf{y}}$
T_3	X^2-Y^2	$\mathbf{T}_5'''$	$(X^2-Y^2)\hat{\mathbf{x}}, (X^2-Y^2)\hat{\mathbf{y}}$
T_4	XY	$\mathbf{T}_5'$	$XY\hat{\mathbf{x}}, XY\hat{\mathbf{y}}$
T_5	iX, iY	$\mathbf{T}_1+\mathbf{T}_2+\mathbf{T}_3+\mathbf{T}_4$:	
		$\mathbf{T}_1$	$\frac{i}{\sqrt{2}}(X\hat{\mathbf{x}}+Y\hat{\mathbf{y}})$
		$\mathbf{T}_2$	$\frac{i}{\sqrt{2}}(Y\hat{\mathbf{x}}-X\hat{\mathbf{y}})$
		$\mathbf{T}_3$	$\frac{i}{\sqrt{2}}(X\hat{\mathbf{x}}-Y\hat{\mathbf{y}})$
		$\mathbf{T}_4$	$\frac{i}{\sqrt{2}}(Y\hat{\mathbf{x}}+X\hat{\mathbf{y}})$



(a) (b)

Figure 2: (a) Coriolis-Zeeman splitting of the $\mathbf{T}_5$ ($\mathbf{T}'_5$) (top panel) and $\mathbf{T}_1$-$\mathbf{T}_4$ (bottom panel) bands as a function of rotation rate Ω. $\Lambda = 4\,\mu$m and $\Delta\varphi = 10^{-4}$. (b) Relative splitting $\frac{\Delta\omega_{L,S}}{\Omega}$ (left axis) and matrix element $\langle \mathbf{T}_1|\mathbf{r}^2|\mathbf{T}_1\rangle^{\frac{1}{2}}$ (right axis) as a function of the lattice contrast $\Delta\varphi$. The lattice constant Λ is 4 (black curves) and 6 μm (gray curves), $n = 3.53$. Other parameters are given in the caption of Fig. 1.

by the orbital symmetry of the Bloch functions. For such states, there is an important angular momentum contribution $\frac{1}{n^2}\hbar\Omega_z\hat{L}_z$ to the energy shift.

The Coriolis-Zeeman splitting of these states is analyzed here using the first-order **kp** expansion in the T point [$\mathbf{k} = (\frac{\pi}{\Lambda}, \frac{\pi}{\Lambda}, k_z)$] of the BZ. The expansion basis is deduced from the empty lattice test. Four scalar plane waves $\exp(\pm\frac{\pi x}{\Lambda}\pm\frac{\pi y}{\Lambda})$, which form the first photonic band of empty lattice, originate from the nearest equivalent T points of reciprocal lattice and provide representation that is reducible under the C_{4v} point group (the symmetry group of $\mathbf{k}$). Their symmetrized combinations of T_1, T_4 and T_5 representations are indicated in the second column of Table 1 with the capitals letters corresponding to the main term of Taylor expansion on parameter $\frac{|\mathbf{r}_\perp|}{\Lambda}$ (e.g., $|S\rangle = \frac{1}{\pi}\cos\frac{\pi x}{\Lambda}\cos\frac{\pi y}{\Lambda}$, $|iX\rangle = \frac{i}{\pi}\sin\frac{\pi x}{\Lambda}\cos\frac{\pi y}{\Lambda}$). The photon's spin transforms as the two-dimensional representation T_5. Therefore, a reduction of the direct product $T_i \otimes T_5$ results in the eight symmetry adapted photonic harmonics of $\mathbf{T}_1$-$\mathbf{T}_5$ and $\mathbf{T}'_5$ representations that constitute a suitable **kp**-expansion basis in the low-order photonic bands. These states are well separated energetically from the other states in the T point [Fig. 1(b)].

For a general state $|\psi_{q\mathbf{k}}\rangle = e^{i\mathbf{k}_\perp\mathbf{r}_\perp}\sum_i c_i\,|\mathbf{T}_i\rangle$ at $\mathbf{k}_\perp$ measured from the T point of the BZ, the 8×8 **kp** Hamiltonian for the coefficients c_i is of the block-diagonal form

$$\hat{H} = \begin{bmatrix} \hat{H}_0 + \hat{H}_{\mathbf{kp}} + \hat{H}_\Omega + \frac{\hbar^2 k_\perp^2}{2m_0} & 0 \\ 0 & \hat{H}_0 + \hat{H}^*_{\mathbf{kp}} - \hat{H}^*_\Omega + \frac{\hbar^2 k_\perp^2}{2m_0} \end{bmatrix} \tag{10}$$

where

$$\hat{H}_0 = \begin{bmatrix} \hbar\omega_{T'_5} & 0 & 0 & 0 \\ 0 & \hbar\omega_{T_1} & 0 & 0 \\ 0 & 0 & \hbar\omega_{T_1} & 0 \\ 0 & 0 & 0 & \hbar\omega_{T_5} \end{bmatrix}, \quad \hat{H}_{\mathbf{kp}} = \frac{\hbar P}{m_0}\begin{bmatrix} 0 & k_- & k_+ & 0 \\ k_+ & 0 & 0 & k_- \\ k_- & 0 & 0 & -k_+ \\ 0 & k_+ & -k_- & 0 \end{bmatrix},$$

$$\hat{H}_\Omega = -\frac{\hbar\Omega}{n^2}\begin{bmatrix} 1 & -M_-\frac{\hbar k_-}{2P} & M_-\frac{\hbar k_+}{2P} & 0 \\ -M_-\frac{\hbar k_+}{2P} & -M+1 & 0 & -M_+\frac{\hbar k_-}{2P} \\ M_-\frac{\hbar k_-}{2P} & 0 & M+1 & -M_+\frac{\hbar k_+}{2P} \\ 0 & -M_+\frac{\hbar k_+}{2P} & -M_+\frac{\hbar k_-}{2P} & 1 \end{bmatrix}$$

in the basis of functions $\frac{1}{\sqrt{2}}(\mathbf{T}'_{5\mathbf{x}}\pm i\mathbf{T}'_{5\mathbf{y}})$, $\frac{\mp i}{\sqrt{2}}(\mathbf{T}_1\mp i\mathbf{T}_2)$, $\frac{\pm i}{\sqrt{2}}(\mathbf{T}_3\pm i\mathbf{T}_4)$, $\frac{\mp i}{\sqrt{2}}(\mathbf{T}_{5\mathbf{x}}\pm i\mathbf{T}_{5\mathbf{y}})$ [1]. The upper (lower) sign refers to the top (bottom) 4×4 block. In (10), $k_\pm = k_x \pm ik_y$, $P = \frac{1}{\sqrt{2}}\langle S|-i\hbar\frac{\partial}{\partial x}|iX\rangle = \frac{\hbar\pi}{\sqrt{2}\Lambda}$ is

[1] The wave functions are $\frac{1}{\sqrt{2}}XY(\hat{\mathbf{x}}\pm i\hat{\mathbf{y}})$, $\pm\frac{1}{2}(X\mp iY)(\hat{\mathbf{x}}\pm i\hat{\mathbf{y}})$, $\mp\frac{1}{2}(X\pm iY)(\hat{\mathbf{x}}\pm i\hat{\mathbf{y}})$, $\mp\frac{i}{\sqrt{2}}S(\hat{\mathbf{x}}\pm i\hat{\mathbf{y}})$.

the interband matrix element of $\hat{\mathbf{p}}$ that defines the band mixing. The eigen solutions of (10) are plotted in Fig. 1(b). Good agreement with the band structure calculated from Eq. (7) justifies the first-order **kp** approximation (10) of the system Hamiltonian.

In Eq. (10), the matrix $\hat{H}_\Omega$ in (10) is a perturbation induced by rotation. It accounts for the Coriolis-Zeeman energy shift, which is of the opposite sign for the left (upper 4×4 block) and right (lower block) handed polarization states of the same orbital symmetry. The rotation lifts degeneracy between such spin states, producing the frequency splitting $\Delta\omega_S = 2\frac{\Omega}{n^2}$, as indicated in Fig. 2(a) [top panel] on example of the $\hat{\mathbf{T}}_5$ ($\hat{\mathbf{T}}'_5$) states. In (10), the parameter $M = M_+ + M_-$ accounts for the orbital part of wave functions. The orbital contribution $\Delta\omega_L = 2\frac{M\Omega}{n^2}$ to the Coriolis-Zeeman energy shift $\pm\frac{1}{2}\hbar\Delta\omega_L \pm \frac{1}{2}\hbar\Delta\omega_S$ in the $\mathbf{T}_1$-$\mathbf{T}_4$ states is shown in the bottom panel of Fig. 2(a) ($\Delta\omega_S$ is not visible at the scale of $\Delta\omega_L$). The energy shift $\pm\frac{1}{2}\hbar\Delta\omega_L$ [the term $-\frac{\mathbf{\Omega}}{n^2}\mathbf{r}\times\hat{\mathbf{p}}$ in Eq. (7)] is evaluated here using the relationship $\mathbf{r}_{mn} = \frac{i\hbar}{m_0}\frac{\mathbf{p}_{mn}}{E_n - E_m}$. Note that the procedures to evaluate the matrix elements of $\hbar\hat{L}_z = \mathbf{r}_\perp \times \hat{\mathbf{p}}_\perp$ in free space [22, 23] and in periodic lattices [24] are different. Here, the f-sum rule $\frac{1}{m_{\alpha\beta}} = \frac{\delta_{\alpha\beta}}{m_0} + \frac{2}{m_0^2}\sum_n \frac{p^\alpha_{mn} p^\beta_{nm}}{E_n - E_m}$ [20, 21] implies that $M_\pm = \mp\frac{1}{2}\left(\frac{m_0}{m_{T_5, T'_5}} - 1\right)$ and

$$\Delta\omega_L = \frac{\hat{\mathbf{z}}\mathbf{\Omega}}{n^2}\left[\frac{m_0}{m_{T'_5}} - \frac{m_0}{m_{T_5}}\right], \qquad \Delta\omega_S = 2\frac{\hat{\mathbf{z}}\mathbf{\Omega}}{n^2}. \tag{11}$$

For an array of square pixels defining the microcavities, $M_\pm = \frac{2nl_z P^2}{\hbar m_0 c F F \Delta\varphi}\left[\frac{\sin\pi\sqrt{FF}}{\pi\sqrt{FF}}\left(1 \pm \frac{\sin\pi\sqrt{FF}}{\pi\sqrt{FF}}\right)\right]^{-1}$. It follows that reducing the effective mass (via the lattice pitch Λ, fill factor FF and contrast $\Delta\varphi$), one can enhance the Coriolis-Zeeman splitting and achieve M of more than 10^3. [Fig. 2(b) shows the ratio $\frac{\Delta\omega_L}{\Omega} = \frac{2M}{n^2}$, left axis]. The enhancement is caused by the weak localization of photonic wave functions to the lattice sites. The intraband matrix element $\langle q\mathbf{k}|\mathbf{r}^2|q\mathbf{k}\rangle = \frac{\hbar^2(M_-^2 + M_+^2)}{2P^2}$ in the $\mathbf{T}_1$-$\mathbf{T}_4$ bands indicates that the photonic wave functions spread over a large PhC crystal domain [Fig. 2(b) right axis], and the frequency splitting is proportional to the characteristic size of this domain

$$\frac{\Delta\omega_L}{\Omega} = \frac{2\pi\sqrt{2\langle\mathbf{r}_\perp^2\rangle}}{n^2\Lambda}\left[1 + \mathrm{sinc}^2\pi\sqrt{FF}\right]^{-1}. \tag{12}$$

Thus, as in the case of a ring cavity, the frequency splitting of optical modes increases with the characteristic modal size in the plane normal to the rotation axis. Finally note that in the bands of $\mathbf{T}_5$ symmetry, the intraband matrix elements of $\mathbf{r}^2$ are nonzero as well. However, the orbital contribution to the Coriolis-Zeeman energy splitting vanishes, in accordance with the group theory selection rules.

5. CONCLUSION

With present experimental techniques, the predicted frequency splitting [Eq. (12), Fig. 2(a)] can be validated by direct measurements.

The analogy between Eq.(7) and electron's (hole's) magnetic moment Hamiltonian suggests that a nonstationary rotation or gravitational field $\mathbf{g}(t)$ will induce quantum transitions between photonic states. This effect might be used for detection of high frequency gravitational waves. Furthermore, Eq. (7) allows formally an inverse process to take place as well. On the basis of such formalism, one can expect that a superposition of nonstationary photonic states, which represent a system undergoing quantum transition, might serve as a source $\mathbf{g}(t)$ for high frequency gravitational waves [25].

ACKNOWLEDGMENT

The author is grateful to Marc-André Dupertuis and Neil J. Gunther for useful comments. This research was partially supported by the Swiss National Science Foundation. The author is grateful for this opportunity to Eli Kapon and Edoardo Charbon.

REFERENCES

1. Heer, C. V., *Phys. Rev.*, Vol. 134, A799, 1964.
2. Heer, C. V., *Proc. of the Third International Conference on Quantum Electronics*, Columbia University Press, New York, 1305, 1963.

3. Khromykh, A. M., *Zh. Eksp. Teor. Fiz.*, Vol. 50, 281, 1966.
4. Post, E. J., *Rev. Mod. Phys.*, Vol. 39, 475, 1967.
5. Boiko, D. L., et al., *Lazernye Novosti — Laser News*, Vol. 2, 11, 1997.
6. Boiko, D. L., et al., *Quantum Electronics*, Vol. 28, 355, 1998.
7. Cao, H., et al., *Appl. Phys. Lett.*, Vol. 86, 041101, 2005.
8. Nojima, S., *J. Phys. Soc. Jpn.*, Vol. 73, 792, 2004.
9. Belonogov, A. M., *Russian Physics Journal*, Vol. 12, 672, 1969.
10. Belonogov, A. M., *Sov. Phys. Tech. Phys.*, Vol. 14, 883, USSR, 1970.
11. Boiko, D. L., *Optics Express*, Vol. 2, 397, 1998.
12. A first account of these results was presented at the NFO-9 conference [D. L. Boiko, *9-th International Conference on Near-field Optics, Nanophotonics and Related Techniques*, Lausanne, Switzerland, September 10–15, 2006, post-deadline paper ThP-78].
13. Orenstein, M., et al., *Appl. Phys. Lett.*, Vol. 60, 1535, 1992.
14. Gourley, P. L., et al., *Appl. Phys. Lett.*, Vol. 58, 890, 1991.
15. Boiko, D. L., et al., *Optics Express*, Vol. 12, 2597, 2004.
16. Landau, L. D. and E. M. Livshits, *The Classical Theory of Fields*, Nauka, Moscow, 1988.
17. Boiko, D. L., "Paraxial Hamiltonian for photons in two-dimensional photonic crystal microstructures," *ArXiv*, 0710.5287, 2007.
18. Erikson, W. L. and S. Singh, *Phys. Rev. E*, Vol. 49, 5778, 1994.
19. Guerrero, G., et al., *Optics Express*, Vol. 12, 4922, 2004.
20. Luttinger, J. M., *Phys. Rev.*, Vol. 102, 1030, 1956.
21. Luttinger, J. M. and W. Kohn, *Phys. Rev.*, Vol. 97, 869, 1955.
22. ONeil, A. T., I. MacVicar, L. Allen, and M. J. Padgett, *Phys. Rev. Lett.*, Vol. 88, 053601, 2002.
23. Courtial, J., et al., *Optics Express*, Vol. 14, 938, 2006.
24. Bir, G. L. and G. E. Pikus, *Symmetry and Strain-induced Effects in Semiconductors*, Nauka, Moscow, 1972.
25. Gertsenshtein, M. E., "Wave resonance of light and gravitational waves," *Sov. Phys. JETP*, Vol. 14, 84–85, 1962.

Surface-Polariton-Enhanced Reflected THz-Field

D. H. Huang[1], **G. Gumbs**[2], **P. M. Alsing**[1], **and D. A. Cardimona**[1]

[1]Air Force Research Laboratory, Space Vehicles Directorate
Kirtland Air Force Base, NM 87117, USA
[2]Department of Physics and Astronomy, Hunter College of the City University of New York
695 Park Avenue, New York, NY 10021, USA

Abstract— The present work predicts the large enhancements at the band edges of a coupled Bloch-surface-plasmon-polariton band in the spectrum of the reflected far electromagnetic field due to anti-crossing gaps induced by the strong coupling between the continuous surface-plasmonpolariton mode and the discrete Bloch-like modes. The existence of these Bloch-like modes is a direct consequence of the nonlocal mixing of specular and diffraction modes of the reflected electromagnetic field by free-electron induced optical polarization and the interference of a pair of surface optical-polarization waves with opposite Bragg order numbers in the presence of a grating.

1. INTRODUCTION

Ebbesen et al. [1, 2] reported an enhancement of the optical transmission through an array of deep sub-wavelength cylindrical holes on a silver film, in which a very intriguing zeroth-order transmission was shown in such a system. Similar phenomena were observed in sub-wavelength metallic gratings [3]. These enhanced optical transmissions are believed to be a result of light coupling to surface-plasmon-polaritons (SPPs) in structured metallic films [4]. The induced Bragg transmission modes are evanescent and acquire resonant coupling to the surface-plasmon mode. As a result, they indirectly contribute to the zeroth-order transmission mode through interference between a pair of surface Bragg waves with opposite diffraction order numbers.

In order to understand and explain the rich physics for the enhanced optical transmission, nearfield studies must be undertaken. The previous methods of calculation are spatially local and adiabatic in time, thereby neglecting the nonlocal nature and the memory of the induced optical polarization in the Maxwell equations. In this paper, we investigate a structure consisting of a half-space of air (region-1) and a semi-infinite conductor (region-2). The surface is covered with a conducting sheet as well as a metal grating at the interface. With the use of our model, we are able to demonstrate electromagnetic-field (EMF) enhancement and the effects due to grating diffraction, nonlocal evanescent and radiation mode mixing, and longitudinal field. In this paper, we present a spatially-nonlocal dynamic theory which includes effects due to retardation (perpendicular to the interface) and Bloch-like states (parallel to the interface). This allows us to determine quantitatively the enhancements of the transmitted near EMF and the reflected far EMF.

The rest of this paper is organized as follows. In Sec. 2, we introduce our spatially-nonlocal dynamic theory for calculating both near and far EMF spectra in a semi-infinite conductive bulk covered by a conducting sheet and a grating on top of it. The paper is concluded in Sec. 3 with a remark on the effect of an absorption loss in a conductive bulk.

2. MODEL AND THEORY

In our model, one half-space is filled with air for $z < 0$, and another half-space $z > 0$ is occupied by a semi-infinite doped semiconductor. The interface at $z = 0$ consists of a current-driven conducting sheet and a metal grating on top of it. Optically, this sheet and the grating can be considered to be in one plane. The reason for this is that the sheet thickness is much smaller than the wavelength and the decay length of the impinging light. They, however, can still be electrically separated from each other by an energy barrier [5]. Light incident from $z < 0$ is diffracted by the grating in both the reflection ($z < 0$) and the transmission ($z > 0$) regions. The diffraction (Bragg) modes of EMFs are excited by the induced optical polarization at the interface from both the conducting sheet and grating. In addition, all the Bragg modes of EMFs are mixed nonlocally with each other by the grating.

The grating is periodic in the x direction. We seek Bloch-like solutions with the use of a planewave expansion. We then take the Fourier transform with respect to x for the electric fields.

Setting $q_y = 0$ and denoting the electric field below the interface by $\mathbf{E}_1$ and above by $\mathbf{E}_2$, we obtain [5, 6]

$$\begin{aligned}
\mathbf{E}_1(q_n,\omega;z) &= \exp(i\eta_1^T z)\begin{bmatrix} A_x^T(q_x;\omega) \\ A_y^T(q_x;\omega) \\ -(q_x/\eta_1^T)A_x^T(q_x;\omega)\end{bmatrix}\delta_{n,0} \\
&\quad + \sum_{n'=-\infty}^{\infty}\exp(-i\beta_{1,n+n'}^T z)\begin{bmatrix} B_x^T(q_{n+n'};\omega) \\ B_y^T(q_{n+n'};\omega) \\ (q_{n+n'}/\beta_{1,n+n'}^T)B_x^T(q_{n+n'};\omega)\end{bmatrix}, \text{ for } z<0, \qquad (1)\\
\mathbf{E}_2(q_n,\omega;z) &= \sum_{n'=-\infty}^{\infty}\exp(i\beta_{2,n+n'}^T z)\begin{bmatrix} C_x^T(q_{n+n'};\omega) \\ C_y^T(q_{n+n'};\omega) \\ -(q_{n+n'}/\beta_{2,n+n'}^T)C_x^T(q_{n+n'};\omega)\end{bmatrix} \\
&\quad + \sum_{n'=-\infty}^{\infty}\exp(i\beta_{2,n+n'}^L z)\begin{bmatrix} C_x^T(q_{n+n'};\omega) \\ 0 \\ (\beta_{2,n+n'}^L/q_{n+n'})C_x^L(q_{n+n'};\omega)\end{bmatrix}, \text{ for } z>0. \qquad (2)
\end{aligned}$$

In this notation, n and n' are integers, the imaginary parts satisfy $\mathrm{Im}[\beta_{1,n+n'}^T], \mathrm{Im}[\beta_{2,n+n'}^T], \mathrm{Im}[\beta_{2,n+n'}^L] \geq 0$, the angular frequency of the incident EMF is ω, the wave vector $q_x = (\omega/c)\sin\theta_i$, $\eta_1^T = (\omega/c)\cos\theta_i$, θ_i is the incident angle, $q_n = q_x + n(2\pi/d)$ where d is the period of the grating. In addition, for the transverse (T) EMF we obtain $\beta_{1,n}^T$ and $\beta_{2,n}^T$ from the dispersion relation

$$\begin{bmatrix}(\beta_{1,n}^T)^2 \\ (\beta_{2,n}^T)^2\end{bmatrix} = \left(\frac{\omega}{c}\right)^2\begin{bmatrix}1 \\ \epsilon_2^T(q_n,\beta_{2,n}^T;\omega)\end{bmatrix} - q_n^2, \qquad (3)$$

where the region-1 ($z<0$) is filled with air. On the other hand, for the longitudinal (L) EMF, we obtain $\beta_{2,n}^L$ as a root of the equation $\epsilon_2^L(q_n,\beta_{2,n}^L;\omega)=0$. Moreover, the summation over n' in Eqs. (1) and (2) represents the contributions from all Bragg modes ($n'\neq 0$) of the reflected and transmitted electric fields in the presence of the conductive grating.

The magnetic field components can be simply obtained through the relation $\mathbf{H} = (-i/\omega\mu_0)\nabla\times\mathbf{E}^T$. When $q_y = 0$, we get from Eq. (1) the magnetic field $\mathbf{H}_1$ below the interface

$$\begin{aligned}
H_1^x(q_n,\omega;z) &= \left(\frac{-i}{\omega\mu_0}\right)\left[-i\eta_1^T\exp(i\eta_1^T z)A_y^T(q_x;\omega)\delta_{n,0}\right. \qquad (4)\\
&\quad \left. + \sum_{n'=-\infty}^{\infty} i\beta_{1,n+n'}^T\exp(-i\beta_{1,n+n'}^T z)B_y^T(q_{n+n'};\omega)\right], \\
H_1^y(q_n,\omega;z) &= \left(\frac{-i}{\omega\mu_0}\right)\left(\frac{\omega^2}{c^2}\right)\left[\left(\frac{i}{\eta_1^T}\right)\exp\left(i\eta_1^T z\right)A_x^T\left(q_x;\omega\right)\delta_{n,0}\right. \\
&\quad \left. + \sum_{n'=-\infty}^{\infty}\left(\frac{-i}{\beta_{1,n+n'}^T}\right)\exp(-i\beta_{1,n+n'}^T z)B_x^T(q_{n+n'};\omega)\right], \qquad (5)\\
H_1^z(q_n,\omega;z) &= \left(\frac{-i}{\omega\mu_0}\right)\left[iq_x\exp(i\eta_1^T z)A_y^T(q_x;\omega)\delta_{n,0}\right. \\
&\quad \left. + \sum_{n'=-\infty}^{\infty} iq_{n+n'}\exp(-i\beta_{1,n+n'}^T z)B_y^T(q_{n+n'};\omega)\right]. \qquad (6)
\end{aligned}$$

In addition, the magnetic field $\mathbf{H}_2$ above the interface can be calculated by means of Eq. (2)

$$H_2^x(q_n,\omega;z) = -\left(\frac{-i}{\omega\mu_0}\right)\sum_{n'=-\infty}^{\infty} i\beta_{2,n+n'}^T \exp(i\beta_{2,n+n'}^T z)C_y^T(q_{n+n'};\omega), \tag{7}$$

$$H_2^y(q_n,\omega;z) = \left(\frac{-i}{\omega\mu_0}\right)\left(\frac{\omega^2}{c^2}\right)\sum_{n'=-\infty}^{\infty}\left(\frac{i}{\beta_{2,\,n+n'}^T}\right)\epsilon_2^T(q_{n+n'},\beta_{2,n+n'}^T;\omega)\times\exp(i\beta_{2,n+n'}^T z)C_x^T(q_{n+n'};\omega), \tag{8}$$

$$H_2^z(q_n,\omega;z) = \left(\frac{-i}{\omega\mu_0}\right)\sum_{n'=-\infty}^{\infty} iq_{n+n'}\exp(i\beta_{2,n+n'}^T z)C_y^T(q_{n+n'};\omega). \tag{9}$$

For an incident field with an s-polarization, we have $A_x^T(q_x;\omega)=0$ and $A_y^T(q_x;\omega)=E_0$, where E_0 is the amplitude of the incident electric field. On the other hand, for an incident field with a p-polarization, we have $A_y^T(q_x;\omega)=0$ and $A_x^T(q_x;\omega)=\eta_1^T H_0/(\omega\epsilon_0)$, where H_0 is the amplitude of the incident magnetic field.

For a conducting bulk medium, we use the random-phase approximation (RPA), [7] to obtain the transverse dielectric function in the long-wavelength limit as $\epsilon_2^T(q_x,q_z;\omega)=\epsilon_b[1-n_{3\mathrm{D}}e^2/\epsilon_0\epsilon_b m^*\omega(\omega+i\gamma_0)]$, where ϵ_b is the background dielectric constant of the bulk, $n_{3\mathrm{D}}$ is the electron concentration, m^* is the effective mass of an electron in the bulk, and γ_0 is the homogeneous broadening describing the finite lifetime of excited electrons. In addition, by using the hydrodynamic model, [7] the nonlocal the nonlocal longitudinal dielectric function in the long-wavelength limit can be expressed as $\epsilon_2^L(q_x,q_z;\omega) = \epsilon_b\{1-n_{3\mathrm{D}}e^2/\epsilon_0\epsilon_b m^*[\omega(\omega+i\gamma_0)-\xi(q_x^2+q_z^2)]\}$, where $\xi=3v_F^2/5$ and $vF=\hbar(3\pi^2 n_{3\mathrm{D}})^{1/3}/m^*$ is the Fermi velocity of electrons in the bulk. The long-wavelength limit requires $v_F q_n/\omega \ll 1$, which can be satisfied with either a high frequency ω (metals) or a low concentration $n_{3\mathrm{D}}$ (doped semiconductors).

As we discussed previously, [5] not all the boundary conditions are independent of each other when $q_y=0$. For transverse EMFs, only four boundary conditions related to the continuity of the x and y components of the $\mathbf{E}$ and $\mathbf{H}$ fields are independent. In the following, we only list the independent boundary conditions.

For the specular ($n=0$) mode, the boundary conditions for the EMFs at the interface require [5]

$$B_x^T(q_x;\omega)-C_x^T(q_x;\omega)-C_x^L(q_x;\omega) = -A_x^T(q_x;\omega), \tag{10}$$

$$B_y^T(q_x;\omega)-C_y^T(q_x;\omega) = -A_y^T(q_x;\omega), \tag{11}$$

$$\begin{aligned}
&-\frac{i\beta_1^T c^2}{\omega^2}B_y^T(q_x;\omega)-\frac{i\beta_2^T c^2}{\omega^2}C_y^T(q_x;\omega)-\sum_{n'=-\infty}^{\infty}\left[\bar{\chi}_s(q_{n'};\omega)+\frac{1}{2}\bar{\chi}_g(q_{n'};\omega)\right]B_y^T(q_{n'};\omega)\\
&= \left[\bar{\chi}_s(q_x;\omega)+\frac{1}{2}\bar{\chi}_g(q_x;\omega)-\frac{i\eta_1^T c^2}{\omega^2}\right]A_y^T(q_x;\omega),
\end{aligned} \tag{12}$$

$$\begin{aligned}
&-\frac{i}{\beta_1^T}B_x^T(q_x;\omega)-\frac{i}{\beta_2^T}\epsilon_2^T(q_x,\beta_2^T;\omega)C_x^T(q_x;\omega)-\sum_{n'=-\infty}^{\infty}\left[\bar{\chi}_s(q_{n'};\omega)+\frac{1}{2}\bar{\chi}_g(q_{n'};\omega)\right]B_x^T(q_{n'};\omega)\\
&= \left[\bar{\chi}_s(q_x;\omega)+\frac{1}{2}\bar{\chi}_g(q_x;\omega)-\frac{i}{\eta_1^T}\right]A_x^T(q_x;\omega),
\end{aligned} \tag{13}$$

$$\begin{aligned}
&\frac{q_x}{\beta_2^T}\left[\epsilon_2^T(q_x,\beta_2^T;\omega)-\epsilon_b\right]C_x^T(q_x;\omega)+\frac{\beta_2^L}{q_x}\epsilon_b C_x^L(q_x;\omega)\\
&+\ iq_x\sum_{n'=-\infty}^{\infty}\left[\bar{\chi}_s(q_{n'};\omega)+\frac{1}{2}\bar{\chi}_g(q_{n'};\omega)\right]\left[C_x^T(q_{n'};\omega)+C_x^L(q_{n'};\omega)\right]=0.
\end{aligned} \tag{14}$$

Here, Eqs. (10) and (11) are related to the continuity of the x and y components of the $\mathbf{E}$ field at the interface. Equations (12) and (13) are associated with the continuity of the x and y components of the $\mathbf{H}$ field. In addition, Eq. (14) comes from the so-called additional boundary condition [5, 7].

The summation over n' in Eqs. (12)–(14) describes the excitation of diffracted EMFs [6] through the specular mode due to an incident EMF by interacting with both the conductive grating ($\bar{\chi}_g$) and the current-driven conducting sheet ($\bar{\chi}_s$). After excluding these Bragg modes, (i.e., by setting $n' = 0$,) in Eqs. (10)–(14), we reproduce the same results derived before [5]. The incident light with an s-polarization or a p-polarization makes a difference in Eqs. (10)–(13) for the generation of diffracted EMFs.

For the Bragg modes with even integers $n \neq 0$, the boundary conditions for the EMFs lead to

$$B_x^T(q_n;\omega) - C_x^T(q_n;\omega) - C_x^L(q_n;\omega) = 0, \tag{15}$$

$$B_y^T(q_n;\omega) - C_y^T(q_n;\omega) = 0, \tag{16}$$

$$\beta_{1,n}^T B_y^T(q_n;\omega) + \beta_{2,n}^T C_y^T(q_n;\omega) = 0, \tag{17}$$

$$\frac{1}{\beta_{1,n}^T} B_x^T(q_n;\omega) + \frac{1}{\beta_{2,n}^T}\epsilon_2^T(q_n,\beta_{2,n}^T;\omega) C_x^T(q_n;\omega) = 0, \tag{18}$$

$$\frac{q_n}{\beta_{2,n}^T}[\epsilon_2^T(q_n,\beta_{2,n}^T;\omega) - \epsilon_b] C_x^T(q_n;\omega) + \frac{\beta_{2,n}^L}{q_n}\epsilon_b C_x^L(q_n;\omega) = 0. \tag{19}$$

Since the Bragg modes with even integers $n \neq 0$ do not directly couple to the conducting sheet and grating due to the mirror symmetry of the system with respect to $x = 0$, Eqs. (15–19) simply describe the continuity of the x and y components of the **E** and **H** fields, as well as the vanishing normal (z) component of the total current at the interface.

For the Bragg modes with odd integers n, the boundary conditions for the EMFs take the form of [5]

$$B_x^T(q_n;\omega) - C_x^T(q_n;\omega) - C_x^L(q_n;\omega) = 0, \tag{20}$$

$$B_y^T(q_n;\omega) - C_y^T(q_n;\omega) = 0, \tag{21}$$

$$-\frac{i\beta_{1,n}^T c^2}{\omega^2} B_y^T(q_n;\omega) - \frac{i\beta_{2,n}^T c^2}{\omega^2} C_y^T(q_n;\omega) - \frac{(-1)^{\frac{|n|-1}{2}}}{|n|\pi} \sum_{n'=-\infty}^{\infty} \bar{\chi}_g(q_{n+n'};\omega) B_y^T(q_{n+n'};\omega) = 0, \tag{22}$$

$$-\frac{i}{\beta_{1,n}^T} B_x^T(q_n;\omega) - \frac{i}{\beta_{2,n}^T}\epsilon_2^T(q_n,\beta_{2,n}^T;\omega) C_x^T(q_n;\omega) - \frac{(-1)^{\frac{|n|-1}{2}}}{|n|\pi} \sum_{n'=-\infty}^{\infty} \bar{\chi}_g(q_{n+n'};\omega) B_x^T(q_{n+n'};\omega) = 0, \tag{23}$$

$$\begin{aligned} &\frac{q_n}{\beta_{2,n}^T}[\epsilon_2^T(q_n,\beta_{2,n}^T;\omega) - \epsilon_b] C_x^T(q_n;\omega) + \frac{\beta_{2,n}^L}{q_n}\epsilon_b C_x^L(q_n;\omega) \\ &+ \frac{i(-1)^{\frac{|n|-1}{2}}}{|n|\pi} q_n \sum_{n'=-\infty}^{\infty} \bar{\chi}_g(q_{n+n'};\omega)[C_x^T(q_{n+n'};\omega) + C_x^L(q_{n+n'};\omega)] = 0. \end{aligned} \tag{24}$$

Here, Eqs. (20) and (21) for the odd integers n are the same as Eqs. (15) and (16) for the even integers $n \neq 0$. The summation over n' in Eqs. (22)–(24) represents the nonlocal mixing among the Bragg modes ($n' \neq -n$) and the specular mode ($n' = -n$) by interacting with the conductive grating through induced sheet currents [6].

Equations (10)–(24) constitute a complete set of linear equations (or a linear matrix equation [5] with a coefficient matrix and a source-term vector) with respect to the independent variables $B_x^T(q_n,\omega)$, $B_y^T(q_n,\omega)$, $C_x^T(q_n,\omega)$, $C_y^T(q_n,\omega)$, and $C_x^L(q_n,\omega)$ for $n = 0, \pm 1, \pm 2, \cdots$. The components of the source-term vector contain $A_x^T(q_n,\omega)$ for the p-polarization or $A_y^T(q_x,\omega)$ for the s-polarization.

For the conducting sheet and grating, their dielectric functions are given by [8] $\bar{\chi}_s(q_x;\omega) = \chi_e(q_x;\omega - q_x v_d)$ and $\bar{\chi}_g(q_x;\omega) = \chi_e(q_x;\omega)$. Here, v_d is the drift velocity of electrons under a bias field and the electron sheet polarizability in the RPA is given by [9]

$$\begin{aligned} \chi_e(q_x;\omega) \;=\; &\frac{2n_{2\mathrm{D}}e^2 m_s^*}{\epsilon_0\hbar^2 k_F|q_x|^3}\left\{\left[2z - C_-\sqrt{(z-u)^2-1} - C_+\sqrt{(z+u)^2-1}\right]\right. \\ &\left. + i\left[D_-\sqrt{1-(z-u)^2} - D_+\sqrt{1-(z+u)^2}\right]\right\}, \end{aligned} \tag{25}$$

where n_{2D} is the electron density, $k_F = \sqrt{2\pi n_{2D}}$ is the Fermi wave number of electrons in the sheet, and m_s^* is the effective mass of electrons. Moreover, we have defined the notations in Eq. (25): $u = m_s^*\omega/\hbar k_F|q_x|$, $z = |q_x|/2k_F$, $C_+ = (z+u)/|z+u|$ and $D_+ = 0$ ($C_+ = 0$ and $D_+ = 1$) for $|z+u| > 1$ ($|z+u| < 1$), and $C_- = (z-u)/|z-u|$ and $D_- = 0$ ($C_- = 0$ and $D_- = 1$) for $|z-u| > 1$ ($|z-u| < 1$). For $\chi_e(q_x;\omega)$, we have to use the full expression in Eq. (25), instead of a Drude-type model, [5] since $|\omega - q_n v_d|$ can be smaller than $\hbar k_F|q_n|/m_s^*$, which invalids the long-wavelength limit.

In the presence of interface roughness, the spatial interference of various diffracted EMF components will be destroyed. In this case, a statistical average is required, i.e., a spatial average for the square of the total EMF including the summation over all diffracted EMF components as shown in Eqs. (1) and (2). The spatially-averaged square of the ratio $\mathcal{R}$ of the reflected to the incident **E**-field amplitude at the interface, which cancels out the interference due to the x dependent relative phase, is given by

$$\mathcal{R}(\omega) = \sum_{n'=-\infty}^{\infty} R_n(q_n;\omega) = \sum_{n'=-\infty}^{\infty} \frac{[1+|q_n/\beta_{1,n}^T|^2]|B_x^T(q_n;\omega)|^2 + |B_y^T(q_n;\omega)|^2}{[1+(q_n/\eta_1^T)^2]|A_x^T(q_n;\omega)|^2 + |A_y^T(q_x;\omega)|^2}, \tag{26}$$

where both the far field ($n = 0$) and near field ($|n| \gg 1$) contributions are included. $\mathcal{R}(\omega)$ in Eq. (26) is the reflection spectrum as a function of incident photon energy at the interface.

3. CONCLUSION

The main prediction in the current research is the enhanced-radiation mode in the spectrum of the reflected electromagnetic field in the presence of a grating in the x direction. This enhanced optical reflection results from the nonlocal mixing of $n = 0$ specular with $|n| \geq 1$ diffraction modes of the reflected electromagnetic field by the free-electron induced optical polarization and the interference of two counter-propagating surface optical-polarization waves. We have demonstrated and explained the effects of grating period and absorption loss on the optical enhancement of the reflected electromagnetic fields.

ACKNOWLEDGMENT

This research was supported by the Air Force Office of Scientific Research (AFOSR).

REFERENCES

1. Ebbesen, T. W., H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, *Nature*, Vol. 391, 667, London, 1998.
2. Ghaemi, H. F., T. Thio, D. E. Grupp, T. W. Ebbesen, and H. J. Lezec, *Phys. Rev. B*, Vol. 58, 6779, 1998.
3. Schröter, U. and D. Heitmann, *Phys. Rev. B*, Vol. 58, 15419, 1998.
4. Barbara, A., P. Quémerais, E. Bustarret, and T. Lopez-Rios, *Phys. Rev. B*, Vol. 66, 161403, 2002.
5. Huang, D. H., C. Rhodes, P. M. Alsing, and D. A. Cardimona, *J. Appl. Phys.*, Vol. 100, 113711, 2006.
6. Gumbs, G. and D. H. Huang, *Phys. Rev. B*, Vol. 75, 115314, 2007.
7. Forstmann, F. and R. R. Gerhardts, *Metal Optics Near the Plasma Frequency*, Springer-Verlag, Berlin, 1986.
8. Kempa, K., P. Bakshi, J. Cen, and H. Xie, *Phys. Rev. B*, Vol. 43, 9273, 1991.
9. Stern, F., *Phys. Rev. Lett.*, Vol. 18, 546, 1967.

An Efficient Approach to Identifying a Complete Photonic Band Gap in Two-dimensional Photonic Crystals with Omnidirectional Light Propagation

Ruei-Fu Jao[1, 2] and Ming-Chieh Lin[1]

[1]NanoScience Simulation Laboratory, Department of Physics, Fu Jen Catholic University
Taipei County 24205, Taiwan, Republic of China
[2]Department of Physics, National Cheng Kung University
Tainan 701, Taiwan, Republic of China

Abstract— Omnidirectional light propagation in two-dimensional (2D) photonic crystals (PCs) is investigated. An efficient approach to identifying a complete photonic band gap (PBG) in 2D PCs has been developed. The in-plane band structure of 2D photonic crystals is calculated by an adaptive finite element method. With adopting the suitable boundary conditions, the eigenvalues can be easily and rapidly calculated no matter how complex the geometric structures are. By symmetry, the omnidirectional photon density of states (PDOS) can be calculated based on the in-plane dispersion relation within the irreducible Brillouin zone. The PDOS corresponding to both the radiation and evanescent waves can be obtained accurately and efficiently. We demonstrate that the "complete band gaps" showed by some previous work due to considering only the radiation modes will be closed by including the contributions of the evanescent modes. These results are of general importance and relevant to the spontaneous emission by an atom, or to dipole radiation in two-dimensional periodic structures.

1. INTRODUCTION

In the past two decades, photonic crystals (PCs) have attracted much attention [1, 2]. According to the dimension of the periodicity, photonic crystals are divided into three categories, namely one-, two-, and three-dimensional (3D) crystals-periodic dielectric materials characterized by photonic band gaps (PBGs). A PBG can prohibit the propagation of electromagnetic (EM) waves whose frequencies fall within the band gap region. These materials are expected to find many applications in optoelectronics and optical communications. It was proposed that the emission of electromagnetic radiation can be modified by the environment [3, 4]. Several environments such as metallic cavities [5], dielectric cavities [6], and superlattices [7–12] have been studied. The environmental effects have been described by the photon density of states (PDOS) which is related to the transition rate of Fermi golden rule. In principle, a complete band gap along all dimensions in space can be best realized in a three-dimensional (3D) system. However, the difficulty in fabricating such 3D crystals with band gaps in the optical regime prohibits the progression of many applications. Many studies in 2D photonic crystals have been mainly focused on the in-plane propagation of EM waves [13]. There exhibits PBGs in various types of period structures. However, for some applications, the investigation of an omnidirectional light propagation is crucial. A few studies show the possibility of some 2D crystal structures to have an omnidirectional absolute band gaps with adopting the off-plane wave vector $k_z = k_0 \sin\theta$, where $k_0 = \omega/c$ [14–15]. In this work, we demonstrate that the "complete band gaps" showed by some previous work due to considering only the radiation modes will be closed by including the contributions of the evanescent modes. Theoretically, there are no band gaps for propagation in the z-direction. As k_z increases, the modes decouple and the bandwidth shrinks to zero [16–21]. The polarization characteristics including both the transverse electric (TE) and transverse magnetic (TM) modes are considered in our simulation model. The contributions of the total PDOS from the radiative and evanescent waves can be distinguished.

2. FORMULATION

The propagation of light in a photonic crystal can be studied by solving the Maxwell's equations. For time-harmonic fields it is convenient to use a phasor notation. The Maxwell's equations lead

to the wave equations, or the master equations:

$$\vec{\nabla} \times \left[\frac{1}{\mu(r)} \times \vec{\nabla} \times \vec{E}(r) \right] - \omega^2 \varepsilon(r) \vec{E}(r) = 0, \tag{1}$$

$$\vec{\nabla} \times \left[\frac{1}{\varepsilon(r)} \times \vec{\nabla} \times \vec{H}(r) \right] - \omega^2 \mu(r) \vec{H}(r) = 0, \tag{2}$$

where $\varepsilon(r)$ and $\mu(r)$ are the permittivity and permeability functions of the PCs, respectively, and ω is the angular eigen-frequency. In a 2D periodic system, the dielectric function is a periodic function of x and y. We assume that the materials are linear, homogeneous, isotropic, lossless, and nonmagnetic. We have

$$\varepsilon_r(x, y) = \begin{cases} \varepsilon_a, & x, y \in \text{air region} \\ \varepsilon_d, & x, y \in \text{dielectric region} \end{cases} \tag{3}$$

where $\varepsilon_r(x, y)$ is the dielectric function profile, and ε_a and ε_d are the dielectric constants of the air and dielectric regions, respectively. The two master equations are reduced to two homogeneous Helmholtz's equations for the air (dielectric) region

$$\nabla^2 \vec{E}(r) + \frac{\omega^2}{c^2} \varepsilon_{a(d)} \vec{E}(r) = 0, \tag{4}$$

and

$$\nabla^2 \vec{H}(r) + \frac{\omega^2}{c^2} \varepsilon_{a(d)} \vec{H}(r) = 0. \tag{5}$$

A two-dimensional photonic crystal is periodic in two directions (x, y) and homogeneous in the third (z). For light propagating in the system, we can separate the modes into two independent polarizations, TE and TM modes, and consider the band structure and photon density of states of each. The propagation properties of TM and TE modes can be characterized by the field components along the homogeneous direction, $E_z(x, y)$ and $H_z(x, y)$, respectively. As the system has translational symmetry along the z-axis, we can assume the longitudinal wave functions as a plane wave, $\exp(ik_z z)$. By using separating of variables, the corresponding Helmholtz's equations, the z components of Eqs. (4) and (5), for the air (dielectric) region can be split into transverse and longitudinal parts and the problem can be simplified as solving the Helmholtz's equation in the xy-plane. We can obtain

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right] E_z(x, y) + \left[\frac{\omega^2}{c^2} \varepsilon_{a(d)} - k_z^2 \right] E_z(x, y) = 0, \tag{6}$$

and

$$\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right] H_z(x, y) + \left[\frac{\omega^2}{c^2} \varepsilon_{a(d)} - k_z^2 \right] H_z(x, y) = 0. \tag{7}$$

To perform the 3D PDOS calculations, we construct two equifrequecy regions $\omega(k_x, k_y, k_z) = \omega$ and $\omega(k_x, k_y, k_z) = \omega + \Delta\omega$, where ω is an arbitrary value of the frequency and $\Delta\omega$ is an infinitesimal increment [12, 13]. The PDOS is $\mathrm{d}N(\omega) \equiv D(\omega)\mathrm{d}\omega$:

$$D(\omega) = \frac{V\sqrt{\mu_r \varepsilon_r}}{8\pi^3 c} \int_{\omega_k} \frac{\omega}{\sqrt{\omega^2 - \omega_c^2}} dS_k. \tag{8}$$

3. RESULTS AND DISCUSSION

We consider the omnidirectional light propagation in an inhomogeneous, linear, and nonmagnetic medium. The calculation of total PDOS is based on the finite element method for a two-dimensional photonic crystal. Using the formulation for the critical angle, $\cos^{-1}(\omega_c/\omega) \geq \sin^{-1}(\sqrt{\varepsilon_a/\varepsilon_d})$, we can calculate the PDOS of the evanescent and radiative modes, separately. Figure 1 shows the comparisons of 3D PDOS calculated by the FEM and the plane wave expansion method (PWEM). These two curves are in good agreement. There is no complete three-dimensional PBG for two dimensional structures. Then, we investigate the 3D PDOS of TE and TM modes, and we employ the same parameters as those in Ref. [16]. Figure 2 and Figure 3 show our calculated 3D PDOS

for the TE and TM modes, respectively. The PDOS for the radiative and evanescent waves are also plotted. As one can see, the 3D PDOS of the evanescent waves is larger than that of the radiative waves for both the TE and TM modes. Although the PDOS for the TE and TM modes exhibit similar behavior, the corresponding contributions from the radiative and evanescent parts are quite different. One can see the 3D PDOS of the radiative waves have a PBG between 0.403–0.434($2\pi c/a$). However, the "complete band gaps" will be closed by including the contribution of the evanescent mode.

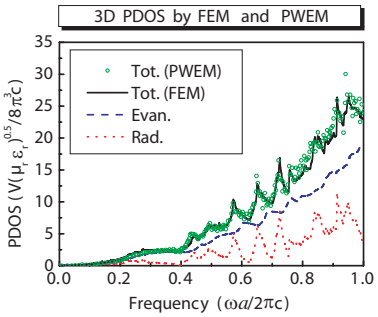


Figure 1: Comparisons of 3D PDOS calculated by the FEM and PWEM [16].

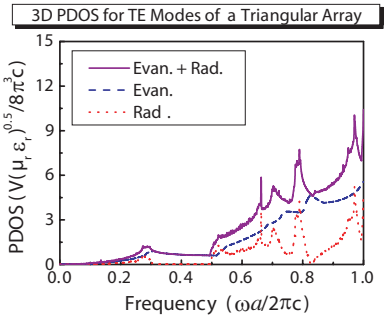


Figure 2: 3D PDOS for TE modes TM modes of a triangular array [14]. The dashed (dotted) lines correspond to the PDOS of evanescent (radiative) waves.

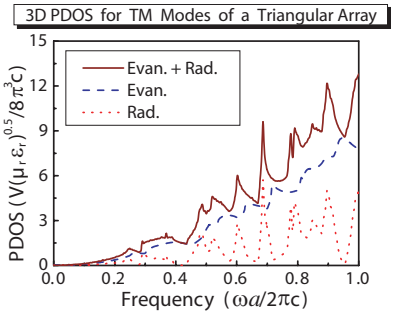


Figure 3: 3D PDOS for TM modes of a triangular array [14]. The dashed (dotted) lines correspond to the PDOS of evanescent (radiative) waves.

4. CONCLUSIONS

In this work, omnidirectional light propagation in 2D photonic crystals is investigated. The polarization characteristics including both the TE and TM modes are considered in our simulation model. The contributions of the total PDOS from the radiative and evanescent waves can be distinguished, respectively. For the quantum optical phenomena, the total PDOS has to be considered. The "complete band gaps" do not exhibit in 2D PCs, if one considers both radiative and evanescent modes. With our approach, an omnidirectional PBG can be determined efficiently. These results are of general importance and relevant to the spontaneous emission by an atom, or to dipole radiation in two-dimensional periodic structures.

ACKNOWLEDGMENT

This work was partially supported by the National Science Council, Taiwan, R. O. C. under Grant No. NSC 96 – 2112 – M – 030 – 004 – MY3, National Center for Theoretical Sciences, and National Center for High-performance Computing, Taiwan, R. O. C. which provides the computing resources.

REFERENCES

1. Joannopoulos, J. D., R. D. Meade, and J. N. Winn, *Photonic Crystals*, Princeton University Press, Princeton, New Jersey, 1995.
2. Joannopoulos, J. D., P. R. Villeneuve, and S. Fan, "Photonic crystals: putting a new twist on light," *Nature*, Vol. 386, No. 13, 143–149, London, 1997.
3. Purcell, E. M., "Spontaneous emission probabilities at radio frequencies," *Phys. Rev.*, Vol. 69, 681, 1946.
4. Kleppner, D., "Inhibited spontaneous emission," *Phys. Rev. Lett.*, Vol. 47, 233–236, 1981.
5. Barut, A. O. and J. P. Dowling, "Quantum electrodynamics based on self-energy: Spontaneous emission in cavities," *Phys. Rev. A*, Vol. 36, 649–654, 1987.
6. Rigneault, H. and S. Monneret, "Modal analysis of spontaneous emission in a planar microcavity," *Phys. Rev. A*, Vol. 54, 2356–2368, 1996.
7. Dowling, J. P. and C. M. Bowden, "Atomic emission rates in inhomogeneous media with applications to photonic band structures," *Phys. Rev. A*, Vol. 46, 612–622, 1992.
8. Suzuki, T. and P. K. L. Yu, "Emission power of an electric dipole in the photonic band structure of the fcc lattice," *Opt. Soc. Am. B*, Vol. 12, No. 4, 570–582, 1995.

9. Kamli, A., M. Babiker, A. Al-Hajry, and N. Enfati, "Dipole relaxation in dispersive photonic band-gap structures," *Phys. Rev. A*, Vol. 55, 1454–1461, 1997.
10. Sánchez, A. S. and P. Halevi, "Spontaneous emission in one-dimensional photonic crystals," *Phys. Rev. E*, Vol. 72, 056609, 2005.
11. Halevi, P. and A. S. Sánchez, "Spontaneous emission in a high-contrast one-dimensional photonic crystal," *Opt. Commun.*, Vol. 251, 109–114, 2005.
12. Lin, M. C. and R. F. Jao, "Quantitative analysis of photon density of states for a realistic superlattice with omnidirectional light propagation," *Phys. Rev. E*, Vol. 74, 046613, 2006.
13. Lin, M. C. and R. F. Jao, "Finite element analysis of photon density of states for two-dimensional photonic crystals with in-plane light propagation," *Opt. Express*, Vol. 15, 207–218, 2007.
14. Li, Z. Y. and Y. Xia, "Omnidirectional absolute band gaps in two-dimensional photonic crystals," *Phys. Rev. B*, Vol. 64, 153108, 2001.
15. Haas, T. and A. Hesse, T. Doll, "Omnidirectional two-dimensional photonic crystal band gap structures," *Phys. Rev. B*, Vol. 73, 045130, 2006.
16. Busch, K. and S. John, "Photonic band gap formation in certain self-organizing systems," *Phys. Rev. E*, Vol. 58, 3896–3908, 1998.
17. Fussell, D. P., R. C. McPhedran, C. M. de Sterke, and A. A. Asatryan, "Three-dimensional local density of states in a finite two-dimensional photonic crystal composed of cylinders," *Phys. Rev. E*, Vol. 67, 045601(R), 2003.
18. Sigalas, M. M., R. Biswas, K. M. Ho, and C. M. Soukoulis, "Theoretical investigation of off-plane propagation of electromagnetic waves in two-dimensional photonic crystals," *Phys. Rev. B*, Vol. 58, 6791–6794, 1998.
19. Foteinopoulou, S., A. Rosenberg, M. M. Sigalas, and C. M. Soukoulis, "In- and out-of-plane propagation of electromagnetic waves in low index contrast two dimensional photonic crystals," *J. Appl. Phys.*, Vol. 89, 824–830, 2001.
20. Rosenberg, A., R. J. Tonucci, and E. L. Shirley, "Out-of-plane two-dimensional photonic band structure effects observed in the visible spectrum," *J. Appl. Phys.*, Vol. 82, No. 12, 6354–6356, 1997.
21. Meade, R. D., K. D. Brommer, A. M. Rappe, and J. D. Joannopoulos, "Existence of a photonic band gap in two dimensions," *Appl. Phys. Lett.*, Vol. 61, No. 4, 495–497, 1992.

Limitation of Spontaneous Emission Enhancement Using Surface Plasmon Polaritons

Greg Sun[1] **and Jacob B. Khurgin**[2]

[1]Department of Physics, University of Massachusetts Boston
100 Morrissey Boulevard, Boston, Massachusetts 02125, USA
[2]Department of Electrical and Computer Engineering, Johns Hopkins University
3400 N. Charles Street, Baltimore, Maryland 21218, USA

Abstract— We develop a rigorous theory of the enhancement of spontaneous emission from a light-emitting device via coupling the radiant energy in and out of surface plasmon polaritons (SPPs) on the metaldielectric interface. We show that while the efficiency of coupling into the SPP mode can be quite high, the radiative efficiency of the SPP itself is relatively low, with a substantial fraction of the energy lost in the metal. Using the GaN/Ag system as an example we obtain easy-to-interpret analytical results that unequivocally indicate that using SPP pays off only for emitters that have medium-to-low luminescence efficiency.

1. INTRODUCTION

Recently, there has been a great deal of interest in enhancing the efficiency of spontaneous emission using the surface plasmon polariton (SPP). The first definite sign of improvement was attained in GaN by placing a thin Ag film atop the GaN [1, 2]. Since then, SPP enhancement has been shown in a large number of different light emitting media [3–8]. The coupling of normally non-radiative SPPs into the radiation mode has been accomplished with 1-D dielectric gratings [3, 4, 7], or 2-D corrugated silver films [6, 9], or more complicated cavity-like structures [10].

There were a number of analytical [11, 12] and numerical calculations [10] that describe the SPP coupling into the radiation mode. Some of the calculations [12] completely disregard the issues of Ohmic losses in SPPs and thus give an overly optimistic prognosis for the SPP enhancement. Others [10] require extensive numerical modeling for each particular case. Recently, we have developed an analytical framework [13] that provides unambiguous answers about the maximum improvement in radiative efficiency that one can expect using SPP and about the optimal configuration in which such improvement can be achieved. The theory relies only upon a very limited number of parameters, namely the intrinsic luminescence efficiency of the emitter and the imaginary part of the dielectric constant of a metal.

Let us consider a given emitter with an internal radiative diative efficiency $\eta_{rad} = \tau_{rad}^{-1}/(\tau_{rad}^{-1} + \tau_{nrad}^{-1})$, where the radiative lifetime τ_{rad} is determined mostly by the density of modes in the dielectric. When the emitter is placed in the vicinity of the metal layer, there appears another decay channel – into the high-density SPP modes with a decay rate $\tau_{SPP}^{-1} = F_P \tau_{rad}^{-1}$ enhanced by the Purcell factor F_P [14]. Thus the energy is very efficiently transferred from the emitter into the SPP mode because of the large values of F_P, but in order for energy to be emitted, it has to be coupled into the continuum of radiation modes by a grating. The radiative coupling into the radiation modes, occurring with a rate κ_{pr} must compete with the nonradiative decay of the SPP itself, κ_{nr}. If one now introduces the efficiency of coupling between SPP and radiation mode as $\eta_{pr} = \kappa_{pr}/(\kappa_{pr}+\kappa_{nr})$, then the radiative efficiency due to SPP is $\eta_{SPP} = \eta_{pr} F_P \tau_{rad}^{-1} / \left(F_P \tau_{rad}^{-1} + \tau_{nrad}^{-1}\right)$. We can then obtain the expression for the overall SPP enhancement of the radiative efficiency as

$$F_{SPP} = \frac{\eta_{SPP}}{\eta_{rad}} = \frac{\eta_{pr}}{\eta_{rad} + (1-\eta_{rad}) F_P^{-1}} \leq \frac{\eta_{pr}}{\eta_{rad}} \tag{1}$$

Thus, no matter how high the Purcell factor, in the end the SPP enhancement is limited by the ratio of radiative coupling efficiencies of the SPP to the original emitting source.

2. SURFACE PLASMON POLARITON-DISPERSION AND PURCELL FACTOR

We consider a typical SPP enhancement scheme that has emission collected on the same side of the metal as shown in Fig. 1.

The SPP mode with a frequency ω localized at the interface of a metal layer with a dielectric function $\varepsilon_M = 1-\omega_p^2/(\omega^2+j\omega\Gamma)$ and a dielectric medium with a dielectric constant ε_D is transverse magnetic (TM) wave. The electric field of the SPP can be written as $\mathbf{E}_p(z,\, x) = A_p(z)\mathbf{e}_p(\beta_p,\, x)e^{j(\beta_p z-\omega t)}$ with $U_p = (\varepsilon_o\varepsilon_D/k_D)\,|A_p|^2$ defined as the surface energy density of the SPP, and the SPP eigen mode

$$\mathbf{e}_p(\beta_p,\, x) = \begin{cases} \dfrac{\sqrt{2}}{\varepsilon w_p^{1/2}\sqrt{2\beta_p^2-1}}\,(j\beta_p\hat{\mathbf{x}}+q_M\hat{\mathbf{z}})\,e^{-q_M x}, & x>0 \\ \dfrac{\sqrt{2}}{w_p^{1/2}\sqrt{2\beta_p^2-1}}\,(j\beta_p\hat{\mathbf{x}}+q_D\hat{\mathbf{z}})\,e^{-q_D x}, & x<0 \end{cases} \tag{2}$$

where w_p is the effective width of the SPP and $\varepsilon = \varepsilon_M/\varepsilon_D$. Here wave vectors β_p, q_D, and q_M, related by $\beta_p^2-q_D^2 = 1$ and $\beta_p^2-q_M^2 = \varepsilon$, are all normalized to $k_D = \sqrt{\varepsilon_D}\omega/c$, while co-ordinates x and z are normalized to $1/k_D$, and are therefore all dimensionless. The SPP dispersion of the real (solid line) and imaginary parts (dashed line) of the propagation constant $\beta_p = \beta_p^{'} + j\beta_p^{''} = \sqrt{\varepsilon/(1+\varepsilon)}$ shown in Fig. 2(a) is obtained using the actual complex dielectric constant for Ag [15] as well as the dispersion of index of refraction for GaN [16]. Also plotted in Fig. 2(a) (dotted line) is the propagation constant β_{p0} obtained under the assumption of no loss: $\mathrm{Im}\,(\varepsilon_M) = 0$. As one can see, the presence of loss in the metal not only engenders a substantial imaginary part $\beta_p^{''}$ of the propagation constant, but also limits the maximum attainable value of the real part $\beta_p^{''}$ and consequently, the Purcell factor F_P (Fig. 2(b)) as the minimum value of group velocity v_g is also

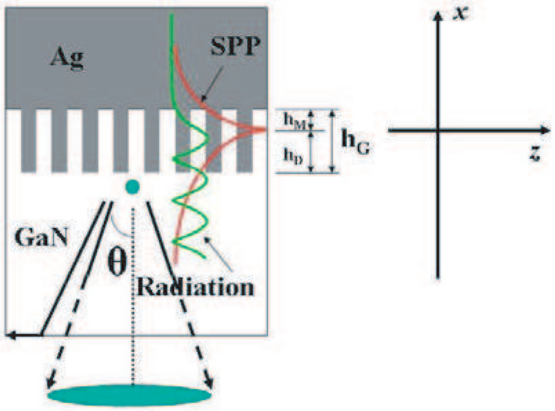


Figure 1: Light emission through coupling of SPP to radiation modes by a grating placed at the interface.

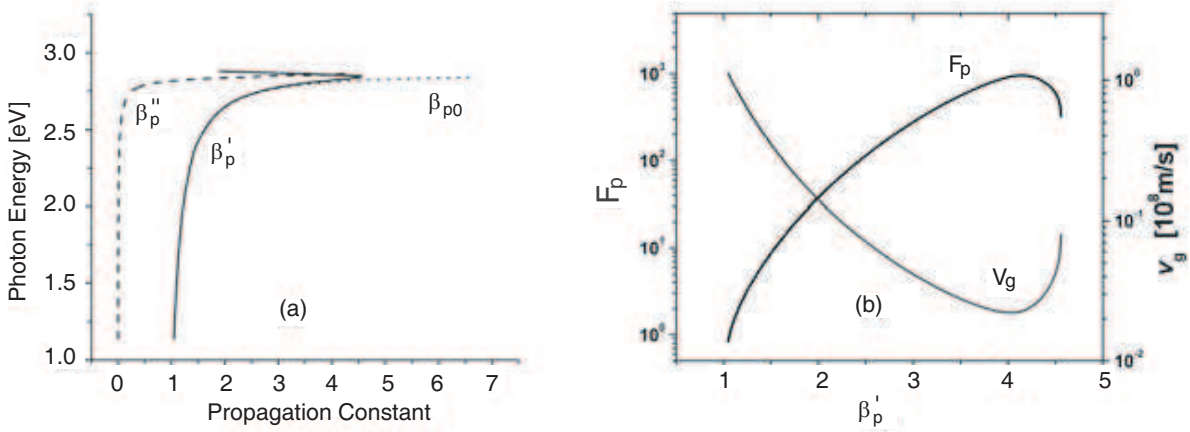


Figure 2: (a) SPP dispersion curve and (b) Purcell factor F_P and group velocity v_g as a function of the propagation constant.

limited.

3. SPP COUPLING INTO THE RADIATION MODES

3.1. Radiation Modes

Consider now the TM radiation modes into which the SPPs get coupled by the grating, We can write the electric field of the TM radiation mode as $\mathbf{E}_r(z,\, x) = A_r(z)\mathbf{e}_r(\beta_r,\, x)e^{j(\beta_r z-\omega t)}$ with $U_r = (\varepsilon_o\varepsilon_D/k_D)\,|A_r|^2$ defined as the surface energy density of the radiation mode and the eigen modes of electric field inside the dielectric normalized to some transverse dimension L_x,

$$e_r = \begin{cases} \dfrac{2k_r}{\sqrt{L_x\,(q_r^2+\varepsilon^2k_r^2)}}(j\beta_r\hat{\mathbf{x}}+q_r\hat{\mathbf{z}})e^{-q_r x}, & x>0 \\ \dfrac{2k_r}{\sqrt{L_x\,(q_r^2+\varepsilon^2k_r^2)}}\left\{j\beta_r\left[\varepsilon\cos\left(k_r x\right)-\left(q_r/k_r\right)\sin\left(k_r x\right)\right]\hat{\mathbf{x}}\right. & \\ \left.+\left[\varepsilon k_r\sin\left(k_r x\right)+q_r\cos\left(k_r x\right)\right]\hat{\mathbf{z}}\right\}, & x<0 \end{cases} \tag{3}$$

where $\beta_r^2+k_r^2=1$ and $\beta_r^2-q_r^2=\varepsilon$.

3.2. SPP-radiation Coupling

We shall treat grating as a perturbation and the dimensionless grating height $h_G = h_M + h_G$ has been divided according to $h_D = 1/4q_D$ and $h_M = 1/4q_M$. Both 1-D (Fig. 3(a)) and 2-D (Fig. 3(b)) gratings have been studied. Consider now the process of coupling from the SPP mode $\boldsymbol{\beta}_p$ into the continuum of the radiation modes with propagation vector $\boldsymbol{\beta}_r = \boldsymbol{\beta}_p - \mathbf{G}$, where $\mathbf{G}$ is the wave-vector of the first order Fourier component of the grating. As shown in Fig. 4(a) the SPP modes are distributed on the ring with a radius $\beta_p(>1)$ while the radiation modes are lying inside a smaller circle with a radius of 1. The angle at which the radiation mode travels respect to the normal of the metal-dielectric interface (Fig. 1) is therefore given as $\theta = \sin^{-1}\beta_r$.

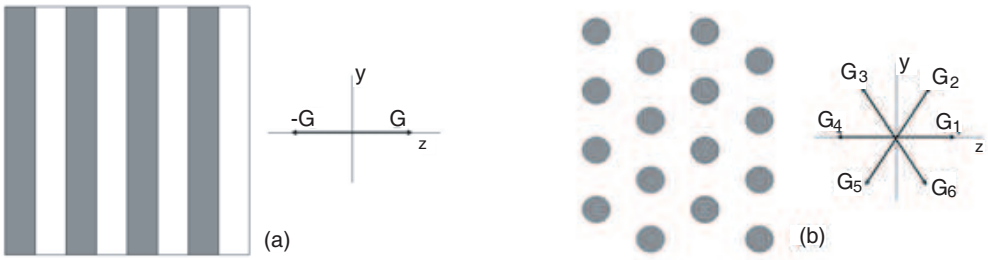


Figure 3: Top view of (a) 1-D and (b) 2-D grating patterns along with their first-order wave-vectors in reciprocal space.

To obtain coupling between the discrete SPP state and the radiation continuum, we first evaluate the coupling between the SPP mode Eq. (2) and one of the radiation modes, Eq. (3), and then integrate over the one-dimensional density of radiation modes $\rho_r(\beta_r,\, \omega) = 1/2\pi\omega\sqrt{1-\beta_r^2}$ to obtain the following equation for the radiative decay of the SPP mode

$$\frac{\partial}{\partial z}\,|A_p(z)|^2 = -\kappa_{pr}(\beta_p,\, \beta_r)\,|A_p(z)|^2 \tag{4}$$

where the coupling strength is given by the following analytical expression

$$\begin{aligned}\kappa_{pr} =& \frac{(\varepsilon-1)^2\,|f_G|^2}{2w_p(2\beta_p^2-1)}\left[\beta_p^2+\frac{(\beta_p^2-1)^2}{\varepsilon_D\beta_p^2}\right]\frac{\beta_p k_r}{q_r^2+\varepsilon^2k_r^2}\left\{\frac{\beta_p\beta_r q_r+\varepsilon k_r^2 q_D}{q_D^2+k_r^2}\right. \\ & \left[1-e^{-q_D h_D}\left[(q_D/k_r)\sin(k_r h_D)+\cos(k_r h_D)\right]\right]+\frac{\varepsilon\beta_p\beta_r q_D-q_D^2 q_r}{q_D^2+k_r^2} \\ & \left.\left[1-e^{-q_D h_D}\left[\cos(k_r h_D)-(k_r/q_D)\sin(k_r h_D)\right]\right]+\frac{\beta_p\beta_r+q_M q_r}{\varepsilon(q_M+q_r)}\left[1-e^{-(q_M+q_r)h_M}\right]\right\}^2\end{aligned} \tag{5}$$

in which the first-order coefficient of the Fourier series of the in-plane ($x = 0$) periodic dielectric perturbation is $f_G = 1/\pi$ for the 1-D grating (Fig. 3(a)) and $f_G = 0.087$ for the 2-D grating (Fig. 3(b)). The coupling coefficient can be obtained as $\eta_{pr} = \chi_{pr}/(1 + \chi_{pr})$ where $\chi_{pr} = \kappa_{pr}/\kappa_{nr}$ and $\kappa_{nr} = 2\beta_p''$.

4. OVERALL RADIATION EFFICIENCY ENHANCEMENT

Depending on the method of collection, the overall radiation efficiency should be integrated within an emission cone (Fig. 4(b)) with a polar angle $\theta_{\max} = \sin^{-1}\beta_{r,\max}$ where the maximum radiation propagation constant $\beta_{r,\max}$ is related to the in-plane angle $\varphi_m = 2\sin^{-1}(\beta_{r,\max}/2\beta_p)$ (Fig. 4(a)). Only those SPP modes with $0 < \varphi < \varphi_m$ coupled into the radiation modes with $0 < \beta_r < \beta_{r,\max}$ are collected in the radiation cone. The overall radiation efficiency should be evaluated as $\eta_{pr,all} = (N_G/2\pi)\int_{-\varphi_m}^{\varphi_m}\eta_{pr}\mathrm{d}\varphi$ where $N_G = 2$ for 1-D grating with two **G**-vectors (Fig. 3(a)), and $N_G = 6$ for 2-D grating with six **G**-vectors (Fig. 3(b)). We consider three different cases of the collection methods that are limited by different maximum polar angles. First, we assume that all radiation modes escape the dielectric eventually by avoiding total internal reflection with the use of surface roughness or a domed structure ($\theta_{\max} = \pi/2$). Next we look at the situation in which the polar angle of the emission cone is limited by total internal reflection ($\theta_{\max} = \sin^{-1}(1/\sqrt{\varepsilon_D})$. Finally for the case where the emission cone is limited by the numerical aperture (NA) of collection lens, the maximum polar angle can be given as $\theta_{\max} = \sin^{-1}\left(\sin[\tan^{-1}(NA/2)]/\sqrt{\varepsilon_D}\right)$. The results of all three cases are shown in Fig. 5(a) for 1-D and in Fig. 5(b) for 2-D gratings along with the Purcell factor.

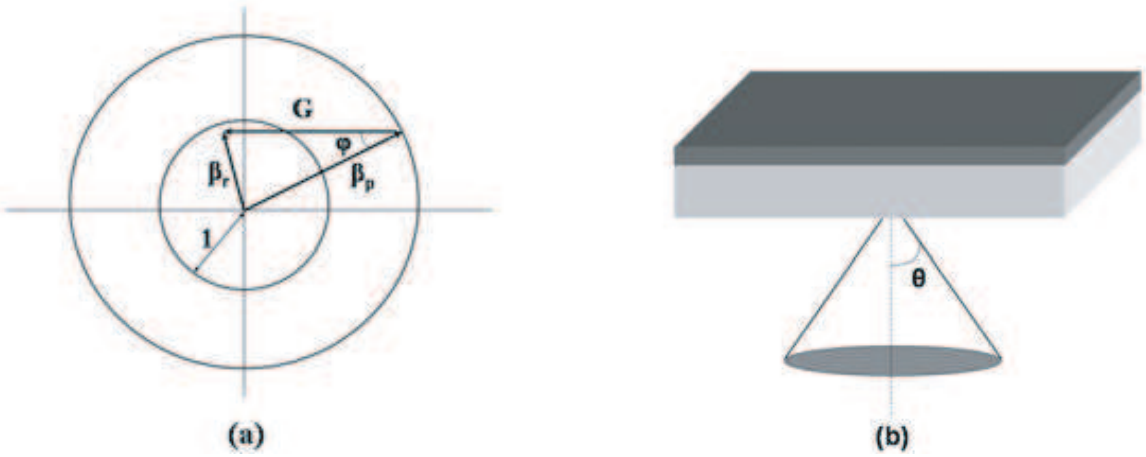


Figure 4: (a) SPP-radiation in-plane propagation constant coupling in reciprocal space, (b) Emission cone with polar angle θ.

A clear pattern is observed: the 1-D grating offers consistently stronger coupling between the SPP and radiation modes than the 2-D grating. As one can see, the coupling efficiency increases as β_p' decreases, but the Purcell factor also increases with β_p', hence overall emission enhancement should reach a maximum at some optimum $\beta_{p,\,opt}'$.

The overall SPP enhancement of the radiative efficiency $F_{SPP} = \eta_{SPP}/\eta_{rad}$ given by Eq. (1) can be calculated as a function of the SPP propagation constant β_p', and the results are shown in Fig. 6 for a wide range of the original radiative efficiency of the emitters in the dielectric under the assumption of escape of all radiation modes ($\theta_{\max} = \pi/2$), which clearly experiences a peak overall enhancement at some optimal values $\beta_{p,\,opt}'$.

Finally, our main results of the optimal overall enhancement factor which is calculated as $F_{SPP,\,opt} = \eta_{SPP}\left(\theta_{\max},\,\beta_{p,\,opt}'\right)/\eta_{rad}\,(1-\cos\theta_{\max})$ at optimal values of the SPP modes $\beta_{p,\,opt}'$ for a given original radiative efficiency η_{rad} with 1-D and 2-D gratings are shown in Fig. 7 for the three different cases of collection methods that limit the maximum polar angle $\theta_{\max}$. Assuming that a rear reflector is used in the absence of the metal layer, for a fair comparison between the different collection methods, we have included a factor $\Omega/2\pi = 1-\cos\theta_{\max}$ in the original radiative efficiency

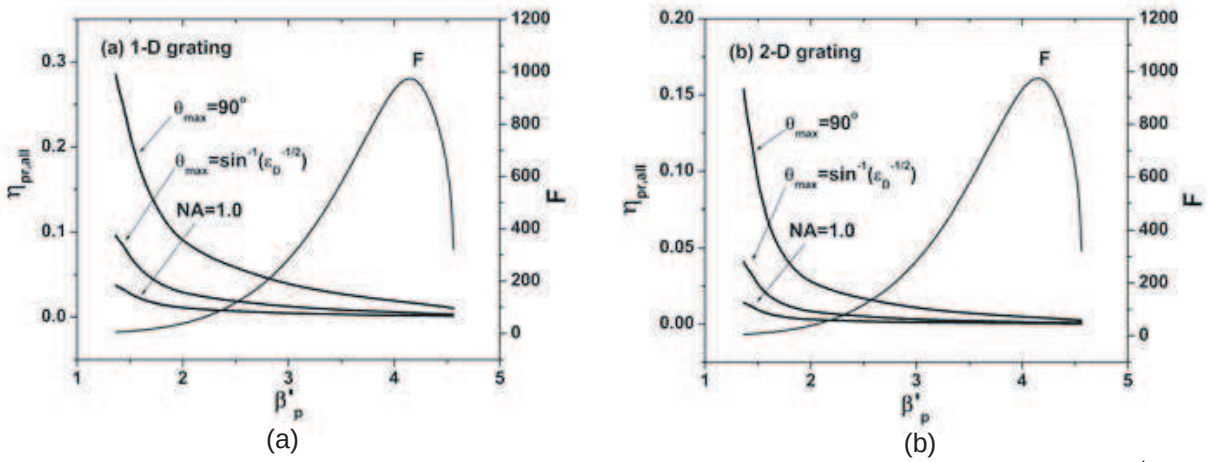


Figure 5: Integrated SPP coupling efficiency by (a) 1-D and (b) 2-D grating as a function of $\beta_p^{'}$ with $\theta_{\max} = \pi/2$, $\theta_{\max} = \sin^{-1}\left(1/\sqrt{\varepsilon_D}\right)$, and $NA = 1$.

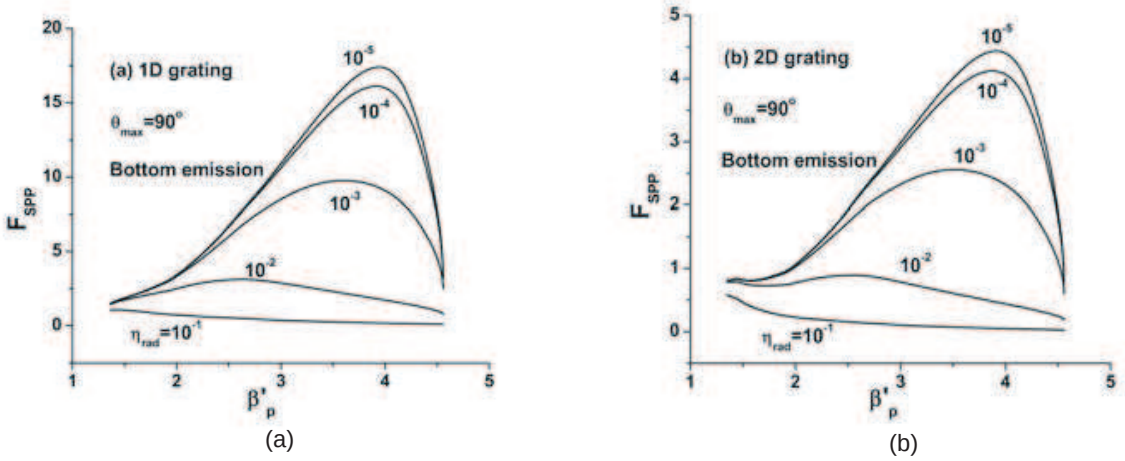


Figure 6: The overall enhancement factor by (a) 1-D and (b) 2-D grating.

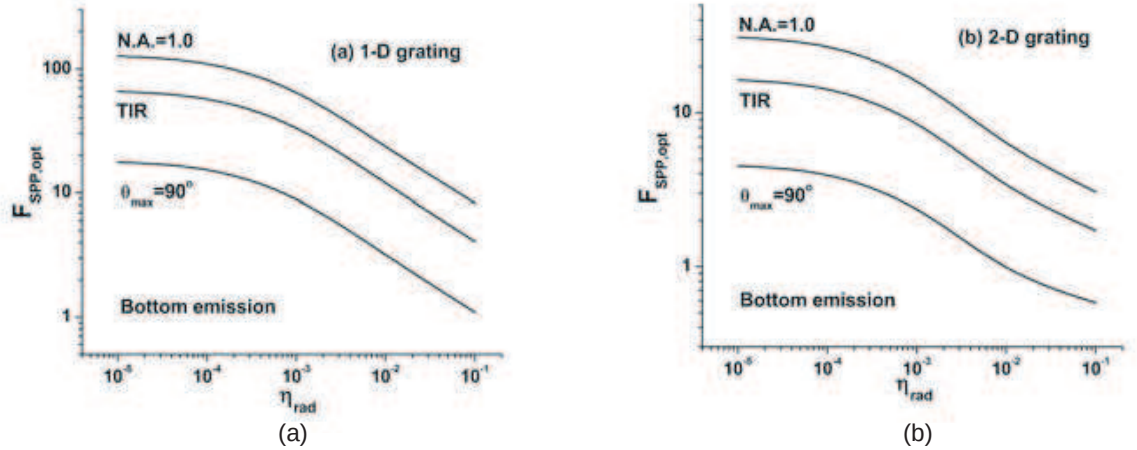


Figure 7: Overall optimal SPP enhancement for bottom emission as a function of the original radiative efficiency of the emitter at optimal value $\beta_{p,\,opt}^{'}$.

η_{rad}, where Ω is the solid angle of the collection cone for either the total internal reflection or lens collection (Fig. 4(b)). It can be seen that the overall enhancement factor is stronger for those emitters with smaller original radiative efficiency, and for those collection methods with smaller collection cones, and the 1-D grating always also couples more of the SPP modes into the radiation modes than the 2-D grating.

5. ANALYSIS AND CONCLUSIONS

Our results shown in Fig. 7 allow us to make a few important observations. It turns out that 2-D gratings that are more difficult to fabricate offer no improvement whatsoever over simple 1-D gratings. This rather welcome result can be traced to the fact that the first order Fourier coefficients in 2-D gratings are much smaller than in 1-D gratings. Substantial SPP enhancement of the overall radiative efficiency can be attained only for the emitters that originally have very low radiative efficiency. Assuming that one can deal with total internal reflection by, say, using surface roughening or having a dome-shaped lens attached to the emitter, the SPP enhancement can be attained only for the emitters with original radiative efficiency η_{rad} of less than 10% and can never be higher than about 20. If, on the other hand total internal reflection remains a factor, the SPP can be more helpful, as it can re-direct the emission towards the normal incidence. Then, for instance, the enhancement by a factor of 5-10 can be attained even if the original efficiency η_{rad} was 5%. This result is in excellent agreement with the experimental results of Ref. 3 for the same InGaN-Ag material system as used in our calculations. We shall therefore make a final conclusion - SPP enhancement of spontaneous radiation is most noticeable only if the original radiative efficiency of the emitter is very small, far less than 0.1%. Even then, the SPP enhancement is not substantially higher than 10 fold. For this reason it does not appear that SPP offers any advantage for light emitting diodes, with the only possibly exception - Si emitters whose original radiative efficiency is very low. The main application of SPP enhancement should remain as improving the efficiency of weak photoluminescence and nonlinear processes for the purpose of detecting small amounts of different substances.

ACKNOWLEDGMENT

REFERENCES

1. Contijo, I., M. Boroditsky, E. Yablonovich, S. Keller, U. K. Mishra, and S. P. DenBaars, *Phys. Rev. B*, Vol. 60, 11564, 1999.
2. Neogi, A., C. W. Lee, H. O. Everitt, T. Kuroda, A. Takemuchi, and E. Yablonovich, *Phys. Rev. B*, Vol. 66, 153305, 2002.
3. Okamoto, K., I. Niki, A. Shvartser, Y. Narukawa, T. Mukai, and A. Scherer, *Nat. Mater.*, Vol. 3, 601, 2004.
4. Okamoto, K., I. Niki, A. Scherer, Y. Narukawa, T. Mukai, and Y. Kawakami, *Appl. Phys. Lett.*, Vol. 87, 071102, 2005.
5. Neal, T. D., K. Okamoto, and A. Scherer, *Opt. Express*, Vol. 13, 5522, 2005.
6. Feng, J., T. Okamoto, and S. Kawata, *Opt. Lett.*, Vol. 30, 2302, 2005.
7. Wedge, S., J. A. E. Wasey, W. L. Barnes, and I. Sage, *Appl. Phys. Lett.*, Vol. 85, 182, 2004.
8. Pillai, S., K. R. Catchpole, T. Trupke, G. Zhang, J. Zhao, and M. A. Green, *Appl. Phys. Lett.*, Vol. 88, 161102, 2006.
9. Worthing, P. T. and W. L. Barnes, *Appl. Phys. Lett.*, Vol. 79, 3035, 2001.
10. Vuckovic, J., M. Loncar, and A. Scherer, *IEEE J. Quantum Electron.*, Vol. 36, 1131, 2000.
11. Barnes, W. L., *IEEE J. Lightwave Technol.*, Vol. 17, 2170, 1999.
12. Paiella, R., *Appl. Phys. Lett.*, Vol. 87, 111104, 2005.
13. Sun, G., J. B. Khurgin, and R. A. Soref, *Appl. Phys. Lett.*, Vol. 90, 111107, 2007.
14. Purcell, E. M., *Phys. Rev.*, Vol. 69, 681, 1946.
15. Johnson, P. B. and R. W. Christy, *Phys. Rev. B*, Vol. 6, 4370, 1972.
16. Ejder, E., *Phys. Stat. Sol. (a)*, Vol. 6, 445, 1971.

Metamaterial Techniques for Automotive Applications

K. Sato, T. Nomura, S. Matsuzawa, and H. Iizuka
Toyota Central Research & Development Labs., Inc., Japan

Abstract— Automotive applications that metamaterials are expected to effectively contribute to are presented, and research activities of metamaterials undertaken at Toyota Central R&D Labs (TCRL) are reviewed. They include development of leaky-wave antennas for future millimeter-wave radar systems, dipoles for UHF band applications, and topology design optimization techniques for periodic structures of metamaterials.

1. INTRODUCTION

Metamaterials are artificially constructed materials that have unusual electromagnetic properties such as backward wave, reduced wavelength with decreasing frequency, nonlinear frequency characteristic of resonances, and so on. In the view point of engineering, metamaterials having such unusual electromagnetic properties are expected to extend significantly the design degrees of freedom for materials, devices, components and systems. Significant research efforts have been expended in the development of microwave and millimeter-wave metamaterials, such as couplers, resonators, small antennas, and beam-scanned leaky wave antennas [1–3]. In addition, there are currently considerable interests in the development of optical metamaterials such as a negative index planar lens or "superlens" [4, 5]. Metamaterials are expected to provide new applications, drastic improvement of performance, simple architectures, low cost, and so on, in future automotive electronics applications, as shown in Fig. 1. The automotive applications that metamaterials are expected to effectively contribute to include beam scanning antenna systems for radars, mobile communication antennas, novel magnetic materials for electric motors, the high-performance absorbing and shielding materials for electromagnetic compatibility. Metamaterials are also expected to be applied to optical devices such as LED headlights and night vision systems using infrared cameras. Research activities undertaken at Toyota Central R&D Labs (TCRL) are reviewed in this paper. Two types of metamaterial-based antennas are presented. One is a leaky-wave antenna for future millimeter-wave radar systems that need wide scanning angle with simple architecture. The other is a dipole for UHF band applications. The dipole provides small size, or opposite polarization to conventional one. Also, design techniques are desired to maximize the performance of metamaterials. A topology design optimization technique for periodic structures of metamaterials is presented with an example model.

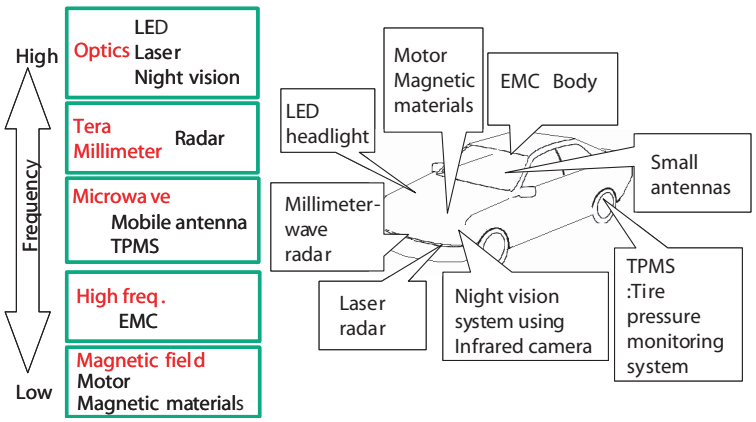


Figure 1: Automotive applications metamaterials are expected to effectively contribute to.

2. TOPOLOGY DESIGN OPTIMIZATION

Topology optimization is a highly flexible optimization method that can simultaneously deal with geometric and topological configuration changes [6]. Topology optimization is being used to develop a novel method for designing the periodic microstructures of electromagnetic materials [7]. The

design algorithm used for the variable change method is based on the density method. The main feature involves representing the shape of a structure by the density of its micropores to allow for the free transformation of the topology of the shape. The density of the material is translated to gradual changes in the physical properties such as the stiffness, weight or electromagnetic permittivity. Then, the density distribution is optimized for a desired specification by applying a mathematical non-linear programming technique. This approach is very attractive because of its simplicity and efficiency.

An example of a periodic electromagnetic band-gap dielectric material designed using topology optimization is shown in Fig. 2. The initial design is of a homogenous material in which the relative permittivity, ε_r, is 5. Fig. 2 shows that the microstructure is generated as the iteration step progresses and finally either a void or a filled/solid material is produced. In this manner, topology optimization can be used to produce new, optimized metamaterial microstructures with a specific band-gap.

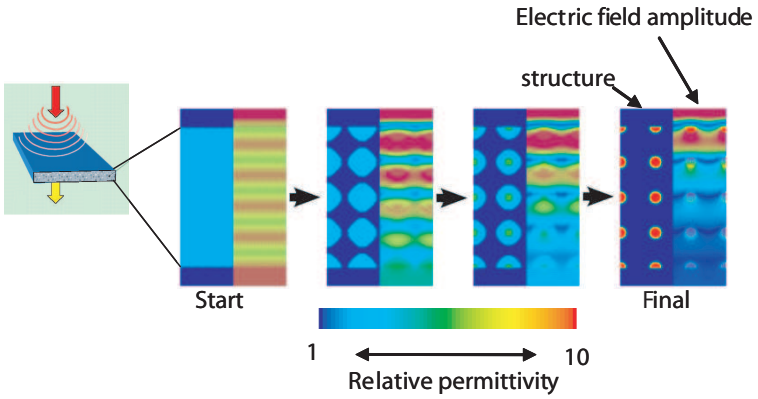
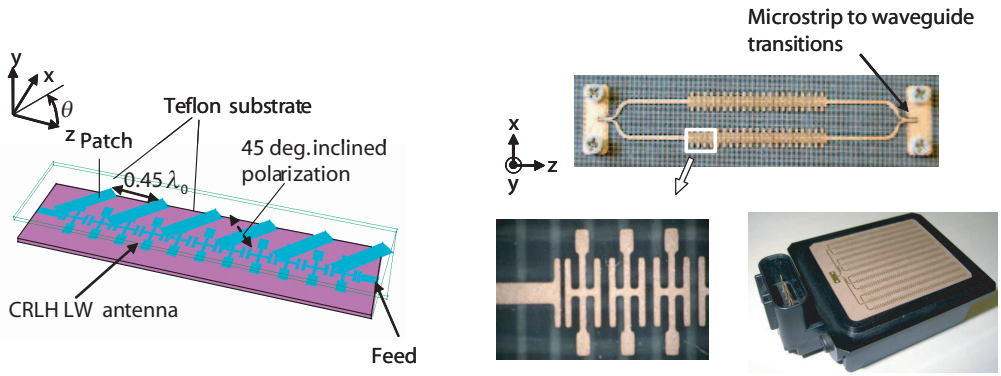


Figure 2: Example of electromagnetic band-gap structure using topology optimization design.

3. COMPOSITE RIGHT/LEFT-HANDED LEAKY WAVE ANTENNA (CRLH LWA)

There is a greatly increased interest in the development of automotive radar sensors for adaptive cruise control and pre-crash safety systems using a millimeter-wave band from 76 to 77 GHz. For these systems, a field of view (FOV) covering about 20° over a range of 150 m is sufficient and can be provided by most sensors on the market today. However, new developments like "stop & go" adaptive cruise control and collision avoidance assistance systems require a broader FOV up to 60°, over a maximum range of 60 m in order to adequately deal with cut-in situations. A novel structure for a frequency-independent steerable composite right/left handed leaky wave antenna for the millimeter-wave band applications is presented [8–9]. This antenna offers the advantages of wide beam scanning, high gain and a simple structure in the millimeter-wave band. The prototype CRLH LWA shown in Fig. 3 was fabricated and tested in the millimeter-wave band. In the near



(a) Configuration (b) Prototype

Figure 3: Millimeter-wave leaky wave antenna.

future, it may be possible to realize automotive radar antenna systems with a high gain in excess of 20 dBi by using the proposed antenna. We believe that the LHLWA is a promising design for automotive millimeter-wave applications.

4. LEFT-HANDED DIPOLE ANTENNAS

A new concept for forming a dipole antenna using a left-handed transmission line is next described [10]. The antenna is composed of a ladder network periodic structure of unit cells. The unit cell has a shunt inductor and two serious capacitors. Adding capacitors to one side of the network leads to out-of-phase currents with different amplitudes that produce high levels of radiation. The antenna has a unique feature in that the wavelength decreases with the frequency. The concept is applied to two antennas. The first is a small dipole, shown in Fig. 4(a). The straight dipole worked in the fundamental mode. The antenna of 0.18 wavelengths in free space provided a gain of -3.9 dBi at 547 MHz and bandwidth of 1.7% for $|S11| < -10$ dB. The second is an orthogonally polarized dipole, shown in Fig. 4(b). The meandered dipole worked in the higher order mode. Polarization orthogonal to a right-handed one was achieved at 643 MHz by the induced current of nine half wavelengths on the meander having 0.77 wavelengths in free space. These novel dipoles offer a great promise for future automotive mobile communications.

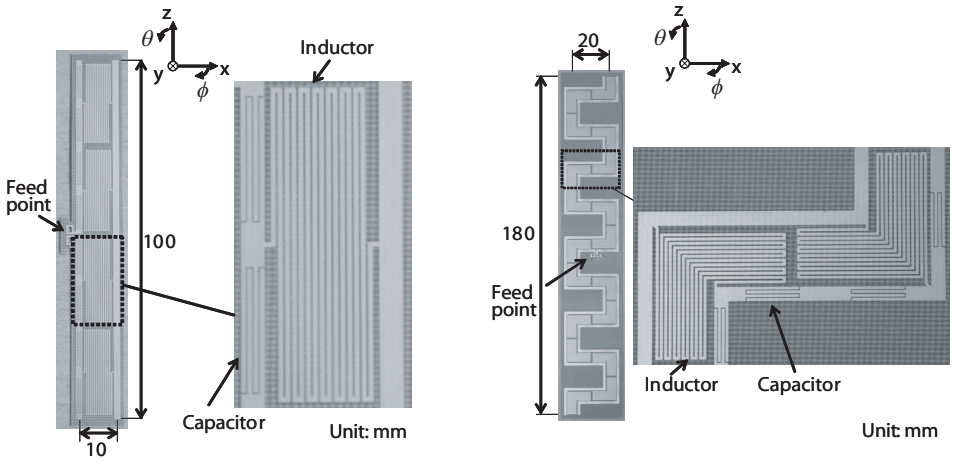


Figure 4: Left-handed dipole antennas.

5. CONCLUSIONS

A topology design optimization technique for electromagnetic materials, a left-handed leaky wave antenna for millimeter-wave applications, and left-handed dipole antennas have been presented along with some thoughts for future investigation. These have been the subjects of recent metamaterial studies at TCRL. Metamaterials will clearly open up a whole new field for automotive electronics applications.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to research collaborators.

REFERENCES

1. Caloz, C. and T. Itoh, *Electromagnetic Metamaterials*, IEEE, John Wiley and Sons, Inc., 2006.
2. Eleftheriades, V. and G. Balmain, *Negative-refraction Metamaterials*, IEEE, John Wiley and Sons., 2005.
3. Smith, D., et al., *Phys Rev. Lett.*, 84–18, 4184, 2000.
4. Pendry, J., *Phys. Rev. Lett.*, 85–18, 3966, 2000.
5. Engheta, N., et al., *Phys. Rev. Lett.*, 95, 095504, 2005.
6. Bendsøe, P. and N. Kikuchi, *Computer Methods in Applied Mechanics and Engineering*, 71, 197–224, 1988.

7. Nomura, T., et al., "Topology optimization of periodic microstructures in electromagnetic material." *Proc. 6th World Congress of Structural and Multidisciplinay Optimization*, Rio de Janeiro, Brazil, 2005.
8. Sato, K., et al., *Proc. Int. Workshop on Antenna Tech.*, New York, 420–424, 2006.
9. Matsuzawa, S., et al., *IEICE Trans Commun.*, Vol. E89-C, No. 9, 1337–1344, 2006.
10. Iizuka, H. and P. S. Hall, *IEEE Trans. Antennas Propag.*, Vol. 55, No. 5, 1246–1253, May 2007.

Force Constants and Dispersion Relations in GaN

D. G. Santiago-Pérez[1], **F. de León-Pérez**[2], **M. E. Mora-Ramos**[3], **and R. Pérez-Álvarez**[3]

[1]Centro Universitario José Martí Pérez, Avenida de los Mártires 360, Sancti Spiritus, Cuba
[2]Departamento de Física de la Materia Condensada, Universidad de Zaragoza, E-50009 Zaragoza, Spain
[3]Universidad Autónoma del Estado de Morelos, Ave. Universidad 1001, 62209 Cuernavaca, México

Abstract— We study the lattice dynamics of the wurtzite phase of bulk group-III nitride using the linear chains models. It is shown that the 3D oscillation problem reduces to uncoupled linear chain equations for high symmetry directions. The relation between the 3D and 1D force constants has been investigated. Our study helps to understand better the richness of the linear chain models. We also show how to fit the 3D bulk force constants (and consequently the whole dynamical matrix) from a few points, either experimental or theoretical.

1. INTRODUCTION

Many properties of solids depend on the dynamics of atoms in the lattice. Some problems demand an appreciable knowledge of the bulk atomic oscillations. Examples are the phonon imaging [1, 2] and the reduction of the thermal conductivity in superlattices in comparison to the bulk materials [3–5].

The three-dimensional (3D) equations of motion for the atoms in the lattice decouple along high symmetry directions resulting in one longitudinal and two transverse oscillations which are described by linear chain models (see [6] and references therein). The values of the bulk force constants are estimated either from experimental dispersion relations or from theoretical calculations, and linear chain models are employed to study the phonon modes. So far, this analysis has been carried out for zincblende and diamond crystal symmetries. However, such an instance is of particular interest in semiconductors with wurtzite structure as is the case of the group-III nitrides. This exact result for the bulk is useful for the study of heterostructures. In fact, for heterostructures grown along high symmetry directions, it is usually assumed that the force constants in each constituent layer are equal to the bulk force constant. In this way the phonon equations of motions are obtained.

In this paper the phonon equations of motion for bulk semiconductors are revisited. As group-III nitride are ionic materials we can separate the contribution of the macroscopic field. Moreover, instead of finding the irreducible representation for a given direction, we consider a given number of atoms and assume harmonic interaction between a limited number of neighbors for interaction of microscopic character [6]. Hence, the dynamical matrix is constructed, taking into account the symmetry of the underlying lattice. Interesting properties of the equations are found in this way. In particular, we check explicitly how the 3D problem reduces to decoupled linear chain equations for high symmetry directions, with the aim of obtaining the relation between the 3D and one-dimensional (1D) force constants. To the best of our knowledge, this relation has not been explored so far for the case of the wurtzite symmetry. We also show how to fit the 3D bulk force constants (and consequently the whole dynamical matrix) from a few points, either experimental or theoretical. We focus our study on wurtzite structure, and particularly on GaN.

The paper is organized as follows: In the next section we enumerate the properties of the equations of motion that are needed in this paper. For details the reader can address to Ref. [7] where a similar calculation are done for the diamond and zincblence structures. In Section 3 we consider the wurtzite lattice. At the end, our main conclusions are summarized.

2. EQUATION OF MOTION AND DYNAMICAL MATRIX

The ionic character of these materials requires the inclusion of the macroscopic electric field into the equations of motion in the reciprocal lattice [8], which are then written as

$$\omega^2 e_\alpha(\kappa) = \sum_{\kappa'\beta} D_{\alpha\beta}\left(\kappa\kappa', \overrightarrow{q}\right) e_\beta\left(\kappa'\right) - \frac{Q_k}{\sqrt{m_k}} E_\alpha. \tag{1}$$

ω is the frequency of the mode, Q_k the amount of charge shared by the κ-atom within the bond. m_k is the mass of the κ-atom, and E_α is the α-component of the electric field given by

$$E_\alpha = -\frac{1}{v\varepsilon_{\shortparallel}} \left(\frac{q_\alpha}{|\overrightarrow{q}|}\right) \sum_\beta \left(\frac{q_\beta}{|\overrightarrow{q}|}\right) \sum_{\kappa'} \frac{Q_{\kappa'}}{\sqrt{m_{\kappa'}}} e_\beta\left(\kappa'\right). \tag{2}$$

Here, υ is the unit cell volume, $\varepsilon_{\shortparallel}$ is the electrical permittivity in such direction, and $\overrightarrow{q}$ is the wavevector. Additionally, $e_\alpha(\kappa)$ is the polarization vector and the dynamical matrix D is given by

$$D_{\alpha\beta}\left(\kappa\kappa', \overrightarrow{q}\right) = \frac{1}{\sqrt{M_\kappa M_{\kappa'}}} \sum_{l'} \Phi_{\alpha\beta}\left(l\kappa, l'\kappa'\right) \exp\left(-i\overrightarrow{q} \cdot \left[\overrightarrow{x}(l) - \overrightarrow{x}(l')\right]\right). \tag{3}$$

$\mathbf{x}(l)$ is the vector position of the elementary cell and $\Phi_{\alpha\beta}\left(l\kappa, l'\kappa'\right)$ are the force constants.

The invariance of the force constants under a symmetry operation S, represented by a unitary matrix, is written in matrix form as $S\Phi S^\dagger = \Phi$. The dagger ($\dagger$) means the hermitian conjugate. From these symmetry relations the specific form of the dynamical matrix can be established. Further details of this approach are discussed in Refs. [8, 9]. In particular several useful symmetry properties of $\Phi_{\alpha,\beta}$ are discussed in these books.

2.1. The Wurtzite Structure

The most common allotropic form of the nitride semiconductors is the wurtzite. This is a close-packed hexagonal structure with 4 atoms per unit cell (two atoms of a group III cation, and two atoms of N). Using the symmetry of the cell it is possible to keep only two equations of motion-one for the cation X (located at $(0, 0, 0)$) and one for the atom of N (located at $(0, 0, uc)$). The equations will be written in the second nearest neighbor approximation. The first four neighbors are invariant under the point group C_3^{yz}. For the transformations we have the following representations

$$C_3 = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \sigma^{yz} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The invariance of the force matrices under these transformations means that

$$\sigma^{yz} C_3 \Phi (\sigma^{yz} C_3)^t = \Phi, \tag{4}$$

from which the first neighbor, l_0, force-constant matrix is obtained

$$\Phi_{l_0} = \begin{pmatrix} \alpha_{11} & 0 & 0 \\ 0 & \alpha_{11} & 0 \\ 0 & 0 & \alpha_{33} \end{pmatrix}. \tag{5}$$

Keeping in mind the symmetry of each atomic position, the particular form of the force-constant matrix corresponding to each neighbor is analogously obtained. The dynamical matrix; that is, the equations of motion in each symmetry direction of the reciprocal lattice, is then derived.

Table 1: The relations between the 1D force constants and the 3D force constant matrix elements.

L Mode		T mode	
Linear Chain	3D Parameters	Linear Chain	3D Parameters
γ_{cc}	$-3\beta_{33}$	γ_{cc}	$-\frac{3}{2}\left(\beta_{11}+\beta_{22}\right)$
γ_{aa}	$-3\gamma_{33}$	γ_{aa}	$-\frac{3}{2}\left(\gamma_{11}+\gamma_{22}\right)$
γ_{ca}	$-\frac{1}{3}\left(8\alpha_{11}+\alpha_{33}\right)-\Omega$	γ_{ca}	$-\frac{1}{3}(5\alpha_{11}+4\alpha_{33})$
γ_{ca1}	$-\alpha_{33}$	γ_{ca1}	$-\alpha_{11}$

3. NUMERICAL RESULTS

The longitudinal and the (twofold degenerated) transversal oscillation modes uncouple. The corresponding equations of motion can be studied as those of linear chains of atoms connected by springs. In particular, for $\overrightarrow{q} = (0, 0, q)$ the symmetry of the atoms in the unit cell makes enough to consider a diatomic linear chain. Consequently, the four oscillation branches are obtained in both the longitudinal and transversal polarizations.

The dispersion relations in the [0001] are given by the expressions

$$\begin{aligned}
\omega_1[q] &= \sqrt{\frac{\Phi_1[q]+\Psi_1[q]}{2} - \sqrt{\frac{(\Phi_1[q]-\Psi_1[q])^2}{2} + \Upsilon_1[q]}} \\
\omega_2[q] &= \sqrt{\frac{\Phi_2[q]+\Psi_2[q]}{2} - \sqrt{\frac{(\Phi_2[q]-\Psi_2[q])^2}{2} + \Upsilon_2[q]}} \\
\omega_3[q] &= \sqrt{\frac{\Phi_2[q]+\Psi_2[q]}{2} + \sqrt{\frac{(\Phi_2[q]-\Psi_2[q])^2}{2} + \Upsilon_2[q]}} \\
\omega_4[q] &= \sqrt{\frac{\Phi_1[q]+\Psi_1[q]}{2} + \sqrt{\frac{(\Phi_1[q]-\Psi_1[q])^2}{2} + \Upsilon_1[q]}}
\end{aligned} \tag{6}$$

where,

$$\begin{aligned}
\Phi_1[q] &= \tfrac{1}{M_X}\left(\gamma_{ca} + \gamma_{ca1} + 4\gamma_{cc}\sin^2[\tfrac{q\pi}{2}]\right) \\
\Psi_1[q] &= \tfrac{1}{M_N}\left(\gamma_{ca} + \gamma_{ca1} + 4\gamma_{aa}\sin^2[\tfrac{q\pi}{2}]\right) \\
\Upsilon_1[q] &= \tfrac{1}{M_X M_N}\left(\gamma_{ca}^2 + \gamma_{ca1}^2 + 2\gamma_{ca}\gamma_{ca1}\cos[q\pi]\right) \\
\Phi_2[q] &= \tfrac{1}{M_X}\left(\gamma_{ca} + \gamma_{ca1} + 4\gamma_{cc}\cos^2[\tfrac{q\pi}{2}]\right) \\
\Psi_2[q] &= \tfrac{1}{M_N}\left(\gamma_{ca} + \gamma_{ca1} + 4\gamma_{aa}\cos^2[\tfrac{q\pi}{2}]\right) \\
\Upsilon_2[q] &= \tfrac{1}{M_X M_N}\left(\gamma_{ca}^2 + \gamma_{ca1}^2 - 2\gamma_{ca}\gamma_{ca1}\cos[q\pi]\right)
\end{aligned} \tag{7}$$

The dispersion functions are written in terms of the force constant matrix elements through the relations shown in Table 1. These are obtained by comparing the linear chain equations with the uncoupled 3D ones.

Table 2: The GaN 3D force constants and their 1D equivalent for linear chains (in 10^5 din.cm^{-1}).

3D Problem		Linear Chain		
α_{11}	7.78	$T[0001]$	γ_{ca}	296.01
α_{33}	-474.42		γ_{ca1}	-7.78
$\beta_{11}+\beta_{22}$	-87.42		γ_{aa}	131.14
$\gamma_{11}+\gamma_{22}$	25.71		γ_{cc}	-38.56
β_{33}	-368.25	$L[0001]$	γ_{ca}	137.40
γ_{33}	57.85		γ_{ca1}	474.42
Ω	323.59		γ_{aa}	65.80
			γ_{cc}	35.49

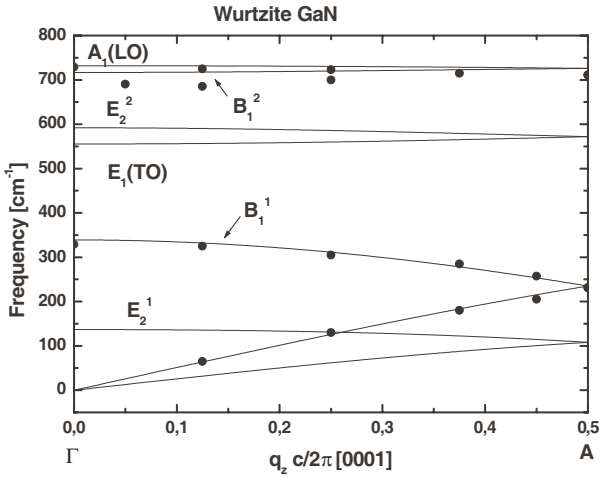


Figure 1: Calculated dispersion relations of the phonon modes along $\Gamma - A$ direction in the Brillouin zone of wurtzite GaN. The filled circles are measured phonon inelastic x-rays scattering data [11].

The dispersion relations in the 3D problem depend upon seven independent parameters: α_{11}, α_{33}, $\beta_{11}+\beta_{22}$, β_{33}, $\gamma_{11}+\gamma_{22}$, γ_{33}, and Ω. For a given material, those functions are evaluated in certain points in which the phonon frequencies of the longitudinal and transversal modes are already known. In GaN, these are the center and boundaries of the Brillouin zone [10]. Hence, one obtains eight independent equations. The parameters α_{11}, α_{33}, $\beta_{11}+\beta_{22}$ and $\gamma_{11}+\gamma_{22}$ are obtained by solving the equations corresponding to the transversal modes. Then, using these values, it is possible to calculate the remaining ones (β_{33}, γ_{33} and Ω), with the help of the equations associated with the longitudinal oscillations. Table 2 shows the numerical values. Substituting these values in (7) analytic expression for the dispersion relations are obtained. Figure 1 shows the obtained dispersion relations. It must be stressed that these curves resemble very closely the results for this direction given in Ref. [10].

4. CONCLUSION

We have obtained — from the 3D bulk problem — the kind of linear chain equations that should be used in order to obtain the phonon dispersion relations along high symmetry directions, for a given number of neighbors in wurtzite polar crystals. In particular, we explain how to take into account the macroscopic electric field, in order to separate the L and T oscillation branches. The most relevant aspect is that we derive analytic expressions for these dispersion relations, which can be of usefulness for the problems that are mentioned in the introduction. The dispersion relations obtained for the GaN in this way reproduce with high accuracy previous experimental and theoretical results, but with a significant simplicity.

ACKNOWLEDGMENT

We thank Professor V. Velasco (CSIC, Spain) for useful comments. M. E. M-R. acknowledges support from CONACyT through Grant 52304.

REFERENCES

1. Narayamurti, V., H. L. Störmer, M. A. Chin, A. C. Gossard, and W. Wiegmann, "Selective transmission of high-frequency phonons by a superlattice: The 'dielectric' phonon filter," *Phys. Rev. Lett.*, Vol. 43, No. 27, 2012–2016, 1979.
2. Tamura, S., "Large-wave-vector phonons in highly dispersive crystals: Phonon-focusing effects," *Phys. Rev. B*, Vol. 28, No. 2, 897–909, 1983.
3. Yao, T., "Thermal properties of AlAs/GaAs superlattices," *Appl. Phys. Lett.*, Vol. 51, No. 22, 1798–1800, 1987.
4. Mahan, G. D. and H. B. Lyon, Jr., "Thermoelectric devices using semiconductor quantum wells," *J. Appl. Phys.*, Vol. 76, No. 3, 1899–1901, 1994.
5. Hyldgaard, P. and G. D. Mahan, "Phonon superlattice transport," *Phys. Rev. B*, Vol. 56, No. 17, 10754–10757, 1997.
6. Fasolino, A., E. Molinari, and K. Kunc, "Planar force-constant method for lattice dynamics of structures," *Phys. Rev. B*, Vol. 41, No. 12, 8302–8312, 1990.
7. Santiago-Pérez, D. G., F. de León-Pérez, and R. Pérez-Álvarez, "Force constants and dispersion relations for the zincblende and diamond structures revisited," *Rev. Mex Fis.*, Vol. 52, No. 2, 163–171, 2006.
8. Born, M. and K. Huang, *Dynamical Theory of Crystal Lattices*, Oxford, ClarendonPress, 1954.
9. Ashcroft, N. W. and N. D. Mermin, *Solid State Physics*, Saunders College, Philadelphia, 1976.
10. Tütüncü, H. M. and G. B. Srivastava, "Phonons in zinc-blende and wurtzite phases of GaN, AlN, and BN with the adiabatic bond-charge model," *Phys. Rev. B*, Vol. 62, No. 8, 5028–5035, 2000.
11. Ruf, T., J. Serrano, and M.Cardona, "Phonon dispersion curves in Wurtzite-Structure GaN determined by inelastic X-ray scattering," *Phys. Rev. Lett*, Vol. 86, No. 5, 906–909, 2001.

Longwave Phonon Tunnelling Using an Impedance Concept

D. Villegas[1], **F. de León-Pérez**[2], **and R. Pérez-Álvarez**[3]

[1]Physics Department, Central University Marta Abreu of Las Villas, Santa Clara, Cuba
[2]Departamento de Física de la Materia Condensada, Universidad de Zaragoza
E-50009 Zaragoza, Spain
[3]Universidad Autónoma del Estado de Morelos
Ave. Universidad 1001, 62209 Cuernavaca, México

Abstract— In this paper we study numerically the transmission, reflection and dwell times of phonons propagating through semiconductor multilayer structures. The tunnelling of optical and acoustic phonon at normal incidence on multiple layers systems is analyzed. We adopt the continuum model valid for long-wavelength oscillations. The isomorphism between acoustical and optical propagation in heterostructures, and electromagnetic waves in transmission lines, is analyzed. We use an impedance concept similar to the quantum-mechanical one introduced by Khondker to calculate the transmission and reflection coefficients. This study could also be useful for the design of phonons filters.

1. INTRODUCTION

The tunnelling of long-wavelength phonons through semiconductor heterostucture has been studied with extent. To describe long-wave modes, we employ a continuum phenomenological model, which has proved useful in both polar [1] and nonpolar [2] structures.

The long-wavelength phonons propagating along high symmetry directions in nonpolar semiconductors are described by a simple one-dimensional differential equation [2]. In the present paper the isomorphism between the equations that describe wave propagation in heterostructures, and electromagnetic waves in transmission lines, is used to define an impedance similar to the quantummechanical impedance one introduced by Khondker [3].

In addition, the tunnelling of optical and acoustic phonon at normal incidence on multiple layers systems is analyzed. The wave impedance allows us to calculate the transmission and reflection coefficients and study numerically the transmission, reflection and dwell times of phonon packets propagating through semiconductor multilayer structures [4, 5].

The generalized concept of quantum-mechanical wave impedance has been formerly used to design an energy band-pass filter [7]. We hope that the present work could encourage further experimental studies and possible application such as gaussian filter for phonons [6].

The paper is organized as follows. In the next section is described our model. Section 3 is devoted to analyse the wave transmission on a transmission line. In Section 4 we study the propagation of phonon waves. Finally we give some conclusions.

2. LONG WAVELENGTH PHONON MODEL

We formulate the problem with the help of its one-dimensional energy density [2]

$$\mathcal{H} = \frac{1}{2}\rho\left|\frac{\partial u}{\partial t}\right|^2 + \frac{1}{2}\rho\omega_0^2|u|^2 + \frac{1}{4}\left[\sigma\frac{\partial u}{\partial z}^* + \sigma^*\frac{\partial u}{\partial z}\right]. \tag{1}$$

The first term in (1) represents the kinetic energy density, the second one the interaction energy density of the phonon field with itself, and the third one the strain energy density that accounts for the dispersive character of the oscillations. These terms depend of the atomic relative displacements u, the linear mass density ρ, the phonon frequency at the center of the Brillouin zone ω_0, the one-dimensional strain tensor $\partial u/\partial z$, and the stress tensor σ, which is equal to $\sigma = -\rho\beta^2\partial u/\partial z$, where β is a parameter that accounts on the behavior of the bulk phonon dispersion relation.

From (1) it is easily obtained the one-dimensional equation of motion [2]

$$\frac{\partial^2 u}{\partial t^2} = -\omega_0^2 u - \beta^2\frac{\partial^2 u}{\partial z^2}, \tag{2}$$

and from Equation (2) it is obtained

$$\frac{\partial^2\sigma}{\partial t^2} = -\omega_0^2\sigma - \beta^2\frac{\partial^2\sigma}{\partial z^2}. \tag{3}$$

The energy density flux, j, is writen in [2] as

$$j = -1/2\left(\sigma\partial u^*/\partial t + \sigma^*\partial u/\partial t\right). \tag{4}$$

The stationary solution of (2) and (3) leads to a Helmholtz equation which is equivalent to both the time-independent Schrödinger equation and the wave propagation on a transmission line. In particular, these equations are

$$\frac{d^2u}{dz^2} + k^2u = 0; \qquad \frac{d^2\sigma}{dz^2} + k^2\sigma = 0. \tag{5}$$

In the frequency range of allowed optical oscillations, with $\omega < \omega_0$, the solution for the equations of motion are

$$u = A^+e^{ikz} + A^-e^{-ikz}; \qquad \sigma = A^+e^{ikz} + A^-e^{-ikz}, \tag{6}$$

where $k = \sqrt{(\omega_0^2 - \omega^2)}$.

For frequency $\omega > \omega_0$ the oscillations are forbidden. The solution for the equations of motion are

$$u = A^+e^{\kappa z} + A^-e^{-\kappa x}; \qquad \sigma = A^+e^{\kappa z} + A^-e^{\kappa z}, \tag{7}$$

where $\kappa = \sqrt{(\omega^2 - \omega_0^2)}$.

3. WAVE PROPAGATION ON A TRANSMISSION LINE

A transmission line is a distributed parameter network, where voltages and currents can vary in magnitude and phase over its length. From the Kirchhoff's voltage law and Kirchhoff's current law the wave equations are obtained. For the steady-state condition these equations are

$$\frac{d^2I}{dz^2} - \gamma^2I = 0; \qquad \frac{d^2V}{dz^2} - \gamma^2V = 0, \tag{8}$$

where $\gamma = \alpha + i\beta$ is the complex propagation constant, which is a function of the frequency.

Travelling wave solution to (8) and (8) can be found as

$$V = V^+e^{-\gamma z} + V^-e^{\gamma z}; \qquad I = \frac{V^+}{Z_0}e^{-\gamma z} + \frac{V^-}{Z_0}e^{\gamma z}, \tag{9}$$

where the $e^{-\gamma z}$ term represents wave propagation in the z direction, and the $e^{\gamma z}$ term represents wave propagation in the $-z$ direction. The characteristic impedance, Z_0, can be defined as $Z_0 = V^+/I^+ = -V^-/I^-$ and results a magnitude which depends of the systems's parameters.

4. LONG WAVELENGTH PHONON TUNNELLING

It is easy to see that the equations that describe the wave propagation on a transmission line, (8) and (8), and the equations that describe the propagation of long wavelength phonons through semiconductor heteroestucture, (5) and (5), are isomorphic. It is also valid for the solutions of these equations.

We consider in the following the phonons falling on a multiple-barrier structure consisting of an alternate stacking of different layers A and B sandwiched between other materials X and Y; X is the substrate of the superlattice and Y denotes the detector. Keeping in mind the suitable analogies, the studied heteroestucture can be represented as a transmission line of phonon waves. In order to introduce and define the concept of wave impedance we will consider a system formed by two half-infinite materials, namely A_1 and A_2, respectively. For this system, the Equations (6), (6), (7) and (7), can be rewritten as

$$u(z) = A^+\left(e^{\gamma z} - re^{\gamma - z}\right), \tag{10}$$

$$\sigma(z) = -\rho\beta^2\gamma A^+\left(e^{\gamma z} + re^{\gamma - z}\right), \tag{11}$$

where $\gamma = i\sqrt{\omega_0^2 - \omega^2}/\beta^2$, and $r = (\rho_2\beta_2^2\gamma_2 - \rho_1\beta_1^2\gamma_1)/(\rho_2\beta_2^2\gamma_2 - \rho_1\beta_1^2\gamma_1)$ is the wave amplitude reflection coefficient. Notice again the isomorphism between Equations (10) and (11) and the correspondig quantum mechanical wave equations [3].

By mean of definitions $\Phi(z) = i\sigma(z)$ and $U(z) = u(z)$, we can identify the characteristic impedance as $Z_0 = i\rho\beta^2\gamma$. Now, the wave impedance in any point on a transmission line is defined as $Z(z) = \Phi(z)/U(z)$ [3]. Then, it is straight to identify that input impedance Z_i is equal to

$$\frac{Z_i}{Z_0} = \frac{Z_T\cosh(\gamma d) - Z_0\sinh(\gamma d)}{Z_0\cosh(\gamma d) - Z_T\sinh(\gamma d)}. \tag{12}$$

If a lossless line of characteristic impedance Z_0 is terminated with a load impedance Z_T, the reflection coefficient at the load can be written as $\Gamma = (Z_T - Z_0)/(Z_T + Z_0)$. Then $r \equiv \Gamma$.

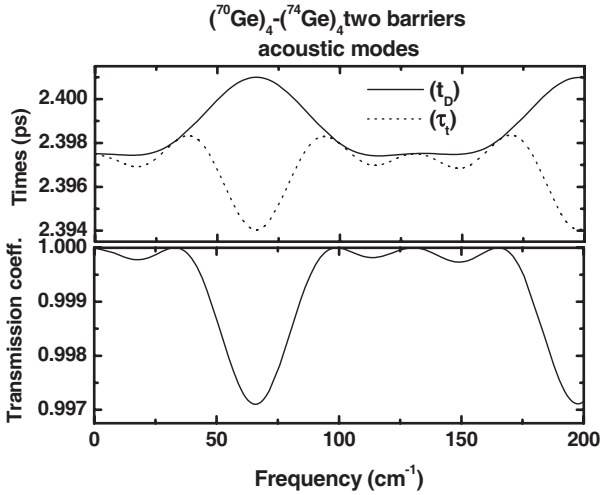


Figure 1: Upper panel: the dwell time (solid line) and the transmission time (dashed line), both in ps, are plotted as a function of the acoustic phonon frequencies (in cm^{-1}). Lower panel: the corresponding transmission coefficient.

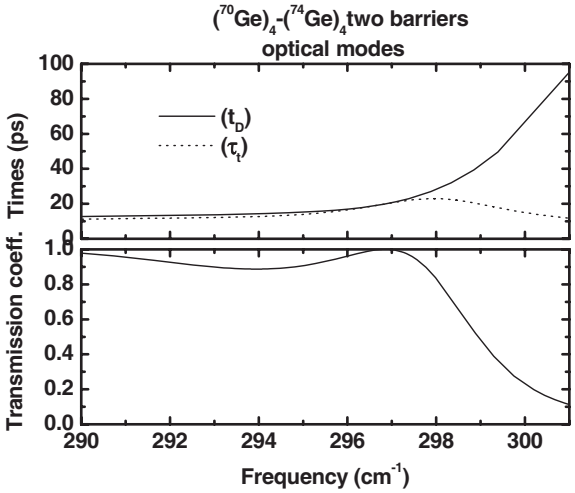


Figure 2: Upper panel: the dwell time (solid line) and the transmission time (dashed line), both in ps, are plotted as a function of the optical phonon frequencies (in cm^{-1}). Lower panel: the corresponding transmission coefficient.

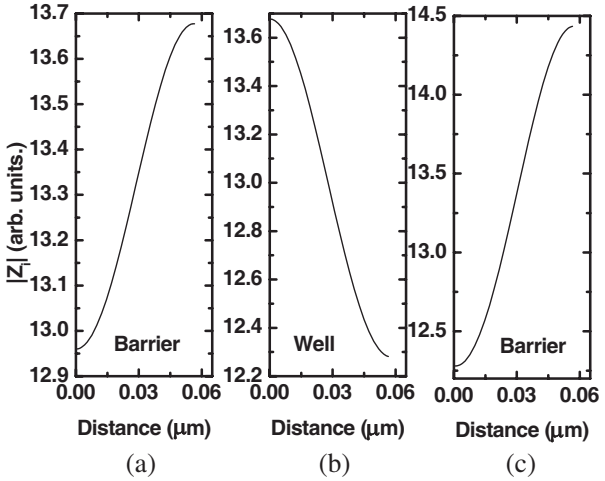


Figure 3: The input impedance (arb. units) is plotted as a function of the distance (in μm) for the barrier in the right end (a), the well in the center (b) and the barrier in the left end (c). A frequency of 65.41 cm^{-1} is used.

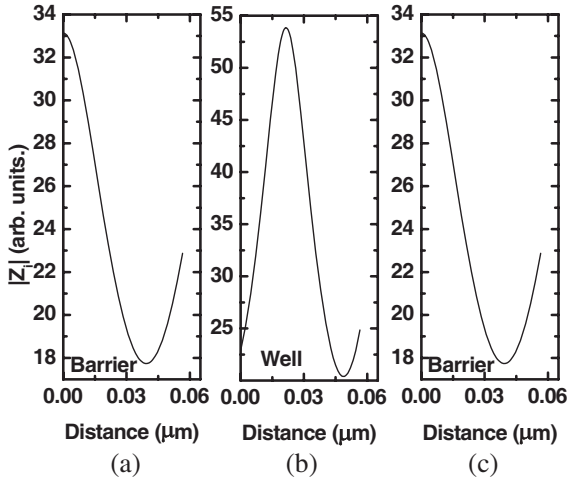


Figure 4: The input impedance (arb. units) is plotted as a function of the distance (in μm) for the barrier in the right end (a), the well in the center (b) and the barrier in the left end (c). A frequency of 294 cm^{-1} is used.

Notice that the stationary solutions of Equations (2) and (3) are $u(z,t) = U(z)e^{i\omega t}$ and $\sigma(z,t) = -i\Phi(z)e^{i\omega t}$; then with these solutions the Equations (1) and (4) are rewritten as

$$\mathcal{H}(z) = \frac{1}{2}\rho\omega^2 |U(z)|^2 + \frac{1}{2}\rho\omega_0^2 U(z)|^2 - \frac{1}{2}Re\left[i\Phi(z)\frac{dU(z)^*}{dz}\right], \quad (13)$$

$$j(z) = \frac{\omega}{2}Re\left[\Phi(z)U^*(z)\right]. \quad (14)$$

The transmission coefficient, transmission time and dwell time, already defined in [4] are calculated for double barriers structures made up of different isotopes of germanium. For further details of the calculation and the parameters employed the reader is addressed to references [2, 4–6]. In Figure 1 these magnitudes are plotted as a function of the phonon frequencies for acoustic modes, while in Figure 2 optical modes are considered. In the same way that in the tunnelling of electrons the maximum transmission is obtained when the load impedance is equal to the characteristic impedance ($\Gamma = r = 0$).

For the sake of completeness the input impedance as a function of the distance is plotted. See Figure 3 for acoustic modes and 4 for the optical ones.

5. CONCLUSION

In this paper we define a characteristic impedance that depends on the parameters of the materials and of the frequency. This impedance is related with the one-dimensional strain and stress tensors.

We took advantage of the isomorphism existing between problems in different fields of Physics in order to transfer concepts and results from one field to another one. It is found particularly useful the concept of impedance to describe tunnelling and reflection in multilayer systems. We have applied this concept to the phonon tunnelling and found the relationships with Elasticity magnitudes.

Given the mathematical analogies it is possible to formulate these problems on the basis of quadrupole theory. It would be useful to design devices like phonon filters.

ACKNOWLEDGMENT

We acknowledge our colleagues from UCLV, UZ and UAEM for many stimulating discussions on this subject.

REFERENCES

1. Trallero-Giner, C., R. Pérez-Alvarez, and F. García-Moliner, *Long Wave Polar Modes in Semiconductor Heterostructures*, Pergamon/Elsevier Science, London, 1998.
2. De León-Pérez, F. and R. Pérez-Álvarez, "Phonon propagation in nonpolar semiconductor heterostructures," *Phys. Rev. B*, Vol. 63, 245304, 2001.
3. Khondker, A. N., M. R. Khan, and A. F. M. Anwar, "Transmission line analogy of resonance tunneling phenomena: The generalized impedance concept," *J. Appl. Phys.*, Vol. 63, No. 10, 5191–5193, 1988.
4. Villegas, D., F. de León-Pérez, and R. Pérez-Álvarez, "Tunneling time of long-wavelength phonons through semiconductors heteroestructures," *Phys. Rev. B*, Vol. 71, 035322, 2005.
5. Villegas, D., F. de León-Pérez, and R. Pérez-Álvarez, "Tunneling time for phonons: Dependence on the system's size," *Phy. Stat. Sol. (b)*, Vol. 242, No. 9, 1767–1770, 2005.
6. Villegas, D., F. de León-Pérez, and R. Pérez-Álvarez, "Gaussian superlattice for phonons," *Microelectronics Journal*, Vol. 36, 411–412, 2005.
7. Tung, H. T. and C. P. Lee, "An energy band-pass filter using superlattice structures," *IEEE Journal of Quantum Electronics*, Vol. 32, No. 3, 507–512, 1996.

Cantor Dielectric Heterostructures Made of Nanostructured Multilayers of Porous Silicon

V. Agarwal[1], **B. Alvarado-Tenorio**[1], **J. Escorcia-Garcia**[1], **and L. M. Gaggero Sager**[2]

[1]Centro de Investigación en Ingeniería y Ciencias Aplicadas, Universidad Autónoma del Estado de Morelos
Av. Universidad 1001, Col. Chamilpa, CP 62209, Cuernavaca, Morelos, México
[2]Facultad de Ciencias, Universidad Autónoma del Estado de Morelos
Av. Universidad 1001, Col. Chamilpa, CP 62209, Cuernavaca, Morelos, Mexico

Abstract— The wave propagation in the nanostructured porous silicon multilayers, where the geometrical length follows the Cantor code, is presented. The total thickness of the multilayered structure is maintained as 9000 nm. The heterostrucutres were fabricated by the porous silicon layers having the refractive indices of 1.9 and 1.2 corresponding to the low and high porosity respectively. The thickness of the high and low porosity layer varied from 1000 to 12 nm for making the Cantor heterostructures up to 7th order. In the reflectance spectra of the Cantor structure of 6th order (63 layers), two major photonic band gaps (PBG) are observed in the visible region with a narrow resonance at 652 nm. In the Cantor structures, with order more than 6, instead of any PBG approximately equidistant fringes are observed.

1. INTRODUCTION

PS has been extensively studied for the last 15 years [1–6]. High reflectivity multilayered structures [2], efficient visible photoluminescence [1], compatibility with standard silicon processes for integrated optoelectronics [3], photonic applications [2, 4–6] and biosensors [7] have been major attractions of this field. The fabrication of porous silicon by electrochemical etching of crystalline silicon (c-Si), gives us the ability of having a wide refractive index contrast within the same material, avoiding the problem of inter-diffusion or lattice mismatch between the layers, through an easy and cheap process. Since we can control both, the layer thickness (through anodization time) and refractive index (through porosity), this method can be used to study experimentally the multilayered structures. Recently, Luigi Moretti et al. [8] compared the sensitivities of resonant optical biochemical sensors, based on both periodic and aperiodic porous silicon structures, such as the Bragg and the Thue-Morse multilayer. The shifts of the reflectivity spectra of these devices on exposure to several chemical compounds have been measured and the aperiodic multilayer is reported to be more sensitive than the periodic one. The optical properties of several other kinds of quasiregular systems have been investigated in different systems [9–12]. In [9] the authors presented theoretical calculation in a Cantor photonic crystal waveguides and the optical spectra of fractal multilayer dielectric structures has been shown to possess spectral scalability, which has been found to be directly related to the structure's spatial (geometrical) self-similarity. Following an example of the work demonstrated in [8], we show our prelimnary experimental results on the relationship between the geometry and the optical properties of the multilayers made of nanostructured porous silicon (taking two different refractive indices), where geometrical length of the layers follows the Cantor code and maintains the total thickness (for different orders) of the multilayered structure as 9000 nm.

2. EXPERIMENTAL DETAILS

We have used boron doped p^{++} type crystalline silicon with resistivity 0.001–0.005 ohm-cm, (100) oriented substrates for fabricating our samples. To have a better control over the interfaces and thickness of the porous layers, anodization was performed by alternating square pulses with a frequency of 100 Hz (Ref. Escorcia et al. PSS), with a 50% of duty cycle. The electrochemical reaction took place at room temperature. The electrolyte with volume ratio of 3:7 of HF (48 wt%): ethanol (98 wt%) was taken for electrochemical anodization process. The current density was controlled by the computer. A high porosity H (current density 50 $\mathrm{mA/cm^2}$, with effective refractive index of 1.35) and low porosity (35%, current density 5 $\mathrm{mA/cm^2}$, with $n_b = 2.0$), were repeated to form the cantor type structures. The refractive indices of the pSi layers have been estimated using reflectivity spectra of 2 microns thick single layers at 1500 nm. In addition, in order to maintain a constant HF concentration over the interface between Si and pSi under chemical attack, during the

etching process a rotator is used to remove the bubbles formed during the reaction and circulate the electrolyte within the Teflon cell. Apart from the pauses during the layer formation (due to the positive square pulses), after the formation of every layer a pause of 3 seconds was given. The optical characterization of porous silicon mirrors was carried out by a Perkin Elmer UV-Vis-NIR spectrophotometer (UV-3101) at 8° incidence. Scanning Electron Microscopy (SEM) was used to examine the structural features of the film.

All the structures discussed in this article, have the total physical thickness as 9 microns. The schematic of the possible 4th to 6th order cantor structures is shown in Fig. 1. The black layer represents the low porosity layer. As shown in the figure, the 4th generation of the cantor structure consists of $2^4-1 = 15$ layers. Similarly 5th, 6th and 7th order have 31, 63 and 127 layers respectively.

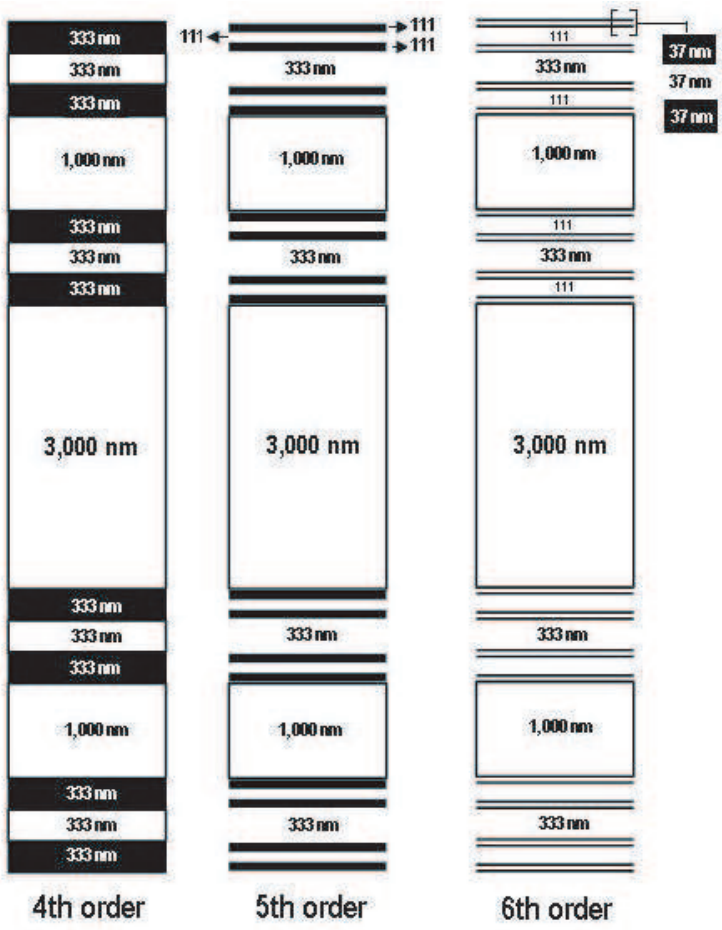


Figure 1: Schematic diagram of the cross-section of the 4th, 5th and 6th series of the Cantor structure. Black and white area represent the low and high porosity layer respectively.

3. RESULTS AND DISCUSSION

Figure 2 shows the measured reflectance spectrum of the 4th–7th order of the Cantor series. In the reflectance spectra of the 4th series of Cantor (Fig. 2(a)) having 15 layers of alternating high and low porosity, a sharp photon localization at 742 nm with a total PBG of 50 nm is observed. The spectra is similar to a multilayer resonator cavity structure (MRC) [13] formed with a periodic sequence of 20 alternating layers of low and high refractive index with equal optical thicknesses ($\lambda/4$) along with a $\lambda/2$ layer in between the 20 layers. The approximate physical thickness of the MRC structure with the similar refractive indices as taken in the present work for the Cantor structures, is 2.49 microns. The PBG is expected to be narrower in case of Cantor structures.

Similarly one can observe that the 5th order (Fig. 2(b)) has a similar localized mode at 651 nm with the total photonic band width of 56 nm. But additionally two more localized modes appear in the near infrared (NIR) region with the 100 nm of full wave half maximum (FWHM) of the

corresponding PBG. The reflectance intensity goes to almost 100% in the NIR region as compared to 90% in the visible region. Making a similar calculation for a periodic structure, replicating the reflectivity spectra of the 5th order Cantor structure, one can find that the total thickness of the corresponding double MRC structure would be around 5 microns.

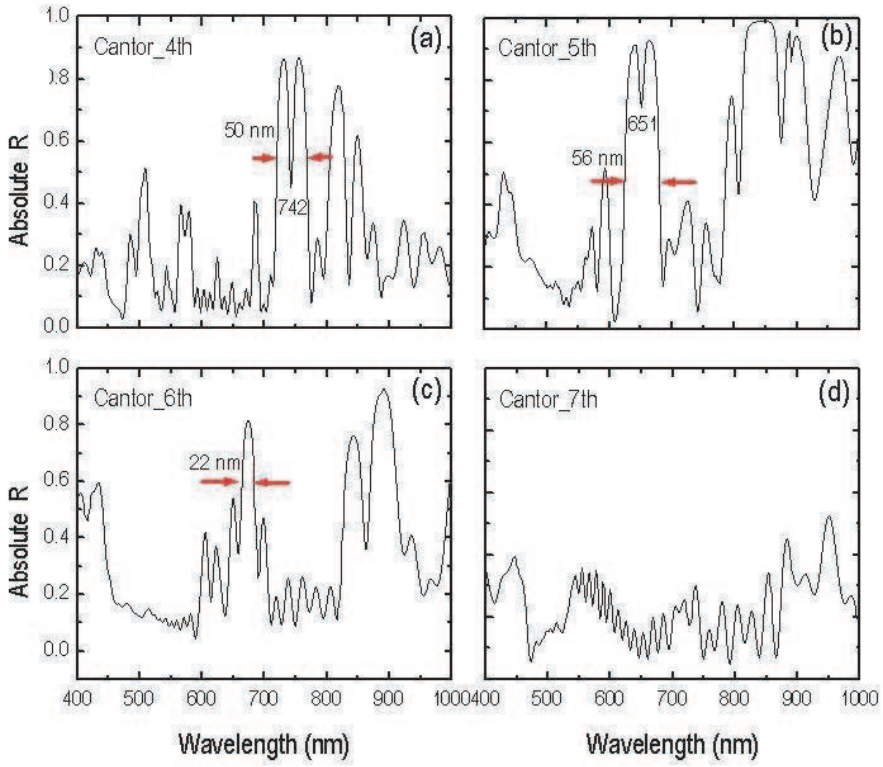


Figure 2: Experimental reflectance for 4 different orders of cantor structures: (a) 4th (15 layers), (b) 5th (31 layers), (c)6th (63 layers), and (d)7th (123 layers).

Figure 2(c) shows the reflectance spectra of the 6th order of the cantor series (63 layers). A PBG of 22 nm of width (with the two shoulders of approximately 50% reflectance) is observed between 650–700 nm. The reflectance of the main peak goes up to 80%. The width of the PBG in the NIR region reduces in this case as compared to the 5th order and appears to be an asymmetric MRC structure. The reflection intensity too is found to reduce to 90% for the higher peak and 80% for the lower peak. Any further increase of Cantor order doesn't show any PBGs. For example, the reflectance spectrum of the seventh order cantor structure (Fig. 2(d)) fails to show any PBG. But a series of equidistant interference fringes, from 550–700 nm, with reflectance of around 30% are observed. The reason of the absence of any PBG is the decrease in the optical thickness of the layers beyond the interaction range of the visible and NIR wavelengths.

Analyzing all the four spectra together, one can observe that the maximum reflectance in the NIR region is found for the fifth series. The possible reason is the physical thicknesses of the layers (as 111 nm), which corresponds to the optical thickness of 222 nm for the refractive index of 2. The optical thickness of 222 nm corresponds to the quarter wavelength of 888 nm. Considering the refractive index of 1.35, the optical thickness of 150 nm, corresponds to the quarter wave length of 600 nm. Hence the reduction in the intensity of the reflected beam for the 6th order can be explained on the basis of the further reduction in the thicknesses of the low and high porosity layers. The absence of PBGs in any series beyond 6th, (e.g., the 7th series in Fig. 2(d)) is justified by the fact that the structure is dominated by the layers having the thicknesses of 12 nm (for the 7th order) or lower, which is too low for demonstrating any PBGs in the visible-NIR region.

4. CONCLUSIONS

The relationship between the geometry and the optical properties of the multilayers made of nanostructured porous silicon were analyzed, where geometrical length of the layers follow the Cantor code maintaining the total thickness of the multilayered structure as 9000 nm. Multiple photonic bandgap structures were observed up to 6th order. The Cantor structures with orders more than 6th (with a physical thickness of 9 microns) show the presence of low intensity equidistant fringes in the visible region. Such structures can be further investigated for the applications in biosensors because of the presence of the multiple PBGs of the reduced width.

ACKNOWLEDGMENT

One of the authors (V. Agarwal) acknowledge the financial support provided by CONACyT (57631).

REFERENCES

1. Canham, L. T., "Silicon quantum wire array fabrication by electrochemical and chemical dissolution of wafers," *Appl. Phys. Lett.*, Vol. 57, 1046, 1990.
2. Agarwal, V. and J. A. del Rio, "Tailoring the photonic bandgap of porous silicon dielectric mirror," *Appl. Phys. Lett.*, Vol. 82, 1512, 2003.
3. Lopez, H. A. and P. M. Fauchet, "Infrared LEDs and microcavities based on erbium-doped silicon nanocomposites," *Mater. Sci. and Eng. B*, Vol. 81, 91, 2001.
4. Ghullinyan, M., C. J. Oton, Z. Gaburro, P. Bettotti, and L. Pavesi, "Porous silicon free standing coupled microcavities," *Appl. Phys. Lett.*, Vol. 82, 1550, 2003.
5. Agarwal, V., J. A del Rio, G. M. Zamfirescu, A. Kavokin, D. Coquillat, D. Scalbert, M. Vladimirova, and B. Gil, "Photon Bloch oscillations in porous silicon optical superlattices," *Phys. Rev. Lett.*, Vol. 92, 097401, 2004.
6. Sapienza, R., P. Costantino, D. Wiersma, M. Ghulinyan, C. J. Oton, and L. Pavesi, "Optical analog of electronic Bloch oscillations," *Phys. Rev. Lett.*, Vol. 91, 263902, 2003.
7. Lin, V. S.-Y., K. Motesharei, K.-P. S. Dancil, M. J. Sailor, and M. R. Ghadiri, "A porous silicon-based optical interferometric biosensor," *Science*, Vol. 278, 840, 1997.
8. Moretti, L., I. Rea, L. de Stefano, and I. Rendina, "Periodic versus aperiodic: Enhancing the sensitivity of porous silicon based optical sensors," *Appl. Phys. Lett.*, Vol. 90, 191112, 2007.
9. Zhukovsky, S. V., A. V. Lavrinenko, and S. V. Gaponenko, "Spectral scalability as a result of geometrical self-similarity in fractal multilayers," *Europhys. Lett.*, Vol. 66, No. 3, 455–461, 2004.
10. Lavrinenko, A. V., S. V. Zhukovsky, K. S. Sandomirsk, and S. V. Gaponenko, "Propagation of classical waves in nonperiodic media: Scaling properties of an optical Cantor filter," *Phys. Rev. E*, Vol. 65, 036621, 2002.
11. Monsoriu, J. A., C. J. Zapata-Rodríguez, E. Silvestre, and W. D. Furlan, "Cantor-like fractal photonic crystal waveguides," *Opt. Commun.*, Vol. 252, 46–51, 2005.
12. Gaggero-Sager, L. M., E. R. Pujals, and O. Sotolongo-Costa, "Self-similarity in a Cantor-like semiconductor quantum well," *Phys. Stat. Sol. (A)*, Vol. 220, 167, 2000.
13. Cazzanelli, M. and L. Pavesi, "Time-resolved photoluminescence of all-porous-silicon microcavities," *Phys. Rev. E*, Vol. 23, No. 56, 264–271, 1997.

Hydrostatic Pressure and Magnetic Field Effects on the Exciton States in Vertically Coupled GaAs-(Ga, Al) As Quantum Dots

M. E. Mora-Ramos[1], **A. H. Rodríguez**[2], **S. Y. López**[3], **and C. A. Duque**[4,5]

[1]Universidad Autónoma del Estado de Morelos, Ave. Universidad 1001, 62209 Cuernavaca, México
[2]Universidad Autónoma de la Ciudad de México, Plantel Iztapalapa, México DF, México
[3]Facultad de Educación, Universidad de Antioquia, AA 1226, Medellín, Colombia
[4]Instituto de Física, Universidad de Antioquia, AA 1226, Medellín, Colombia
[5]Instituto de Física, Unicamp, CP 6165, Campinas-SP, 13083-970, Brazil

Abstract— The variational procedure, in the effective-mass and parabolic-band approximations, is used in order to investigate the combined effects of hydrostatic pressure and in-plane-direction-applied magnetic field on the exciton states in vertically coupled GaAs-(Ga, Al) As quantum dots. Calculations are performed for two cylindrical-shape quantum dots. The exciton envelope wave function is obtained through a variational procedure using a hydrogenic 1 s-like wave function and an expansion in a complete set of trigonometric functions for the electron and hole wave functions. The anticrossing effects on the dispersion with applied magnetic field and hydrostatic pressure of the photoluminescence peaks associated with direct and indirect excitons have been considered.

1. INTRODUCTION

A symmetric/asymmetric coupled double quantum well (DQW) is made of two identical/different quantum wells (QW) that are separated by a thin barrier. For the symmetric case, in flat-band conditions, i.e., without applied electric field in the growth direction of the heterostructure, the eigenfunctions of the DQW have well-defined symmetries. These are broken in the asymmetric case. In this case, only transitions between electron and hole states with the same symmetry are optically allowed. Whenever the maximum probability of the electron and hole wave functions are distributed in the same well, the transitions are known as spatially direct transitions. The intensity of these optical transitions is essentially given by the overlap integral of the electron and hole single-particle envelope wave functions and temperature-dependent populations of electrons and holes in the subbands. Also, it is certainly necessary to take into account the electron-hole ($e-h$) Coulomb interaction for an appropriate description of the optical transitions in semiconducting heterostructures. Of course, effects of the $e-h$ Coulomb interaction are essential whenever the fine structure of the optical spectra shows features which are within the range of the exciton binding energy. On the other hand, the application of hydrostatic pressure results in changes of the dielectric constant and of band structure parameters such as the energy gap and the conduction band mass. This may result in modifications of the interband optical transitions in GaAs-based QW's (see Ref. [1], and references therein).

By applying an in-plane magnetic field in coupled QWs, it is possible to induce strong changes in the excitonic-related photoluminescence (PL) spectra due to field-induced displacement of the interwell exciton dispersion in momentum space, which leads to a transition from the momentum-space direct exciton ground state to the momentum-space indirect exciton ground state [2–4]. The indirect exciton lifetime in coupled DQW heterostructures under applied magnetic fields has been studied by Butov *et al* [2–4] and they attribute the observed results to an increase in the magnetoexciton mass. Also, Butov *et al* [2–4] have studied long-lifetime indirect excitons in coupled QWs and, at low temperatures and high exciton densities, strong deviations of the indirect exciton PL kinetics from monoexponential PL rise/decay were observed. Parlangeli *et al* [5] have studied the indirect exciton dispersion in k space by considering the simultaneous effect of in-growth direction applied electric field and in-plane magnetic field in DQW heterostructures and found that the PL spectra increase with the magnetic field following a quadratic behavior. Additionally, they present measurements of the PL peak positions of both direct and indirect excitons in biased $GaAs/Ga_{1-x}Al_xAs$ coupled DQWs under in-plane applied magnetic fields.

In the present work we are concerned with a theoretical study of the effects of applied hydrostatic pressure and in-plane magnetic fields on the exciton direct and indirect states in $GaAs/Ga_{1-x}Al_xAs$ vertically coupled double quantum dots (CDQD). The potential function profile of this kind of systems is mathematically treated in a way that is very similar to that of the double QW. The

theoretical framework is outlined in Section 2. Results and discussion are presented in Section 3 and finally in Section 4 we outline our conclusions.

2. THEORETICAL FRAMEWORK

The theoretical approach assumes the envelope-function and parabolic-band approximations [6]. We choose the reference system at the barrier center, with the z-axis along the growth direction of the structure, the in-plane magnetic field in the x-direction, $\vec{\mathbf{B}} = B\widehat{x}$, and use the Landau gauge $\vec{\mathbf{A}}(\vec{\mathbf{r}}) = -Bz\widehat{y}$. The Hamiltonian for the exciton then takes the following form [7–9]

$$\begin{aligned}\hat{H} &= \frac{1}{2m_e^*}\left(\hat{\vec{\mathbf{p}}}_e + \frac{e}{c}\vec{\mathbf{A}}_e\right)^2 + \frac{1}{2m_h^*}\left(\hat{\vec{\mathbf{p}}}_h - \frac{e}{c}\vec{\mathbf{A}}_h\right)^2 \\ &\quad + V_e(z_e) + V_h(z_h) + V_e(\rho_e) + V_h(\rho_h) - \frac{e^2}{\epsilon|\vec{\mathbf{r}}_e - \vec{\mathbf{r}}_h|}, \end{aligned} \tag{1}$$

where $\vec{\mathbf{A}}_e = \vec{\mathbf{A}}(\vec{\mathbf{r}}_e)$, $\vec{\mathbf{A}}_h = \vec{\mathbf{A}}(\vec{\mathbf{r}}_h)$, and $\hat{\vec{\mathbf{p}}}_i$, $\vec{\mathbf{r}}_i$, m_i^* and V_i, with $i = e, h$, are the momentum operators, electron and hole coordinates, effective masses and corresponding CDQD confining potentials, respectively, e is the absolute value of the electron charge and ϵ is the GaAs dielectric constant. For simplicity, the dielectric constant and the effective masses are considered the same as in GaAs throughout the GaAs-$Ga_{1-x}Al_xAs$ CDQW. The dependencies on hydrostatic pressure of conduction band mass, barrier heights, and the dielectric constant are introduced according to Ref. [1].

In our model we consider two cylindrical quantum dots vertically coupled by a finite potential barrier. In the external walls of each quantum dot we have considered an infinite potential barrier. In the case of the radial confinement, the potentials that we consider are also infinite. It is for that reason that in the Eq. (1), both for the electron and the hole, the potential that confines them can be written as a sum of two potentials the first one in the z-direction and the second one in the ρ-direction. In the case in which we consider finite barriers in all direction, this separation of the potential in a sum of potentials only will be valid for large quantum dots where the wave functions of each particle have a little contribution in the regions of the barriers.

In order to obtain the exciton eigenfunctions for the GaAs-$Ga_{1-x}Al_xAs$ CDQD, we adopt the variational scheme which consists in minimizing the functional

$$E(\Phi) = \langle\Phi|\hat{H}|\Phi\rangle \tag{2}$$

by using the variational wave functions as

$$\Phi(\vec{\rho}, z_e, z_h) = N f(z_e)F(z_h)g(\rho_e)g(\rho_h)e^{-\lambda r}\,, \tag{3}$$

where $r = \sqrt{\rho^2 + (z_e - z_h)^2}$, λ is a variational parameter, $f(z_e)$ and $F(z_h)$ are, in general, linear combinations of the z-dependent part of the electron $f_i(z_e)$ and hole $F_j(z_h)$ eigenfunctions of the total Hamiltonian neglecting the Coulomb interaction [10], and $g(\rho_e)$ and $g(\rho_h)$ are the corresponding in-plane wave functions. The coefficients $a_i^{(e)}$ and $b_j^{(h)}$ of above mentioned linear combinations are also variational parameters satisfying the usual normalization conditions. Finally, in order to obtain the non-correlated $f_i(z_e)$ electron and $F_j(z_h)$ hole eigenfunctions, it is convenient to use the method by Xia and Fan [11] of expansion in terms of sine functions, used in the study of electron states in semiconductor superlattices in the presence of in-plane magnetic fields. In the variational approach described above, the effect of the Coulomb interaction is to mix the GaAs-$Ga_{1-x}Al_xAs$ CDQD electron and hole-wave functions $f_i(z_e)$ and $F_j(z_h)$, respectively. Here we are interested in excitons associated to the GaAs-$Ga_{1-x}Al_xAs$ CDQD ground state, and limit ourselves to the cases for which only the mixing between the CDQD electron ground state $f_0(z_e)$ and electron first-excited state $f_1(z_e)$ is important, whereas mixing effects for the CDQD hole states are disregarded. The corresponding variational exciton wave functions then take the form

$$\Phi_+(\vec{\rho}, z_e, z_h) = \left[\alpha\, f_0(z_e) + \sqrt{1-\alpha^2} f_1(z_e)\right] F_0(z_h)g(\rho_e)g(\rho_h)\, e^{-\lambda_+ r} \tag{4}$$

and

$$\Phi_-(\vec{\rho}, z_e, z_h) = \left[-\sqrt{1-\alpha^2} f_0(z_e) + \alpha\, f_1(z_e)\right] F_0(z_h)g(\rho_e)g(\rho_h)\, e^{-\lambda_- r}\,, \tag{5}$$

where $F_0(z_h)$ is the CDQD hole ground state, α, λ_+ and λ_- are variational parameters, and we follow the procedure by Fox *et al* [10] in the process of minimizing $E(\Phi)$ [cf. Eq. (2)], using the wave functions (4) and (5).

3. RESULTS AND DISCUSSION

In Figure 1 we present our results for the in-plane applied magnetic field dependence of the calculated PL peak transitions for two vertically coupled cylindrical quantum dots. Results are for two different values of the hydrostatic pressure. Clearly, there is a quadratic behavior as a function of the applied magnetic field. The vertices of the parables are shifted in energy as an effect of the vertical and radial confinement. As the magnetic field grows the influence of the potential barrier, which separates the two quantum dots, decreases, and the energy curves go to a linear behavior determined by the first Landau level. This level is displaced in a value of the energy determined by the radial confinement. We note that the energy curves grow with the increasing width of the potential barriers. This is due to the fact that barrier width increasing implies that the two quantum dots are more and more isolated and the carriers become essentially confined in the region of a single quantum dot. Accordingly, the energy of each particle (electron or hole) becomes higher and, as a result, the energy of the PL-peak increases.

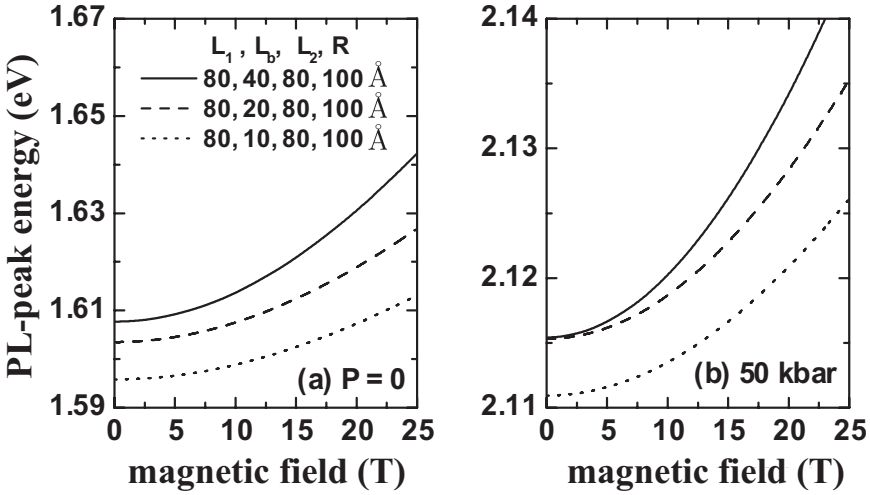


Figure 1: In-plane applied magnetic field dependence of the calculated PL peak for the transitions between the first coulomb-perturbed electron states and the first heavy-hole level for a system of two couple cylindrical quantum dots with radius R, heights L_1 and L_2, separated by a barrier of width L_b. Hydrostatic pressure. (a) $P = 0$, and (b) 50 kbar.

In the Figure 1(b) we have considered the dimensions of the structures, but with the additional effect of a pressure of 50 kbar. In essence, the effect of the latter is seen as a blue-shift in the PL-peaks. This large blue-shift is mainly due to the dependence with the pressure of the band-gap of the quantum dot material (GaAs). The dependencies with the pressure of the dimensions of the structure modify in an non-significant value the PL-peak energies (changes are smaller than 1%). The changes in the effective masses and the static dielectric constant with the pressure modify the binding energy leading to its increasing with the pressure, which manifests as a red-shift in the PL-peak. Indeed, this shift is overlapped by the dependence with the pressure of the energy gap of the GaAs. Finally, we observe that the applied pressure makes the effects of the magnetic field to become greater. This fact is associated with the decrease of the height of the central barrier and with the increasing in the effective mass of the electrons.

4. CONCLUSIONS

The variational procedure, in the effective-mass and parabolic-band approximations, have been used in order to investigate the combined effects of hydrostatic pressure and in-plane-direction-applied magnetic field on the exciton states in vertically coupled GaAs-(Ga,Al) As quantum dots. Calculations are performed for two cylindrical-shape quantum dots. We have observed a quadratic dependence with the pressure of the PL-peak energy. However, for large magnetic field values, a linear behavior can be predicted, in accordance with the variation of the Landau levels due to the quantum confinement. Additionally, we observe that the magnetic field effects are magnified when a hydrostatic pressure is considered on the structure.

ACKNOWLEDGMENT

M. E. M. R. acknowledges support from Mexican CONACYT through grant No. 52304. C. A. D. and S. Y. L. would like to thank the Colombian CODI-Univ. de Antioquia for partial financial support. This work was also partially financed by the Excellence Center for Novel Materials and COLCIENCIAS under contract No. 043-2005.

REFERENCES

1. Duque, C. A., S. Y. López, and M. E. Mora-Ramos, "Hydrostatic pressure effect on the $\Gamma - X$ conduction band mixing and the binding energy of a donor impurity in GaAs-$Al_xGa_{1-x}As$ quantum wellds," *Phys. Stat. Sol. (b)*, Vol. 244, No. 6, 1964, 2007.
2. Butov, L. V., A. Imamoglu, A. V. Mintsev, K. L. Campman, and A. C. Gossard, "Magneto-optics of the spatially separated electron and hole layers in GaAs/$Al_xGa_{1-x}As$ coupled quantum wells," *Phys. Rev. B*, Vol. 60, 8753, 1999.
3. Butov, L. V., A. A. Shashkin, V. T. Dolgopolov, K. L. Campman, and A. C. Gossard, "Photoluminescence kinetics of indirect excitons in GaAs/$Al_xGa_{1-x}As$ coupled quantum wells," *Phys. Rev. B*, Vol. 59, 1625, 1999.
4. Butov, L. V., A. V. Mintsev, Yu. E. Lozovik, K. L. Campman, and A. C. Gossard, "From spatially indirect excitons to momentum-space indirect excitons by an in-plane magnetic field," *Phys. Rev. B*, Vol. 62, 1548, 2000.
5. Parlangeli, A., P. C. M. Christianen, J. C. Maan, I. V. Tokatly, C. B. Soerensen, and P. E. Lindelof, "Optical observation of the energy-momentum dispersion of spatially indirect excitons," *Phys. Rev. B*, Vol. 62, 15323, 2000.
6. Dzyubenko, A. B. and A. L. Yablonskii, "Intrawell and interwell magnetoexcitons in InxGa1-xAs/GaAs coupled double quantum wells," *Phys. Rev. B*, Vol. 53, 16355, 1996.
7. Vol, E. D. and S. I. Shevchenko, "Magnetic exciton in a two-layer system," *Low Temp. Phys.*, Vol. 26, No. 8, 577, 2000.
8. Chang, K. and F. M. Peeters, "Bright-to-dark exciton transition in symmetric coupled quantum wells induced by an in-plane magnetic field," *Phys. Rev. B*, Vol. 63, 153307, 2001.
9. Lozovik, Yu. E., I. V. Ovchinnikov, S. Yu. Volkov, L. V. Butov, and D. S. Chemla, "Quasi-two-dimensional excitons in finite magnetic fields," *Phys. Rev. B*, Vol. 65, 235304, 2002.
10. Fox, A. M., D. A. B. Miller, G. Livescu, J. E. Cunningham, and W. Y. Jan, "Excitonic effects in double quentum wells," *Phys. Rev. B*, Vol. 44, 6231, 1991.
11. Xia, J.-B. and W.-J. Fan, "Electronic structures of superlattices under in-plane magnetic field," *Phys. Rev. B*, Vol. 40, 8508, 1989.
12. Orlita, M., R. Grill, M. Zvára, G. H. Döhler, S. Malzer, M. Byszewski, and J. Soubusta, "Luminescence of coupled quantum wells: Effects of indirect excitons in high in-plane magnetic fields," *Phys. Rev. B*, Vol. 70, 75309, 2004.
13. Orlita, M., M. Byszewski, G. H. Döler, R. Grill, S. Malzer, J. Soubusta, and M. Zvára, "Luminescence of indirect excitons in high in-plane magnetic fields," *Physica E*, Vol. 30, 1, 2005.
14. Li, E. H., "Material parameters of InGaAsP and InAlGaAs systems for use in quantum well structures at low and room temperatures," *Physica E*, Vol. 5, 215, 2000.

Internal Mobility Edge in Doped Graphene: Frustration in a Renormalized Lattice

Gerardo G. Naumis
Departamento de Física-Química, Instituto de Física
Universidad Nacional Autónoma de México (UNAM)
Apartado Postal 20-364, México DF 01000, Mexico

Abstract— We show that an internal localization mobility edge can appear around the Fermi energy in graphene by introducing impurities in the split-band regimen, or by producing vacancies in the lattice. The edge appears at the center of the spectrum and not at the band edges, in contrast with the usual picture of localization. Such result is explained by showing that the bipartite nature of lattice allows to renormalize the Hamiltonian, and the internal edge appears because of frustration effects in the renormalized lattice. The size in energy of the spectral region with localized states is similar in value to that observed in narrow gap semiconductors.

Only very recently a two dimensional form of carbon was obtained [1]. This material, known as graphene has attracted a lot of research due to its amazing electrical and mechanical properties [2]. For example, electrons in graphene behave as massless relativistic fermions that satisfy the Dirac equation [4]. Such property is a consequence of the bipartite crystal structure [5], in which a linear dispersion relationship appears at the center of the electronic spectrum. Also, one can cite the high mobility of its charge carriers that remains higher even at high electric-field induced concentration, that translates into ballistic transport on a submicron scale [3] at 300° K. These and other unusual electronic properties of graphene makes it a promising material for building electronic devices. However, from the point of view of applications, the use of pure graphene pose some problems. The transmission probability of electrons across a potential barrier is always unity, irrespective of the height and width of the barrier. This behavior is related to the Klein paradox in relativistic quantum mechanics [2]. As a result, conductivity can not be changed by an external gate voltage, a feature required to build a FET transistor, although a quantum dot has been used to perform the required task [6]. In spite of all this research in pure graphene, at the moment there is not so much knowledge in the effects of impurities in the electronic properties and its potential use to produce gates. In a previous work [7], the density of states (DOS) of graphene with Anderson type of disorder revealed that the linear dispersion relationship was affected [7], and recently many electrical properties of graphene with disorder have been obtained [8]. However, the existence of a mobility edge has not addressed. Here we show that graphene doped with impurities or with vacancies presents a very unusual property; instead of having a localization mobility edge at the band limits as in the usual Anderson localization, the localized states appear at the center of the spectrum, around the Fermi energy. As we will show, this is a simple consequence of the bipartite crystal structure, which produces a frustration effect in a renormalized. Hamiltonian that removes one of the bipartite lattices. The observed effect can be used in certain applications, since the mobility edge can be tuned with a given concentration of impurities.

Let us start by considering the tight-binding Hamiltonian of graphene with disorder, which can be written as $H = H_0 + H_1$, where H_0 is the pure graphene Hamiltonian given by [9],

$$H_0 = E_0 \sum_i |i\rangle\langle i| + \gamma_0 \sum_{<i,j>} |i\rangle\langle j| + H_1. \tag{1}$$

E_0 is the self-energy of carbon and γ_0 is the carbon-carbon resonance integral, as given in Ref. [9]. H_1 is the Hamiltonian of the perturbation due to defects,

$$H_1 = \delta E \sum_i |i\rangle\langle i| + \delta\gamma_0 \sum_{<i,j>} |i\rangle\langle j|, \tag{2}$$

where we define $\delta E \equiv E_I - E_0$ and $\delta\gamma_0 \equiv \gamma_I - \gamma_0$. Here E_I is the self-energy of the defects, and γ_I the transfer integral between impurities (which are basically isolated in the dilute limits). When $\delta E \gg E_0$, the spectrum is divided in two parts, one centered around E_0 and the other at $E_0 + \delta E$. This case is known as the split-band limit. The states in the sub-band around the carbon self

energy E_0, that we call the C-band, are strongly confined on carbon atoms. Furthermore, in the limit $\delta E \gg E_0$, it has been shown that impurity atoms can be formally removed in a tight-binding Hamiltonian [10], and thus the C-band can be studied by using a Hamiltonian restricted to C sites only,

$$H_{CC} = E_0 \sum_{i \in C} |i\rangle\langle i| + \gamma_0 \sum_{i,\, j \in C} |i\rangle\langle j|. \tag{3}$$

This Hamiltonian describes an electron that can hop from one site to its neighbors only if both are carbon atoms (C). Furthermore, the problem for the C sub-band is similar to a lattice with holes, because impurity atoms act as perfect barriers in the limit of infinite self-energy. As a result, the results presented here are also valid for vacancies in the lattice.

Now let us study the spectrum of H_{CC}. First it is convenient to work on a renormalized Hamiltonian H_{CC}, which takes advantage of the bipartite nature of the C lattice, once the I atoms are removed. The bipartite character of the C lattice means that it can be separated in two inter-penetrating sublattices, A and B. We define two orthogonal operators that project the wavefunctions into each sublattice,

$$P_A = \sum_{i \in A} |i\rangle\langle i|, \quad \text{and} \quad P_B = \sum_{j \in B} |j\rangle\langle j| \tag{4}$$

Therefore, any eigenvector $|\phi\rangle$ of H_{CC} can be written in terms of these projectors,

$$H_{CC}(P_A + P_B)|\phi\rangle = E(P_A + P_B)|\phi\rangle. \tag{5}$$

Since H_{CC} produces a hopping in the wavefunction between the A and B sublattices,

$$H_{CC}P_A|\phi\rangle = EP_B|\phi\rangle, \quad \text{and} \quad H_{CC}P_B|\phi\rangle = EP_A|\phi\rangle. \tag{6}$$

From these equations, one can see that the spectrum is symmetric around $E = E_0$, since if $(P_A + P_B)|\phi\rangle$ is an eigenvector with eigenvalue E, $(P_A - P_B)|\phi\rangle$ is also an eigenvector with eigenvalue $-E$. We can decuple the sublattices by further applying H_{CC} to Eqs. (6),

$$H_{CC}(H_{CC}(P_i|\phi\rangle)) = H^2_{CC}(P_i|\phi\rangle) = E^2(P_i|\phi\rangle), \tag{7}$$

where $i = A,\, B$. Thus, the projection of an eigenvector in each sublattice is a solution of the squared Hamiltonian. Observe that the eigenvalues of H^2_{CC} are positive definite, and their eigenstates are, at least, doubly degenerate. This spectrum can be regarded as the folding of the original spectrum of H_{CC} around $E = 0$, in such a way that the two band edges of H_{CC}, are mapped into the highest eigenvalue of H^2_{CC}, while the states at the center of the original band are now at the minimum eigenvalue of the squared Hamiltonian (E^2). The important property of the renormalized Hamiltonian H^2_{CC} is that the states at the bottom of the spectrum have an antibonding nature (the phase between neighbors is π), and we can expect that the frustration of the wavefunction can prevent the spectrum from reaching its minimum eigenvalue in a continuous form when frustration is present [11, 12]. In fact, frustration acts as an effective potential which leads to localization since the wave-function tends to avoid regions of higher-frustration. The mobility edge appears when the energy cost in localization is less than that of having amplitude in frustrated bonds. As we will show next, this frustration augments with disorder. To do this, we observe that the Hamiltonian H^2_{CC} is equivalent to a renormalization of sites B in the lattice, which leads to a triangular lattice with an effective interaction, as shown in Fig. 1(a). The new lattice contains odd rings, and when impurities are present, there are holes, as indicated in Fig. 1(b). The corresponding Schröedinger equation derived from H^2_{CC} is,

$$\left((E - E_0)^2 - Z_i\gamma_0^2\right) c_i(E) = \gamma_0^2 \sum_{(j,\, i) \in A} c_j(E), \tag{8}$$

where $c_i(E)$ is the amplitude of the wave-function at site i for an eigenenergy E, and the notation $(j,\, i) \in A$ means that the sum is taken only for carbon atoms which are first neighbours in the new triangular lattice, i.e., those carbon atoms that were second neighbours in the original lattice. Notice that such atoms belong to only one of the bipartite sublattices A or B. Due to the symmetry of the problem, we can solve for any sublattice, say for example sublattice A. Finally, Z_i is the

coordination number at site i. This number goes from 0 when a carbon atom is surrounded by impurities, to 3 in the lattice without defects. Then we can perform a variational procedure to estimate the ground state of Eq. (8). After multiplying Eq. (8) by $c_i^*(E)$ and summing over i,

$$(E-E_0)^2 = \sum_i Z_i \gamma_0^2 |c_i(E)|^2 + \gamma_0^2 \sum_i \sum_{(j,i)\in A} c_i^*(E) c_j(E), \tag{9}$$

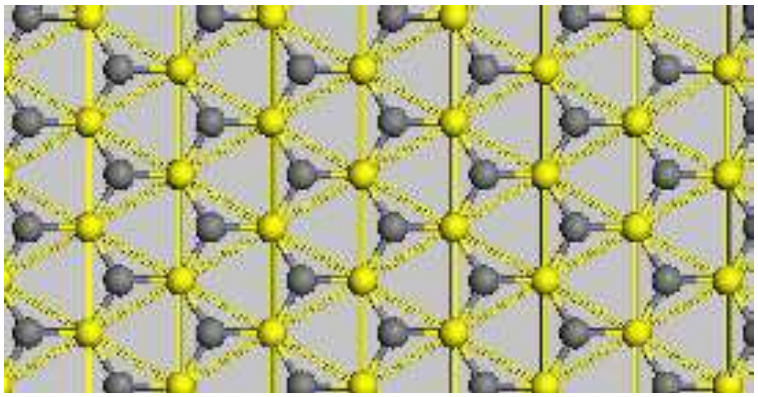


Figure 1: Renormalization of the graphene lattice. Atoms in the A sublattice are shown with different color than those in the B sublattice. The new lattice that appears after renormalizing the B, is represented with double bonds.

The first contribution is an effective self-energy while the second depends on the number of bonds and the amplitude and phase of the wave-function. For example, in pure carbon $Z_i = 3$. Also, the lattice is periodic from where we can write $c_j(E) = ce^{i\phi_j}$ where c is an amplitude (in fact $c = 1/\sqrt{N}$ where N is the number of atoms), and ϕ_j is a phase. The minimal eigenvalue is thus obtained from Eq. (8) when the phase difference between sites in the A sublattice is π. Thus, the ground state has an antibonding nature and $c_i^*(E)c_j(E) = -1/N$. Using that there are three second neighbours for each atom, it follows that $E = E_0$. As a consequence, this shows that there is no gap for pure graphene, as expected. However, the previous case reveals an interesting fact, the zero gap is obtained due to the balance between the positive renormalized self-energy Z_i and the antibonding contribution. In pure graphene, both contributions match exactly to produce a gapless spectrum.

Now consider the case of a finite concentration x of impurities or holes. Since an impurity belongs to one of the bipartite sublattices, say A, there are two effects. The first is a reduction in the average coordination number and the second is that some bonds are deleted. This coordination effect is estimated as follows. The first term of Eq. (9) can be written as an average term plus a correlation of amplitude-coordination,

$$\sum_i Z_i \gamma_0^2 |c_i(E)|^2 = \langle Z \rangle \gamma_0^2 + V \gamma_0^2 \sum_i \delta Z_i \delta c_i^2(E) \tag{10}$$

where it was used that Z_i can be written as an average $\langle Z \rangle$ plus a fluctuation part δZ_i. A similar procedure can be made for $|c_i(E)^2 \equiv \langle c^2(E) \rangle + \delta c_i^2(E)$. The average coordination number can be obtained by observing that around a given carbon atom, there are four possible configurations: it can be surrounded by one, two and three impurities, or it can be completely surrounded by carbon atoms. For each configuration, there is a different coordination number Z, since impurities act as holes. As a result, the coordination number Z has a binomial probability distribution $P(Z) = C_Z^3 x^Z (1-x)^{3-Z}$ where C_Z^3 is a combinatorial factor. It follows that $\langle Z \rangle$ is the first moment of the binomial distribution: $\langle Z \rangle = \sum_{Z=0}^{Z=3} ZP(Z) = 3(1-x)$. The contribution of the last term in Eq. (10) leads to the production of impurity states, since it is the correlation between amplitude and self-energy fluctuations. Thus, the system has a mobility edge when this term lowers the energy compared with the energy required for having an extended state with amplitude in frustrated bonds.

The other effect is the removal of bonds that changes the second term of Eq. (9). We can estimate this effect as follow. For low concentration of impurities $x \ll 1$, most of them are isolated,

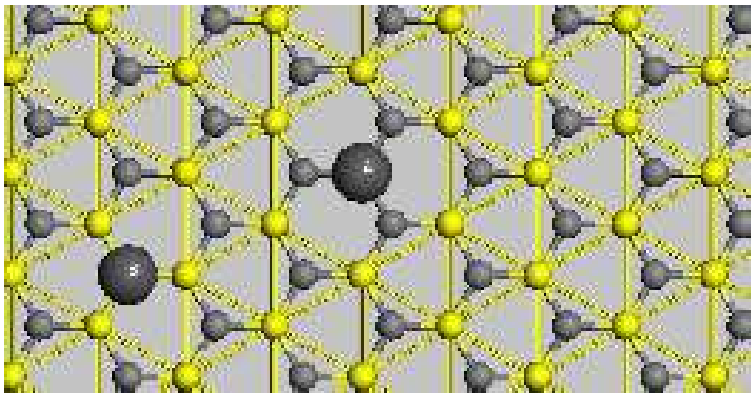


Figure 2: Renormalization of the lattice with defects. The impurities are shown with dark color. There are two cases: the impurities can fall in the A sublattice or in the B, as indicated in the figure. In the first case, six bonds are deleted in the renormalized sublattice, while only three disappear in for other case.

since the probability of having two impurities as neighbours goes as x^2. Thus, we will consider that impurities are isolated. Two situations are possible. Either an impurity belongs to the renormalized sublattice, or it can remain as shown in Fig. 2. For each impurity site that is renormalized, 3 bonds are lost. In the other case, 6 bonds are lost for each impurity. Since they are randomly distributed in sublattices A and B, the concentration of impurities is x on each sublattice. As a result, the number of missing bonds is $(6+3)xN$, from a previous total of $3N$. Using this count in Eq. (9), and assuming no self-energy amplitude correlation Eq. (10) for an antibonding trial state, we obtain the approximate position of the mobility edge (E_d),

$$(E_d - E_0)^2 \approx 3\gamma_0^2(1-x) - \gamma_0^2(3-9x) = 6\gamma_0^2 x,$$

which leads to a symmetric mobility edge separated an energy Δ from the center of the band,

$$\Delta \approx \pm\sqrt{6x}\gamma_0. \tag{11}$$

As a check of these ideas, in Fig. 3 we present the normalized logarithm of the inverse participation

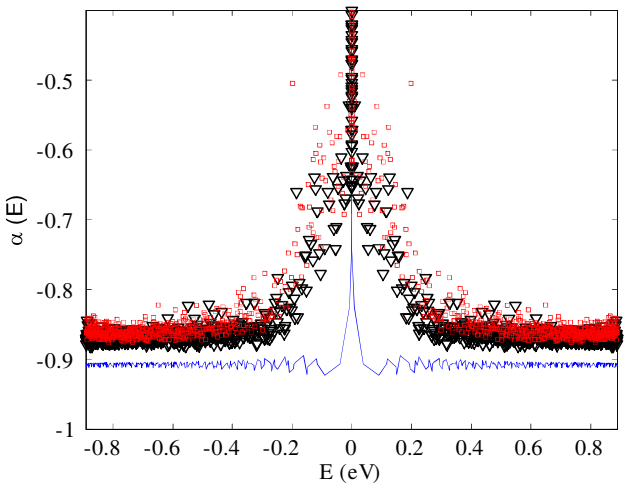


Figure 3: Logarithm of the inverse participation ratio as a function of the energy for pure graphene (line) compared with the dope case with $x = 0\,01$ (triangles) and $x = 02$ (squares) around the center of the spectrum of the carbon sub-band, for a lattice with 5184 sites. Observe the rise at the center of the spectrum for the doped case. A band of degenerate states is also observed for pure graphene. The zero corresponds to the Fermi energy.

ratio,

$$\alpha(E) = \frac{\log IP(E)}{\log N}$$

where $IP(E)$ is the inverse participation ratio, defined as $IP(E) = \sum_{i=1}^{N} \|c_i(E)\|^4$, which is a well-known measure of localization. For extended states, $\alpha(E) \approx -1$, while it tends to be bigger values when localization is present. Fig. 3 shows a comparison between pure graphene case and the doped cases, for a tigth-binding simulation using an average of 10 lattices with $N = 5184$ sites. It is worthwhile mentioning that a band of degenerated states appears in the center of pure graphene, which has not been reported previously by other workers. They are a consequence of the local topology of the lattice, as also happens in the square [10] and Penrose lattices, and are due to a decoupling of the A and B sublattice at the center of the spectrum. Fig. 3 shows that the $IP(E)$ is in general bigger for the doped case, but at the center of the spectrum there is a clear rise in its value, indicating a greater degree of localization. In Fig. 4 we compare Eq. (11) with the numerical value of Δ obtained from the localization plot, which shows a good agreement with the predicted value.

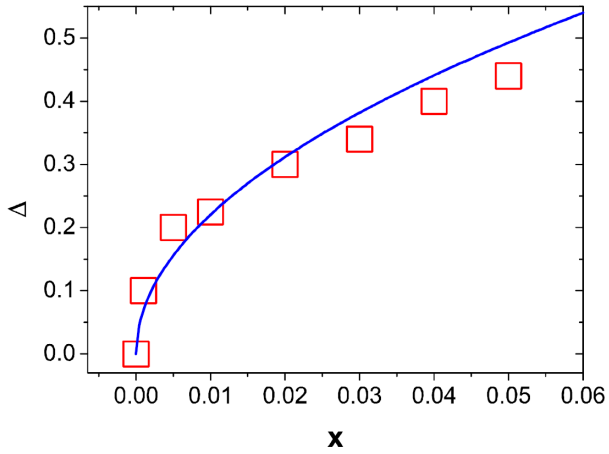


Figure 4: Theoretical value of the mobility edge predicted from Eq. (11), indicated with a solid line, and the value obtained from a direct diagonalization of the Hamiltonian (squares). The numerical results were obtained from an average of 10 lattices with $N = 5184$ sites.

The value of γ_0 is around [9] 0.9 eV or $\gamma_0 = 20$ Kcal/mol. For a 1% doping, the size of the whole localized region is around $2\Delta \approx 0.44$ eV. Since light absorbed when the band-gap energy is in the limit of the visible spectrum 1.77 eV (700 nm), the localized region in doped graphene can be considered as similar in size as the energy gap in narrow-band-gap semiconductors.

In conclusion, we have shown that doped graphite in the split band regimen presents a mobility edge at the center of the spectrum, an this can be useful for many devices since the position of the mobility edge can be controlled by doping.

ACKNOWLEDGMENT

I would like to thank M. and H. Terrones for useful suggestions. This work was supported by DGAPA UNAM project IN108502, and CONACyT 48783-F and 50368.

REFERENCES

1. Novoselov, K. S., et al., *Science*, Vol. 306, 666, 2004.
2. Katsnelson, M. I., *Materials Today*, Vol. 10, 20, 2007.
3. Geim, A. K. and K. S. Novoselov, *Nature Materials*, Vol. 6, 183, 2007.
4. Semenoff, G. W., *Phys. Rev. Lett.*, Vol. 53, 2449, 1984.
5. Slonczewski, J. C. and P. R. Weiss, *Phys. Rev.*, Vol. 109, 272, 1958.
6. Novoselov, K. S., et al., *Science*, Vol. 306, 271, 2006.
7. Hu, W. M., J. D. Dow, and C. W. Myles, *Phys. Rev. B*, Vol. 30, 1720, 1984.
8. Peres, N. M. R., F. Guinea, and A. H. Castro-Neto, *Phys. Rev. B*, Vol. 73, 125411, 2006.
9. Wallace, P. R., *Phys. Rev.*, Vol. 71, 622, 1947.
10. Kirkpatrick, S. and T. P. Eggarter, *Phys. Rev. B*, Vol. 6, 3598, 1972.
11. Cohen, M., *Topological Disorder in Condensed Matter, Springer Series in Solid State Sciences*, edited by F. Yonezawa and T. Ninomiya, Vol. 46, 122, Springer, New York, 1983.
12. Naumis, G. G., R. A. Barrio, and Ch. Wang, *Phys. Rev. B*, Vol. 50, 9834, 1994.

Small Nano-dot Incorporated High-efficiency Phosphorescent Blue Organic Light-emitting Diode

Jwo-Huei Jou[1], **Wei-Ben Wang**[1], **Mao-Feng Hsu**[1], **Chi-Ping Liu**[1], **Cheng-Chung Chen**[1], **Chun-Jan Wang**[1], **Yung-Cheng Tsai**[1], **Jing-Jong Shyue**[2], **Sung-Cheng Hu**[3], **Chung-Che Chiang**[4], **and He Wang** [5]

[1]Department of Materials Science and Engineering, National Tsing Hua University
Hsin-Chu, Taiwan 30013, China
[2]Research Center for Applied Sciences Academia Sinica
128 Academia Rd., Sec. 2 Nankang, Taipei 115, Taiwan, China
[3]Chung-shan Institute of Scince and Technology, Armament Bureau, M. N. D.
No. 481, Sec. Chia An, Zhongzheng Rd., Longtan Shiang, Taoyuan County 325, China
[4]Department of Applied Chemistry, National Chi Nan University
Nantou Hsien, Taiwan 545, China
[5]Department of Materials Science and Engineering, Tsinghua University
Beijing 100084, China

Abstract— High efficiency phosphorescent blue organic light-emitting diode (OLED) was obtained by incorporating small amino or hydroxyl functional group-modified polymeric nano-dot (APND or HPND) in the hole transporting layer (HTL), poly (ethylenedioxythiophene): poly (styrene sulfonic acid) (PEDOT:PSS). The device comprised a 1250 Å anode layer of indium tin oxide, a 350 Å HTL of PEDOT:PSS doped with APND or HPND, a 400 Å blue emissive layer composed of a molecular host of 4,4'-bis (carbazol-9-yl) biphenyl doped with 14 wt% blue dye of bis (3,5-difluoro-2-(2-pyridyl)-phenyl-(2-carboxypyridyl) iridium (III), a 320 Å electron-transporting layer of 2, 2′, 2″-(1, 3, 5-benzenetriyl)-tris (1-phenyl-1-H-benzimidazole), a 7 Å electron-injection layer of lithium fluoride and a 1500 Å cathode layer of aluminum. The resultant power efficiency at 100 cd/m^2, for example, was increased from 12.0 to 25.9 lm/W, an increase of 116%, as 7 wt% APND of 8 nm in size was added. By employing 7 wt% HPND, the power-efficiency was 21.7 lm/W. The resultant luminance markedly increased with the incorporation of the PND. Whilst,the corresponding current density continuously decreased. These results indicate that the marked efficiency improvement may be attributed to a better balance of carrier-injection resulted from the hole-blocking-function possessed APND and the hole-trapping-function possessed HPND, which respectively exhibited positive and negative charge on the surface. Moreover, the chromaticity coordinate at 100 cd/m^2, for example, was (0.19, 0.34), barely changed in the presence of the nano-dots. Importantly, since the nano-dot was not employed in the emissive layer, the same concept may be applied to fluorescent blue or other OLEDs.

1. INTRODUCTION

Organic light-emitting diodes (OLEDs) are increasingly attracting interest because of their high potential as flat-panel displays and for liquid-crystal-display backlighting and area illumination. [1–4] These applications require highly efficient OLEDs. Numerous approaches have been reported to improve the efficiency, such as the use of electroluminescence (EL) efficient phosphorescent and/or fluorescent materials [4], coupled with appropriate device architectures. Efficient devices typically possess optimized device-thickness, low carrier-injection-barrier, effective carrier/exciton-confinement, highly efficient host-to-guest energy-transfer and balanced carrier-injection. [4–14] Recently, the incorporation of quantum- or nano-dot in the emissive or another layer has been found to be effective for some OLED devices. [13–18] However, the mechanism of this improvement is not yet clear. A homogeneous distribution of the embedded nano-dots may also be crucial, which restrains the use of a dry-process for their incorporation. In order to obtain high efficiency, OLED devices must frequently be kept relatively thin, which would consequently limit the use of large nano-dots.

In this letter, we present phosphorescent blue OLEDs with marked efficiency-improvement obtained by incorporating small amino or hydroxyl functional group-modified polymeric nano-dot (APND or HPND) in the hole transporting layer (HTL), poly(ethylenedioxythiophene): poly(styrene sulfonic acid) (PEDOT:PSS). The effect of the concentration of these two polymeric nano-dot (PND) on the electroluminescent (EL) characteristics of the resultant devices was examined. The

resultant power efficiency at 100 cd/m^2, for example, was increased from 12.0 to 25.9 lm/W, an increase of 116%, as 7 wt% APND of 8 nm in size was added. By employing 7 wt% HPND, the power-efficiency was 21.7 lm/W.

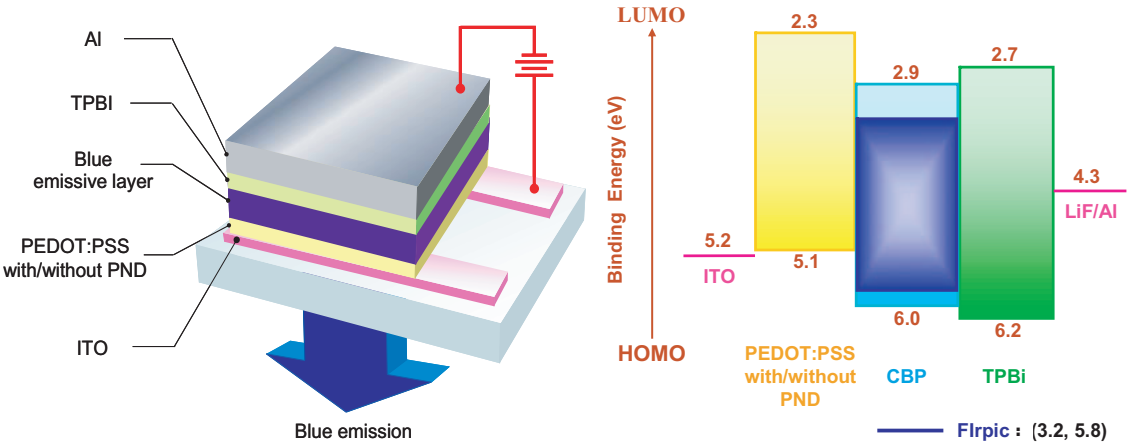


Figure 1: The schematic device structure and energy-level diagram of the phosphorescent blue OLEDs.

2. EXPERIMENTAL

Figure 1 shows the schematic device structure and energy-level diagram of the phosphorescent blue OLEDs studied. The device comprises a 1250 Å anode layer of indium tin oxide (ITO), a 350 Å hole-injection layer of PEDOT:PSS doped with PNDs, a 400 Å blue emissive layer, a 320 Å electron-transporting layer of 2, 2′, 2″-(1, 3, 5-benzenetriyl)-tris(1-phenyl-1-H-benzimidazole) (TPBi), a 7 Å electron-injection layer of lithium fluoride (LiF) and a 1500 Å cathode layer of aluminum (Al). The blue emissive layer was composed of a molecular host of 4, 4'-bis(carbazol-9-yl) biphenyl (CBP) doped with 14 wt% blue dye of bis(3, 5-difluoro-2-(2-pyridyl)-phenyl-(2-carboxypyridyl) iridium (III) (FIrpic). The emission area of all the resultant devices was 25 mm^2 and only the luminance in the forward direction was measured.

The resultant electro-luminescent characteristics were determined by using Minolta CS100A luminance meter and KEITHLEY 2400 source meter. All the measurements were carried out at the ambient condition.

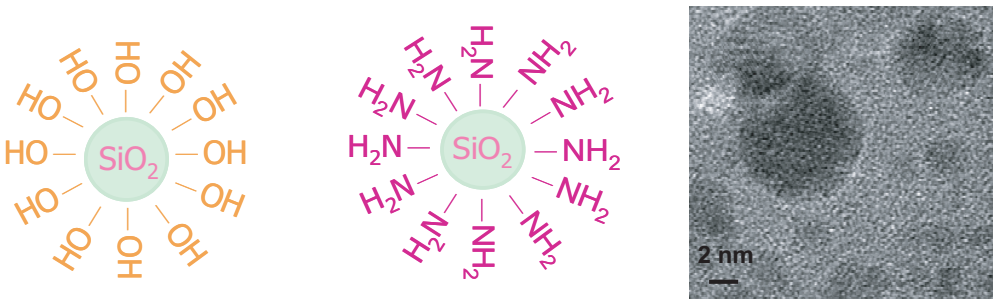


Figure 2: The schematic molecular structures of the studied HPND and APND also shown is the TEM image of the HPND.

The PNDs were prepared by hydrolysis and condensation of sodium metasilicate. [19] To examine the doping effect, APND and HPND with size of 8 nm were synthesized. Figure 2 shows a schematic illustration of the molecular structure and transmission electron microscopic (TEM) image of the synthesized PND. The resultant APND or HPND also exhibited positive or negative charge as determined by the value of their zeta potential measured with a Nano ZS ZEN-3600.

3. RESULT AND DISCUSSION

Figure 3 shows the power efficiency of the blue OLEDs with and without the incorporation of the 8 nm APND or HPND. The power efficiency increased as the PND was employed. Without the

incorporation of PND, the power efficiency at 100 cd/m^2, for example, was 12.0 lm/W. The power efficiency became 25.9 lm/W, an increase of 116%, as APND was added. By employing HPND, the power efficiency was 21.7 lm/W.

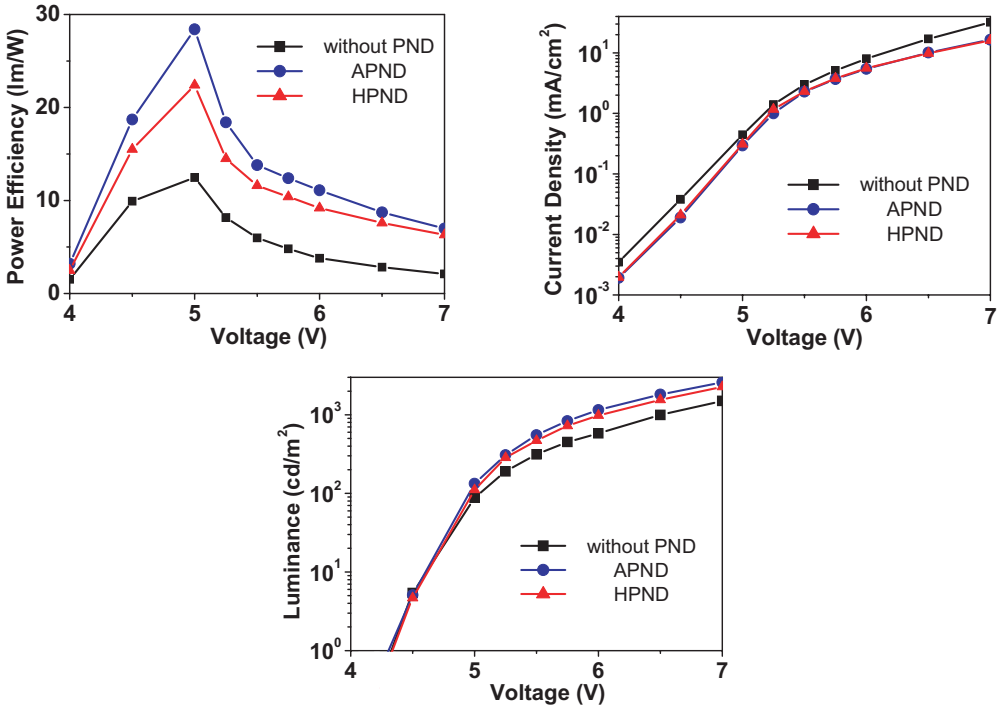


Figure 3: Doping effects of the 8 nm PND on the power efficiency, current density and luminance of the blue OLEDs.

Figure 3 also shows the effects of the employed PND on the current density and luminance of the blue OLEDs. The current density decreased as APND or HPND was added, indicating that the PND had effectively reduced the injection of hole-carrier. The incorporation of titanium oxide nano-dot in a separated layer of a green OLED was found to enhance the injection of hole caused by tunneling effect as revealed by the marked increase of current density and decrease of turn-on voltage. [11–13] However, in the present work the turn-on voltage did not change much with the incorporation of PND with various different concentrations, revealing the absence of tunneling effect. The size of the PND, 8 nm, was much smaller than the 35 nm thickness of the PEDOT:PSS HTL, so that the PND was presumably well embedded within the HTL. The turn-on voltage described herein was defined as the voltage at which the luminance is equal to or greater than 10 cd/m^2. The resultant luminance, especially at voltage between 4.5 to 6 V, did not decrease, but increased obviously with the incorporation of APND or HPND. This indicates that higher carrier-recombination efficiency was resulted from the addition of the PND, since its corresponding current density was comparatively lower than that of its counterpart without PND incorporation.

The effect of concentration of PND on the EL characteristics of the blue OLEDs was shown in Table 1. The power efficiency at 100 cd/m^2, for example, increased from 12.0 to 20.3 lm/W as 0.7 wt% APND was added. It was further increased to 25.9 lm/W as 7.0 wt% APND was incorporated. By increasing the PND concentration to 70 wt%, the power efficiency dropped to 15.4 lm/W. Similarly, the power efficiency was strongly depended on the concentration of the incorporated HPND. Moreover, the chromaticity coordinate at 100 cd/m^2, for example, was (0.19, 0.34), barely changed in the presence of the PND, as also shown in Table 1.

These results indicate that the marked efficiency improvement may be attributed to a better balance of carrier-injection resulted from the hole-blocking-function possessed APND and the hole-trapping-function possessed HPND, which respectively exhibited positive and negative charge on the surface. Importantly, since the nano-dot was not employed in the emissive layer, the same concept may be applied to fluorescent type OLEDs.

Table 1: The effects of size and concentration of PND and thickness of ETL on the EL characteristics of the blue OLEDs.

Type of PND (wt%)	Concentration of PND (wt%)	Driving voltage (V)	Power efficiency (lm/W)**		CIE 1931 chromatic coordinates (x, y)	
			at 100 cd/m^2	max.	at 100 cd/m^2	at 1000 cd/m^2
–	0.0	4.6	12.0	12.5	(0.19, 0.34)	(0.18, 0.34)
	0.7	4.5	20.3	21.1	(0.19, 0.34)	(0.18, 0.34)
APND	7.0	4.5	25.9	28.4	(0.19, 0.34)	(0.18, 0.34)
	70.0	4.6	15.4	18.9	(0.19, 0.34)	(0.18, 0.34)
	0.7	4.6	18.2	18.4	(0.19, 0.34)	(0.18, 0.34)
HPND	7.0	4.5	21.7	22.4	(0.19, 0.34)	(0.18, 0.34)
	70.0	4.6	13.6	15.3	(0.19, 0.34)	(0.18, 0.34)

4. CONCLUSIONS

In conclusion, a novel small PND was synthesized and added in the hole transporting layer, PEDOT:PSS, to markedly improve the efficiency of phosphorescent blue OLEDs. The device efficiency was strongly dependent on the concentration of the PND. The resultant power efficiency at 100 cd/m^2, for example, was increased from 12.0 to 25.9 lm/W, an increase of 116%, as 7 wt% APND of 8 nm in size was added. By employing 7 wt% HPND, the power-efficiency was 21.7 lm/W. These results indicate that the marked efficiency improvement may be attributed to a better balance of carrier-injection resulted from the hole-blocking-function possessed APND and the hole-trapping-function possessed HPND. Importantly, since the nano-dot was not employed in the emissive layer, the same concept may be applied to fluorescent type OLEDs.

ACKNOWLEDGMENT

This work was financially supported under NSC95-2221-E-007-128-MY3, AFOSR-AOARD-05-0488 and BD96013P.

REFERENCES

1. Kido, J., M. Kimura, and K. Nagai, *Science*, Vol. 267, 1332, 1995.
2. Duggal, R., J. J. Shiang, C. M. Heller, and D. F. Foust, *Appl. Phys. Lett.*, Vol. 80, 3470, 2002.
3. Forrest, S. R., *Org. Electron.*, Vol. 4, 45, 2003.
4. D'Andrade, W. and S. R. Forrest, *Adv. Mater.*, Vol. 16, 1585, 2004.
5. Lee, M. T., C. H. Liao, C. H. Tsai, and C. H. Chen, it Adv. Mater., Vol. 17, 2493, 2005.
6. Hung, L. S. and C. H. Chen, *Mat. Sci. Eng. R.*, Vol. 39, 143, 2002.
7. Adamovich, V. I., S. R. Cordero, P. I. Djurovich, A. Tamayo, M. E. Thompson, B. W. D'Andrade, and S. R. Forrest, *Org. Electron.*, Vol. 4, 77, 2003.
8. Xie, Z. Y., L. S. Hung, and S. T. Lee, *Appl. Phys. Lett.*, Vol. 79, 1048, 2001.
9. Jou, J. H., Y. S. Chiu, C. P. Wang, R. Y. Wang, and H. C. Hu, *Appl. Phys. Lett.*, Vol. 88, 193501, 2006.
10. Poon, C. O., F. L. Wong, S. W. Tong, R. Q. Zhang, C. S. Lee, and S. T. Lee, *Appl. Phys. Lett.*, Vol. 83, 1038, 2003.
11. Zhu, F., B. Low, K. Zhang, and S. Chua, *Appl. Phys. Lett.*, Vol. 79, 1205, 2001.
12. Deng, Z. B., X. M. Ding, S. T. Lee, and W. A.Gambling, *Appl. Phys. Lett.*, Vol. 74, 2227, 1999.
13. Zhang, Z. F., Z. B. Deng, C. J. Liang, M. X. Zhang, and D. H. Xu, *Displays*, Vol. 24, 231, 2003.
14. Caruge, J.-M., J. E. Halpert, V. Bulovic, and M. G. Bawendi, *Nano Lett.*, Vol. 6, 2991, 2006.
15. Carter, S. A., J. C. Scott, and P. J. Brock, *Appl. Phys. Lett.*, Vol. 71, 1145, 1997.
16. Bliznyuk, V., B. Ruhstaller, P. J. Brock, U. Scherf, and S. A. Carter, *Adv. Mater.*, Vol. 11, 1257, 1999.

17. Kim, Y. K., K. Y. Lee, O. K. Kwon, D. M. Shin, B. C. Sohn, and J. H. Choi, *Synth. Met.*, Vol. 207, 111–112, 2000.
18. Oey, C. C., A. B. Djurisic, C. Y. Kwong, C. H. Cheung, W. K. Chan, J. M. Nunzi, and P. C. Chui, *Thin Solid Films*, Vol. 492, 253, 2005.
19. Hsu, Y. G., K. H. Lin, and I. L. Chiang, *Mater. Sci. Eng.*, B87, Vol. 31, 2001.
20. Tsai, Y. C., J. H. Jou, *Appl. Phys. Lett.*, Vol. 89, 243521, 2006.
21. Aziz, H. and Z. D. Popvic, *Chem. Mater.*, Vol. 16, 4522, 2004.

Linear Sampling Method: Physical Interpretation and Guidelines for a Successful Application

Ilaria Catapano[1], **Lorenzo Crocco**[1], **and Tommaso Isernia**[2]

[1]National Research Council - Institute for Electromagnetic Sensing of Environment, CNR-IREA Naples, Italy

[2]Mediterranea University of Reggio Calabria, Reggio Calabria, Italy

Abstract— The Linear Sampling Method (LSM) is an effective method to tackle the problem of reconstructing the shape of unknown metallic or dielectric scatterers from the knowledge of single frequency multi-view/multi-static data. Notably, as it just requires to solve a linear problem, its implementation is straightforward and its computational burden almost negligible. However, no results are available in the literature to explain under which operating conditions (e.g., the number of incident waves and receivers which has to be considered) LSM properly works. With respect to the case of dielectric scatterers, in this communication, starting from the physical interpretation of LSM, we provide some guidelines for its successful application. These results are then confirmed processing experimental data from the "Marseille" data-set.

1. INTRODUCTION

The capability of retrieving in an effective way the geometrical features of unknown scatterers from the knowledge of the scattered fields is an important issue in many applications. As a matter of fact, not only location and shape are often the survey's final goal, but they may also represent a valuable starting point for quantitative imaging procedures [1].

An effective way to cope with the shape reconstruction problem is the Linear Sampling Method (LSM) [2], which can provide reconstructions for dielectric and metallic objects without *a priori* information on their physical nature and without any approximation on the scattering model. Moreover, LSM requires a negligible computational burden, as its numerical implementation only involves a matrix, whose size is determined by the amount of available data [2, 3].

However, despite the above mentioned advantages and the large number of examples which assess its reconstruction capabilities against synthetic and experimental data [2–4], it is not completely clear *why* and *how* LSM works and which are its limitations. In particular, no theoretical demonstration has been (yet?) given that the method can work in general [5], whereas only recently some results concerning the resolution which can be achieved have been published [6].

A way to address these questions from a different point of view is that of reasoning on the physical meaning of LSM. In this respect, an interpretation of LSM as applied to perfect electric conductor targets has been provided by Shelton and Warnick [7], while the case of dielectric scatterers has been studied in [4], where it has been shown that LSM is related to the possibility or the impossibility of focusing an electromagnetic wave in presence of an obstacle (i.e., the scatterer). In this contribution, by exploiting such an analogy, we propose some guidelines to fix the number of transmitting and receiving probes which are needed to ensure a successful application of LSM, while reducing redundant measurements.

The paper is organized as follows. In the next section the basic concepts of the LSM are briefly recalled as well as its physical interpretation; in Section 3 the guidelines to fix the number of illuminations are discussed and given and, finally, in Section 4 they are proved by processing experimental data coming from the "Marseille" data-set [9].

2. A BRIEF REVIEW OF THE LINEAR SAMPLING METHOD

Let us consider the reference scenario sketched in Fig. 1. The targets are supposed in free space and invariant along the z-axis, while a TM polarization is assumed for the electromagnetic fields. Ω denotes the investigated region and $\Sigma \subset \Omega$ the (possibly not connected) support of dielectric scatterers. Measurement probes and primary sources are displaced on a circumference Γ located in the far-zone with respect to Ω.

For a fixed frequency, in order to apply LSM, one first fixes a set of arbitrary sampling points $r_p \in \Lambda \subseteq \Omega$ and then solves, in each r_p, the linear integral equation [2]:

$$\int_{\Gamma} E_{\infty}(\phi,\theta) g(\theta, r_p) ds(\theta) = -\frac{j}{\sqrt{8\pi kR}} \exp\left[-jk\rho\cos(\phi-\alpha) + j\frac{\pi}{4}\right] = f(r_p,\phi), \qquad (1)$$

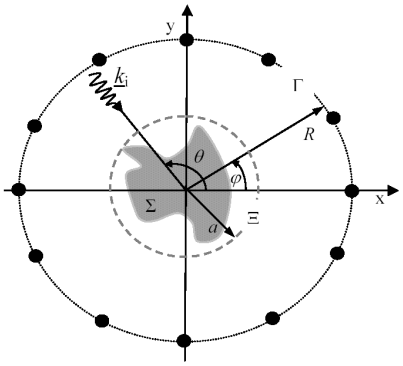


Figure 1: Reference geometry.

k being the background wave-number and $E_\infty(\phi,\theta)$ being the far field pattern of the scattered field as measured on Γ in the direction ϕ when a plane wave impinges from the direction θ [10].

The claim and the numerical observation is that the norm of the regularized (i.e., approximated) solution of Eq. (1) will become large when r_p approaches the boundary of the scatterer from inside and will stay large when r_p is outside of Σ [2]. Therefore, the geometrical features of the scattering system (i.e., its location, size and shape) can be easily reconstructed by solving Eq. (1) $\forall r_p \in \Lambda$ and then by plotting the L^2 norm $||g||$ of the solution.

When the scatterers are dielectric, it has been proved [4] that LSM is interpretable as the problem of focusing an electromagnetic field in presence of an obstacle. In particular, apart from special cases (see [4] for a detailed discussion), solving Eq. (1) is equivalent to induce a current focused in r_p by combining the primary sources according to a function $g(r_p)$, whose L^2 norm is hence proportional to the energy of the resulting source. Since the induced current is generally expressed as $J(r_p) = j\omega\chi(r_p)E(r_p)$, χ being the contrast function and E the total internal field, any induced current would be necessarily null if $r_p \notin \Sigma$, as $\chi(r_p \notin \Sigma) = 0$. Therefore, the divergent behavior of $||g||$ in those points corresponds to the impossible task of combining the primary sources in such a way to induce a current in a location where this latter is *de facto* null. On the other hand, due to intrinsic resolution limitations, the indicator $||g||$ results to be actually unbounded everywhere, as the realization of an induced current *exactly* localized in r_p at a fixed frequency requires a primary source of infinite energy. Therefore, in order to prevent primary sources from having an infinite energy regardless of the sampling point, it is necessary to solve Eq. (1) in a regularized form, as also demonstrated by Arens [5]. As a result, the desired focusing is pursued in an *approximated* way since the induced current will not be exactly located in the sampling point but in an approximately circular neighbor of it [4].

3. GUIDELINES FOR SUCCESSFUL APPLICATION OF THE LSM

Given the analogy between LSM and focusing problems, it follows that an important role in successful reconstruction of shapes is played by the number of primary sources and measurement points. As a matter of fact, if the number of transmitters is low, one could not achieve, even in the full aperture case, the desired focused field, while if the number of receivers is low, one could not properly control that the synthesized scattered field matches to the desired one. On the other hand, an arbitrary large number of illuminations and measurements, which would overcome these problems, can be largely redundant and may unnecessarily increase the measurements costs. In the following, we derive some guidelines to fix in a convenient way the number of transmitters and receivers.

Let us suppose that the radius a of the minimum circle Ξ enclosing the scatterers (and centered at the origin of the coordinate system, see Fig. 1) is known or has been estimated using the simple strategy described in [8]. Then, for any primary source J_p on Γ, the incident field E_{inc} inside Ξ can be expressed through the linear operator:

$$A_p : J_p(\theta) \in L^2(\Gamma) \rightarrow E_{inc}(r) = \frac{-j}{\sqrt{8\pi kR}} \int_{-\pi}^{\pi} J_p(\theta) \exp\left\{-jk[R + r\cos(\theta - \alpha)]\right\} d\theta \in L^2(\Xi) \quad (2)$$

and useful indications on the number and kind of primary sources can be gained from the Singular Value Decomposition (SVD) of the operator A_p. As a matter of fact, by means of this expansion,

we can express any incident field, as:

$$E_{inc}(r) = A_p[J_p] = \sum_{n=-\infty}^{\infty} \sigma_n \left[\int_{-\pi}^{\pi} J_p(\theta) u_n^*(\theta) d\theta \right] v_n(r) \tag{3}$$

wherein σ_n are the singular values, v_n the left singular functions and u_n are the right singular functions, which form a basis for the primary sources.

For the considered geometry, quantities appearing in Eq. (3) are known in a closed form [11] and, in particular:

$$\hat{\sigma}_n^2 = \frac{\sigma_n^2}{\sigma_0^2} = \frac{[J_n(ka)]^2 - J_{n-1}(ka) J_{n+1}(k_a)}{[J_0(ka)]^2 + [J_1(ka)]^2} \tag{4}$$

wherein J_n is the n-th Bessel function and

$$u_n(\theta) = \frac{1}{\sqrt{(2\pi)}} \exp(jn\theta). \tag{5}$$

For several values of the radius a, the singular values σ_n are plotted in Fig. 2. Note, they are a fast decreasing function of $|n|$, if $n > k_a$. Such a circumstance has two interesting consequences:

- provided an upper bound is given on the energy of primary sources (as in our regularized framework), only a finite number of terms in (3) is needed to achieve a given incident field with any required accuracy. Fig. 2 suggests such a number is $M = 2\gamma ka$, with $\gamma > 1$. The value of γ has to be fixed on the basis of the accuracy of the measurement set-up;
- the resulting M singular functions $u_1, \ldots, u_M$ provide the convenient set of primary sources; however, since any linear combination of exponential functions can be turned into a Dirichelet sampling series, an equivalent and simpler set of convenient primary sources is given by M elementary filamentary sources uniformly spaced in angle along Γ.

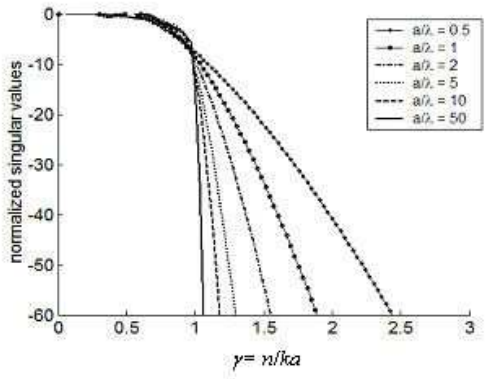


Figure 2: Behavior of the normalized singular values of A_p.

In a dual fashion, to fix the number and kind of measurement probes, we can inspect the SVD of the operator A_e, which relates the current induced in Ξ to the scattered field on Γ. However, provided transmitters and receivers are located in the same positions, from reciprocity it follows that the singular values of A_e are still given by Eq. (4) and its (left) singular functions are simply u_n^*. Therefore, $N_m = M$ filamentary probes uniformly spaced in angle along Γ represent the convenient choice for the receiving probes.

As a consequence of all the above, from the estimate of the radius of the minimum circle a, one can foreseen success or failure of the regularized LSM, as N_v or N_m have to be at least larger than $2ka$ in order to properly discretize the far-field equation. On the other hand, an upper bound to the number of probes can also be given reasoning on the factor γ. As a matter of fact, as it can be deduced from Fig. 2, if the scatterers' convex envelope is large with respect to the wavelength (i.e., in the high frequency regime) the singular values steeply decay, so that $\gamma \approx 1$ is a convenient choice, regardless of the required accuracy. Conversely, for scatterer enclosed in a convex envelope whose

size is comparable to the wavelength (i.e., in the "resonant region"), the singular values exhibit a slower decay and a larger value of γ will be needed in order to take into account the singular values belonging to the transition region $ka < n < 2ka$. In particular, while the optimal choice for γ depends on the desired accuracy, $\gamma \approx 2$ is a generally convenient one.

4. NUMERICAL EXAMPLES

Let $\mathbf{E}_s$ be the $N_v \times N_m$ (N_v and N_m being the number of transmitters and receivers, respectively) multistatic data matrix and $\{\mu_n, \lambda_n, \rho_n\}$ its SVD, in each sampling point $r_p \in \Lambda$, the support indicator is given by:

$$||\mathbf{g}(r_p)||^2 = \sum_{n=1}^{N} \left(\frac{\lambda_n}{\lambda_n^2 + \alpha} \right)^2 |\langle \mathbf{f}, \rho_n \rangle|^2 \tag{6}$$

where $N = \min\{N_v, N_m\}$, f is the N_m dimensional vector containing the values of the right hand side of Eq. (1) and α is the Tikhonov regularizing parameter. According to the proposed physical interpretation, such a parameter is the same for all sampling points [4] and it is heuristically fixed equal to $0.01\lambda_1$.

To provide an example of the above discussed guidelines, let us consider the "Marseille" experimental data-set concerning of two plexiglass cylinders ($\varepsilon = 3$) of radius $R_c = 0.015\,\mathrm{m}$. This data-set is related to an aspect-limited configuration in which the primary source is moved along a circumference with a 10 angular step ($N_v = 36$), and for each illumination, the measurement probe is moved with an angular step of 5° ($N_m = 49$) along a 240° arc, which excludes the 120° angular sector centered around the incidence direction. Data collected in the (1–8) GHz frequency range with 1 GHz step and in the (8–16) GHz range with a 4 GHz step are available [9].

The radius of the minimum circle containing the scatterers is equal to $a = 0.06\,\mathrm{m}$, therefore the electric size of the convex envelope a/λ ranges from 0.2 at 1 GHz to 3.2 at 16 GHz. From Fig. 2 one can thus infer that LSM is expected not work at frequencies higher than 10GHz, being the number of available sources lower than the one fixed according to the criteria in Section 3. Such an expectation is confirmed by the result drawn in Fig. 3(a), which shows that LSM badly fails. On the other hand, when considering data at 8 GHz, as the given guidelines are better matched ($\gamma \approx 1.8$), the LSM indeed reconstructs the unknown shapes, Fig. 3(b), although a slight overestimation is

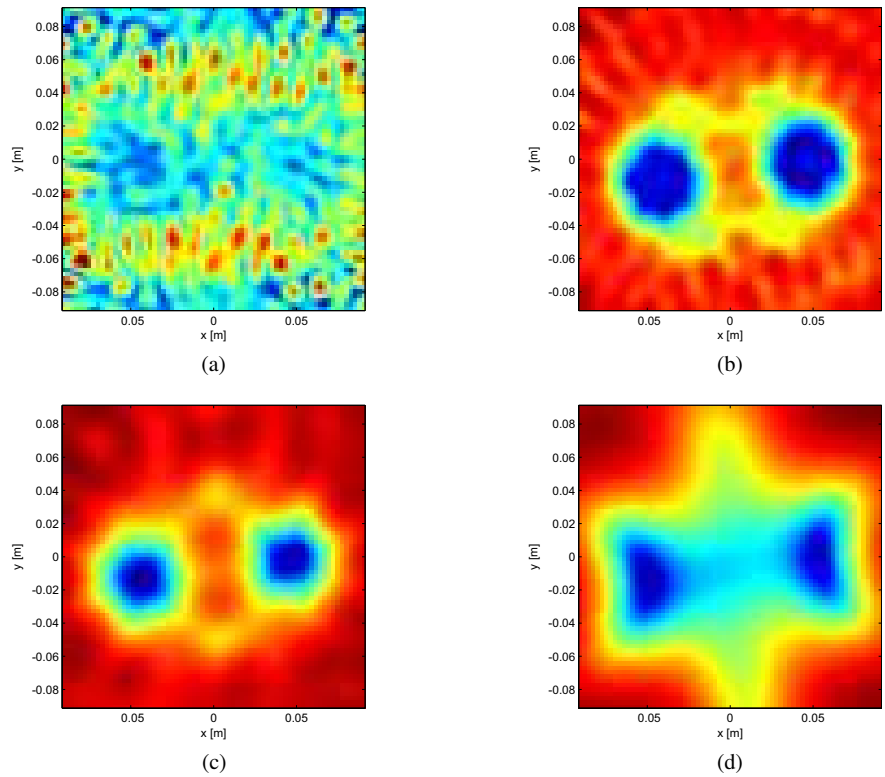


Figure 3: LSM reconstructed profile. (a) 12 GHz, (b) 8 GHz, (c) 5 GHz, (d) 2 GHz.

observed, possibly due to measurement noise. However, a further improvement is obtained when considering data at the "resonant" frequency of 5 GHz, wherein, being $\gamma > 2$ and being lower the measurement error, a very accurate reconstruction is achieved, see Fig. 3(c). Finally, it is interesting to observe, see Fig. 3(d), the result obtained at 2 GHz, when the reconstruction worsens, as the electrical distance between the two objects is lower than the achievable LSM resolution as defined by [6].

5. CONCLUSION

In this paper the basic concepts of the LSM and its physical interpretation have been briefly reviewed and some guidelines for its successful application have been traced. The validity of these guidelines has been verified against experimental data concerning multiple targets.

REFERENCES

1. Catapano, I., L. Crocco, M. D. Urso, and T. Isernia, "On the effect of support estimation and of a new model in 2D inverse scattering problems," *IEEE Trans. Antennas Propagat.*, Vol. 55, 1895–1899, 2007.
2. Colton, D., H. Haddar, and M. Piana, "The linear sampling method in inverse electromagnetic scattering theory," *Inv. Probl.*, Vol. 19, 105–137, 2003.
3. Collino, F., M. B. Fares, and H. Haddar, "Numerical and analytical studies of the linear sampling method in electromagnetic inverse scattering problems," *Inv. Probl.*, Vol. 19, 1279–1298, 2003.
4. Catapano, I., L. Crocco, and T. Isernia, "On simple methods for shape reconstruction of unknown scatterers," *IEEE Trans. Antennas Propagat.*, Vol. 55, 1431–1436, 2007.
5. Arens, T., "Why linear sampling works," *Inv. Probl.*, Vol. 20, 163–173, 2004.
6. Aramini, R., M. Brigone, and M. Piana, "The linear sampling method without sampling," *Inv. Probl.*, Vol. 22, 2237–2254, 2006.
7. Shelton, N. and K. F.Warnick, "Behavior of the regularized sampling inverse scattering method at internal resonance frequencies," *Progress In Electromagnetics Research*, Vol. 38, 29–45, 2002.
8. Crocco, L., M. D. Urso, and T. Isernia, "Testing the contrast source extended born method against real data: the TM case," *Inv. Probl.*, Vol. 21, S33–S50, 2005.
9. Belkebir, K. and M. Saillard, "Special section: testing inversion algorithms against experimental data," *Inv. Probl.*, Vol. 17, 1565–2028, 2001.
10. Colton, D. and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer-Verlag, Berlin, Germany, 1992.
11. Bucci, O. M., L. Crocco, and T. Isernia, "Improving the reconstruction capabilties in inverse scattering problems by exploitation of close-proximity setup," *J. Opt. Soc. Am. A*, Vol. 16, 1788–1798, 1999.

Soil Dielectric Model Accounting for Contribution of Bound Water Spectra through Clay Content

V. L. Mironov, L. G. Kosolapova, and S. V. Fomin
Kirensky Institute of Physics SB RAS, Russia

Abstract— In this paper, there was studied correlation of dielectric predictions for moist soils with the measured values, regarding the well known and prevalent in common usage semiempirical mixing dielectric model (SMDM) proposed by M. C. Dobson et al., on the one hand, and the generalized refractive mixing dielectric model (GRMDM) proposed by V. L. Mironov et al., on the other hand. The analysis is based on the measured dielectric data borrowed from the Technical Report EL-95-34, by J. O. Curtis et al., in which the assemblage of soils measured included all of grain-size distributions that are observed in nature, with measurements being performed over the frequency range from 45 MHz to 26.5 GHz at the moistures spanning from the nearly dry samples to the ones saturated close to total moisture capacity. The GRMDM was found to provide for much more accurate dielectric predictions. To make the GRMDM as convenient for the users, as the semiempirical model is, which uses only clay and sand contents as the only input parameters, the spectroscopic parameters of the GRMDM were linked to the clay content. As a result, a new type of mineralogy based soil dielectric model (MBSDM) was developed to provide for much better accuracy of dielectric predictions over a broader variety of soil types, as compared to those of the SMDM.

1. INTRODUCTION

At present, the SMDM proposed in [1–3] has become a most preferable mean for predicting the dielectric spectra of moist soils in the microwave band. To account for frequency dispersion of moist soil permittivity, the SMDM uses the Debye relaxation spectrum of liquid water, which is located out of soil. Assuming soil water permittivity independent on soil type, it takes into account the impact of soil mineralogy on moist soil dielectric spectra by modifying the dependence of soil permittivity on volumetric moisture with the use of specific regression parameters dependent on clay and sand contents.

Recently, a generalized refractive mixing dielectric model (GRMDM) for moist soils was developed in [4], which employs the refractive mixing dielectric model (RMDM) suggested in [5]. In comparison with [5], the RMDM was modified [4, 6] so as to distinguish between two types of soil water (bound and free), with their permittivities being separately determined from dielectric measurements. In addition, a methodology was suggested to derive the maximum bound water fraction (MBWF) regarding a specific type of soil, which was first treated in [7] as the transition soil moisture. In [4], the microwave dielectric spectra of both bound and free soil water were shown to follow the Debye relaxation formulas, with a procedure of deriving Debye relaxation parameters and ohmic conductivity for both bound and free types of soil water having been proposed.

As a result, dielectric spectra of moist soils appeared to be predictable for a given type of soil as a function of moisture and temperature on the basis of spectroscopic parameters [8] relating to dry soil and both types of soil water. In this paper, the GRMDM spectroscopic parameters were correlated with the values of clay content, covering the whole variety of soil types presented in [1], [9], and [10]. As a result, there was developed a mineralogy based soil dielectric model (MBSDM), which similar to the SMDM provides for dielectric spectra predictions for moist soils using only granulometric mineralogy data, that is, clay content by mass. This model was found to provide for less error of predictions, compared to the SMDM, over a broader assemblage of soils.

2. THE SMDM CONCEPT

The semiempirical mixing dielectric model was developed in [1–3] on the bases of dielectric data covering five soil types, a wide range of moisture conditions, and two frequency ranges extending from 0.3 to 1.3 GHz and from 1.4 to 18 GHz [1]. The SMDM has the following form:

$$\varepsilon'_m = \left[1 + \frac{\rho_b}{\rho_s}(\varepsilon'^{\alpha}_s - 1) + m_v^{\beta'}\varepsilon'^{\alpha}_{fw} - m_v\right]^{1/\alpha}, \quad \text{and} \quad \varepsilon''_m = \left[m_v^{\beta''}\varepsilon''^{\alpha}_{fw}\right]^{1/\alpha}, \tag{1}$$

where ε'_m and ε''_m are respectively the dielectric constant (DC) and loss factor (LF) of moist soil, ε'_s is the DC of soil solids, m_v is the volumetric moisture content, ρ_b is the bulk density in g/cm^3, ρ_s is the specific density of the soil solids. The quantities ε'_{fw} and ε''_{fw} are the DC and LF of free water, which can be calculated with the Debye like dielectric relaxation formulas. The values $\alpha = 0.65$, β', and β'' are empirically determined constant. To account for moist soil conductivity [1–3], the dielectric relaxation formulas were taken in the form

$$\varepsilon'_{fw} = \varepsilon_{w\infty} + \frac{\varepsilon_{w0} - \varepsilon_{w\infty}}{1 + (2\pi f \tau_w)^2}, \quad \text{and} \quad \varepsilon''_{fw} = \frac{2\pi f \tau_w (\varepsilon_{w0} - \varepsilon_{w\infty})}{1 + (2\pi f \tau_w)^2} + \frac{\sigma_{eff}}{2\pi\varepsilon_0 f} \frac{(\rho_s - \rho_b)}{\rho_s m_v} \tag{2}$$

where $\varepsilon_0 = 8.854 \times 10^{-12}$ F/m is the permittivity of free space, f is the frequency in Hz, ε_{w0} and $\varepsilon_{w\infty} = 4.9$ are the low and high frequency limits for the free water dielectric constants, respectively, τ_w is the relaxation time for free water, σ_{eff} is the effective conductivity parameter. Expressions for ε_{w0} and τ_w are given as a function of temperature by Ulaby et al. [11, Appendix E-2]. At room temperature (20°C), $2\pi\tau_w = 0.58 \times 10^{-10}$ s and $\varepsilon_{w0} = 80.1$.

All the parameters in (1), (2), which were empirically determined in [1–3], are presented in Table 1.

Table 1.

$\alpha = 0.65$	$\varepsilon'_s = (1.01 + 0.44\rho_s)^2 - 0.062$	$\beta' = 1.275 - 0.519S - 0.152C$	$\beta'' = 1.338 - 0.603S - 0.166C$
For $0.3 < f < 1.3$ GHz	$\sigma_{eff} = 0.0467 + 0.2204\rho_b - 0.4111S + 0.6614C$		
For $1.4 < f < 18$ GHz	$\sigma_{eff} = -1.645 + 1.939\rho_b - 2.25622S + 1.594C$		

In Table 1, S and C represent in percentage the mass fraction of sand and clay, respectively. In addition, in the frequency range from 0.3 to 1.3 GHz. the DC of moist soil, $\varepsilon'_{m,\,0.3-1.3}$, is to be calculated using the equation $\varepsilon'_{m,\,0.3-1.3} = 1.15\varepsilon'_m - 0.68$, in which ε'_m is given by expression (1).

It is worth noticing that all the SMDM input parameters relating to a given soil type are available from the soil granulometric mineralogy data, which can be borrowed from soil databases. This is a major advantage of the SMDM over many other models, resulting in a wide employment of that model in the radar and radio thermal remote sensing.

3. THE MINERALOGY BASED SOIL DIELECTRIC MODEL CONCEPT

In contrast to the SMDM, exclusively employing dielectric relaxation spectrum valid for the water located out of soil, the GRMDM suggested in [4] employs the dielectric spectra explicitly related to either bound or free soil water. The description of this concept is given below.

In accordance with [4], the DC, ε'_m, and LF, ε''_m, as a function of volumetric moisture, m_v, can be represented in the form of the RMDM:

$$\varepsilon'_m = n_m^2 - \kappa_m^2, \qquad \varepsilon''_m = 2n_m\kappa_m \tag{3}$$

$$n_m = \begin{cases} n_d + (n_b - 1)\, m_v, & m_v \le m_{vt} \\ n_d + (n_b - 1)\ m_{vt} + (n_u - 1)\ (m_v - m_{vt}), & m_v \ge m_{vt} \end{cases}$$

$$\kappa_m = \begin{cases} \kappa_d + \kappa_b\, m_v, & m_v \le m_{vt} \\ \kappa_d + \kappa_b\, m_{vt} + \kappa_u\ (m_v - m_{vt}), & m_v \ge m_{vt}. \end{cases} \tag{4}$$

where n_m, n_d, n_b, n_u, and κ_m, κ_d, κ_b, κ_u are the values of refractive index and normalized attenuation coefficient, which is understood here as a proportion of the standard attenuation coefficient to the free space propagation constant. The subscripts m, d, b, and u in (3), (4) and further on stand for moist soil, dry soil, bound soil water (BSW), and free soil water (FSW), respectively, and m_{vt} is a value of the maximum bound water fraction (MBWF) in a given type of the soil. The latter depends on the soil mineral contents [4, 8]. The DC and LF for bound, ε'_b, and free, ε'_u, water components are presented with the Debye relaxation equations

$$\varepsilon'_{b,\,u} = \varepsilon_\infty + \frac{\varepsilon_{b0,\,u0} - \varepsilon_\infty}{1 + (2\pi f \tau_{b,\,u})^2}, \qquad \varepsilon''_{b,\,u} = \frac{\varepsilon_{b0,\,u0} - \varepsilon_\infty}{1 + (2\pi f \tau_{b,\,u})^2} 2\pi f \tau_{b,\,u} + \frac{\sigma_{b,\,u}}{2\pi\varepsilon_0 f}. \tag{5}$$

In formulas (5), the value f designates wave frequency, while the values $\sigma_{b,u}$, $\tau_{b,u}$, and $\varepsilon_{0b,u}$ are conductivities, relaxation times, and low frequency limit dielectric constants, relating to either BSW or FSW components. The value ε_0 is the DC for free space, while $\varepsilon_\infty = 4,9$ represents the DC in the high-frequency limit for both types of soil water. As can be seen from Equations (3)–(5), a certain type of moist soil, in terms of its dielectric spectra, can be completely determined via a set of the following spectroscopic parameters: DC, $\varepsilon\prime_d$, for dry soil; LF, ε''_d, for dry soil; value of maximum bound water fraction, m_{vt}; low frequency limit dielectric constants, ε_{0b} and ε_{0u}, for bound and free soil water; relaxation times, τ_b, and τ_u, for bound and free soil water; conductivities, σ_b, and σ_u, for bound and free soil water.

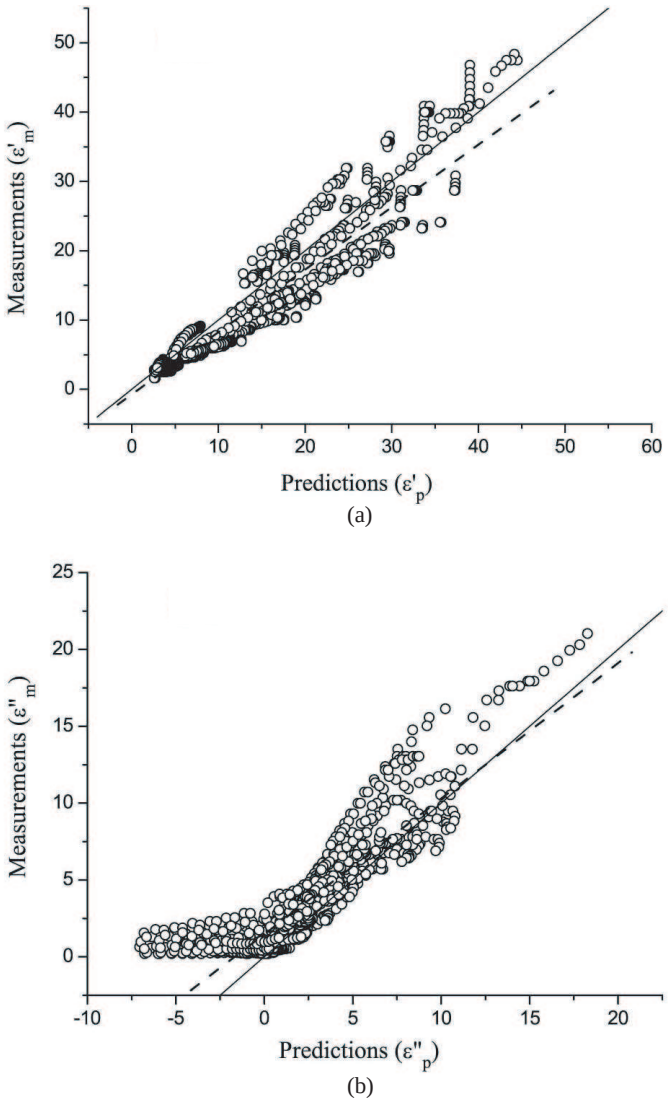


Figure 1: Correlation of the SMDM predictions, ε'_p, ε''_p, for DCs (a) and LFs (b) with the measured ones, ε'_m, ε''_m, in the case of soils studied in [9]. Solid and dotted lines represent bisectors and linear fits, respectively. Correlation coefficients, R_{DC} and R_{LF}, and standard deviations, SD_{DC} and SD_{LF}, appeared to be as follows: $R^2_{DC} = 0.942$, $R^2_{LF} = 0.882$, $SD_{DC} = 3.391$, $SD_{LF} = 1.695$. The linear fits were found to be: $\varepsilon'_m = -0.753 + 0.902\varepsilon'_p$, $\varepsilon''_m = 1.483 + 0.881\varepsilon''_p$.

For a specific type of soil, all of these parameters can be derived with the use of conventional dielectric measurements regarding moist soils, as given in [4]. Therefore, to be employed in the microwave remote sensing, this model requires a set of prior dielectric measurements for the soils involved to be carried out. This is a much more laborious way to have dielectric predictions, as compared with the SMDM, which uses for that purpose only conventional granulometric mineralogy characteristics, namely, gravimetric clay and sand contents.

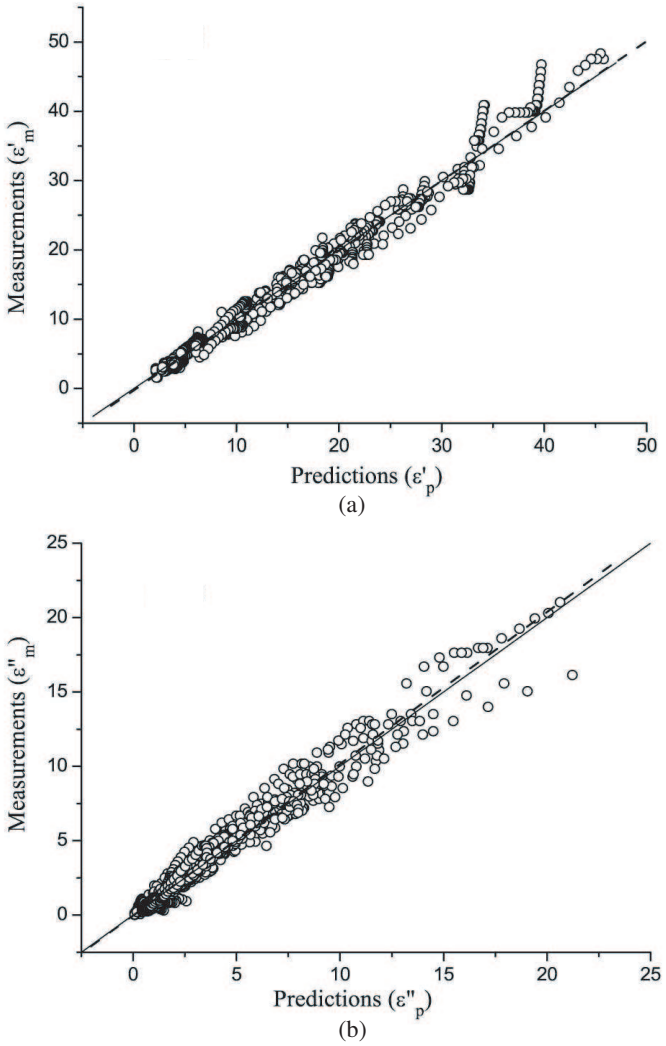


Figure 2: Correlation of the MBSDM predictions, $\varepsilon'_p,\ \varepsilon''_p$, for DCs (a) and LFs (b) with the measured ones, $\varepsilon'_m,\ \varepsilon''_m$, in the case of soils studied in both [9] and [1], [10]. Solid and dotted lines represent bisectors and linear fits, respectively. Correlation coefficients, $\mathrm{R_{DC}}$ and $\mathrm{R_{LF}}$, and standard deviations, $\mathrm{SD_{DC}}$ and $\mathrm{SD_{LF}}$, appeared to be as follows: $\mathrm{R^2_{DC}} = 0.992,\ \mathrm{R^2_{LF}} = 0.983,\ \mathrm{SD_{DC}} = 1.259,\ \mathrm{SD_{LF}} = 0.64$. The linear fits were found to be: $\varepsilon'_m = -0.216 + 1.007\varepsilon'_p,\ \varepsilon''_m = -0.01765 + 1.019\varepsilon''_p$.

In this research, the GRMDM spectroscopic parameters were obtained, using the dielectric data available in [9] and [1], [10]. These dielectric data were measured over the frequency ranges from 45 MHz to 26.5 GHz, in [9], and from 1.4 GHz to 18 GHz, in [1, 10], with the moistures spanning from the nearly dry samples to the ones saturated close to total moisture capacity. The spectroscopic parameters were derived with the use of dielectric data in the frequency ranges from 0.3 to 26.5 GHz and from 1.4 to 18 GHz, relating to [9] and [1], [10], respectively. In the analysis conducted, there were involved 11 and 4 types of soil measured in [9] and [1], [10], respectively. The clay content in the soils varied from close to 0% to 54%. All the data available were measured in the temperature range from 20°C to 22°C.The spectroscopic parameters obtained were correlated to clay contents regarding the respective soil types. The result of correlation analysis are given in Table 2 in a form of regression equations where the clay content, C, and relaxation times, τ_b and τ_u, are expressed in percentages and seconds, respectively. According to the granulometric classification of soils by the USDA, the value C was determined as a fraction by mass of the soil solids having effective size less then 2 microns.

Further on, the formulas (3)–(5) in conjunction with the equations presented in Table 2 is identified as the mineralogy based soil dielectric model (MBSDM), while all the values involved are referred to as the dielectric spectroscopic parameters.

Table 2.

$n_d = 1.634 - 0.539 \cdot 10^{-2}C$ $+0.2748 \cdot 10^{-4}C^2$	$\varepsilon_{0b} = 79.8 - 85.4 \cdot 10^{-2}C + 32.7 \cdot 10^{-4}C^2$	$\sigma_u = 0.3631 + 1.217 \cdot 10^{-2}C$
$k_d = 0.03952 - 0.04038 \cdot 10^{-2}C$	$\tau_b = 1.062 \cdot 10^{-11} + 3.450 \cdot 10^{-12} \cdot 10^{-2}C$	$\varepsilon_{0u} = 100$
$W_t = 0.02863 + 0.30673 \cdot 10^{-2}C$	$\sigma_b = 0.3112 + 0.467 \cdot 10^{-2}C$	$\tau_u = 8.5 \cdot 10^{-12}$

4. THE MBSDM AND SMDM COMPARATIVE VALIDATION

To estimate the correlation between DCs and LFs predicted, using both the SMDM and MBSDM, with the ones measured, the experimental DCs and LFs were plotted in Figs. 1 and 2 versus predicted ones for the assemblages of soils available in [9] and [1], [10], respectively. The correlation analysis presented in Figs. 1 and 2 signifies much better accuracy of dielectric predictions, in terms of both the standard deviation (SD) and correlation coefficient (R^2), provided by the MBSDM as compared with those of the SMDM.

5. CONCLUSIONS

Summing up, it can be stated that the MBSDM proved to be a dielectric model capable to predict permittivity of moist soils with good accuracy, using a clay content in the soil as the only input parameter of granulometric mineralogy origin. From this view, the MBSDM retained accuracy of the GRMDM, complimenting this feature with the simplicity of its application, which is as simple as that of the SMDM.

ACKNOWLEDGMENT

The work was supported by the Siberian Branch of the Russian Academy of Sciences, Project 70, and by the RFBR grant No. 05-02-97712-r-yenisei-a.

REFERENCES

1. Dobson, M. C., F. T. Ulaby, M. T. Hallikainen, and M. A. El-Rayes, "Microwave dielectric behavior of wet soil — Part II: Dielectric mixing models," *IEEE Trans. Geosci. Remote Sensing*, Vol. 23, No. 1, 35–46, 1985.
2. Peplinski, N. A., F. T. Ulaby, and M. C. Dobson, "Dielectric properties of soils in the 0.3–1.3 GHz range," *IEEE Trans. Geosci. Remote Sensing*, Vol. 33, No. 3, 803–807, 1995.
3. Peplinski, N. A., F. T. Ulaby, and M. C. Dobson, "Correction to 'Dielectric properties of soils in the 0.3–1.3 GHz range'," *IEEE Trans. Geosci. Remote Sensing*, Vol. 33, No. 6, 1340, 1995.
4. Mironov, V. L., M. C. Dobson, V. H. Kaupp, S. A. Komarov, and V. N. Kleshchenko, "Generalized refractive mixing dielectric model for moist soils," *IEEE Trans. Geosci. Remote Sensing*, Vol. 42, No. 4, 773–785, 2004.
5. Birchak, J. R., C. G. Gardner, J. E. Hipp, and J. M. Victor, "High dielectric constant microwave probes for sensing soil moisture," *Proc. IEEE*, Vol. 62, No. 1, 92–98, 1974.
6. Mironov, V. L., S. A. Komarov, N. V. Rychkova, and V. N. Kleshchenko, "Study of the dielectric properties of wet grounds at microwave frequencies," *Earth Obs. Rem. Sens.*, Vol. 12, No. 4, 495–504, 1995.
7. Wang, J. R. and T. J. Shmugge, "An empirical model for the complex dielectric permittivity of soils as a function of water content," *IEEE Trans. Geosci. Remote Sensing*, Vol. 18, No. 4, 288–295, 1980.
8. Mironov, V. L., "Spectral dielectric properties of moist soils in the microwave band," *Proceedings IGARSS*, Vol. 5, 3474–3477, Anchorage, USA, 2004.
9. Curtis, J. O., C. A. Weiss, Jr., and J. B. Everett, "Effect of soil composition on dielectric properties," Technical Report EL-95-34, December 1995.
10. Hallikainen, M. T., F. T. Ulaby, M. C. Dobson, M. A. El-Rayes, and L.-K. Wu, "Microwave dielectric behavior of wet soil — Part I: Empirical model and experimental observations," *IEEE Trans. Geosci. Remote Sensing*, Vol. 23, No. 1, 25–34, 1985.
11. Ulaby, F. T., R. K. Moor, and A. K. Fung, *Microwave Remote Sensing, Active and Passive*, Vol. 3, Artech House, Dedham, MA, 1986.

Monitoring of Satellite Thermal Patch Formed by A Wave Facet Ocean Surface Water Waves

S. Nakamura
Kyoto University, Japan

Abstract— In a case of monitoring a satellite thermal patch, frequently found a wave facets on the sea surface. Assuming a concave facet of a part of the sea surface waves, a simple model is introduced in order to realize that the facet acts an infrared beam out of the ocean surface as a black body. Supporting satellite thermal patterns has been obtained by a direct receiving of the satellite signals. A note is given in relation to special spectrum of the sea surface waves.

1. INTRODUCTION

In a case of monitoring a satellite thermal patch which is reduced from a satellite signal of an infrared band directly received at a station settled on the coast facing the ocean, it is frequently found a set of concave wave facets on the sea surface is formed by a coupling of the two water waves. When several conditions are satisfied, the wave facet on the sea surface can be a radiator of infrared beam out of the ocean as a black body. The beam is possibly to have a radius of curvature approximately same to the distance between the sea surface of the interested area and the sensor mounted on a satellite.

In this work, a satellite thermal patch problem is concentrated to one case of coupling ocean surface wave facet which has a focusing point or a caustics formation of the infrared beam out of the wave just neighbor the satellite.

Supporting satellite thermal patterns has been obtained by a directly receiving of satellite signals at a station settled on the coast facing the ocean.

2. ASSUMPTION OF OCEAN SURFACE WAVES

In order to realize a satellite thermal patch, it is assumed that an ocean surface water wave for convenience.

When an arbitrary reference is taken in a Cartesian co-ordinates on the still sea surface as a horizontal plane, it can be assumed to have the x and y co-ordinates on the plane and the z axis as vertical upward positive.

Then, it is possible to consider a coupling of two water surface waves to form a set of two waves crossing each other on the sea surface.

Now, the radiation beam out of the ocean surface as the black body can concentrate at a thermal sensor mounted on an interested satellite. And then, the height or altitude of the satellite above the reference H is given and the sensor axis θ referring to the vertical for a plane wave $F = F(x;\, t)$, as follows cf. Figure. 1,

$$H = H\cos\theta + F. \tag{1}$$

A simplest model of the wave facet as a part of the sea surface wave is to consider a water surface wave propagating only along or parallel to the x axis. The effect of the wave facet's curvature is expressed by the derivative of x for a function of the sea surface F'.

That is,

$$(Z - z) = -[l/(F')](X - x), \tag{2}$$

when $F = F(x;\, t)$, and the orthogonal of the tangent F' at $(X,\, Z) = (x,\, z)$ is given.

In the expression of (2), the function F can be written in a form a spectral function. For example, a Fourier transform of F at $t = t$ can be wtitten by a function S as follow;

$$S = \int_{-\infty}^{\infty} F\exp\left(-ikx\right)\, dx, \tag{3}$$

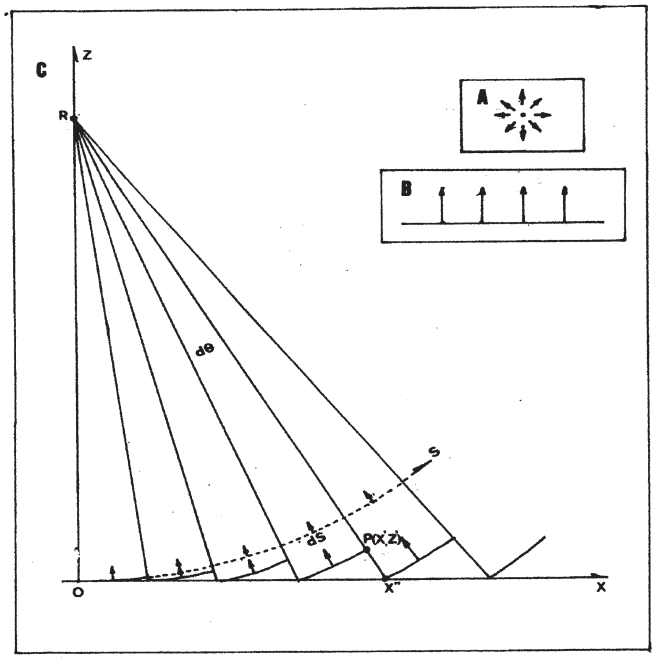


Figure 1: Model for concave facet of sea surface water wave (ds) and position of a satellite (R). (1) Inset A for a point black body. (2) Inset B for a flat surface of the black body. (3)Main frame C for concave facet of sea surface water wave.

At considering a coupled effect of two crossing waves, it can be seen that the effective concentrated beam condition at the sensor of the satellite have to be satisfied the following relation, i.e.,

$$(Z - z) = -[l/(G')](Y - y), \tag{4}$$

when $G = G(y;\, t)$, and the orthogonal of the tangent G' at $(Y,\, Z) = (y,\, z)$.

3. PHYSICAL BACK GROUND OF MODELING

We have a systematic net work of the ocean in a global scale formed by the nations in order to monitor directly continuously by using the research vessels, ships and boats under an international co-operative operation projects.

The scientific oceanographers have learned about the ocean surface waves for more than one hundred years by this time. On the other hand, the field of "Hydrodynamics" has a theoretical knowledge about the water surface waves in the deep or shallow waters. This knowledge helps us to understand what is processes of the water surface waves in the sea.

This leads us to consider a concave facet of the sea surface waves which can radiate an infrared beam out of the ocean as a black body. That is, i.e., the ocean surface is assume as a radiator of infrared beam.

When all of the infrared beams out of a concave facet of the interested water surface waves concentrates at the sensor mounted on a satellite, the sensor on the satellite finds a concentrated beam signal to reduce an apparently high temperature on the sea surface.

The satellite monitored beam signal and an anomaly of the sea surface temperature can be well related at an applying Stefan-Boltzmann's criteria for a black body problem.

In this work, what is important is the following relation between total radiation intensity B and surface temperature T of a considering black body, i.e.,

$$B = \sigma\, T^4, \tag{5}$$

where the notion σ is a constant.

The following expression can be obtained what relation is between dB and dT;

$$dB/B = 4dT/T. \tag{6}$$

What above noted is useful for the author's model introduing in this work.

4. SUPPORTING THERMAL PATTERNS

There have been obtained several monitored sea surface thermal patterns at a station on the coast facing the ocean. This station can only receive the satellite thermal signal at the time of the passage above the station on the polar orbit. Nevertheless, we have no data of the sea surface temperature at that time though a series of the routine observations gives us a collection of the directly observed result during the time period covering that time.

At present, it should be referred to the available data. In future, a data obtained by a more advanced technique. If so, we may expect a much more advanced reference might be taken as a reference.

5. APPARENTLY ANOMALY OF SEA SURFACE TEMPERATURE

In this work, the author introduces a model for a concave facet on the sea surface. This concave facet can be a part of the sea surface water waves. Assuming the ocean surface as a black body, the facet can be an infrared beam radiator which has a focus just neighbor the sensor mounted on the interested satellite. When the concentrated infrared beam out of the facet with its focus just neighbor the satellite, the concentrated beam is related well related to an apparently high temperature which looked as an anomary in a thermal pattern on the sea surface. Then, this work tells that the concentration of the infrared beam out of the water surface wave is effectively controlled by the function F and G. In other word, the interested spectral pattern for the function F and G governs the curvature of the facet with a concave shape as a part of the sea surface waves.

6. CONCLUSION

The author notes what effect of a concave facet of the sea surface waves is. For this purpose, a simple model is introduced with some assumptions. The ocean surface is taken as a black body radiating infrared beam. The satellite polar orbital track hit just neighbor the focus of the facet of the sea surface waves. The curvature of a concave wave facet in a part of the sea surface waves can be described by a special derivative of the function F expressing the sea surface pattern. Then, it can be seen that the spectral pattern of the sea surface waves controls the waves facet's normal line to hit the satellite located near the focus of the wave facet. With the consideration of the curvature of the facet, the model in this work can be expressed by an application of Stefan-Boltzmann's criteria. That is, the infrared beam concentration controls the anomaly of the sea surface temperature in the satellite thermal pattern reduced from the directly monitored at a station on the coast.

Monitoring of Satellite Thermal Patch on the Ocean Surface Generated by Strong Wind Duration in Mid-night

S. Nakamura
Kyoto University, Japan

Abstract— This work concerns monitoring thermal plateau on the ocean surface generated by a strong wind duration in mid-night time of a cold season. This problem was raised first at a satellite thermal monitoring of a data set directly received signals at the satellite passing time just above the station of a system for receiving the satellite signal directly. There has been introduced a physical model for our understanding of a thermal high in a part of the reduced thermal pattern. The author has considered to introduce an application of Stefan-Boltzmann's criteria for a reasonable model with a radiation of an infrared beam out of the earth's surface.

1. INTRODUCTION

A problem of a satellite thermal patch on the earth surface is introduced, especially, on the sea surface. One of the effective monitoring techniques is a satellite thermal monitoring of the earth's surface by using a polar orbital satellite. It seems in the author's understand that the recent scientific researches are strongly concentrated for the global processes of the earth's natural environmental factors as well as of the artificial factors reduced referring to the satellite monitored data set in a scale of a month or a year for "Climate Change".

For these twenty years, the author has reported about a sporadic satellite thermal pattern on the earth which have had been monitored by a simple system for receiving signals from a satellite directly at a station settled on the earth's surface.

2. THERMAL PATTERN ON THE SEA SURFACE

In this work, the author introduces what had been found on the thermal patterns of the earth's surface which have been reduced by processing the satellite signal directly received.

Referring the data set of the earth's surface thermal pattern covering the station settled on the coast facing on the northwestern Pacific, it is easily found some apparently abnormal high temperature in some thermal pattern. The pattern has had been monitored even at a normal operation of the system.

3. MODELING AND STEFAN-BOLTZMANN'S CRITERIA

The author has considered that there might be some physical conditions for reducing a patch of an apparently high temperature on the earth's surface in a thermal pattern reduced out of the directly received satellite signals processing.

The author here introduces a physical model for realizing these thermal patterns. These thermal patterns are, for example, the thermal pinnacle, thermal plateau and thermal basin.

Briefly, an application of Stefan-Boltzmann's theoretical criteria for electromagnetic beam radiation out of a black body is introduced in order to demonstrate that this criteria can be helpful at physical understanding of the thermal pattern on the earth surface.

The application of Stefan-Boltzmann's criteria for the considering thermal pattern should be effective for the earth's surface to be assumed as a black body, when a coupled effect of a concave earth's surface which concentrates the beam radiation out of the earth surface, i.e., the sea surface or the land surface at a sensor mounted on a satellite.

Reflection of the solar beam in the infrared band out of the earth is effective for this work, and the beam in the visible band is not effective in fact.

4. STEFAN-BOLTZMANN'S CRITERIA FOR RADIATION

When the earth's surface can be assumed to be black body, the surface of the black body radiates an electromagnetic radiation beam normal to the surface.

An application of Kirchhoff's physical criteria for radiation is helpful us to realize the model at our physically reasonable understand.

Planck's radiation law tells us that several physical factors help to formulate the law to describe the relation between the intensity of radiation out of the black body's surface in a unit time for a

unit area. The specific factors are Planck's constant h, propagation speed c of electromagnetic ray as a wave, Boltzmann constant k and absolute temperature T of the black body. Then, Planck's formula gives a relation for a interested wave length λ of the wave length and the radiation intensity B_1, that is,

$$B_1 = [2hc^2/\lambda^5]\left[(\exp(F_1) - 1)^{-1}\right], \tag{1}$$

where,

$$F_1 = \left[hc(\lambda kT)^{-1}\right]. \tag{2}$$

Total radiation intensity B is obtained by integration of B_1 for $0 > \lambda < \infty$.

Considering $\lambda\,\nu = c$ (the notation ν is frequency of the interested beam), total radiation flux B is obtained as follow;

$$B = \pi B_1 = \sigma T^4. \tag{3}$$

This relation of B and T is called as Stefan-Boltzmann's law for what the author interested in is simply expressed as noted above process, the relation of B and T is written a simple form.

Now, for the author's interested problem, what is important is to know a relation of $\mathrm{d}B$ and $\mathrm{d}T$. This relation can be reduced easily and written as,

$$\mathrm{d}B/B = 4\mathrm{d}T/T. \tag{4}$$

5. CONCAVE FACET OF SEA SURFACE WAVE

On the ocean ($\mathrm{O}xy$), it is observe usually a complicate form of the sea surface waves.

Assuming an arbitrary function of the sea surface at time t as,

$$F = F(x,\, y;\, t), \tag{5}$$

where,the sea surface height above a reference is expressed by the notation z.

Then, a mathematical manipulation gives what the author aimed to get the spectral form. Assuming a Fourier transform of the Function F, as written following form,

$$F = F_x(x,\, k_x;\, t)F_y(y,\, k_y;\, t), \tag{6}$$

with wave number k_x and k_y for the x and y component respectively, and,

$$[S_x,\, S_y] = \int_0^\infty F_x(x,\, k_x;\, t)\exp(ik_x)\mathrm{d}x \int_0^\infty F_y(y,\, k_y;\, t)\exp(ik_y)\,\mathrm{d}y. \tag{7}$$

Then, the spectral expression S of the sea surface pattern can be described, i.e.,

$$S = \int_0^\infty \int_0^\infty S_x(x,\, k_x;\, t)S_y(y,\, k_y;\, t)\exp(-ik_x)\exp(-ik_y)\mathrm{d}x\mathrm{d}y, \quad \text{at} \quad t = t \tag{8}$$

Now, the radiation beam out of the black body as the ocean surface can concentrate at a thermal sensor for the infrared band, when the height of the satellite above the reference H is given and the sensor axis angle is θ and $F = F(x;\, t)$ for a plane wave,

$$H = H\cos\theta + F \tag{9}$$

Considering a radius of geometrical curvature for a concave sea surface facet as a part of the interested sea surface water waves in the thermal pattern, here, the concave facet of the sea surface at $(x,\, z)$ in an orthogonal frame of OZX can be expressed as a function of F' (writing $F' = \partial F/\partial x$), and can be expressed in the frame of OZX as following:

$$(Z - z) = -[1/(F')][X - x] \tag{10}$$

for a simplest case of $F = F(x;\, t)$ instead of the case of (5).

This helps us to realize that a concave facet of sea surface water wave is effective to concentrate the infrared beam out of the sea surface as a black body at the sensor mounted on the satellite.

This concave wave facet was introduced first for a pixel at a thermal pinnacle. Thermal patch group in a satellite thermal pattern is found as a thermal plateau or a thermal basin in a satellite thermal pattern under a natural condition.

6. MONITORED THERMAL PATTERNS

Now, the author introduces an example of the sea surface thermal patterns obtained by the author's directly monitored sea surface thermal patterns at the couple of the NOAA satellite passage above the monitoring station.

In Figure 1, the pattern at the top shows a thermal plateau (30 °C at "P") appeared just after the sea surface waves grown or matured after several hours strong winds. The sea surface waves were generated by the wind effects on the sea surface water. In other words, the winds caused to make an ensemble of the concave facets of the sea surface waves. This must be a key to form a thermal plateau in the thermal pattern on the sea surface. One of the thermal pinnacles was at the location of the red mark off the Offshore Tower settled for the operation of Oceanographic Observation of Kyoto University.

A reference of the sea surface temperature 15 °C was obtained at "T". This is surely a main cause to form the thermal plateau even in the mid-night of the cold season.

As for the case shown at the top of Figure 1, the relation $\mathrm{d}B/B = 4\mathrm{d}T/T$ is for $\mathrm{d}T = 30 - 15$ and $T = 273 + 15$, then, $\mathrm{d}B/B = 4 \times (15/288) = 0.208$. This means that the concentration rating of the infrared beam out of the ocean surface is evaluated as 0.208, or that the value of the beam intensity exceeds as much as 0.208 than that of the reference. This means 1.208 time at the sensor mounted on the satellite comparing to the beam intensity on the sea surface.

In the case of Figure 1, decay of the sea surface water waves might be several hours. The thermal pattern sown at the bottom is normal, so that the next satellite thermal pattern in morning could be taken to be consistent to the sea surface temperatures observed by the research ships in the interested ocean area.

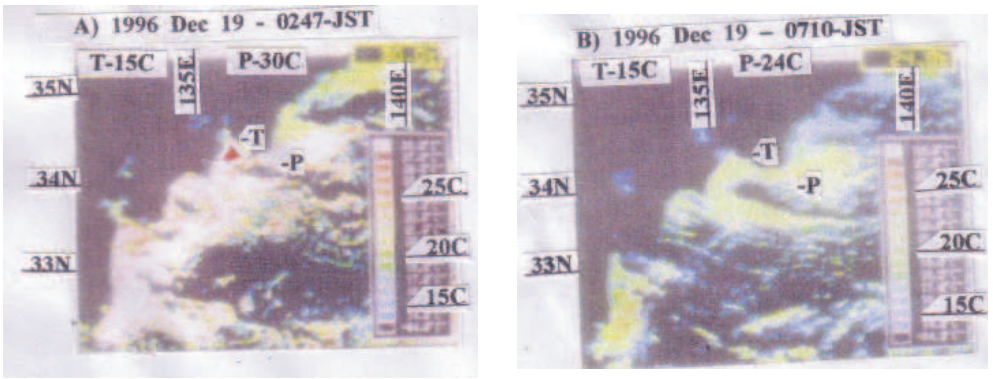


Figure 1: Sea surface thermal patterns obtained by directly satellite monitoring. Top: at 0247-JST on 1996 December 19 (NOAA-14), [showing a thermal plateau of the beams in the infrared band]. Bottom: at 0710-JST on 1996 December 19 (NOAA-12), [showing a thermal pattern found usually in the infrared band]. Note: the local time JST = GMT(UT)+9 hours.

In Figure 1, the atmospheric conditions on the sea surface were caused to find patches of the thermal pattern on the sea surface, even though it is understood that the coast of the northwestern Pacific can be demonstrated as a thermal boundary in the thermal pattern of the coastal zone.

The monitored satellite thermal pattern after the direct receiving of the satellite signals is one of the supporting example for the model introduced in this work. Even though, there are left many problems to be solved at considering a more detailed evaluation of a thermal pattern on the sea surface in the shadow side zone as well as that in the sunny side zone.

The model supports well the satellite thermal plateau under a condition of a strong wind duration mid-night in cold season, when the atmospheric column between the ocean surface and the sensor mounted on the satellite. In this work, no interest of polarity or interferrometic problem is for the beam radiated out of the ocean surface. The author has taken that polarity and interference as the minor factors of considered. Hence, no consideration is in the sight of the author for a system or a problem as seen for SAR data.

7. CONCLUSIONS

Monitoring of satellite thermal pattern, especially a thermal plateau found mid-night in, a cold season is introduced. A model introduced in this work is a model of radiation out of the sea surface which is assumed as a black body with an assumed condition of curvature of a concave wave facet on the sea surface water. Combined effect of radiation out of the ocean and the wave facet curvature on the sea surface water, it easy to see that a satellite thermal plateau monitored directly is well supported by the model even when the thermal plateau is found under a condition of a strong wind duration mid-night in cold season.

Study of a Circular Disc Monopole Ultrawide-band Miniature Antenna

Lingling Zhong, **Bo Sun**, **Jinghui Qiu**, **and Ning Zhang**
Department of Electronics and Communication Engineering
Harbin Institute of Technology, Harbin 150001, China

Abstract— Study of a novel planar circular disc monopole ultrawide-band antenna fed by coaxial line is presented in this paper. The radiator, ground plane and the feeder equipment of the antenna are placed on the same plate. The special structure reduces the spatial volume, and it is used to realize the miniaturization of the antenna. The basic theory and design method are analyzed, and detailed exploration is conducted to determine the antenna's properties. The current distributions, return loss, radiation patterns and gain of the antenna are discussed. In addition, the time domain performance of the proposed antenna is also evaluated in simulations. The research results show that this kind of planar antenna can radiate and receive short pulse signals without distortion. The 10 dB return loss bandwidth is from 2.5 GHz to 55 GHz. And it can realize near omnidirectional pattern in the H-plane. It is a real planar structure and can really reduce the spatial volume. On this basis, the structure of the planar circular disc monopole antenna is improved. The radiator moves around the upper edge of the ground plane. The new structure reduces the antenna height to a certain degree and it is used to achieve further miniaturization. Analyses are conducted to determine the novel antennas' properties. The result of study indicates that the improved antenna can realize good bandwidth performance as the planar circular disc monopole antenna, and it has low-cost, simple structural characteristics. The novel miniature antenna and the improved type are suitable for the wireless communications, satellite communications and mobile communications systems with good prospects.

1. INTRODUCTION

With the development of RF (Radio Frequency) technology and UWB equipment, there has been considerable research effort put into ultrawide-band antenna. In recent years, monopole antennas are the focus of UWB antenna. Several broadband monopole configurations, such as circular, square, elliptical, pentagonal and hexagonal, have been proposed so far [1–4]. However, they are not planar structures as the ground planes are perpendicular to radiators.

In this paper, a novel circular disc monopole antenna fed by coaxial line is proposed. It is a real planar structure. The radiator, ground plane and the feeder equipment of the antenna are placed on the same plate. This reduces the spatial volume to a great degree. And then the structure of the planar circular disc monopole antenna is improved. Its structure reduces the antenna height and further miniaturization is realized.

2. BASIC THEORY OF THE PLANAR CIRCULAR DISC MONOPOLE ANTENNA

The proposed planar monopole is illustrated in Figure 1. Copper can be chosen as the material, and its thickness $d \approx 1$ mm. The sheet is cut as a disk with radius r. Ground plane is composed of two pieces of rectangle metal (length of side is m and n), and they are placed on the same plane with the circular radiator. The antenna is fed by a coaxial line. Core of the coaxial is exposed out as the feed point, and the outer metal is connected with ground planes. The lower end of the disk is adopted as the feed point. The height of the feed gap is h.

Theoretically, circular disc monopole with radius r could be equivalent as cylindrical monopole with height l and radius a. Its first resonant frequency can be determined by equating the area of the circular disk to that of the cylindrical monopole, and equating the height of the disk $2r$ to the height of the cylindrical monopole l, i.e., $\pi r^2 = 2\pi al$, $2r = l$ [5]. The length of monopole for real input impedance is given by

$$l = 0.24\lambda F \tag{1}$$

where $F = (l/a)/(1 + l/a)$.

From the above equations, the first resonant frequency of the circular disc monopole is given by

$$f = \frac{3.2}{r} \tag{2}$$

where l and r are in centimeters.

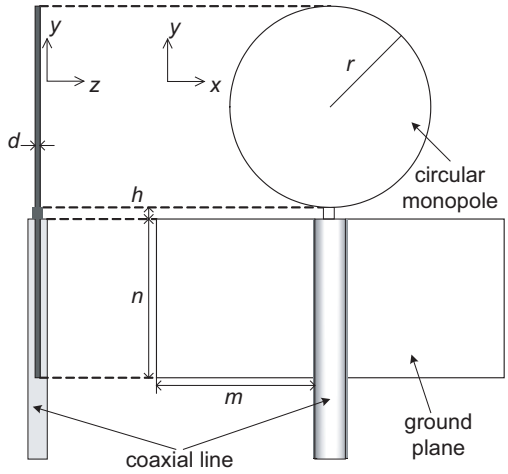


Figure 1: Geometry of the planar circular disc monopole.

3. SIMULATED RESULTS AND ANALYSIS OF THE PLANAR CIRCULAR DISC MONOPOLE ANTENNA

A dimension of the proposed circular disc monopole antenna is chosen, i.e., $r = 10.7\,\text{mm}$, $m = n = 16.4\,\text{mm}$, $h = 0.6\,\text{mm}$, and $d = 1\,\text{mm}$. And the corresponding first resonant frequency is 3 GHz. The simulation of the antenna is performed using the CST Microwave Studio package which utilizes the finite integration technique for electromagnetic computation [6].

3.1. Current Distributions

Current distribution at 3 GHz is showed in Figure 2. The current is mainly distributed along the edge of the disc. As circular disc monopole could be equivalent as infinite dipole antennas, and all of them have the same feed point, the disc dimension mainly defines the first resonant frequency [7]. Besides the disc, current is distributed on the upper edge of the ground plane. The portion of the ground plane close to the disc acts as the radiating structure. Figure 3 shows the return loss for different dimensions of the transverse length of the ground plane m (return loss at 3 GHz is showed in the block). With increase of m, the first resonant frequency decreases.

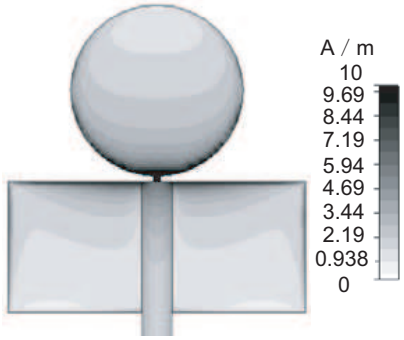


Figure 2: Simulated current distributions at 3 GHz.

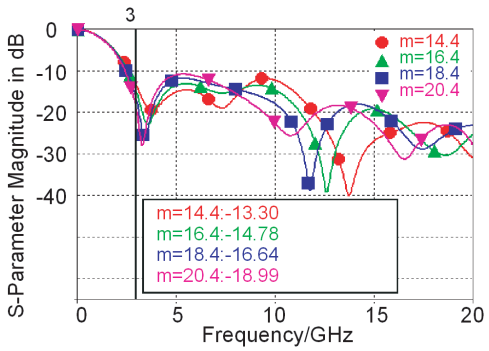


Figure 3: Simulated return loss curves for different m.

3.2. The Effect of the Height of the Feed Gap Between the Feed Point and the Ground Plane

The simulated return loss curves for different feed height h are presented in Figure 4. Lots of simulations demonstrate that h would influence the return loss and bandwidth to a great extent. For circular disc monopole, the ground plane serves as an impedance matching circuit. Consequently, it tunes the input impedance and hence the 10 dB return loss bandwidth by changing h [8]. With the increase of h, the first resonant frequency decreases at lower frequencies. When $f = 3\,\text{GHz}$, the return loss is less than $-15\,\text{dB}$.

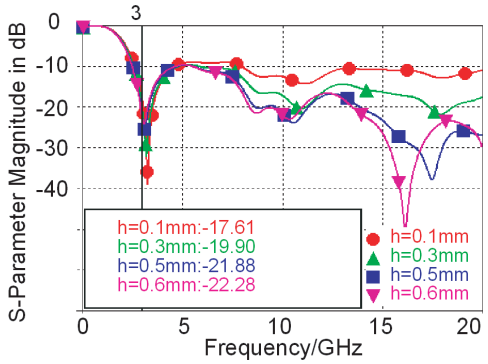


Figure 4: Simulated return loss curves for different h.

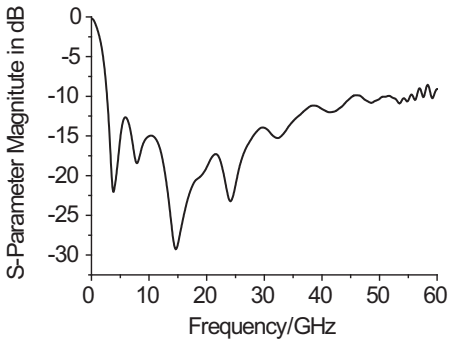


Figure 5: Simulated return loss of the planar antenna.

3.3. Return Loss

Simulated return loss of the circular disc monopole from 0 GHz to 60 GHz is showed in Figure 5. The 10 dB return loss bandwidth is from 2.5 GHz to 55 GHz, and 15 dB bandwidth is from 10 GHz to 28 GHz. As a result, the circular disc monopole antenna has bandwidth ratio of 20 : 1 at least.

3.4. Radiation Patterns and Gain

Radiation patterns and gain are important factors to evaluate the performance of the UWB antenna. The simulated normalized radiation patterns at 3 GHz, 6 GHz and 10 GHz are plotted in Figure 6. It is noticed that the simulated E-plane (x-y plane) pattern is like a traditional monopole, and the H-plane (x-z plane) pattern is near omnidirectional. The radiation patterns are steady within the bandwidth 3–10 GHz.

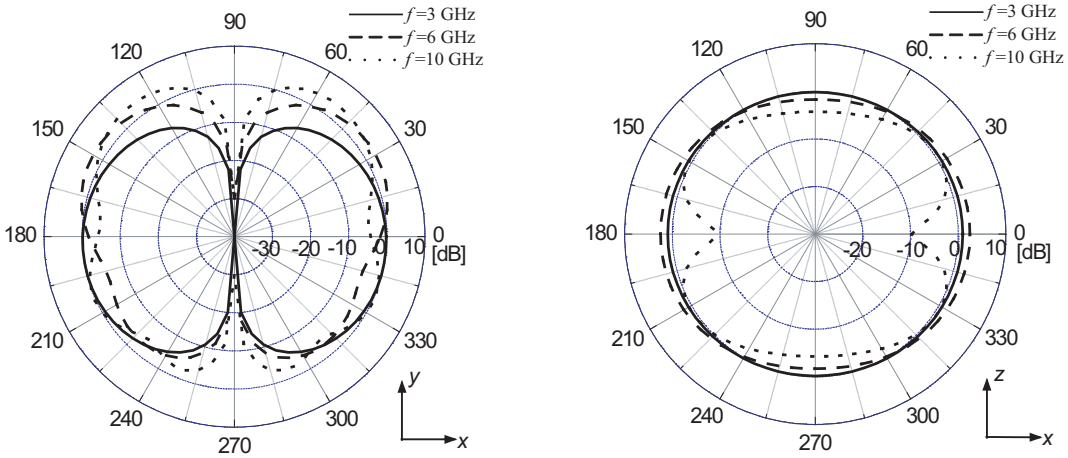


Figure 6: Simulated normalized radiation patterns of the planar antenna. (a) E-plane, (b) H-plane.

Figure 7 illustrates the simulated maximum gains of the proposed antenna. It is shown that the maximum variation of the gain is 6 dB within the frequency range 3–60 GHz. With the increase of the frequency, the maximum gain increases too. When $f = 50$ GHz, the maximum gain is greater than 5 dB.

3.5. Time Domain Performance

Apart from the consideration of the 10 dB return loss bandwidth and radiation patterns, a good time domain characteristic is an essential requirement for an UWB antenna. Flat group delay time indicates all frequencies are delayed for the same time. This is conducive for the maintenance of the signal waveform. Figure 8 is the simulated group delay curve of the planar antenna. The curve is nearly constant throughout all the bands. This ensures the planar circular disc monopole can radiate and receive short pulse signals without distortion.

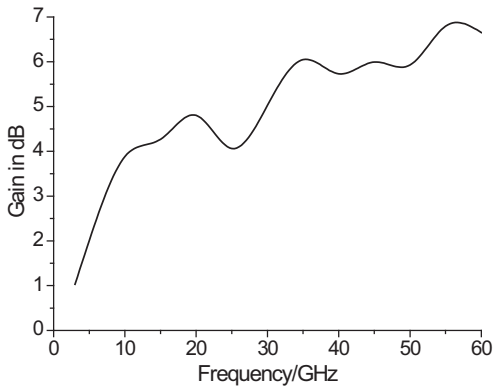


Figure 7: Maximum gains of the planar antenna.

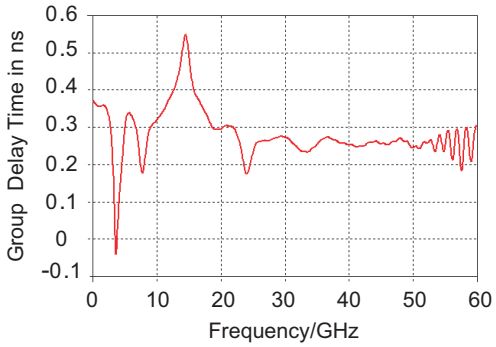


Figure 8: Group delay curve of the planar antenna.

4. IMPROVED CIRCULAR DISC MONOPOLE MINIATURE ANTENNA

On the basis of the foregoing statements, the structure of the planar circular disc monopole antenna is improved. The radiator moves around the upper edge of the ground plane ϕ, as Figure 9 shows. This reduces the height of the antenna to a certain degree. It is favorable for the miniaturization of the antenna. The return loss for $\phi = 10°$, $\phi = 30°$ and $\phi = 45°$ are plotted in Figure 10 respectively. It can be seen from the figure that the improved antenna is able to achieve good bandwidth characteristics as the planar circular disc monopole antenna.

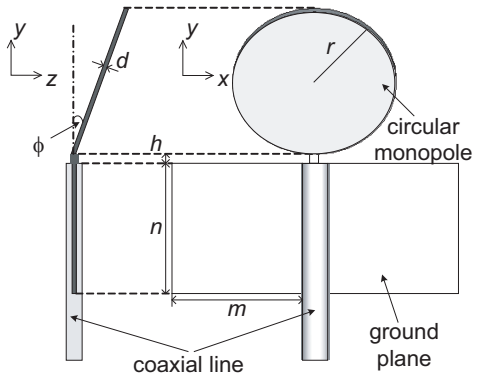


Figure 9: Geometry of the improved UWB antenna.

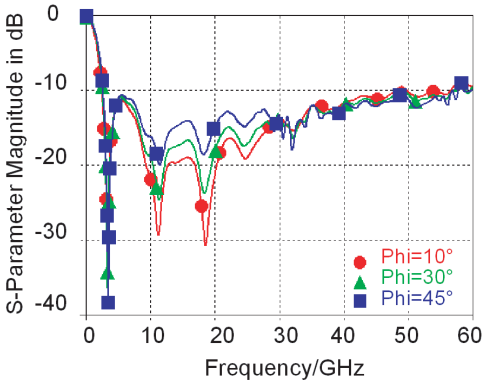


Figure 10: Return loss of the improved UWB antenna.

5. CONCLUSIONS

The novel planar circular disc monopole antenna is investigated detailedly in this paper. The basic theory and design method are analyzed and research results show that this kind of planar antenna can radiate and receive short pulse signals without distortion. The 10 dB return loss bandwidth is from 2.5 GHz to 55 GHz. And it can realize near omnidirectional pattern in the H-plane. It is a real planar structure and can reduce the spatial volume. On the basis, the disc of the planar circular monopole moves around the upper edge of the ground plane a certain angle. The result of study indicates that the improved antenna reduces the height and can realize good bandwidth performance, and it has low-cost, simple structural characteristics. The novel miniature antenna is suitable for wireless and satellite communication system.

REFERENCES

1. Hans, G. S., "Bottom fed planar elliptical UWB antennas," *Proceedings of IEEE Conference on Ultra Wideband System and Technology*, 219–223, VA, USA, November 2003.
2. Narayan, P. A., K. Girish, and K. P. Ray, "Wide-band planar monopole antennas," *IEEE Trans. Antennas Propag.*, Vol. 46, No. 2, 294–295, 1998.
3. Ammann, M. J. and Z. N. Chen, "Wideband monopole antennas for multi-band wireless systems," *IEEE Antennas Propag. Mag.*, Vol. 45, No. 2, 146–150, 2003.

4. Liang, J. X., C. C. Choo, C. X. Dong, and C. G. Parini, "Study of a printed circular disc monopole antenna for UWB systems," *IEEE Trans. Antennas Propag.*, Vol. 53, No. 11, 3500–3504, 2005.
5. Balanis, C. A., *Antenna Theory: Analysis and Design*, Harper and Row, New York, 1982.
6. Min, Z., *CST Microwave Studio® 5 User's Manual*, University of Electronic Science and Technology of China Press, Cheng Du, 2003.
7. Honda, S., M. Ito, H. Seki, and Y. Jingo, "A disc monopole antenna with 1 : 8 impedance bandwidth and omni-directional radiation pattern," *Proceedings of ISAP '92*, 1145–1148, Sapporo, Japan, September 1992.
8. Chen, Z., X. Wu, H. Li, N. Yang, and M. Y. Chia, "Considerations for source pulses and antennas in UWB radio systems," *IEEE Trans. Antennas Propag.*, Vol. 52, No. 7, 1739–1748, 2004.

Virtual Sources for a Sinh-Gaussian Beam

Y. C. Zhang, **Z. R. Chen**, **Z. X. Shi**, **J. Q. Dong**, **Z. L. Wu**, **and N. Zhang**
Department of information science and technology, Xichang College, Xichang 615500, China

Abstract— On the basis of the superposition of beams, a group of virtual sources that generate a sinh-Gaussian wave is identified. A closed-form expression is derived for the sinh-Gaussian wave, which, in the appropriate limit, yields the paraxial approximation for the sinh-Gaussian beam. From this expression, the paraxial approximation and the nonparaxial corrections of all orders for the corresponding paraxial sinh-Gaussian beam are determined.

1. INTRODUCTION

The propagation of electromagnetic beams in optical systems has long been of interest and in some cases exact analytical solutions of the paraxial wave equation have been obtained. For many years the most general analytical solutions in Cartesian coordinates have been in the form of Hermite-Gaussian functions in which the arguments of both the Hermite-polynomial factor and the Gaussian factor are complex [1], and the real-argument Hermite-Gaussian beams are a special case of the complex-argument beams [2]. About ten years ago, an alternative set of complex Cartesian beam solutions of the paraxial wave equation was obtained in the form of sinusoidal-Gaussian beams which reduce to conventional rectangular waveguide modes in an appropriate limit [3]. Sinusoidal-Gaussian beams can be used as an alternative basis set for propagation of arbitrary electromagnetic beams and as solution modes in novel laser resonators, and exhibit different propagation behavior from real-argument beams [3–5]. Thus such beams are needed for the more general propagation studies. Virtual source method, a rigorous method of treating the beams in and beyond the paraxial limit, which was pointed out by Deschamps and systematically developed by Shin and Felsen [6, 7], is widely applied to investigate characterization and propagation of beams within and beyond the paraxial regime [8–12]. Here we choose a sinh-Gaussian (ShG) beam from sinusoidal-Gaussian beams as an example and develop this approach to study the propagation properties of a ShG beam in and beyond the paraxial region. The key to this problem is how to determine a configuration of the suitable virtual sources to yield the desired ShG beam. In this paper, based on superstition of beams, a group of virtual sources required for generation of the scalar ShG wave is introduced, a closed-form expression is derived for the ShG wave and from this expression the paraxial approximation and all-order nonparaxial corrections are determined.

2. THE VIRTUAL SOURCES TO YIELD THE SHG BEAM

Assume that the field distribution of ShG beam at the $z = 0$ plane is characterized by [3]

$$E_p(\boldsymbol{\rho}, 0) = \exp\left[-\frac{\boldsymbol{\rho}^2}{w_0^2}\right] \sinh(\Omega_0 x) \sinh(\Omega_0 y), \tag{1}$$

where subscript p stands for the paraxial approximation, $\boldsymbol{\rho} = (x, y)$ is a position vector at the transverse plane, w_0 is the waist width of the Gaussian amplitude distribution, and Ω_0 is the width parameter associated with the sinh part at the $z = 0$ plane. Eq. (1) can be written alternatively as

$$\begin{aligned} E_p(\boldsymbol{\rho}, 0) = {} & \frac{\exp\left(\frac{a^2}{2w_0^2}\right)}{4} \left\{ \exp\left[-\left(\left(\frac{x - a/2}{w_0} \right)^2 + \left(\frac{y - a/2}{w_0} \right)^2 \right) \right] \right. \\ & + \exp\left[-\left(\left(\frac{x + a/2}{w_0} \right)^2 + \left(\frac{y + a/2}{w_0} \right)^2 \right) \right] - \exp\left[-\left(\left(\frac{x - a/2}{w_0} \right)^2 + \left(\frac{y + a/2}{w_0} \right)^2 \right) \right] \\ & \left. - \exp\left[-\left(\left(\frac{x + a/2}{w_0} \right)^2 + \left(\frac{y - a/2}{w_0} \right)^2 \right) \right] \right\}, \end{aligned} \tag{2}$$

with $a = \Omega_0 w_0^2$, which implies that a ShG beam can be produced simply by superposition of four decentered Gaussian beams with the same waist width and the first two being in-phase and anti-phase with the last two, whose centers are located at the positions $(a/2, a/2, 0)$, $(-a/2, -a/2, 0)$,

$(a/2, -a/2, 0)$, and $(-a/2, a/2, 0)$, respectively. From this point of view, in the physical space $z > 0$, in order to construct a scalar ShG beam that propagates along the z axis of a rectangular coordinate system (x, y, z), we can reasonably assume that the beam is generated by four point sources of strength S_{ex} with the first two being in-phase and anti-phase with the last two situated at $(b,\ b,\ z_{ex})$, $(-b,\ -b,\ z_{ex})$, $(b,\ -b,\ z_{ex})$, and $(-b,\ b,\ z_{ex})$, respectively. For $z > 0$, the scalar wave function $E(\boldsymbol{\rho}, z)$ that describes the beam, satisfies the homogeneous Helmholtz equation. Proper choice of S_{ex}, z_{ex}, and b yields the desired beam. The wave function $E(\boldsymbol{\rho}, z)$ satisfies the inhomogeneous Helmholtz equation

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + k^2\right) E(\boldsymbol{\rho}, z) = -S_{ex}\delta(z - z_{ex}) [\delta(x - b)\delta(y - b) + \delta(x + b)\delta(y + b) - \delta(x - b)\delta(y + b) - \delta(x + b)\delta(y - b)], \tag{3}$$

where k is the wave number. From Eq. (3), the Fouriertransform $\overline{E(\mathbf{q}, z)}$ of the wave function $E(\boldsymbol{\rho}, z)$ obey

$$\left[\frac{d^2}{dz^2} + \left(k^2 - u^2 - v^2\right)\right] \overline{E(\mathbf{q}, z)} = -S_{ex}\delta(z - z_{ex}) \{\exp[-i(u + v)b] + \exp[i(u + v)b] - \exp[-i(u - v)b] - \exp[i(u - v)b]\}, \tag{4}$$

where

$$E(\boldsymbol{\rho}, z) = \frac{1}{4\pi^2} \iint \overline{E(\mathbf{q}, z)} \exp(i\boldsymbol{\rho} \cdot \mathbf{q})\, d\mathbf{q}, \tag{5a}$$

$$\overline{E(\mathbf{q}, z)} = \iint E(\boldsymbol{\rho}, z) \exp(-i\boldsymbol{\rho} \cdot \mathbf{q})\, d\boldsymbol{\rho} \tag{5b}$$

are the Fouriertransform pair, $\mathbf{q} = (u, v)$ is a spatial frequency vector, $d\boldsymbol{\rho} = dxdy$ and $d\mathbf{q} = dudv$. Unless otherwise stated, all integrations extend from $-\infty$ to $+\infty$.

The solution of Eq. (4) is given by [13]

$$\overline{E(\mathbf{q}, z)} = \frac{iS_{ex}}{2\xi} \exp[i\xi(z - z_{ex})] \{\exp[-i(u + v)b] + \exp[i(u + v)b] - \exp[-i(u - v)b] - \exp[i(u - v)b]\} \tag{6}$$

for $\mathrm{Re}(z - z_{ex}) > 0$, where $\xi = (k^2 - u^2 - v^2)^{1/2}$. On substituting Eq. (6) into Eq. (5a), we obtain

$$E(\boldsymbol{\rho}, z) = \frac{1}{4\pi^2} \iint \exp(i\boldsymbol{\rho} \cdot \mathbf{q}) \frac{iS_{ex}}{2\xi} \exp[i\xi(z - z_{ex})] \{\exp[-i(u + v)b] + \exp[i[u + v]b] - \exp[-i(u - v)b] - \exp[i(u - v)b]\} d\mathbf{q} \tag{7}$$

for $\mathrm{Re}(z - z_{ex}) > 0$. As Seshadri pointed out [12], the domains of u and v are extended to complex values, and the radiation condition is used to select the proper branch of ξ. Also, $\overline{E(\mathbf{q}, z)}$ is an analytic function of the complex variable $(u^2 + v^2)^{1/2}$.

Let us first recover the standard paraxial description. To obtain the contributions for $u^2 + v^2 << k^2$, we expand ξ for small $u^2 + v^2$ and retain the leading term for the amplitude factor and the first two terms for the phase factor in Eq. (7). In this approximation, Eq. (7) becomes

$$E(\boldsymbol{\rho}, z) = \frac{1}{4\pi^2} \iint \exp(i\boldsymbol{\rho} \cdot \mathbf{q}) \frac{iS_{ex}}{2k} \exp[ik(z - z_{ex})] \exp\left[-i\frac{\mathbf{q}^2}{2k}(z - z_{ex})\right] \{\exp[-i(u + v)b] + \exp[i(u + v)b] - \exp[-i(u - v)b] - \exp[i(u - v)b]\} d\mathbf{q}. \tag{8}$$

Evaluating the integrals in Eq. (8), we obtain

$$E(\boldsymbol{\rho}, z) = \frac{S_{ex} \exp[ik(z - z_{ex})]}{4\pi(z - z_{ex})} \left\{\exp\left[ik\frac{(x - b)^2 + (y - b)^2}{2(z - z_{ex})}\right] + \exp\left[ik\frac{(x + b)^2 + (y + b)^2}{2(z - z_{ex})}\right] - \exp\left[ik\frac{(x - b)^2 + (y + b)^2}{2(z - z_{ex})}\right] - \exp\left[ik\frac{(x + b)^2 + (y - b)^2}{2(z - z_{ex})}\right]\right\}. \tag{9}$$

External to $z > 0$ and at its boundary is the input plane $z = 0$. To generate the ShG beam for $z > 0$, in the paraxial approximation, we assume the input distribution to be Eq. (1) or (2). We determine S_{ex}, z_{ex}, and b by the requirement that $E(\boldsymbol{\rho}, z)$ given by Eq. (9) for $z = 0$ reduce to $E_p(\boldsymbol{\rho}, 0)$ as given by Eq. (1) or (2), with the result that

$$b = a/2, \tag{10}$$

$$z_{ex} = \frac{ikw_0^2}{2} = iz_0, \tag{11}$$

$$S_{ex} = -\pi z_{ex} \exp\left(\frac{a^2}{2w_0^2} + ikz_{ex}\right) = -i\pi z_0 \exp\left(\frac{a^2}{2w_0^2} - kz_0\right). \tag{12}$$

Substituting Eqs. (10)–(12) into Eq. (9) yields the paraxial approximation to $E(\boldsymbol{\rho}, z)$:

$$E_p(\boldsymbol{\rho}, z) = -\frac{iz_0}{z - iz_0} \exp\left(\frac{a^2}{2w_0^2} + ik\left(z - \frac{a^2}{4\left(z - iz_0\right)}\right)\right) \exp\left(ik\frac{\boldsymbol{\rho}^2}{2\left(z - iz_0\right)}\right) \sinh\left(\frac{ikax}{2\left(z - iz_0\right)}\right) \times \sinh\left(\frac{ikay}{2\left(z - iz_0\right)}\right). \tag{13}$$

Further, from Eqs. (1) and (13), we can obtain the transformation equations of the parameters that characterize the propagation of the ShG beam, as follows:

$$q(z)^{-1} = \frac{q(0)^{-1}}{1 + zq(0)^{-1}}, \tag{14}$$

$$\Omega(z) = \frac{\Omega_0}{1 + zq(0)^{-1}}, \tag{15}$$

$$w(z) = w_0\left[1 + (z/z_0)^2\right]^{1/2}, \tag{16}$$

$$R(z) = z\left[1 + (z_0/z)^2\right], \tag{17}$$

where $q(z) = \frac{1}{R(z)} + \frac{2i}{kw(z)}$ and $\Omega(z)$ are, respectively, the q-parameter associated with the Gaussian part and the width parameter associated with the sinh part of the beam at the plane z, and $w(z)$ and $R(z)$ are the spot size and the radius of curvature of the phase front of the Gaussian part at the plane z, respectively. The above paraxial results are consistent with the ones in Ref. [3].

3. THE NONPARAXIAL CORRECTION OF ALL ORDERS FOR THE CORRESPONDING PARAXIAL SINH-GAUSSIAN BEAM

In a similar manner, by use of Eqs. (10)–(12), the exact expression for $E(\boldsymbol{\rho}, z)$ is obtained from Eq. (7):

$$E(\boldsymbol{\rho}, z) = \frac{z_0}{4\pi} \iint \exp\left(i\boldsymbol{\rho} \cdot \mathbf{q}\right) \frac{\exp\left(\frac{a^2}{2w_0^2} - kz_0\right)}{2\xi} \exp\left[i\xi\left(z - iz_0\right)\right] \left\{\exp\left[-i\frac{(u+v)a}{2}\right] + \exp\left[i\frac{(u+v)a}{2}\right] - \exp\left[-i\frac{(u-v)a}{2}\right] - \exp\left[i\frac{(u-v)a}{2}\right]\right\} d\mathbf{q}. \tag{18}$$

Eqs. (18) and (13) represent the ShG wave and the corresponding ShG beam, respectively. In view of Eq. (11), the sources in Eq. (3) are outside $z > 0$. Therefore, for $z > 0$, Eq. (18) is the exact solution to the homogeneous equation corresponding to Eq. (18). The exact solution reduces to the correct paraxial approximation in the appropriate limit but contains all the nonparaxial contributions and the evanescent waves. All the contributions can be summed up, that is, the integral in Eq. (18) can be evaluated analytically.

Recalling the Fouriertransform formulas [14]

$$\frac{i}{2\pi} \iint \frac{\exp\left[ik\left(k^2 - u^2 - v^2\right)^{1/2} z\right]}{\left(k^2 - u^2 - v^2\right)^{1/2}} \exp\left[i\left(xu + vy\right)\right] dudv = \frac{\exp\left[ik\left(x^2 + y^2 + z^2\right)^{1/2}\right]}{\left(x^2 + y^2 + z^2\right)^{1/2}}, \tag{19}$$

$$\iint \exp\left[-i\left(ux_0 + vy_0\right)\right] F\left(u, v\right) \exp\left[i\left(ux + vy\right)\right] dudv = f\left(x - x_0, y - y_0\right), \tag{20}$$

with $F(u,v)$ being the Fouriertransform of $f(x,y)$, evaluation of the integral in Eq. (18) yields

$$E(\boldsymbol{\rho},z)=-\frac{iz_0}{4}\exp\left(\frac{a^2}{2w_0^2}-kz_0\right)\left[\frac{\exp(ikr_1)}{r_1}+\frac{\exp(ikr_2)}{r_2}-\frac{\exp(ikr_3)}{r_3}-\frac{\exp(ikr_4)}{r_4}\right] \tag{21}$$

with

$$r_1 = \left[\left(x-\frac{a}{2}\right)^2+\left(y-\frac{a}{2}\right)^2+(z-iz_0)^2\right]^{1/2}, \tag{22a}$$

$$r_2 = \left[\left(x+\frac{a}{2}\right)^2+\left(y+\frac{a}{2}\right)^2+(z-iz_0)^2\right]^{1/2}, \tag{22b}$$

$$r_3 = \left[\left(x+\frac{a}{2}\right)^2+\left(y-\frac{a}{2}\right)^2+(z-iz_0)^2\right]^{1/2}, \tag{22c}$$

$$r_4 = \left[\left(x-\frac{a}{2}\right)^2+\left(y+\frac{a}{2}\right)^2+(z-iz_0)^2\right]^{1/2}, \tag{22d}$$

which implies that a ShG wave can be regarded as superposition of four complex-source-point spherical waves corresponding to four decentered paraxial Gaussian beams. In other words, the exact solution of the ShG wave given by Eq. (21) can be separated into a paraxial part and a nonparaxial part for all $z>0$ by expanding such complex-source-point spherical waves near the z axis in power series and grouping together equal powers of parameter f, namely,

$$E(\boldsymbol{\rho},z)=E^{(0)}(\boldsymbol{\rho},z)+\sum_{n=1}^{\infty}f^{2n}E^{(2n)}(\boldsymbol{\rho},z), \tag{23}$$

where $f=1/(kw_0)$, $E^{(0)}(\boldsymbol{\rho},z)=E_p(\boldsymbol{\rho},z)$ standing for the paraxial ShG beam, $E^{(2n)}(\boldsymbol{\rho},z)\,(n\geq 1)$ standing for the nth-order correction to the ShG beam. Couture and Belanger showed that the paraxial Gaussian beam became the complex-source-point spherical wave when all-order corrections were made [15]. Following the simila but inverse reasoning as Couture and Belanger in appendices A abd B of Ref. [15], however, here, applying the generating function for the spherical Bankel function of the first kind, from this exact solution of the ShG wave, the all-order corrections to the ShG beam can be obtained immediately as follows:

$$E^{(0)}(\boldsymbol{\rho},z) = E_1^{(0)}(\boldsymbol{\rho},z)+E_2^{(0)}(\boldsymbol{\rho},z)-E_3^{(0)}(\boldsymbol{\rho},z)-E_4^{(0)}(\boldsymbol{\rho},z) \tag{24}$$

$$E^{(2n)}(\boldsymbol{\rho},z) = E_1^{(2n)}(\boldsymbol{\rho},z)+E_2^{(2n)}(\boldsymbol{\rho},z)-E_3^{(2n)}(\boldsymbol{\rho},z)-E_4^{(2n)}(\boldsymbol{\rho},z) \tag{25}$$

$$E_j^{(0)}(\boldsymbol{\rho},z) = -\frac{i}{4}\exp\left(\frac{a^2}{2w_0^2}+ikz\right)Q\exp\left(iQ\rho_j^2\right), \tag{26}$$

$$E_j^{(2n)}(\boldsymbol{\rho},z) = (-1)^n(Q\rho_j)^{2n}L_n^n\left(-iQ\rho_j^2\right)E_j^{(0)}(\boldsymbol{\rho},z), \tag{27}$$

valid for $|Q\rho_j f|\langle 1$, where $j=1,\,2,\,3,\,4$, the dimensionless variables are defined as

$$Q = \left(\frac{z}{z_0}-i\right)^{-1}, \tag{28}$$

$$\rho_1^2 = \frac{(x-a/2)^2+(y-a/2)^2}{w_0^2}, \tag{29}$$

$$\rho_2^2 = \frac{(x+a/2)^2+(y+a/2)^2}{w_0^2}, \tag{30}$$

$$\rho_3^2 = \frac{(x+a/2)^2+(y-a/2)^2}{w_0^2}, \tag{31}$$

$$\rho_4^2 = \frac{(x-a/2)^2+(y+a/2)^2}{w_0^2}, \tag{32}$$

and $L_n^n(\alpha)$ is the associated Laguerre polynomial

$$L_n^n(\alpha) = (2n)! \sum_{m=0}^{n} \frac{(-\alpha)^m}{m!(n-m)!(n+m)!}. \tag{33}$$

It is easy to verify that $E^{(0)}(\boldsymbol{\rho}, z) = E_p(\boldsymbol{\rho}, z)$ satisfies the paraxial wave equation, and successive correction $E^{(2n)}(\boldsymbol{\rho}, z)$ is related to the previous one $E^{(2n-2)}(\boldsymbol{\rho}, z)$ through a differential recurrence relation obtained by Agrawal and Lax [16]. Furthermore, setting $x = y = 0$ in Eqs. (24)–(27) yields the all-order corrections to the ShG beam on the z axis as

$$E^{(0)}(0, 0, z) = 0, \tag{34}$$

$$E^{(2n)}(0, 0, z) = 0. \tag{35}$$

Eq. (35) reveals that the nonparaxial corrections to the ShG beam on the z axis do vanish.

4. CONCLUSIONS

In this paper, based on superposition of beams, a group of virtual sources required for generation of a ShG wave has been obtained. A closed-form expression for this ShG wave is derived. From the expression of this ShG wave, in the appropriate limit, the paraxial approximation and the all-order nonparaxial corrections for the corresponding ShG beam are determined. Because the complex-source-point spherical wave has a very simple mathematical form, it should be used more widely to deal with very strong focusing of laser beams. Finally, we mention that this treatment method can be employed to study propagation properties of the other Sinusoidal-Gaussian beams in and beyond paraxial region.

REFERENCES

1. Siegman, A. E., "Hermite-gassian functions of complex argument as optical-beam eigenfunctions," *J. Opt. Soc. Am.*, Vol. 63, 1093–1094, 1973.
2. Siegman, A. E., *Lasers*, University science, Mill Valley, Calif., 1986.
3. Casperson, L. W., D. G. Hall, and A. A. Tover, "Sinusoidal-Gaussian beams in complex optical systems," *J. Opt. Soc. Am. A*, Vol. 14, 3341–3348, 1997.
4. Casperson, L. W. and A. A. Tover, "Hermite-sinusoidal-Gaussian beams in complex optical systems," *J. Opt. Soc. Am. A*, Vol. 15, 954–961, 1998.
5. Tover, A. A. and L. W. Casperson, "Production and propagation of Hermite-sinusoidal-Gaussian laser beams," *J. Opt. Soc. Am. A*, Vol. 15, 2425–2432, 1998.
6. Deschamps, G. A., "Gaussian beam as a bundle of complex rays," *Electron. Lett.*, Vol. 7, 684–685, 1971.
7. Shin, S. Y. and L. B. Felsen, "Gaussian beam modes by multipoles with complex source points," *J. Opt. Soc. Am.*, Vol. 67, 699–700, 1977.
8. Cullen, A. L. and P. K. Yu, "Complex source-point theory of the electromagnetic open resonator," *Proc. R. Soc. London, Ser. A*, Vol. 366, 155–171, 1979.
9. Sheppard, C. J. R. and S. Saghafi, "Beam modes beyond the paraxial approximation: A scalar treatment," *Phys. Rev. A*, Vol. 57, 2971–2979, 1998.
10. Seshadri, S. R., "Virtual source for the Bessel-Gauss beam," *Opt. Lett.*, Vol. 27, 998–1000, 2002.
11. Seshadri, S. R., "Virtual source for a Laguerre-Gauss beam," *Opt. Lett.*, Vol. 27, 1872–1874, 2002.
12. Seshadri, S. R., "Virtual source for a Hermite-Gauss beam," *Opt. Lett.*, Vol. 28, 595–597, 2003.
13. Byron, F. W. and R. W. Fuller, *Mathematics of Classical and Quantum Physics*, Vol. 2, Addison-Wesley, 1969.
14. Stamnes, J.-J., *Waves in Focal Regions: Propagation, Diffraction and Focusing of Light, Sound and Water Waves*, Adam Hilger, Bristol and Boston, 1986.
15. Couture, M. and P.-A. Belanger, "From Gaussian beam to complex-source-point spherical wave," *Phys. Rev. A*, Vol. 24, 355–359, 1981.
16. Agrawal, G. P. and M. Lax, "Free-space wave propagation beyond the paraxial approximation," *Phys. Rev. A*, Vol. 27, 1693–1695, 1983.

Method for Magnetic Field Approximation in MR Tomography

Michal Hadinec[1], **Pavel Fiala**[1], **Eva Kroutilová**[1]
Miroslav Steinbauer[1], **and Karel Bartušek**[2]

[1]Institute of Scientific Instruments of the ASCR, v.v.i
Královopolská 147, 612 64, Brno, Czech Republic
[2]Department of Theoretical and Experimental Electrical Engineering
Faculty of Electrical Engineering and Communication, Brno University of Technology
Kolejní 2906/4, 612 00, Brno, Czech Republic

Abstract— This paper describes a method, which can be used for creating map of magnetic field. Method has a great usage in magnetic resonance tomography, when we need to get information about homogeneity and characteristics of magnetic field inside the working space of the MR tomograph. The main purpose of this article is to describe basic principles of magnetic resonance phenomenon and mathematical method of Legendre polynoms which can be used for signal processing of FID (Free Induction Decay) signal obtained from tomograph detection coils. In the end of my article is experimental solution of magnetic field and models of magnetic field created by Matlab.

1. INTRODUCTION

Magnetic resonance tomography is an imaging technique used primary in medical setting to produce high quality images of the human body [2]. Lately, medical science lays stress on the measuring of exactly defined parts of human body, especially human brain. If we want to obtain the best quality images we have to pay attention to homogeneity of magnetic fields, which are used to scan desired samples inside the tomograph. We should know how to reduce inhomogeneity, which can cause misleading information at the final images of samples [5]. Our theory was proved during experiments solved by NMR team at Institute of Scientific Instruments of Academy of Sciences in Brno. Generally, inhomogeneity of magnetic fields at magnetic resonance imaging cause contour distortion of images. This paper presents the experimental method, which can easily create the map of electromagnetic induction at any defined area inside the magnetic resonance tomograph. The coefficients of Legendre multinomial, which are computed using measured values of magnetic induction at exactly defined discrete points are used for creating map of magnetic field. If we know these coefficients, we are able to compute magnetic induction at any point of defined area. At the ideal case, there should be no difference between measured data and approximated data.

2. MR PRINCIPLES

In quantum mechanics, spin [2] is important for systems at atomic length scales, such as individual atoms, protons or electrons. One of the most remarkable discoveries associated with quantum physics is the fact, that elementary particles can possess non zero spin. The spin carried by each elementary particle has a fixed value that depends only on the type of particle, and cannot be altered in any known way. Particles with spin can possess a magnetic dipole moment, just like a rotating electrically charged body in classical electrodynamics. The main principle of magnetic resonance spectroscopy and magnetic resonance imagining is, that radiofrequency fields (RF pulses) excite transitions between different spin states in a magnetic field.

The information content can be retrieved as resonance frequency, spin to spin couplings and relaxation rates. We can imagine, that protons are rotating along their axes and there is also a wobbling motion called precession, that occurs when a spinning object is the subject of an external force. Thanks to the positive charge of protons and its spin, protons generate a magnetic field and gets a magnetic dipole moment. If the protons are placed in a magnetic field, the magnetic moment will do precessional motion about the direction of magnetic field with specific frequency. This frequency is called Larmor frequency and can be described by the Larmor Equation [2]

$$\Omega = \gamma \cdot B \tag{1}$$

where Ω [MHz] is the frequency of precession, [MHz/T] is the gyro magnetic ratio and B is strength of external magnetic field. Pulse sequence is a pre-selected set of defined RF and gradient

pulses, usually repeated many times during a scan. Pulse sequences control all hardware aspects of the measurement process. At the x-y plain, there is scanning coil, which is used for scanning of FID signal

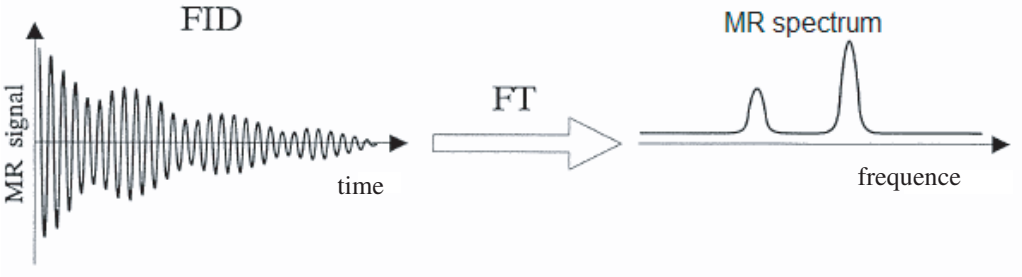


Figure 1: Fid signal and MR spectrum.

After excitation pulses, the spins has tendency to minimize transverse magnetization and to maximize longitudinal magnetization. The transverse magnetization decays toward zero with characteristic time constant T2 and the longitudinal magnetization returns towards maximum with a characteristic time constant T1.

3. INHOMOGENEITY CORRECTION

Magnetic resonance, which has a great performance in soft tissue imaging, is very sensitive to magnetic field inhomogeneities. The large value of gyro magnetic coefficient causes a significant frequency shift even for few parts per million (ppm) field inhomogeneity, which causes distortions in both geometry and intensity of the MR images. In this article, we will mainly concentrate on mapping of basic magnetic field inside the working space of tomograph, which is distorted by itself.

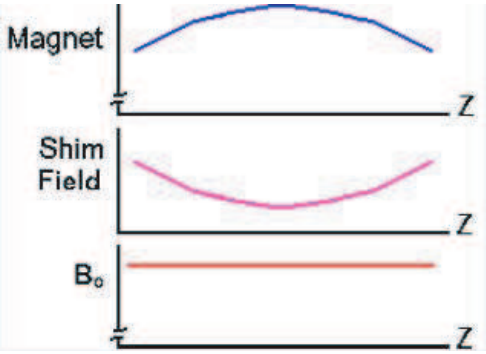


Figure 2: Principle of inhomogeneity correction.

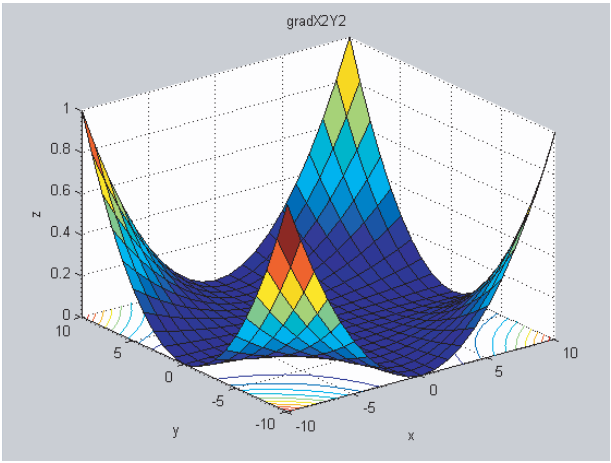


Figure 3: Correction gradient X^2Y^2.

In fact the principle of inhomogeneity correction is very simple, as we can see on Fig. 2. The blue trace shows a magnetic field of coil, which is used for creating homogeneous magnetic field (theoretically homogeneous), but this field is in praxis inhomogeneous. If we will have a map of this inhomogeneous field, we will apply shim field, which is opposite to inhomogeneous field. In best case, we will gain homogeneous field as a result.

Real map of the inhomogeneous field is much more complicated than our example. In this case we need to create a set of shim fields, which will be able to trace real map. We can see an example of this correction shim field generated in Matlab on Fig. 3.

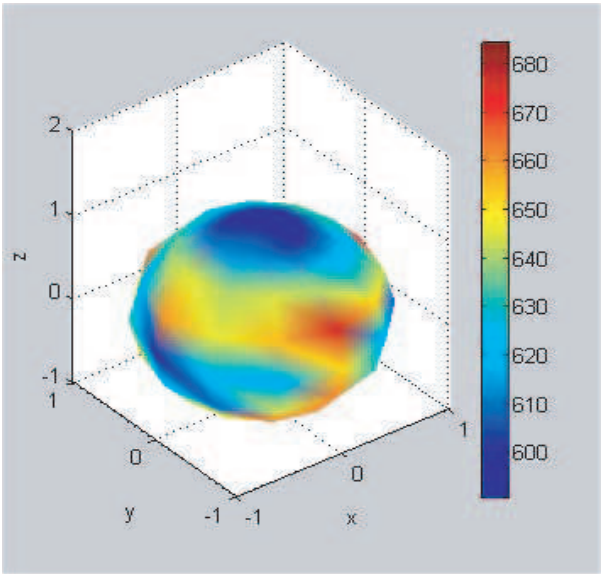


Figure 4: 3D map on the surface (measured values).

4. LEGENDRE POLYNOMS THEORY

If we want to determine the magnetic induction values in the specific points of measured area, we should use Legendre multinomial [4] which are defined according to Equation (2)

$$P_n(z) = \frac{1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n-1)}{n!} \cdot \left[z^n - \frac{n(n-1)}{2(2n-1)} z^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2 \cdot 4(2n-1)(2n-3)} z^{n-4} - \ldots \right] \quad (2)$$

Legendre multinomial of zero and first order

$$P_0(z) = 1 \quad (3)$$

$$P_1(z) = z = \cos \upsilon \quad (4)$$

Legendre multinomial of higher order are defined according to recursion formula

$$P_{n+1}(z) = [(2n+1) \cdot z \cdot P_n(z) - n \cdot P_{n-1}(z)]/(n+1) \quad (5)$$

Now we can define functions associated to Legendre multinomial, which are derivation of Legendre multinomial

$$P_{n+1}^m(z) = [(2n+1) \cdot z \cdot P_n^m(z) - (n+m) \cdot P_{n-1}^m(z)]/(n-m+1) \quad (6)$$

$$P_n^{m+2}(z) = \frac{2 \cdot (m+1) \cdot z}{\sqrt{1-z^2}} \cdot P_n^{m+1} - (n-m) \cdot (n+m+1) \cdot P_n^m(z) \quad (7)$$

Magnetic field induction can be approximated at any point of measured area. These points can be selected using spherical coordinates $[r, \theta, \varphi,]$, so we can define approximation formula as follows

$$B_a(r, \theta, \varphi) = \sum_{k=0}^{N_K} \cdot \sum_{m=0}^{m=k} \cdot r^k \cdot P_{m,k}(\cos \theta) \cdot [A_{m,k} \cos m \cdot \varphi + B_{m,k} \sin m \cdot \varphi] \quad (8)$$

where N_K is the highest order of Legendre multinomial for chosen approximation, A_{mk} a B_{mk} are unknown coefficients. N_K is defined according to sampling theorem and depends on number of measured points N_b:

$$N_k = \frac{N_b}{2} - 1 \quad (9)$$

Coefficients A_{mk} a B_{mk} is then possible to find like the minimum value of this formula:

$$\Psi = \min \sum_{i=1}^{N_m} (B_{im} - B_{ia})^2 \quad (10)$$

where B_m are measured values of magnetic induction at the desired area (circle, sphere, cylinder) and B_a are approximated values of magnetic induction. This method is known as Least Square method.

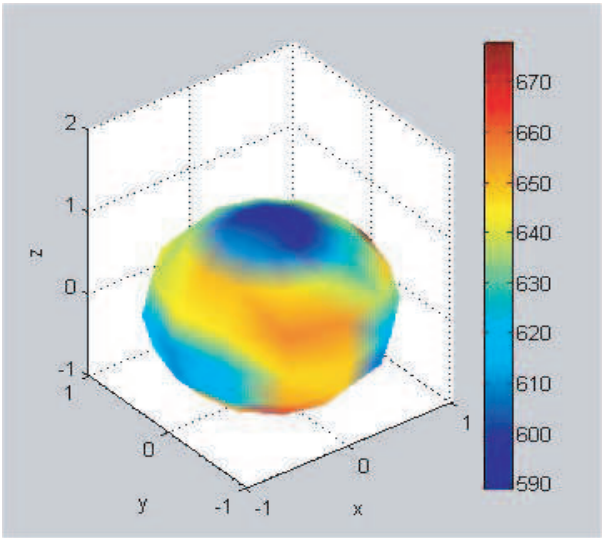


Figure 5: 3D map on the surface (computed values).

5. EXPERIMENTAL RESULTS

All values of magnetic induction on following figures are presented at [μT] unit. Fig. 2 shows three-dimensional map of the field on the surface of sphere, which is created only from measured values. Fig. 3 is map created from computed values, it means values which were computed during minimum searching (least square method) in Matlab. At the Fig. 4 and Fig. 5 is a comparison using 2D contour plot.

Finally if we want to get values of magnetic field induction inside the sphere, we use computed coefficients. Then we generate new coordinates of desired points inside the sphere and compute map of the field. This map can be any slice through the sphere as we can see in the Fig. 6.

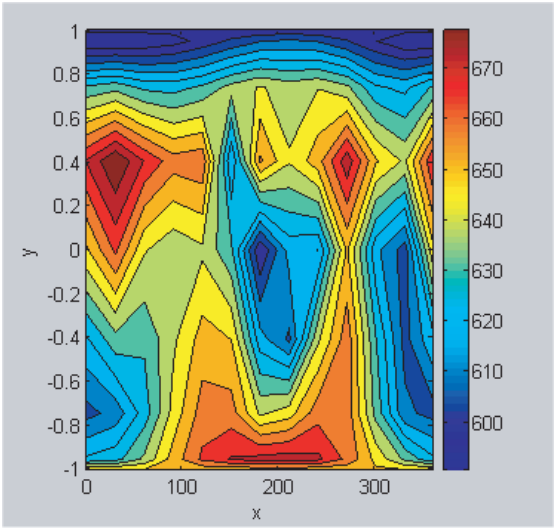


Figure 6: Contour map of the field (measured values).

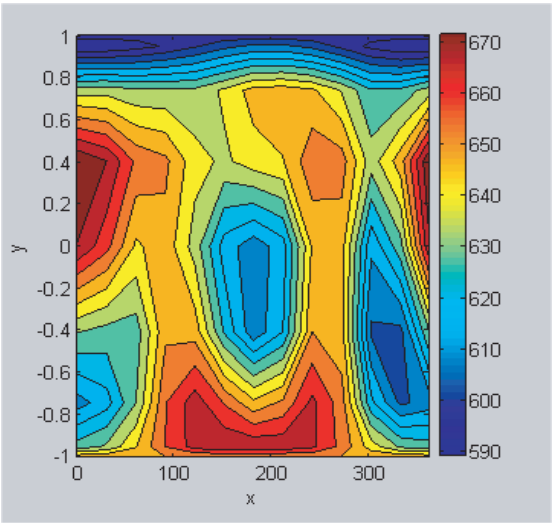


Figure 7: Contour map of the field (computed values).

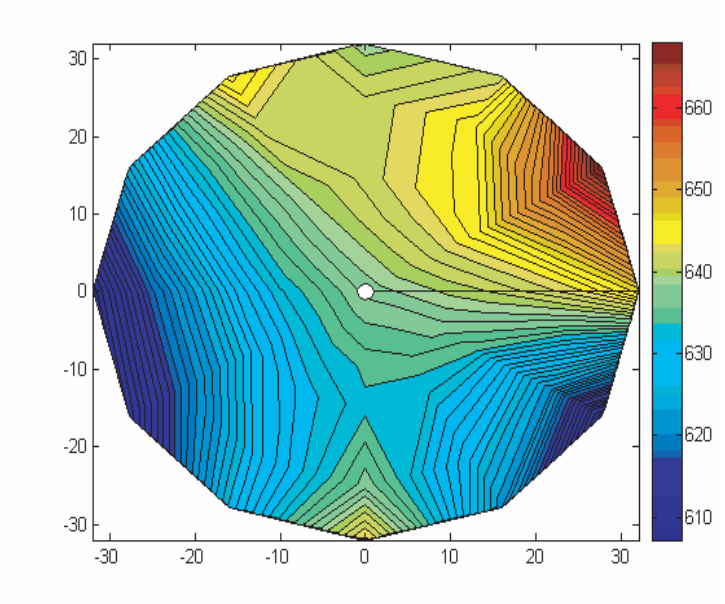


Figure 8: Slice through the sphere (measured values).

6. CONCLUSION

We proposed a method for magnetic field mapping and approximation on the basis of measured values along specific area. As we can see from results, the map created from measured and computed values are quite similar. Future work can be directed towards minimization of differences between measured values of magnetic induction and approximated values, so we can obtain exact coefficients for approximation inside the sphere.

ACKNOWLEDGMENT

The research described in the paper were financially supported by FRVŠ by research plan No. MSM 0021630513 ELCOM, No. MSM 0021630516 and grant GAAV No. B208130603. of the Ministry of Education, Youth and Sports of the Czech Republic.

REFERENCES

1. Fiala, P., E. Kroutilová, and T. Bachorec, Modelování elektromagnetickhých polí, počítačová cvičení vyd. Brno: VUT v Brně, FEKT, Údolní 53, 602 00, Brno, 2005.
2. Haacke, E. M., R. W. Brown, M. R. Thomson, and R. Venkatesan, *Magnetic Resonance Imaging-physical Principles and Sequence Design*, John Wiley & Sons, ISBN 0-471-48921-2, 2001.
3. Stratton, J. A., *Teorie Elektromagnetického pole*, SNTL Praha, 1961.
4. Angot, A., Used mathematics, Státní nakladatelství technické literatury, Praha 1972.
5. Bartušek, K., Z. Dokoupil, and E. Gescheidtová, "Magnetic field mapping using an asymmetric spin echo MRI sequence around metal implants," *Measurement Science and Technology*, roč. 17, č. 12, ISSN 0957-0233, 2006.

Design Simulation and Optimization the Source of Light

Eva Kroutilová, **Tomáš Kříž**, **Pavel Fiala**, **and Michal Hadinec**
Department of Theoretical and Experimental Electrical Engineering, Brno University of Technology
Kolejni 4, 612 00 Brno, Czech Republic

Abstract— This paper presents information about design of light sources, which is intended for commercial use. Required properties were continuous spectral characteristic with respect to active wavelength area, 3D light characteristics. Design of light source on basis R-FEM method [5] consists of classical used light source, as well as experimental results, are presented. The light source was designed, optimized and tested for the research activity.

1. INTRODUCTION

In the past, in Lighting Engineering, the elementary projection method and other geometrical optical methods were used for the modeling of illumination [6–9]. There are described a radiosity methods well suited to dynamic changes while requiring less memory compared to classical hierarchical radiosity. The light originating at emitters and arriving at a receiving input surface is stored on separate meshes, each one corresponding to an emitter-receiver pair. These meshes are called interaction meshes and facilitate the handling of dynamic changes since for each input surface the surfaces illuminating it can be determined very quickly. The interesting application of radiosity method is in a new three-dimensional model applicable for characterizing the propagation characteristics for microcellular communications. For microcellular communication, the working frequency is about 1.7–2.0 GHz. The wavelength is about 15–18 cm. In essence, these wavelengths are comparable with the dimensions of cavities such as windows, balconies, etc. on buildings. The source [8] get a computer graphics algorithm for simulating the propagation of light and its interaction with matter on a massively parallel computer is presented. This algorithm, called the tagged shooting method, is designed for a virtual machine containing a great number of simple communicating processors arrayed into a cubical three-dimensional lattice. The algorithm is similar in spirit to the classical progressive refinement radiosity method designed for more conventional computers but is not an adaptation of that technique to massive parallelism. Its known the methods [9] which implements radiosity using stochastic raytracing, which affords both instruction-level and data parallelism.

Today, new directions in the modeling of lighting systems are emerging. These new directions start from the Computer Graphics modeling methods and from new methods of numerical modeling used for the CAD/CAE solution of technical problems from different industrial sectors. The main applications are in architectural engineering and domestic architecture. Computer Graphics allows us to use the method of Global illumination. The goal of global illumination is to generate realistic-looking images of objects and scenes, which may or may not exist in reality, by simulating the way light is transported. The Global illumination method is based on the light transport mechanism in the real world. The light transport mechanism can be expressed in terms of BRDF (Bi-directional Reflectance Distribution Function) of each element, and the most common and practical way is dividing the BRDF into seculars and diffuse components. Mathematically, Global illumination is a problem of solving numerical equations concerned with the convergence, convergence speed and whether the solution converges the right answer. The Radiosity and Ray-tracing methods are used to calculate energy propagation in each iteration step. The BRDF of the geometrical element is important as regards the efficiency of each method. If the BRDF is ideally diffused, the Radiosity method will converge and it will converge to right the answer. However, the Ray-tracing algorithm works more efficiently in calculating seculars reflection, refraction and caustic surfaces. The Radiosity and Ray-tracing algorithms can be measured in two aspects, namely the accuracy and efficiency in BRDF simulation, and the rendering speed. (for more information see reference [1]). Into paper is used the R-FEM method, which was created within the framework dissertation thesis [4] as novel numerical simulation method.

2. THE DESIGN BY R-FEM METHOD

The R-FEM method is a new direction in the modeling of lighting systems [3–5]. It utilizes the similarity between physical models. This paragraph demonstrates the usage of analogy between

different physical models for the modeling of light problems, detailed described in PhD thesis [4]. The R-FEM method is able to solve tasks that fulfill the condition $\lambda_S \ll \max(D) \wedge \lambda_S < 10.\max(D)$, where λ_S is the source of light wavelength and D is one of the geometrical dimensions of the modeling task. It can be used for to model more complicated physical problems than the methods mentioned up to now. An example of a more complicated physical problem, which we can solve by the R-FEM method, is the modeling of light intensity distribution in interior or exterior spaces with non-homogeneous environment, where the light has passed through some impure air (e.g., filled with smoke, fog, mist, vapour, dust, etc.).

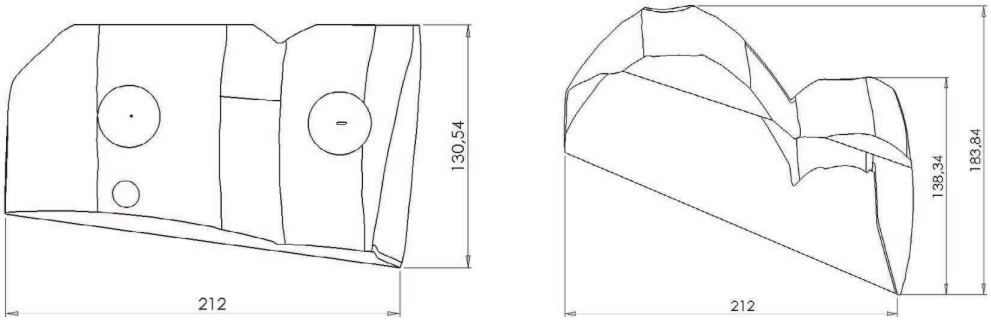


Figure 1: Geometrical configuration source of the light.

In technical praxis we often encounter conjugate problems. A necessary part of the design process during the development and measurement of light sources is the modeling and experimental verification of results. The most accurate mathematical models of the sources of light include models based on the radiation principle. One possibility is to use standard one — purpose programs while another possibility offers the usage of sophisticated numerical methods, among them the finite element method, for example the ANSYS program. The ANSYS program uses standard program tools such as modeling, discretization into a net of elements, solvers, evaluation, and interpretation of the results. The crux of the whole problem lies in the transformation of thermal field quantities into optical quantities. This can be done using the general rules described in [4]. In the following text the basics of modeling the primitive light problem are described. The verification of the model of light source is done via experiment and then it continues to the hollow light guide problems and it was also verified by experiment (for more information, see references [3–5]). The geometrical situations that were modeled and verified is shown in the Fig. 1.

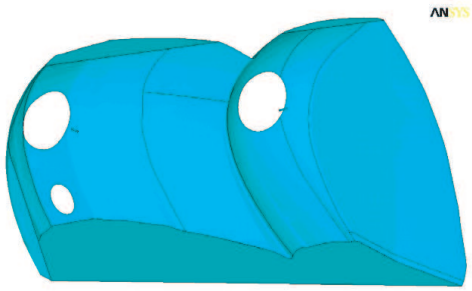


Figure 2: Geometrical configuration source of the light.

3. NUMERICAL MODELING

The formulation of the basic thermal model is based on the first law of thermodynamics

$$q + \rho\, c\, v \cdot \mathrm{div}T - \mathrm{div}(k\ \mathrm{grad}T) = \rho c \left(\frac{\partial T}{\partial t}\right) \tag{1}$$

where q is the specific heat, ρ is the specific weight, c is the specific solidification heat, T is the temperature, t is the time, k is the coefficient of calorific conduction, v is the velocity of flow.

This model can, with respect to the application of Snell's principles and according to the Stefan-Boltzmann principles, heat transfer by way of radiation between surfaces with relative indexes i, j is formulated as be simplified into the form

$$q_{ri} = \sigma\ \varepsilon_i\ A_{i,j}\ S_i(T_i^4 - T_j^4) \tag{2}$$

where q_{ri} is the specific heat transferring from surface with index i, σ is the Stefan-Boltzmann constant, ε_i is the emissivity of surface, $A_{i,j}$ is the projection factor of surface with index i to surface with index j, S_i is the area of surface with index i, T_i, T_j are the temperature of surfaces i, j.

When the projection factor is determined, it is possible to use the Galerkin principles for converting this problem into model (1). The mathematical problem is solved by classical numerical method, finite element method (FEM) [3]. The result of modeling by the R-FEM method is shown in the Fig. 3.

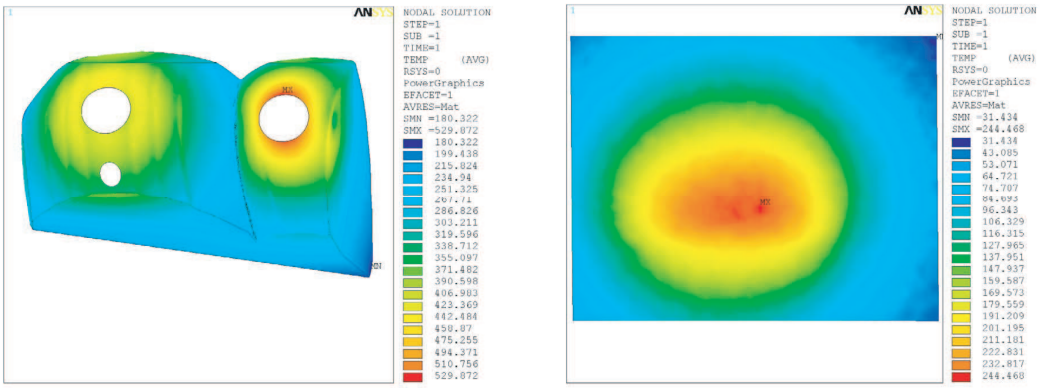


Figure 3: The result of the R-FEM method-lighting intensity $E(lx)$.

4. VERIFICATION THE R-FEM BY EXPERIMENT

The results of the experimental verification by LED sensor is given in Fig. 4. There are differences between the values obtained by modeling, Fig. 3 and experimental measurement, Fig. 5, ranging from 5–15%, depending on the distribution of the net of elements. When the elements of the net are of a lower density, the differences are also lower. This problem requires the net of elements to be optimized.

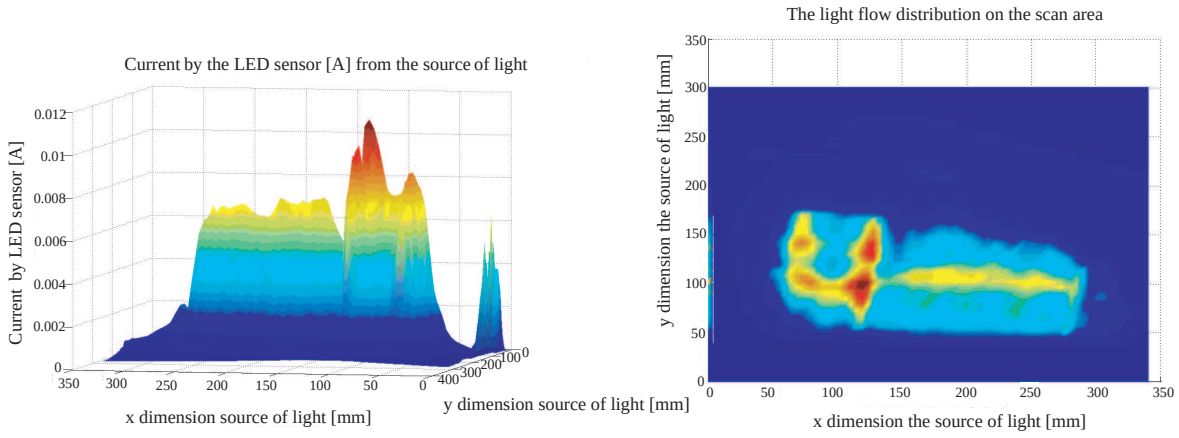


Figure 4: The shape of distribution of the light intensity by experiment with the LED.

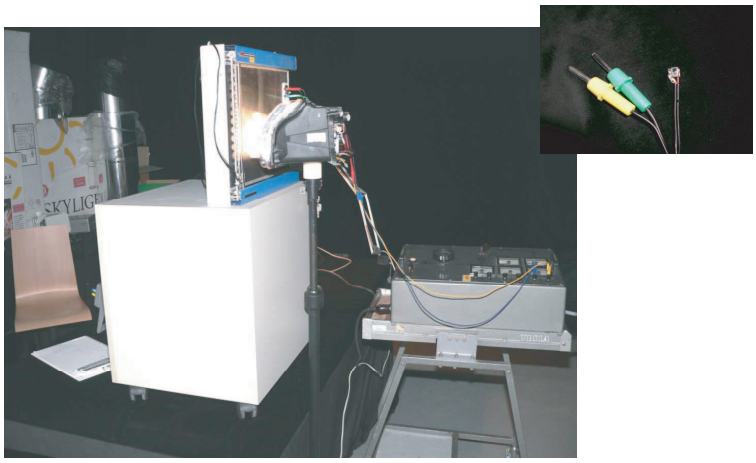


Figure 5: The experiments with light system.

5. CONCLUSION

This article describes design source of light, which is intended for commercial use. Required properties were continuous spectral characteristic with respect to active wavelength area, 3D light characteristics. Design of light source consists of classical used source of light, as well as experimental results, are presented. The light source was designed, optimized and tested for the research activity.

ACKNOWLEDGMENT

The research described in the paper were financially supported by FRVŠ by research plan No. MSM 0021630516 and grant GAAV No. B208130603.

REFERENCES

1. http://www-geog.bu.edu/brdf/brdfexpl.html.
2. Kroutilová, E., M. Bernard, and P. Fiala, "Illumination of interiors by the Hollow Light Guides," *14th International Conference Light 2003*, Bratislava, 2003.
3. Kroutilová, E. and P. Fiala, "Light guide modeling," *Energy Forum 2004*, 338–341, Sofia, Technical University Sofia, Bulgaria, 2004.
4. Kroutilová, E., "Automated system of calculation of reflecting surface of light sources," *Ph.D. thesis*, VUT v Brně, FEKT, Brno srpen, 2004.
5. Fiala, P. and E. Kroutilová, "Numerical modeling of the special light source with novel R-FEM method," *PIERS Online*, Vol. 2, No. 6, 644–647, 2006.
6. Carre, S., J. M. Deniel, E. Guillou, and K. Bouatouch, "Handling dynamic changes in hierarchical radiosity through interaction meshes," *Proceedings of the Eighth Pacific Conference on Computer Graphics and Applications*, 40–436, Oct. 3–5, 2000.
7. Tsang, K.-F., W.-S. Chan, D. Jing, K. Kang, S.-Y. Yuen, and W.-X. Zhang, "Radiosity method: a new propagation model for microcellular communication," *IEEE Antennas and Propagation Society International Symposium*, Vol. 4, 2228–2231, Jun. 21–26, 1998.
8. Kochevar, P., "Too many cooks don't spoil the broth: light simulation on massively parallel computers," *Proceedings of the 3rd Symposium on the Frontiers of Massively Parallel Computation*, 100–109, Oct. 8–10, 1990.
9. Baker, P., T. Todman, H. Styles, and W. Luk, "Reconfigurable designs for radiosity," *13th Annual IEEE Symposium on Field-Programmable Custom Computing Machines (FCCM 2005)*, 95–104, Apr. 18–20, 2005.

Inversion Reconstruction of Signals Measured by the NMR Techniques

Eva Kroutilova[1], **Miloslav Steinbauer**[1], **Premysl Dohal**[1]
Michal Hadinec[1], **Eva Gescheidtova**[1], **and Karel Bartušek**[2]

[1]Brno University of Technology, Czech Republic
[2]Academy of Science of the Czech Republic, Czech Republic

Abstract— The paper describes the magnetic resonance imaging method applicable mainly in MRI and MRS in vivo studies. There is discussed and solved the effect of changes of magnetic fields in MR tomography. This article deals with the reverse reconstruction results obtained from the numerical simulation of MR signals by various techniques, which will be usable for the experimental results verification.

1. INTRODUCTION

Nuclear Magnet Resonance (NMR) is well known diagnostic non-destructive and non-invasive method [2–6]. It is used to investigation of the materials properties. One of the often use of NMR is medical application [8, 11]. The magnetic resonance imaging (MRI) is used to the propel a ferromagnetic core. The concept was studied for future development of microdevices designed to perform minimally invasive interventions in remote sites accessible through the human cardiovascular system. A mathematical model is described [3] taking into account various parameters such as the size of blood vessels, the velocities and viscous properties of blood, the magnetic properties of the materials, the characteristics of MRI gradient coils, as well as the ratio between the diameter of a spherical core and the diameter of the blood vessels. Other paper [4] present a MRI-compatible micromanipulator, which can be employed to provide medical and biological scientists with the ability to concurrently manipulate and observe micron-scale objects inside an MRI gantry. The micromanipulator formed a two-finger micro hand, and it could handle a micron-scale object using a chopstick motion. The material engineering use the NMR too. For example NMR logging is an advanced method in formation evaluation and oil field production. It can provide the porosity, permeability, bound water volume, free fluid volume and oil viscosity. NMR logging is playing more and more important roles in oil and gas exploitation.

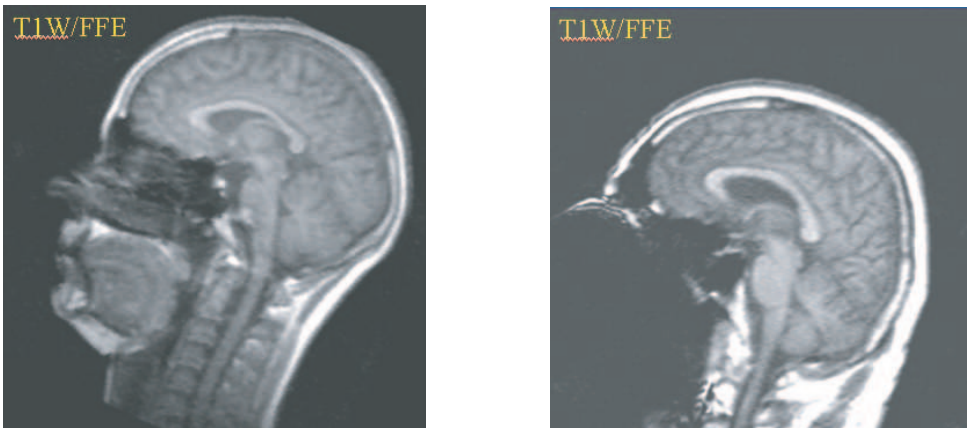


Figure 1: The sample of two MRI results, the right results (a), the results with imaging loss.

The results of the NMR method and signal processing are sensitive to used materials with extreme differences like a susceptibility, conductivity and also permeability from the macroscopic view to the mater. It is going to deformations and it makes the NMR can not be use. The Fig. 1 shows information losses effect.

The paper brings the design of the hybrid numerical-experimental inverse method, which can reduce influence of the materials with extreme differences of the properties. The inverse methods

are used in different applications [1, 7]. There is problem with task nonlinearity. It is described in [1]. The authors develop two nonlinear inverse methods to reconstruct the conductivity profile from electromagnetic induction (EMI) measurements: the improved two-step inverse method based on the extended Born approximation (EBA) and the combination of the EBA and the contrast source inversion (CSI) method. In the first method, the nonlinear problem is recast as a two-step linear inversion and is solved by using the extended Born approximation. The application of inverse method in the NMR techniques is in the temperature measurements, for the direct identification of the surface heat flux. The temperature measurements are not always possible considering an aggressive environment or an inaccessible zone. That is why an inverse method has been developed: the direct problem with the unknown boundary condition (wall heat flux) is solved by adding an observation equation given by temperature measurements on the opposite face of the wall [7].

The next step is to find simply model for numerical modeling and NMR experiment-verification. There can be changed material properties and next experimentally measured. The idea of increase of MRI is in the hybrid experimental and numerical inverse method. The numerical results are used in the MRI experimentally obtained data. The application of numerical results to the NMR post-processing can much more increase the final images.

2. GEOMETRICAL MODEL

Fig. 2 describes the simply geometry for the numerical modeling. On both sides, the sample is surrounded by the referential medium. During the real experiment, the reference is represented by water, which is ideal for obtaining the MR signal. As shown in Fig. 2, in the model there are defined four volumes with different susceptibilities. The materials are defined by their permeabilites : material No. 1 – the medium outside the cube (air), $\chi = 0$, material No. 2 – the cube walls (sodium glass), $\chi = -11,67.10^{-6}$, material No. 3 is the sample material (sodium glass), $\chi = -11,67.10^{-6}$, quartz glass, $\chi = -8,79.10^{-6}$, the simax glass (commercial name), $\chi = -8,82.10^{-6}$, material No. 4 is the medium inside the cube (water with nickel sulfate solution NiSO4, $\chi = -12,44.10^{-6}$). The permeability rate was set with the help of the relation $\mu = 1 + \chi$. For the sample geometry according to Fig. 1, the geometrical model was built in the system. In the model there was applied the discretization mesh with 133584 nodes and 126450 elements, type Solid96 (Ansys). The boundary conditions (1) were selected for the induction value of the static elementary field to be $B_0 = 4,7000\,\mathrm{T}$ in the direction of the z coordinate (the cube axis) – corresponds with the real experiment carried out using the MR tomograph at the Institute of Scientific Instruments, ASCR Brno.

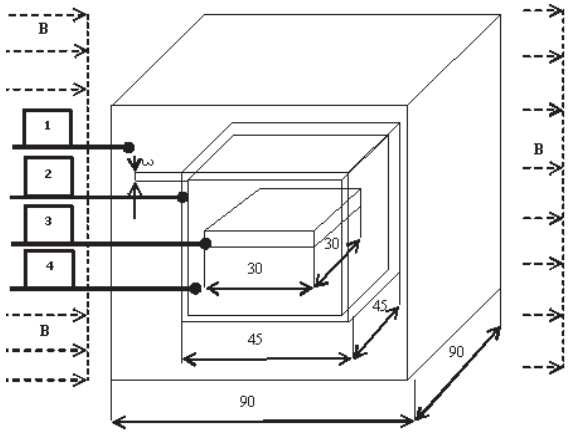


Figure 2: The sample geometry for numerical modeling.

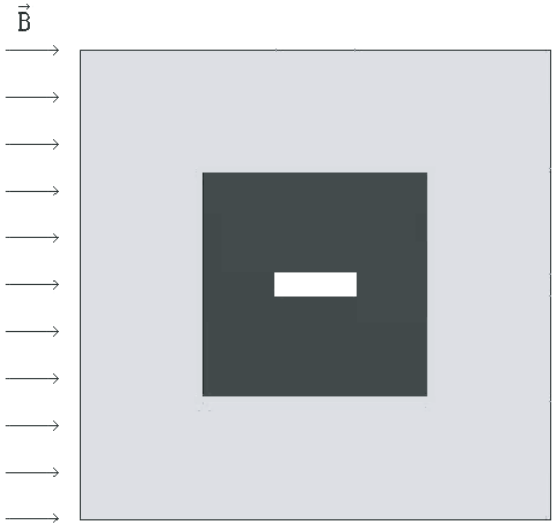


Figure 3: The geometrical model in the system ansys.

3. NUMERICAL ANALYSIS

The numerical modelling was realized using the finite element method (FEM) together with the Ansys system and described in paper [11]. As the boundary condition, there was set the scalar

magnetic potential φ_m by solving Laplace's equation on the geometrical model in Fig. 3

$$\Delta\varphi_m = \operatorname{div}\mu(-\operatorname{grad}\varphi_{\mathrm{m}}) = 0 \tag{1}$$

together with the Dirichlet boundary conditon

$$\varphi_m = konst. \quad \text{on the areas} \quad \Gamma_1 \text{ and } \Gamma_2 \tag{2}$$

and the Neumann boundary condition

$$\boldsymbol{u}_n \cdot \operatorname{grad}\varphi_{\mathrm{m}} = 0 \quad \text{on the areas} \quad \Gamma_3 \text{ and } \Gamma_4. \tag{3}$$

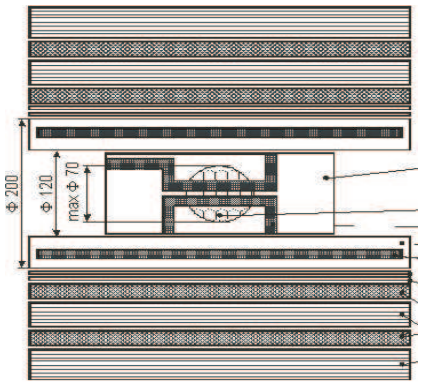


Figure 4: Elementary configuration of the MR magnet for the 200 MHz tomograph, ISI ASCR.

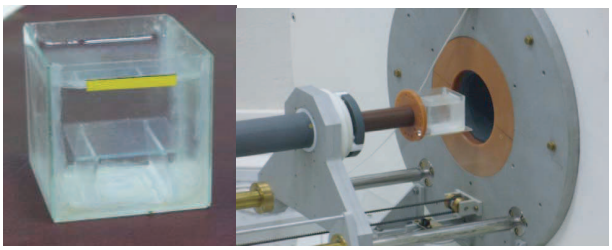


Figure 5: The measured preparation. The preparation seating in the tomograph.

4. NUMERICAL MODEL

The numerical modelling results are represented in Fig. 3 and Fig. 7. The numerical modelling results were then used for the representation of the module of magnetic induction **B** along the defined path. For the model meshing, the element size selected as optimum was $0,5.10^{-3}$ m. The boundary conditions $\pm\varphi/2$ were set to the model edges, to the external left and right boundaries of the air medium, as represented in Fig. 1. The excitation value $\pm\varphi/2$ was set using again the relation (21). This is derived for the assumption that, in the entire area, there are no exciting currents, therefore there holds for the *rot* $\boldsymbol{H} = 0$ and the field is irrotational.

Consequently, for the scalar magnetic potential φ_m holds

$$H = -\operatorname{grad}\varphi_{\mathrm{m}} \tag{4}$$

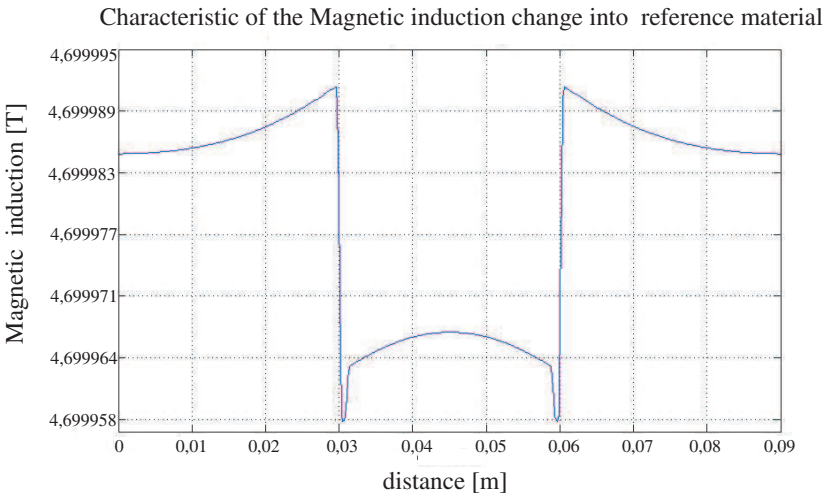


Figure 6: The magnetic flux density $\boldsymbol{B}$ pattern, without sample material.

The potential of the exciting static field with intensity $\boldsymbol{H}_0$ is by applying (4)

$$\varphi_m = \int \vec{H}_0 \cdot \vec{u}_z dz = H_0 \cdot z \tag{5}$$

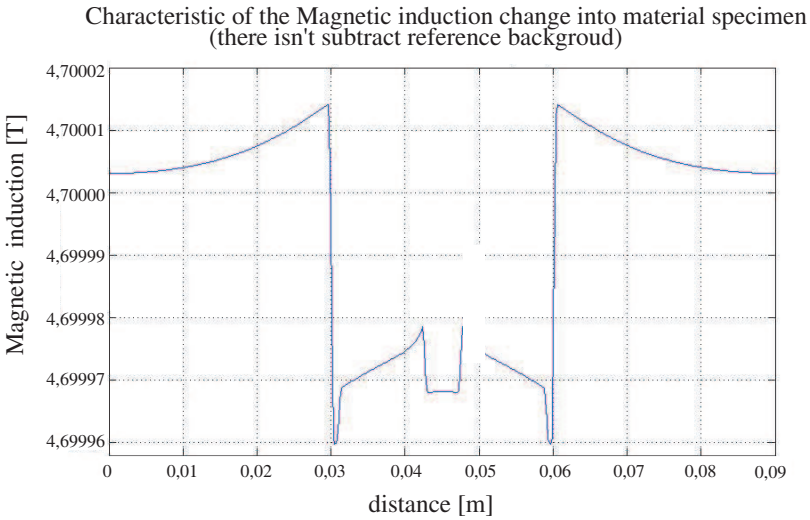


Figure 7: The magnetic flux density $\boldsymbol{B}$ pattern, with sample material.

5. EXPERIMENTAL VERIFICATION

The experimental measuring was realized using the MR tomograph at the Institute of Scientific Instruments, ASCR Brno. The tomograph elementary field $B_0 = 4,7000\,\mathrm{T}$ is generated by the superconductive solenoidal horizontal magnet produced by the Magnex Scientific company. The corresponding resonance frequency for the 1 H cores is 200 MHz. The Fig. 6 bring the numerical results of simply model solution without tested material. In the Fig. 7. is shoed result of numerical model with the tested material. The different of these results is showed in Fig. 8. The final result very good correspond with the theoretical conditions publicated in the thesis [10].

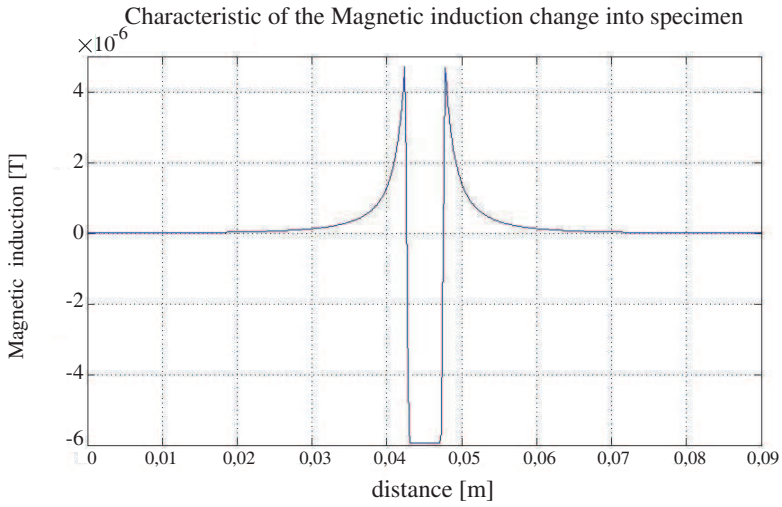


Figure 8: The magnetic flux density $\boldsymbol{B}$ pattern, differential result.

6. CONCLUSION

The numerical modelling and analysis of the task have verified the experimental results and, owing to the modificability of the numerical model, we have managed to advance further in the experimental qualitative NMR image processing realized at the ISI ASCR. This numerical and experimental models showed the possibility of inverse hybrid use.

ACKNOWLEDGMENT

The research described in the paper were financially supported by FRVŠ by research plan No. MSM 0021630513 ELCOM, No. MSM 0021630516 of the Ministry of Education, Youth and Sports of the Czech Republic and grant GAAV No. B208130603 and GA102/07/0389.

REFERENCES

1. Zhang, Z. Q. and Q. H. Liu, "Two nonlinear inverse methods for electromagnetic induction measurements," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 6, 1331–1339, June 2001.
2. Trakic, A., H. Wang, F. Liu, H. S. Lopez, and S. Crozier, "Analysis of transient eddy currents in MRI using a cylindrical FDTD method," *IEEE Transactions on Applied Superconductivity*, Vol. 16, Issue 3, 1924–1936, ISSN: 1051-8223, Sept. 2006.
3. Mathieu, J.-B., G. Beaudoin, and S. Martel, "Method of propulsion of a ferromagnetic core in the cardiovascular system through magnetic gradients generated by an MRI system," *IEEE Transactions on Biomedical Engineering*, Vol. 53, Issue 2, 292–299, ISSN: 0018-9294, Feb. 2006.
4. Yoshihiko, K., T. Tanikawa, and C. Kiyoyuki, "MRI-compatible micromanipulator design and implementation and MRI-compatibility tests," *29th Annual International Conference of the IEEE*, Engineering in Medicine and Biology Society, EMBS 2007, 465–468, Lyon, France, ISSN: 1557-170X, Aug. 22–26, 2007.
5. Arima, H., S. Kato, K. Maehata, K. Ishibashi, T. Nakamoto, and T. Shintomi, "Multipole analysis for absolute magnetic field measured by multi-probe pulsed-NMR method," *IEEE Transactions on Applied Superconductivity*, Vol. 10 , No. 1, 1450–1453, March 2000.
6. Long, H. F., Z. Li, Z. Xiao, and Y. Zhang, "Numerical simulation of NMR logging tools antenna," *7th International Symposium on Antennas, Propagation & EM Theory, ISAPE'06*, China, 1–3, Oct. 2006.
7. Reulet, P., D. Nortershauser, and P. Millan, "Inverse method using infrared thermography for surface temperature and heat flux measurements," *20th International Congress on Instrumentation in Aerospace Simulation Facilities, ICIASF'03*, 118–126, Aug. 25–29, 2003.
8. Vaughan, J. T., D. N. Haupt, P. J. Noa, J. M. Vaughn, and G. M. Pohost, "RF front end for a 4.1 tesla clinical NMR spectrometer," *IEEE Transactions on Nuclear Science*, Vol. 42, No. 4, 1333–1337, August 1995.
9. Fiala, P., E. Kroutilová, and T. Bachorec, Modelování elektromagnetických polí, počítačová cvičení, FEKT, VUT v Brně, Údolní 53, 602 00, Brno, s. 1–69, 2005.
10. Steinbauer, M., Měření magnetické susceptibility technikami tomografie magnetické rezonance, PhD Thesis, FEKT, VUT v Brně, Údolní 53, 602 00, Brno, 2006.
11. Fiala, P., E. Kroutilova, M. Steinbauer, M. Hadinec, and K. Bartusek, "The effect of non-homogenous parts into materials," *PIERS Online*, Vol. 3, No. 5, 1245–1249, 2007.

Numerical Modeling of Electromagnetic Field a Tornado

P. Fiala, V. Sadek, and T. Kriz
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejni 2906/4, 612 00 Brno, Czech Republic

Abstract— This study deals with the numerical model of both the physical and the chemical processes in the tornado. Within the paper, a basic theoretical model and a numerical solution are presented. We prepared numerical models based on the combined finite element method (FEM) and the finite volume method (FVM). The model joins the magnetic, electric and current fields, the flow field and a chemical nonlinear ion model. The results were obtained by means of the FEM/FVM as a main application in ANSYS software.

1. INTRODUCTION

The full electromagnetic-hydro-dynamic (EMHD) model of a tornado is a coupled problem in which there are coupled the electric, magnetic, fluid flow fields, electric circuit and chemical (dynamical ions) models. This model was built with the combined finite element methods (FEM) and the finite volume methods (FVM).

A more complete understanding of tornado-genesis must be developed before the feasibility of mitigation by heating fine structure, such as cold downdraft regions, can be determined. Also, the present severe storm diagnostic capability and numerical simulation codes are not yet suitable for real-time assessment of electromagnetic heating results. It is suggested that their be used in a tornado-genesis mitigation system:

- Real-time calculations. It would be desirable to predict the development faster than real-time to be able to provide better targeting.
- Continuously updated with fresh data from diagnostic systems. Validation using extensive field data.
- Nested grid calculations down to 58 meters or less in horizontal grid dimensions.
- Ability to calculate heating patterns with complex electromagnetic heating beam geometry.
- Inclusion of important microphysics considerations.

The diagnostics must also be capable of real-time operation, a one-second or less response time. It is generally known [1–7] that the tornado produces characteristic sferics of its own. In this regard, Jones [1] reported that the 10 kHz sferics associated with the tornado, but observed prior to the occurrence of the tornado, were much more intense than those associated with ordinary thunderstorms. Furthermore, after the tornado forms, a shift in the sferics frequency response is reported from low frequencies (app. 10 kHz) to high frequencies. Jones [2] reported that tornados have characteristic waveforms, and that there is a significant correlation between the number of sferic flashes at 150 kHz and the occurrence of tornados. Jones [3] also concludes that the flashing rate is a good indication of the intensity of the storms, and has given a tentative storm classification system based upon the flashing rate. Huebner et al. [6] were able to show that a variation of the sferic frequency spectra of tornados with respect to storms does exist. However, they were unable to confirm all of the predictions made by Jones.

According the last work of Kikuchi [9, 10] is possible to build model EMHD for numerical simulation of transient effects in the tornado. The examples of different tornado are showed in Fig. 1.

2. MATHEMATICAL AND NUMERICAL MODEL

Electromagnetic part is derived from Maxwell equations

$$rot\boldsymbol{H} = \boldsymbol{J}_T, \tag{1}$$

$$div\boldsymbol{B} = 0, \quad div\boldsymbol{D} = \rho \tag{2}$$

Figure 1: Different type of tornado body.

where $\boldsymbol{H}$ is the vector of magnetic field intensity, $\boldsymbol{B}$ is the magnetic flux density, $\boldsymbol{J}_T$ is the vector of total current density, $\boldsymbol{D}$ is the electric flux density , ρ is the electric charg density.

$$rot\boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \tag{3}$$

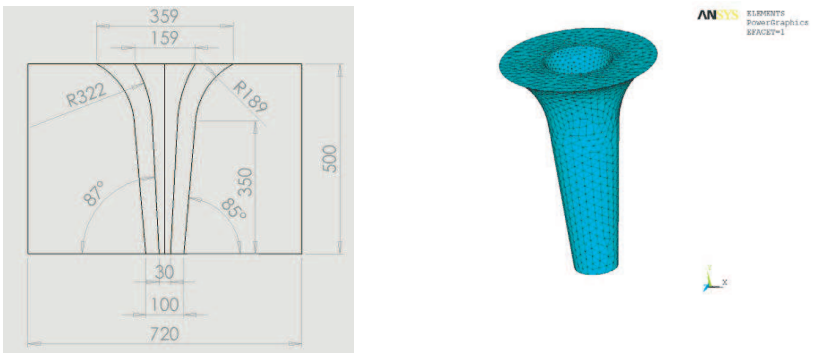
$$div\boldsymbol{J}_T = -\frac{\partial \rho}{\partial t}, \tag{4}$$

where $\boldsymbol{E}$ is the vector of electric field intensity. Vector functions of electric, magnetic field are expressed by means of a scalar electric ϕ_e and vector magnetic potentials $\boldsymbol{A}$. Final current density from (4) $\boldsymbol{J}_T$ is influenced by velocity $\boldsymbol{v}$ of the flowing ion solution and outer magnetic field

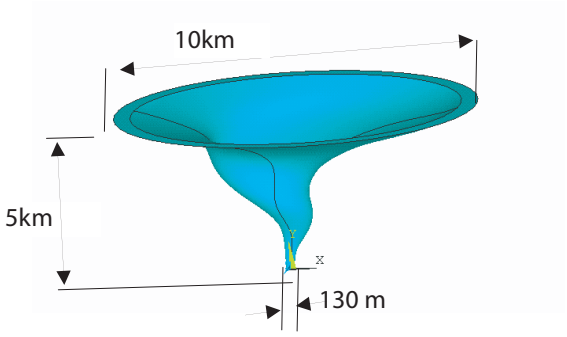
$$\boldsymbol{J}_T = \gamma(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) - \frac{\partial(\varepsilon \boldsymbol{E})}{\partial t} + \gamma \frac{m d\boldsymbol{v}}{q\, dt}. \tag{5}$$

where m is particle mass, q is the electric charge, γ is the conductivity of parts tornado body from the macroscopic view.

The model from relations (1) to (5) was solved by finite element methods (FEM) [11] with



(a)



(b)

Figure 2: (a) Basic geometrical dimensions of the first tornado model (left) in meters and its FEM model (right); (b) Basic geometrical dimensions of the second tornado model.

ANSYS system. The geometrical model was built with two modifications. First of them is simply geometrical model, showed in Fig. 2(a).

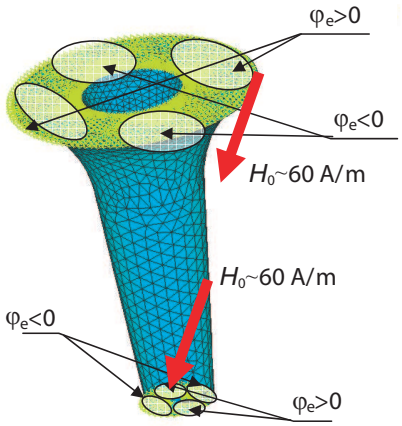
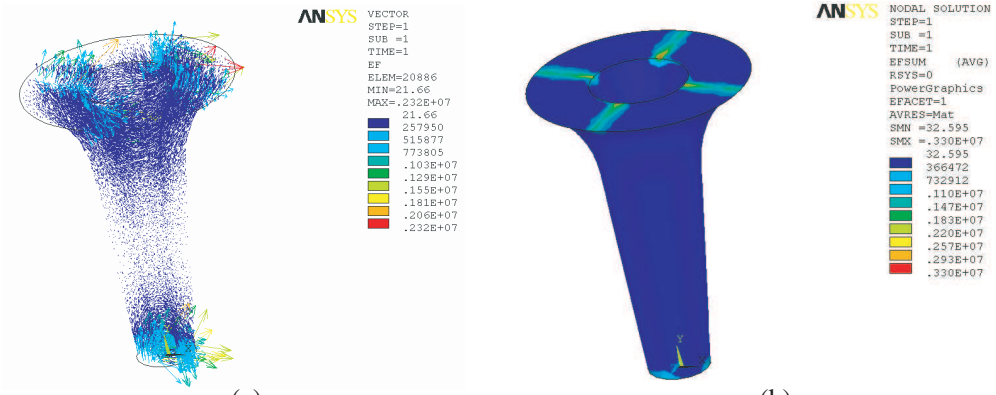


Figure 3: Boundary and initial condition of EHD and EMHD tornado model, scalar electric potential φ_e, magnetic field intensity $\boldsymbol{H}_0$.

The parametrs of the first tornado can be written as a several parameters. We have used the bold signed parameters. Vnitrni rychlost: 100–190 km/h, rychlost otaceni: 72–720 km/h, vnejsi prumer: 10–100 m, vnitrni prumer: 15–30.48 m, rychlost pohybu: 0–120 km/h, výška: 50–500 m.

3. NUMERICAL SOLUTION FEM/FVM

The numerical model was prepared by means of ANSYS tools [11] and main FEM/FVM solution was solved with APDL program over ANSYS system. In the Fig. 3 we can see boundary condition



(a) (b)

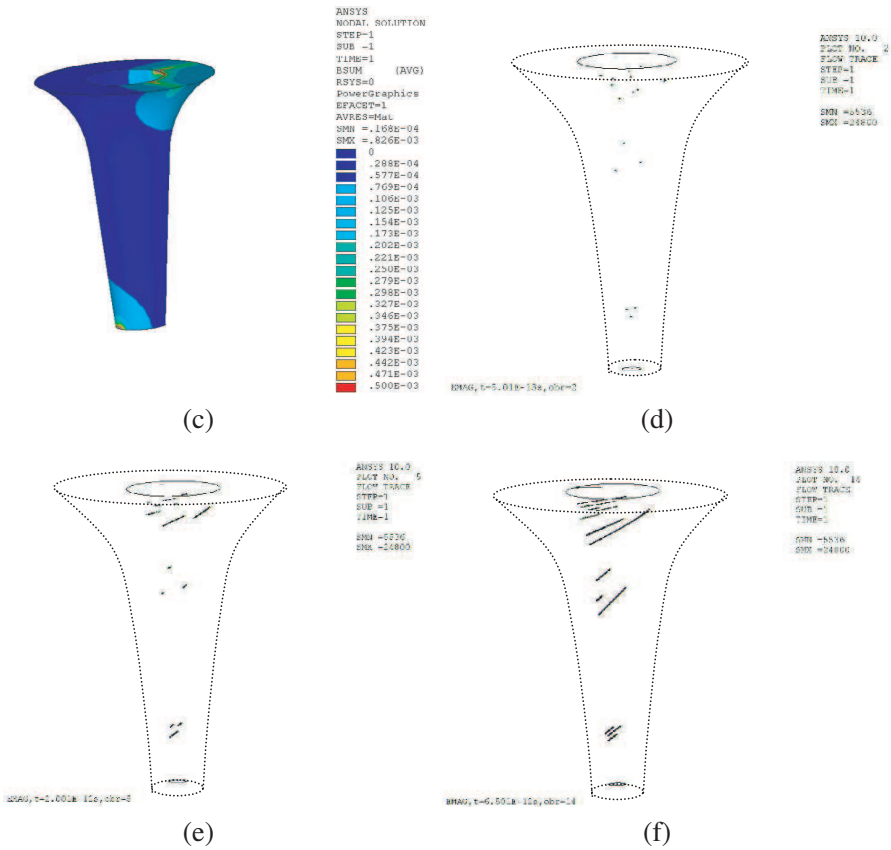


Figure 4: Simply solution of EHD tornado model, displacement of a vector electric field intensity (a) and its module (b), displacement of a magnetic field intensity module (c), positive charge moving, $t = 0.5$ ps (d), positive charge moving, $t = 2.0$ ps (e), positive charge moving, $t = 6.5$ ps (f).

for basic solution of electric charge moving. This model definition is not so perfect expression of tornado state. There are only 50 points where were different electric charge placed and started the transient EHD effect. A first results from the model are showed in the figure Fig. 4. Therefore was built full EMHD tornado model according of (1) to (5) condition. The model is under the test just now. Both model was solved for time interval from $t \in \langle 1.10^{-12}, 1.10^{-2} \rangle$ sec. Then we can observe the ions moving trend and changing of electromagnetic field in the tornado body during the time interval.

4. CONCLUSION

This work deals with EHD and EMHD numerical tornado model. There is basic mathematical and numerical description in the article and first results of electromagnetic field distribution and ions moving in the critical part of tornado body. Such analysis could be used for understanding of time state and time dependent of tornado effects in the breadth time interval observation. On the two different tornado geometrical model was tested the EHD and EMHD numerical model.

ACKNOWLEDGMENT

The research described in the paper were financially supported by FRVŠ by research plan No. MSM 0021630513 ELCOM, No. MSM 0021630516 and grant GAAV No. B208130603. of the Ministry of Education, Youth and Sports of the Czech Republic.

REFERENCES

1. Eastlund, B. J., "Systems considerations of 'Weather modification experiments using high power electromagnetic radiation," *Workshop on Space Exploration and Resources Exploitation — Explospace*, European Space Agency, October 20–22, 1998.
2. Glaser, P. E., "Feasibility study of a satellite solar power station," NASA Contact. Rep. CR-2357, NTiS N741 7784, *Nat. Tech. Inform. Serv.*, Springfield, VA, USA, 1974.

3. *Policy Implications of Greenhouse Warming: Mitigation, Adaptation and the Science Base*, National Academy Press, Washington, DC, 1992.
4. Xue, M. and K. K. Droegemeier, The Advanced Regional Prediction System (ARPS), March 1997, "A multiscale nonhydrostatic atmospheric simulation and prediction tool: Model dynamics," *Monthly Weather Review, IEEE A & E Systems Magazine*, June 2007.
5. Matsumoto, H., et al., "A feasibility study of power supplying satellite (PSS)," *Space Power*, Vol. 12, 1, 1993.
6. Mankins, J., "A fresh look at the concept of space solar power, SF5 '97," *Energy and Space for Humanity*, August 26, 1997.
7. Eastlund, B. J. and L. M. Jenkins, "Space-based concepts for taming tornadoes, *51' International Astronautical Congress*, Rio de Janeiro, Brazil, October 2, 2000.
8. Xue, M., et al., "Simulation of tornado vortices within a supercell storm using adaptive grid refinement technique," *1i' Conference on Severe Local Storms*, Saint Louis, Missouri, American Meteorological Society, October 4, 1993.
9. Kikuchi, H., *Electrohydrodynamics in Dusty and Dirty Plasmas, Gravito-electrohydrodynamics and EHD*, Kluwer Academic Publishers, 2001.
10. Kikuchi, H., "Helicity or vortex generation in hydrodynamic (HD), magneto-hydrodynamic (MHD), and electrohydrodynamic (EHD) regimes," *Progress In Electromagnetics Research Symposium Proceedings*, 996–999, Beijing, China, March 26–30, 2007.
11. *Ansys. User's Manual*, Svanson Analysys System, Inc., Huston, USA, 2003.
12. Fiala, P., "Model of induction flowmeter DN100," Research report No. 2/01, Laboratoř modelování a optimalizace polí v elektromechanických systémech, FEI VUT Brno, Brno, Czech Republic, June 21, 2001.
13. Lehnert, B., "Extended electromagnetic theory: New solutions of old problems," *Progress In Electromagnetics Research Symposium 2004*, Pisa, Italy, March 28–31, 2004.
14. Kikuchi, H., "Roles of EHD, MHD and (T)HD in tornadic thunderstorms," *Progress In Electromagnetics Research Symposium 2004*, Pisa, Italy, March 28–31, 2004.
15. Callebaut, D. K. and G. K. Karugila, "Plasma up to highers including its magnetic effects," *Progress In Electromagnetics Research Symposium 2004*, Pisa, Italy, March 28–31, 2004.

The Numerical Modeling and Conformal Mapping Method Applied to the Strip-centered Coaxial Line Analysis

V. Šádek, P. Fiala, and M. Hadinec
Faculty of Electrical Engineering and Communication, Brno University of Technology
Kolejni 2906/4, Brno 612 00, Czech Republic

Abstract— This contribution compares five different solutions of the characteristic impedance of the strip-centered coaxial line (SCCL) - two different strategies of boundary element method (BEM-1, BEM-2), numerical evaluating of the Schwarz-Christoffel integral (NSC) and two different solutions based on conformal mapping method (CMM-1, CMM-2).

All these methods are quasi-static; therefore here presented results (all methods) are valid for wavelengths longer than the longest electrical dimension of the SCCL – diameter of the surrounding electrode. On higher frequencies occur not-TEM modes.

The aim of this work is to compare different methods for one structure analysis – to learn about their properties, advantages and disadvantages. The results and gained experiences should be applied to other projects in the future.

1. INTRODUCTION

A lot of different electronic equipments have to work together. Unfortunately the differences of their power levels are over 200 dB very often. Coaxial structures are widely used because of their good shielding effect, which suppress the fields around strong distortion sources (e.g., transmitting antenna feeder) and protect sensitive parts of receivers, measurement inputs of the equipments etc.

Whereas coaxial line (two concentric cylindrical electrodes) is widely known, strip-centered coaxial line (SCCL, see Fig. 1.) is mentioned rarely (founded only in very special literature like [1–3]).

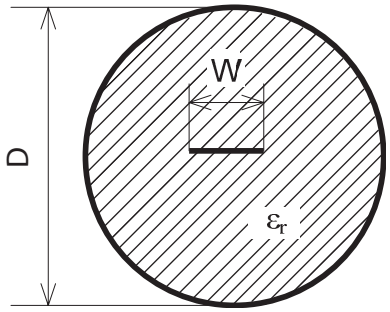


Figure 1: Strip-centered coaxial line – cross section.

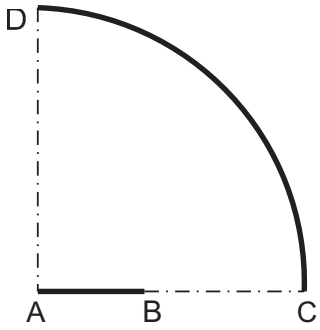


Figure 2: Analyzed part of SCCL.

Bongianni [2] also highlights a number of microwave applications which require the miniaturization of coaxial cable. In some medical, delay line or microelectronic application, the radial size is a major constraint and must be minimized. As a cable dimensions decreases, however, the loss increases because of skin effect. The central strip has a larger effective area as an equivalent circular inner conductor of the coaxial line.

The SCCL structure also offers a very attractive occasion of matching to microstrip line, coplanar waveguide, etc.

Stronger and stronger numerical methods consequently confine an application of the analytical methods on basic technical praxis. The main advantage of the analytic method occurs during the synthesis of the transmission lines, when we can directly specify dimensions of the structure from its required electrical parameters.

In the other hand, numerical methods are universal contrary to analytical solutions, each limited only to one strictly defined type of the transmission line.

The SCCL structure has two planes of symmetry – one perpendicular and one parallel to central strip. One quarter of cross section (Fig. 2.) is sufficient for all our analysis. The characteristic

impedance of this quarter is 4 times bigger than the characteristic wave impedance of the whole SCCL structure.

2. BOUNDDARY ELEMENT METHOD (BEM-1, BEM-2)

The boundary element method (BEM) [4] is based on integral Maxwell's equations (Laplace's equation is solved along boundary of the structure only). We need values of the field intensity just along PEC's for characteristic impedance determination according formula below:

$$Z_0 = \frac{1}{4}\frac{\int_D^A \mathrm{d}E}{\int_A^B \mathrm{d}H} = \frac{120\pi}{4}\frac{U}{\int_A^B E_y \mathrm{d}x}, \tag{1}$$

where U is voltage between conductors and E_y is electric field y-component along central electrode.

The boundary $ABCD$ (Fig. 2) is divided into boundary elements, with Dirichlet's condition along PEC (AB and CD) and Neumann's condition along PMC (BC and DA). The vectors of potentials in all boundary elements and the vectors of their gradients (both vectors partly known and partly not known) are mutually tied by two matrices H and G, built up according the rules of BEM.

The task was solved in $Matlab^{®}$ again. Two different strategies were tested.

BEM-1 is based on constant number of boundary elements ($n = 50$) per one basic element. The flat inner electrode is divided onto $n = 50$ elements, the round outer electrode is divided also onto $n = 50$ elements and both the magnetic walls are also divided each into $n = 50$ pieces. That means total sum of all boundary elements is $4n = 200$. The lengths of element are not the same, especially in case of $|AB| \neq |BC|$.

The strategy BEM-2 is limited on rational ratios $k = |AB|/|AC|$ and based on equidistant division of the border (except outer electrode CD). The inner electrode AB is divided to k times n pieces, the continuing magnetic wall BC into $(1-k)$ times n pieces, the outer electrode CD into n elements and wall DA also into n elements. Of course, the product kn must be a integer number. On the ground of calculation in wide range of k, the number of elements was $n = 100$.

3. NUMERICAL SOLUTION OF THE SCHWARZ-CHRISTOFFEL INTEGRAL (NSC)

Schwarz-Christoffel method is a conformal mapping which transforms upper half complex plane $\boldsymbol{z} = x + jy$ (and in some modifications also 1st quadrant of the complex plane, circle, rectangle or infinite strip) to an interior of a polygon in complex plane $\boldsymbol{w} = u + jv$. The mapping is given as

$$\frac{\mathrm{d}\boldsymbol{w}}{\mathrm{d}\boldsymbol{z}} = \prod_{i=1}^{n}(\boldsymbol{z} - \boldsymbol{z}_i)^{\frac{\alpha_i}{\pi} - 1}, \tag{2}$$

where $\boldsymbol{z}_i$ are points on real axis in $\boldsymbol{z}$-plane belong to vertices of the polygon in $\boldsymbol{w}$-plane, α_i are internal angles of the polygon and n is a number of polygon vertices.

A formula for the transform is given as Schwarz-Christoffel integral

$$\boldsymbol{w} = \int_0^{\boldsymbol{z}} \prod_{i=1}^{n}(\boldsymbol{z} - \boldsymbol{z}_i)^{\frac{\alpha_i}{\pi} - 1}\mathrm{d}\boldsymbol{z}, \tag{3}$$

but only rare cases are possible analytically integrated, integral (3) is too complicated usually. Hence, this integral is necessary to evaluate numerically. There is a free $Matlab^{®}$ toolbox for it on the web [5].

This method is based on polygons, so the arc CD must be approximated by a piecewise broken straight line with N vertices. The accuracy is better for higher N, but the time consumption rapidly grows simultaneously. $N = 20$ is very good compromise. The main part of the m-file is

```
p = polygon([0; k; exp(linspace(0, -i*pi/2, N)]);
f = rectmap(p, [1 2 3 length(p)]);
m = modulus(f);
Zsc = 30*pi/m;
```

4. CONFORMAL MAPPING METHOD (CMM-1, CMM-2)

CMM-1 is borrowed from [1–3]. All these sources refer on the same method splited up two models, one for characteristic impedances higher then $Z_0 = 30\pi/\sqrt{\varepsilon_r}$, is the formula given as

$$Z_0 = \frac{60}{\sqrt{\varepsilon_r}} \log \frac{2D}{W} \tag{4a}$$

and for lower impedances is formula

$$Z = \frac{15\pi^2}{\sqrt{\varepsilon_r}} \frac{1}{\log\left(2\dfrac{D+W}{D-W}\right)}. \tag{4b}$$

All the symbols are described on Fig. 1.

There is a bit chicken-to-egg problem here because it is necessary to know characteristic impedance in order to choose between (4a) and (4b). Both equations pass quite close according Wadell [3], but about the accuracy report neither Hilberg [1] nor Wadell [3], unfortunately. Bongianni [2] compares this method with his experimental results. His measured values in a graph are very close to the theoretical curve.

Second conformal mapping method CMM-2 is originated in [6]. Method is based on analytical solution of Schwarz-Christoffel integral according [7] and [8]. The result of basic mapping is degraded for higher ratio k by a shape distortion, so a conformal correction of the curved part is described in [6]. The wave characteristic impedance of the SCCL is here for whole the range given as

$$Z_0 = \frac{60}{\sqrt{\varepsilon_r}} \operatorname{argtanh} \sqrt[4]{\frac{D^2 - W^2}{D^2 + W^2}}. \tag{5}$$

Table 1: Strip-centered coaxial line – results of 5 different types of analysis.

	$Z_0[\Omega]$				
k	BEM-1	BEM-2	NSC	CMM-1	CMM-2
0.01	651.46	323.41	317.86	317.89	317.89
0.05	273.26	223.21	221.31	221.33	221.33
0.1	206.83	180.75	179.71	179.74	179.74
0.2	152.71	138.64	138.11	138.16	138.14
0.3	122.83	114.10	113.73	113.83	113.75
0.4	96.53	96.63	96.34	96.57	96.32
0.5	82.89	82.91	82.66	82.63	82.58
0.6	71.42	71.39	71.17	71.19	70.95
0.7	60.98	61.15	60.94	60.98	60.47
0.8	51.48	51.38	51.17	51.22	50.25
0.9	41.03	40.89	40.64	40.70	38.78
0.95	34.44	34.24	33.89	33.98	31.05
0.99	26.14	25.39	24.57	24.73	19.67

5. CONCLUSIONS

Results of all here mentioned methods are shown in Tab. 1. Characteristic impedances are calculated for free space ($\varepsilon_r = 1$).

BEM-1 with constant number of element along every part of closed boundary curve $ABCDA$ is in good agreement with other methods in range $k \in <0.4; 0.8>$, outside this interval grows the deviation - extremely for $k < 0.2$. The reason is in great disproportion between lengths of elements along central strip AB and the rest part of the boundary curve $BCDA$. For improvement of the result is necessary to pay greater number of the boundary elements and very longer computer time. Strategy BEM-1 is not very successful and authors do not recommend it to future application.

BEM-2 with dynamically allocated number of element according its dimension and constant (approximately) length of the elements is markedly better, results copy other method results with only marginal differences. The disadvantage of this method is a various number of elements along central strip AB and along magnetic wall BC, which complicates the computer program (task was solved in $Matlab$®. Advantage is a reliability of the result.

NSC method based on numerical solution of the Schwarz-Christoffel integral also very well copies the values of both CM methods. The integration step is very small, standard value is 10^{-5} and is a optional parameter, so the result error is marginal. The main reason of small deviation is the approximation of the arc CD by a piecewise broken line. If the number of parts along this arc is higher than 20, the time consumption and computer memory waste will enormously grow without significant influence on the accuracy, what was been experimentally recognized. This method combines the advantages of conformal mapping method (direct description of the global parameters like characteristic impedanc.) and numerical method (ability to solve fast every structure). Disadvantage of this method is its limitation for 2D tasks. That's why only quasi-static problems with quasi-TEM wave can be solved.

CMM-1 is the only one method taken from the independent literature [1–3]. Bongianni [2] compares this analytical method with his consistent experimental results. This method has only one disadvantage – chicken-and-the-egg problem with splitting into two partly-valid models mentioned above.

CMM-2 is derived by author on the basis of the Schwarz-Christoffel integral. Results are in great agreement with other methods (except BEM-1). Compared to CMM-1, the main advantage of this method is one model for the whole range of shape parameter k. Only algebraic methods are necessary for the derivation of the synthesis equation. The only one disadvantage is more complicated mathematical model than CMM-1.

In the end we can come to the conclusion that better results we obtain with equidistant elements with BEM than with constant number of them along variable electrodes. Analytical models based on CMM are such accurate as the method used in its derivation. NSC method collects advantages, but sometimes disadvantages of both, analytical and numerical methods.

ACKNOWLEDGMENT

This work was supported by the research plan MSM 0021630513 "Advanced Electronic Communication Systems and Technologies (ELCOM)" of the Brno University of Technology.

REFERENCES

1. Hilberg, W., *Electric Characteristics of Transmission Lines*, Artech House Books, 122, Dedham, MA, 1979.
2. Bongianni, W. L., "Fabrication and performance of strip-centered microminiature coaxial cable," in *Proceeding of the IEEE*, Vol. 72, No. 12, 1810–1811, December 1984.
3. Wadell, B. C., *Transmission Line Design Handbook*, Artech House, Boston, London, 1991.
4. Paris, F. and Canas, J., *Boundary Element Method*, Oxford University Press, Oxford, 1997.
5. Driscoll, T. A., SC-Tolbox, for MATLAB, http://www.math.edu/∼driscoll/SC/.
6. Šádek, V., "The conformal correction in the strip-centered coaxial line," *Radioengineering*, Vol. 16, No. 3, 68–71 September 2007.
7. Svačina, J., *Investigation of Hybrid Microwave Integrated Circuits by a Conformal Mapping Method*, Publications of technical and scientific papers of Brno University of Technology, Vol. A-46, Brno, Czech, 1991.
8. Šádek, V., "Analysis of the two-wire transmission line by a conformal mapping method," in *Proc. of Radioelektronika 2002*, Bratislava, Slovakia, 2002.

A Novel Hypothesis for Quantum Physics, Model with Telegraphs Equation

P. Fiala[1], **K. Bartusek**[2], **and M. Steinbauer**[1]

[1]Department of Theoretical and Experimental Electrical Engineering
University of Technology Brno, Kolejni 4, 612 00 Brno, Czech Republic
[2]Institute of Scientific Instruments, Academy of Sciences of the Czech Republic
Královopolská 147, 612 64 Brno, Czech Republic

Abstract— The article describes numerical model of basic structures in quantum physics. It was tested on the basic configuration of the electron beam and it was verified in Institute of Scientific Instruments Academy of Sciences of the Czech Republic experimentally. We had prepare the numerical model which was based on the particle theory. We prepared model which respects classical Electrodynamics. Numerical results were evaluated. The second model was prepared in respect to theory of wave packet (Louis de Broglie and Material Wave Theory (MWT)) and solved again. Results of both models were the similar, same in their quality, we evaluated electric field intensity $\boldsymbol{E}$ on the electron impact area, and they corresponded with results from experiments.

1. INTRODUCTION

R. P. Feynman once said that, according to him, it was reasonably possible to presume that nobody could understand quantum physics [1]. In another one of his many comments, the physicist expressed the opinion that quantum physics was merely a working hypothesis and it would remain so.

In order to attempt — in modeling the H_2O elementary molecule and the subsequent high quantity of H_2O in the electric field — to obviate the problems resulting from quantum mechanics [1], a simple task was created and solved using the finite element method [2]. The main reason consists in the fact that, using the above hypothesis [1], it is not possible to formulate the problem with a finite number of the degrees of freedom (the problem of a variable quantity of particles) or the problem of relation between the electric field and the particles (a Lorentz force effect and transitory stages of elementary particles, dynamic behaviour of particles and other items). Referring now back to the main task, the first step consisted in creating a numerical model based on the quantum mechanics hypothesis with an elementary model in the form of a diffusion equation — the Schrodinger equation. The numerical model simulated the reality that had been experimentally verified in the 1960s. Then, we created a numerical model that is rather more complicated in terms of formulation: the formulation in question is realized using the telegraph equation, therefore it is a diffusion equation complemented with an element with second partial derivative. This model — based on the Material Wave Theory (MWT) first published by de Broglie — is however capable of solving transient conditions and the dynamics of the elementary structure of matter. The testing [1] has proved that the model is applicable. Nevertheless, there were certain variations in the testing due to the failure to simulate for identical conditions the task formulated with the help of quantum mechanics and the telegraph equation. Here, the conditions were merely approximate. Yet, based on the surprisingly proximate results, these models are being refined and they are expected to reach a substantially higher degree of conformity.

At this point we would like to attempt to briefly describe the formulated hypothesis, the model of the basic elements of matter behaviour, which is based on solving partial differential equations of a dimension higher than the diffusion equation — the Schrödinger equation.

2. HYPOTHESIS

The working hypothesis for quantum mechanics, based on the reduced diffusion equation

$$\Delta u = C_{s1}\frac{\partial u}{\partial t} + C_{s2}u + C_{s3} \tag{1}$$

where u is the function or the functional, $C_{s1,2,3}$ the constants. The modified relation and well-known as the Schrödinger equation

$$\Delta u = C_{s1}\frac{\partial u}{\partial t} + C_{s2}u \tag{2}$$

is suitable for solving tasks in quantum mechanics. This working hypothesis is applicable in solving stationary conditions of models. It does not, in itself, concern dynamics or transient processes of the u function. The resulting form of solution, after satisfying the boundary and the starting conditions plus the relation (2) solvability conditions, can be written as, for example,

$$u = C_{sv1} e^{F} \tag{2a}$$

where F is the function or the functional, C_{sv1} is the constant. The relation (2a) is therefore a mere form of non-periodical solution (damped solution — diffusion). It is not the real wave form, despite being asserted to be such for the quantum wave character of the existing hypothesis.

For any case of analysis of transient processes in the conditions of quantum physics particle position shift, the model according to relation (1) is not suitable for the description of the transient processes. The numerical solution gives a robust model, which does not render in a wholly objective manner the change of condition, dynamics or other effects that are currently measurable or recordable.

For that reason, a working hypothesis was proposed for the field of numerical modeling of the dynamic condition of basic elements of matter. The hypothesis is based on a model with a higher dimension of time variation of the functional u, namely on the telegraph equation formed as

$$\Delta u = C_{t0} \frac{\partial^2 u}{\partial t^2} + C_{t1} \frac{\partial u}{\partial t} + C_{t2} u + C_{t3} \tag{3}$$

This model shows more respect for the system dynamics and the simplified form of the model is the wave equation (without damping)

$$\Delta u = C_{w0} \frac{\partial^2 u}{\partial t^2} \tag{4}$$

The above hypothesis based on relation (3) or, alternatively, relation (4), quite paradoxically offers — for the numerical field of modeling — an easier solution, albeit of a more complicated model, using the numerical apparatus. After satisfying the boundary and starting conditions plus the relation (3) solvability conditions, the solution can be written in the form of a periodical damped wave written as

$$u = C_{tv1} e^{F(t)} e^{F(\Omega)} \tag{4a}$$

where Ω is the space for which the relation (3) was formulated, t is time.

As a result of the proposed hypothesis there occurs a shift in the field of quantum physics (QF). Based on relations (3) and (4) it is possible to assume that the elementary phenomenon of quantum physics solution is an electromagnetic wave. According to the current QF hypothesis, a particle is a consequence of the wave interference phenomenon. It is possible to assume that the basic element of matter is an electric charge and its motion. Further, it is also possible to claim — according to [2] and [4] — that

$$\rho_e^2 = \rho_g \tag{5}$$

where ρ_e is the electric charge density and ρ_g is the specific mass.

3. EXPERIMENTS

For the proposed QF hypothesis of relation (3), a task is sought to either prove some of the conjectures resulting from the solution of relation (3) or to overcome some of the working solutions to the QF hypothesis based on the model according to relation (2).

The QF hypothesis based on the solution of the Schrodinger equation asserts that particles are of a dual character, namely they at times behave as particles and, at other times, as a wave.

Therefore, we will utilize the assertion that a beam of emitting electrons actually is the particles. If we used yet another source of the beam of electrons with different kinetic energy and if both the beams intersected in space and time (interaction, intersection), then, according to the laws of Newtonian physics, there should occur the interaction of particles, and the deflection in those with

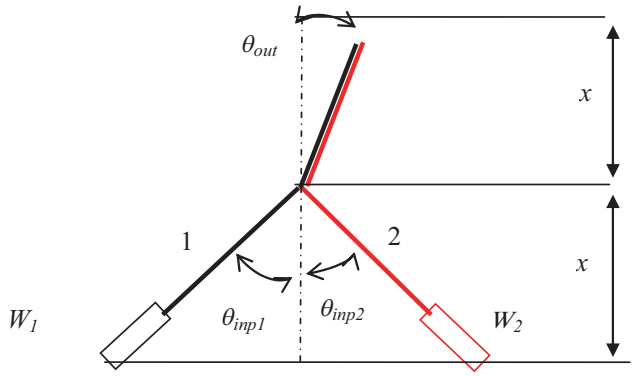


Figure 1: Interaction of two beams of electrons.

lower energy should be more prominent that in those with higher energy. For the task drawing see Fig. 1.

Figure 1 shows the source of a beam of electrons with higher dynamics particles (energy W_1); the source of a beam of electrons with lower dynamics particles than in beam 1 is marked in red and the energy W_2. The coincidence angles are marked as θ_{inp} and the beam number index 1 or 2. The resulting beam is supposed to show deflection of θ_{out}. According to the current QF hypothesis, the energy of the impinging particle is

$$W = \frac{1}{2}mv^2 \tag{6}$$

where m is the particle mass, v its medium velocity. Let us assume that the particle features constant weight in different velocity of movement. Then, in coincidence according to relation (6) there occurs the deflexion of beam in the direction of the ray with the higher-energy particles. If $m_e = 9.10938188e\text{-}31\,\text{kg}$ is assumed, then the impact velocity of the first beam will be

$$v_1 = \sqrt{\frac{2 \cdot W_1}{m_e}} \tag{7}$$

for the voltage on the electrodes of the electron beam $U_1 = 1\,\text{kV}$, $W_1 = 1.602176462e\text{-}16\,\text{eV}$, $v_1 = 18.755373\,\text{Mm/s}$, $U_2 = 1.5\,\text{kV}$, $W_2 = 2.403264693e\text{-}16\,\text{eV}$, $v_2 = 22.970547\,\text{Mm/s}$. Assuming that $\theta_{inp1} = \theta_{inp2}$ is the angle

$$\theta_{out} = \arctan\left(\frac{v_1 \cdot \sin\theta_{inp1} - v_2 \cdot \sin\theta_{inp2}}{v_1 \cdot \cos\theta_{inp1} + v_2 \cdot \cos\theta_{inp2}}\right) \tag{8}$$

where x is the electron beam source distance from the coincidence point. In this case, the impact angle is $\theta_{out} = 5.76°$, $\theta_{inp1} = \theta_{inp2} = 45°$.

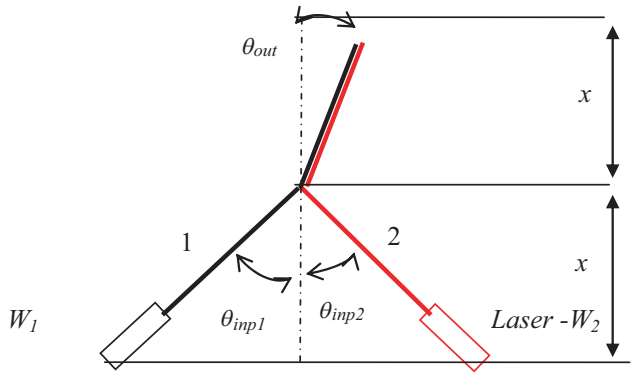


Figure 2: Interaction between a beam of electrons and an electromagnetic wave.

Let us assume an identical task configuration according to Fig. 1 and propose a different experiment, with one of the sources being a continuous monochromatic laser — the source of

the electromagnetic wave. The laser features the wave length of $\lambda = 633\,\mathrm{nm}$, the frequency of $f = 4.7361e14\,\mathrm{Hz}$. According to the current QF theory, relation (2) is the frequency of the electron beam 1 and 2. According to the relation

$$f = \frac{W}{h} \tag{9}$$

where h is the Planck constant $h = 6.62606876e\text{-}34\,\mathrm{Js}$. Then, the beam 1 frequency is $f_1 = 2.417989e17\,\mathrm{Hz}$ and the beam 2 frequency is $f_2 = 3.62698e17\,\mathrm{Hz}$. According to the classical theory there occurs the interference of waves. Laser with the energy of $W = 3.13817e\text{-}19\,\mathrm{eV}$ ought to perform the deflexion of the beam of electrons 1 to the angle $\theta_{out} = 42.46°$. With the help of this experiment it is possible to clarify the character of both sources. Therefore, a beam of electrons is the source of electromagnetic waves. If the task assumptions are proved to be correct, the laser version should be recorded on frequency $f_1 - f = 2.41325e17\,\mathrm{Hz}$.

4. CONCLUSION

On the basis of the above described experiments, the model of basic particles will be built. This study described and analyzed the experiments to provide the verification of the novel theory. The activity is connected with the research plan FRVS CR No. MSM 0021630513 ELCOM, No. MSM 0021630516. The proposed experiments are verified and the conclusion suggested should clarify the problem of whether the proposed working hypothesis — according to relation (3) — can be used for modelling the dynamics of the basic elements of matter. The aim of this research is description and numerical modelling of the water molecule behaviour in the electric field, with known effect intensity $\boldsymbol{E}$. The behaviour is experimentally tested the NMR spectrometer.

ACKNOWLEDGMENT

The research described in the paper were financially supported by FRVŠ by research plan No. MSM 0021630513 ELCOM, No. MSM 0021630516 and grant GAAV No. B208130603.

REFERENCES

1. http://en.wikipedia.org/wiki/Richard_Feynman.
2. Fiala, P. and K. Bartusek, "A simple numerical simulation of internal structure of particles test," *PIERS Online*, Vol. 2, No. 6, 653–656, 2006.
3. Van Vlaenderen, K. J. and A. Waser, "Electrodynamics with the scalar field," *Physics*, Vol. 2, 1–13, 2001.
4. Kikuchir, H., *Electrohydrodynamics in Dusty and Dirty Plasmas, Gravito-electrodynamics and EHD*, Kluwer Academic Publishers, Dordrecht, Boston, London, 2001.
5. Van Vlaenderen, K. J., "A charge space as the orogin of sources, fields and potentials," *Physics*, Vol. 1, 1–13, Oct. 16, 1999.
6. Hofer, W. A., "A charge space as the origin of sources, fields and potentials," *Physics*, Vol. 3, 1–13, Apr. 17, 1997 .
7. Prosser, V a kolektiv, "Experimentální metody biofyziky," *Academia*, Praha, 1989.
8. Delong, A., "Verbal information," Czech Academy of Science, ISI Brno, Brno, Jul. 2, 2006.

Extending the Concept of Debye Length for Chasmas

D. K. Callebaut[1] and H. Kikuchi[2]

[1]Physics Dept., CDE, University of Antwerp, Antwerp, B-2610, Belgium
[2]Institute for Environmental Electromagnetics, Tokyo 170, Japan

Abstract— Chasmas are a generalization of plasmas, i.e., the condition of quasi-neutrality is dropped. That means that in chasmas the quasi-neutrality may be (strongly) violated over distances many times the Debye length which requires special circumstances (double layers, electric fields, ...). The question arises what the meaning is of a shielding length in chasmas. It was demonstrated that the so-called chasma (angular) frequency has an expression similar to the plasma frequency: $\omega_c^2 = |n_- - n_+|e^2/\epsilon m_-$. However, this chasma frequency plays a role as well in the equilibrium (or steady state) as in the stability. Similarly for the chasma equilibrium we obtained a kind of Debye length $\lambda_c^2 = 2\epsilon k_B T/|n_- - n_+|e^2$, where the difference of the densities occurs in contrast to the shielding length $\lambda_D^2 = \epsilon k_B T/(n_- + n_+)e^2$, for a chasma equilibrium, supposing that the temperature of electrons and ions is the same, that the ions are only once ionized and that the kinetic energy is much larger than the potential energy. This means that the chasma shielding length is much the same as the Debye length and that $(\lambda_D \omega_p)^2 \approx (\lambda_c \omega_c)^2 \approx k_B T/m_-$. We expect those features to be relatively universal, up to factors of order unity, in spite of the great variety of chasmas.

1. INTRODUCTION

In a plasma one has by definition quasi-neutrality in volumes which have dimensions larger than the Debye length or at least a few times the Debye length. In some situations one has no quasi-neutrality over many times the Debye length. E.g., in certain discharges or in the multipactor effect [1–3] (secondary electron resonance discharge) in the cavities of linear accelerators. It is expected that chasmas occur too in certain extended double layers with currents and magnetic fields playing a role in the solar convective zone and solar atmosphere or in the terrestrial atmosphere. Callebaut and Knuyt [4–7] investigated theoretically the steady state, using an approach based on a singular integro-differential equation, cf. [8–10] and [11–19] as well.

However, the question of the relevant parameters for chasmas arises, in particular: is the shielding length the same as for plasmas?

We investigate first the chasma equilibrium (Section 3). Next we deal with the shielding length (Section 4), followed by the conclusion (Section 5).

2. CHASMA EQUILIBRIUM

Most chasmas occur as a steady state or a state in motion, as the net charge requires special conditions to be contained. But here we consider the chasma equilibrium without motions or supplementary forces. Suppose a box with isolating walls in which electrons and ions are enclosed. Suppose the electrons and the ions have the same uniform temperature T and that recombination may be neglected. How is the spatial distribution of the densities and the corresponding electric potential for this chasma equilibrium?

The basic equations here are the Boltzmann and the Poisson equations:

$$n_- = n_{-0} \exp\left(\frac{e\varphi}{k_B T}\right), \quad n_+ = n_{+0} \exp\left(\frac{-e\varphi}{k_B T}\right), \tag{1}$$

$$\epsilon \Delta\varphi = e(n_- - n_+). \tag{2}$$

ϵ is the electric permittivity and equals in vacuum 8.8542×10^{-12} C/Vm; k_B is the Boltzmann constant (1.3807×10^{-23} J/K). We assume that ϵ is constant which is a reasonable hypothesis. However, as some other quantities are varying in space an extension may be to have ϵ in tensorial form. The number density of the electrons is n_-, which depends on the space coordinates, at the origin it is n_{-0}, where the potential φ is supposed to vanish; the electron charge is $-e$ and their mass is m_-. For the ions, supposed ionized once only and all of the same mass, the same notation applies, but with a + sign replacing everywhere the − sign corresponding to the electrons. Pressure terms are neglected here.

We now consider for simplicity the 1-D case: a plane-parallel chasma, infinite in 2 directions. The x-axis is chosen perpendicular to the walls. Equations. (1)–(2) yield upon elimination of n_- and n_+:

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} = \frac{e}{\epsilon}\left(n_{-0}\exp\left(\frac{e\varphi}{k_BT}\right) - n_{+0}\exp\left(\frac{-e\varphi}{k_BT}\right)\right). \tag{3}$$

This differential equation is highly nonlinear.

2.1. One Kind of Charges

In the case $n_{+0} = 0$ we found the following solution:

$$\varphi = \frac{-2k_BT}{e}\ln\left(1+\frac{x}{\lambda_-}\right), \quad n_- = \frac{\lambda_-^2 n_{-0}}{(\lambda_- + x)^2}, \quad \lambda_- = \left(\frac{2\epsilon k_BT}{n_{-0}e^2}\right)^{1/2}, \tag{4}$$

λ_- looks like the Debye length in plasmas (up to a factor 2.) However λ_- appears already in the chasma equilibrium itself, without external influence of a charge or potential. (Moreover, in Section (3.2) it becomes clear that the real structure of λ_- is different from the one of λ_D.) The potential is zero at the origin, but its derivative is not. To avoid singularities we must require $x + \lambda_- > 0$ on the other hand for $x > 0$ there is no limit where the (isolated) wall is placed. Similarly when $n_{-0} = 0$ the solution given in Equation (4) applies with e replaced by $-e$. Note the singular solution of Equation (3) with $n_{+0} = 0$ (very similar to the regular one):

$$\varphi = \frac{-2k_BT}{e}\ln\frac{x}{\lambda_-}, \quad n_- = \frac{\lambda_-^2 n_{-0}}{x^2}. \tag{5}$$

2.2. Both Species: Series Development

We choose the origin in the middle between the plane-parallel (isolated) walls, which we suppose to be at the same potential, so that the situation is symmetric with respect to the middle plane. The extension is easy but requires more algebra. We try a series development for φ in Equation (3). Put:

$$\frac{e\varphi}{k_BT} = a_2x^2 + a_4x^4 + a_6x^6 + \ldots, \tag{6}$$

$$A = \frac{e^2n_{-0}}{\epsilon k_BT}, \quad B = \frac{e^2n_{+0}}{\epsilon k_BT}. \tag{7}$$

This yields

$$a_2 = \frac{A-B}{2}, \quad a_4 = \frac{A^2-B^2}{24}, \quad a_6 = \frac{A-B}{720}[(A+B)^2 + 3(A-B)^2], \tag{8}$$

$$a_8 = \frac{A^2-B^2}{56\times 720}[(A+B)^2 + 33(A-B)^2], \tag{9}$$

etc. If $n_{-0} = n_{+0}$ then $\varphi = 0$ and we recover the plasma case. If $n_{-0} \approx n_{+0}$, then the series converges rather fast. We may rewrite a_2 as

$$a_2 = \frac{e^2(n_{-0}-n_{+0})}{2\epsilon k_BT} \equiv \lambda_c^{-2}, \tag{10}$$

where λ_c is similar in structure to the Debye length for plasmas, but involves the difference of the charge densities. However, it appears already in the equilibrium. It has clearly to do with the non-quasi neutrality, but is not a proper shielding length in the same sense as the Debye length itself. λ_c is a measure of the distance over which the chasma is roughly uniform. Note that λ_c is infinite for a plasma.

3. SHIELDING

Suppose e.g., that an extra isolated charge is introduced in the chasma: how will the charge distributions alter themselves? We take here the simple case that a grid at a given potential Φ_0 is introduced e.g., at the origin. Equation (3) becomes now

$$\frac{\mathrm{d}^2(\varphi+\Phi)}{\mathrm{d}x^2} = \frac{e}{\epsilon}\left[n_{-0}\exp\left(\frac{e(\varphi+\Phi)}{k_BT}\right) - n_{+0}\exp\left(\frac{-e(\varphi+\Phi)}{k_BT}\right)\right], \tag{11}$$

or, provided $\frac{e\Phi}{k_BT} << 1$, to a fair approximation:

$$\frac{\mathrm{d}^2(\varphi+\Phi)}{\mathrm{d}x^2} = \frac{e}{\epsilon}\left[n_-\left(1+\frac{e\Phi}{k_BT}\right) - n_+\left(1+\frac{-e\Phi}{k_BT}\right)\right]. \tag{12}$$

Using Equation (3) for the equilibrium in (12) yields:

$$\frac{\mathrm{d}^2\Phi}{\mathrm{d}x^2} = \frac{e^2(n_-+n_+)\Phi}{\epsilon k_BT}. \tag{13}$$

In the vicinity of the origin the density is constant. The condition for this is that $x << \lambda_c$. In this approximation we have the solution for equation

$$\Phi = \Phi_0 \exp\left(-|x|/\lambda_D\right). \tag{14}$$

The shielding length is given by

$$\lambda_D^2 = \frac{\epsilon k_BT}{e^2(n_{-0}+n_{+0})}. \tag{15}$$

λ_D becomes the λ_D for plasmas when the densities of the positive and negative particles are the same, so a different notation seems unnecessary. The condition to treat $n_- - n_+$ as a constant is $x << \lambda_c$ according to Equation (10). We usually have $\lambda_c >> \lambda_D$ so that the approximation is valid in a reasonable range.

3.1. Homogeneous Chasma in Steady State

We know from the studies using the singular integro-differential Equations (2–5) that steady states exist in which both n_+ and n_- are constant. Clearly the above analysis of the shielding can be repeated. However, this is not an equilibrium but a steady state with flows of electrons and ions, requiring a somewhat adapted treatment. Preliminary calculations show that the essence of the results above remains valid. Still the question of the universality of the expression for λ_D remains as chasmas exits in widely different situations. In [11] we obtained the so-called chasma (angular) frequency:

$$\omega_c^2 = \frac{e^2|n_{+0}-n_{-0}|}{\epsilon m_-}. \tag{16}$$

Although similar to the plasma frequency, its meaning is different: this is an entity occurring in the steady state, and as such occurring in the stability analysis too, while the plasma frequency, although constituted by equilibrium quantities, appears in the perturbation analysis only.

4. CONCLUSION

We studied the chasma equilibrium. A new parameter, the chasma length, λ_c, is introduced. Over this distance the chasma may be approximated as uniform. It matches the previously introduced chasma frequency, ω_c. Both are similar to λ_D and ω_p, but use the difference of the densities instead of the sum. For the shielding length we obtained $\lambda_D^2 = \epsilon k_BT/(n_- + n_+)e^2$, which generalizes the Debye length, but hardly differs from it. We have $(\lambda_D\omega_p)^2 \approx (\lambda_c\omega_c)^2 \approx k_BT/m_-$.

REFERENCES

1. Callebaut, D. K., *Physica*, Vol. 29, 784–802, 1963.
2. Callebaut, D. K., *Physica*, Vol. 31, 1177–1192, 1965.
3. Callebaut, D. K., *Physica*, Vol. 32, 1151–1158, 1966.
4. Callebaut, D. K. and G. K. Knuyt, *Phys. Lett.*, Vol. 29A, 72, 1969.
5. Callebaut, D. K. and G. K. Knuyt, *Plasma Phys.*, Vol. 20, 511–523, 1978.
6. Callebaut, D. K. and G. K. Knuyt, *Plasma Phys.*, Vol. 20, 524–550, 1978.
7. Callebaut, D. K. and G. K. Knuyt, *Relation between Laboratory and Space Plasmas*, (Ed. H. Kikuchi), 207–230, Reidel Publ. Co., Dordrecht, The Netherlands, 1981.
8. Callebaut, D. K. and G. J. Bex, *Workshop on Nonlinear Stability*, (Ed. D. K. Callebaut), 154 and 161, UIA Press, University of Antwerp, Antwerp, Belgium, 1989.

9. Callebaut, D. K., *Second URSI-ICPIG-RIKEN Symposium on Critical Problems of Discharge and Plasma Physics and Electro-Plasmo-Meteorology*, 40–41, Riken (The Institute of Physical and Chemical Research), 351-01, Wako, Saitama, Japan, 1995.
10. Callebaut, D. K. and A. H. Khater, *International Symposium on Nonconventional Plasmas*, (Eds. H. Kikuchi and D. K. Callebaut), 1–2, Niigata, Japan, 1997.
11. Callebaut, D. K., G. K. Karugila, and A. H. Khater, "Chasma perturbations," *Proc. Progress In Electromagnetics Research Symposium 2005*, 720–723, August 22–26, Hangzhou, China, 2005.
12. Callebaut, D. K. and H. Kikuchi, "Debye shielding in chasmas," *Progress In Electromagnetics Research Symposium 2007 in Prague Abstracts*, August 27–30, Prague, 2007.
13. Kikuchi, H., (ed.), *Relation between Laboratory and Space Plasmas*, Reidel Publ. Co., Dordrecht, The Netherlands, 1981.
14. Kikuchi, H., *Handbook of Atmospheric Electrodynamics*, Vol. 1, H. Volland (ed.), CRC Press, Bonn, 167, 1995.
15. Kikuchi, H., *Electrodynamics in Dusty and Dirty Plasmas (Gravito-electrodynamics and EHD)*, ASSL, Vol. 258, Kluwer Acad. Press, Dordrecht, 2001.
16. Schrittwieser, R. W. and G. Eder, (eds.), *Proc. Second Symposium on Plasma Double Layers and Related Topics*, Institut fr Theoretische Physik, Univ. of Innsbruck, A-6020 Innsbruck, Auqstria, 1984.
17. Schrittwieser, R. W., (ed.), *Fourth Symposium on Double Layers and other Nonlinear Potential Structures in Plasmas*, (Innsbruck, Austria), World Scientific, Singapore, 1993.
18. Verheest, F., M. Goossens, M. A. Hellberg, and R. Bharuthram, (eds.), *Waves in Dusty, Solar and Space Plasmas*, AIP Conference Proc. 537, Leuven, Belgium, 2000.
19. Verheest, F., *Waves in Dusty Space Plasmas*, Vol. 245, Kluwer Acad. Press, Dordrecht, 2000.

Further Results on Post-MHD

D. K. Callebaut[1] **and G. K. Karugila**[2]

[1]Physics Department, CDE, University of Antwerp, Antwerp, B-2610, Belgium
[2]Sokoine University, Morogoro 3038, Tanzania

Abstract— The exact solution for the evolution equation of ideal magnetohydrodynamics (MHD) for a given velocity profile and a given initial magnetic field was obtained by Callebaut and coworkers [1–4] This allows the exact calculation of the other electromagnetic quantities, in particular of the neglected term, the displacement current, and thus to verify the MHD-approximation and possibly to make an iteration i.e., to pass from MHD to Post-MHD. This paper continues previous work on Post-MHD [5] where spherical coordinates and an involved differential rotation profile were used in view of the application to the Sun. Here we give the relevant solution of the evolution equation in Cartesian coordinates. Two very simple illustrations are considered to give a clear insight. The drift velocity is calculated too. It turns out that this represents a kind of Lenz law, i.e., it opposes the driving motion. The drift velocity is maximal and equal but opposite to the driving velocity when the magnetic field component in the direction of the velocity vanishes. As a side result the question of quasi-neutrality is considered (cf. the concept of non-quasi-neutral plasmas or "chasmas").

1. INTRODUCTION

Callebaut and coworkers [1–4] solved the equation of evolution for the magnetic field, $\mathbf{H}$, in ideal (i.e., assuming perfect conductivity) magnetohydrodynamics (MHD), together with the conservation of flux:

$$\partial_t \mathbf{H} = rot(\mathbf{v} \times \mathbf{H}), \tag{1}$$

$$div\mathbf{H} = 0, \tag{2}$$

in spherical coordinates (r, ϑ, φ) in the case that the velocity, $\mathbf{v}$, has an azimuthal component only and depends on r and ϑ only. (Meanwhile the last restriction can be dropped.) We have $\mathbf{B} = \mu\mathbf{H}$ with μ , the magnetic permeability, taken as constant; $\mu = 4\pi 10^{-7}\,\mathrm{kgm/C^2}$ (or henry/m) in vacuum; SI units. The use of spherical coordinates in [1–3] was obvious: the aim was to generate magnetic field from the differential rotation in the Sun and other heavenly bodies. However, the spherical coordinates and the complex differential rotation obscure somehow what is really happening. Cartesian coordinates, (x, y, z), are much simpler to handle and may give a better insight. Moreover, in view of testing the theory by experiments it is probably easier in a plane-parallel case or rectangular tubes. Of course cylinders have other advantages too [4] and the sphere has no problems with the 'ends', so all cases have to be investigated.

Having an exact and simple analytic solution allows to verify the consistency of the MHD theory. It allows to calculate easily the remaining electromagnetic quantities like the (MHD) current density $\mathbf{j} = rot\mathbf{H}$, the electric field $\mathbf{E}$, the displacement field $\mathbf{D}$, the electric charge ρ and the displacement current $\partial_t \mathbf{D}$. The ratio of the magnitudes of $\mathbf{j}$ and $\partial_t \mathbf{D}$ is crucial to know the approximation involved in MHD and possibly to apply a Post-MHD analysis [5], using e.g., an iteration.

The study is interesting too in view of recent research on non-quasi-neutral plasmas or chasmas [6–8]. See [9–11] too. There are two mayor kinds of chasmas: more or less permanent ones like equilibria and steady states (e.g., some discharges and double layers) and rather transient ones, e.g., occurring during instabilities or changing situations. It is clear that a change in e.g., frequency due to some non-quasineutrality gives different information about the plasma/chasma quantities like density, etc. In particular a slight change in the resonant heating frequency can decrease the heating by an order of magnitude and more.

The plan of the paper is as follows. In Section 3 we give the relevant solution of Eqs. (1) and (2) in Cartesian coordinates followed by simple illustrations. In Section 4 we calculate the remaining electromagnetic quantities and check the consistency of MHD. Section 5 gives the conclusions.

2. RELEVANT SOLUTION OF THE IDEAL MHD EVOLUTION EQUATION IN CARTESIAN COORDINATES

2.1. General Case when v Has One Component

The x-axis is chosen perpendicular to the (isolating) walls, plane-parallel or of rectangular cross-section. The velocity is parallel to the walls in the y-direction $\mathbf{v} = (0,\, v(x,\, y), 0)$. We suppose that v is independent of time although the generation of magnetic field requires energy; a pump may compensate for this loss. However, the generalization to $v(x,\, y,\, z,\, t)$ is possible but is left away in view of the clarity. Omitting infinite and periodic solutions we obtained the following solution for $\mathbf{H}$:

$$H_x = \partial_z\psi, \quad H_y = t\frac{\partial(v,\psi)}{\partial(x,z)} + H_y(0), \quad H_z = -\partial_x\psi, \tag{3}$$

with $H_y(0) = H_y(x,\, z,\, t = 0)$. Hence H_x and H_z are derived from a stream function $\psi(x,\, z)$. They do not depend on time. On the other hand H_y grows linearly with time as can be expected on physical grounds (the stretching of the field lines does not change in time). Moreover the growing part of H_y contains the Jacobian

$$J = (\partial_x v)\partial_z\psi - (\partial_z v)\partial_x\psi = H_x\partial_x v + H_z\partial_z v = \frac{\partial(v,\, \psi)}{\partial(x,\, z)}, \tag{4}$$

J contains the derivatives of v and ψ. The product of tJ is a general feature in all coordinate systems (up to possibly an additional factor, e.g., r^{-1} in spherical coordinates).

2.2. The Simplest Case

We take the initial magnetic field uniform and perpendicular to the plane-parallel walls $\mathbf{H}_0 = (H_0,\, 0,\, 0)$ and $\psi = H_0 z$. We obtain

$$\mathbf{H} = (H_0,\ tH_0\partial_x v + H_y(0),\ 0). \tag{5}$$

Hence in this case the growth of H_y is purely proportional to the initial magnetic field and the gradient of the velocity. Physically speaking the expression for H_y is obvious. It is the purest case possible. For an experiment one may take $\partial_x v =$ constant to make the situation easy and the results clear.

2.3. Both Gradients Involved

Consider $\psi = H_1 x + H_2 z$ with H_1 and H_2 constants. The solution reads

$$\mathbf{H} = [H_2,\ t(H_2\partial_x v - H_1\partial_z v),\ H_1]. \tag{6}$$

It is obvious that both contributions to the growth can cooperate or counteract each other. There is no growth when $H_2\partial_x v - H_1\partial_z v = 0$, which happens when $v = v(H_1 x + H_2 z)$. The particular case $v = C(H_1 x + H_2 z)$ with C a constant may be interesting for an experiment.

3. CHECKING THE CONSISTENCY OF MHD

For the MHD current density we simply have

$$\mathbf{j} = rot\mathbf{H} = [-t\partial_z J,\ (\partial_{xx} + \partial_{zz})\psi,\ t\partial_x J.] \tag{7}$$

Note that the x- and z-components of $\mathbf{j}$ grow linearly with time, while the y-component remains constant; for $\mathbf{H}$ it is just the opposite.

For the Lorentz-force density we obtain

$$\mathbf{L} = \mathbf{j}\times\mathbf{B} = -\mu\left((\partial_x\psi)\Delta\psi + t^2 J\partial_x J,\ \ t[(\partial_x\psi)\partial_z J + (\partial_z\psi)\partial_x J],\ \ t^2 J\partial_z J + (\partial_z\psi)\Delta\psi\right). \tag{8}$$

$\mathbf{L}$ contains terms which increase quadratically and linearly with time unless $J =$ constant.

For the electric field we obtain:

$$\mathbf{E} = -\mathbf{v}\times\mathbf{B} = \mu(-vH_z,\ 0,\ vH_x) \tag{9}$$

Thus

$$rot\mathbf{E} = -\mu[0,\ \partial_x(vH_x) + \partial_z(vH_z),\ 0] = (0,\ -\mu J,\ 0), \tag{10}$$

which equals $-\partial_t \mathbf{B}$ as it should be. Next we have $\mathbf{D} = \epsilon \mathbf{E}$ and

$$div\mathbf{D} = -\epsilon\mu[\partial_x(vH_x) - \partial_z(vH_z)] = \rho, \tag{11}$$

with ϵ the electric permittivity which equals 8.8542×10^{-12} C/Vm) in vacuum. The charge density is of the order of $\epsilon\mu = c^{-2}$ with c the velocity of light. Thus ρ is usually negligible except for very fast changes as Alfvén observed in funding MHD. Reconsidering the simplest case (Section 3.2) we obtain $\rho = \epsilon\mu\partial_x v$ which is proportional to the field and to the gradient of v. However, reconsidering Section 3.3 we may end up with both signs in the contributions to ρ or even zero according to the specific situation. Usually the MHD approach turns out to be excellent. Next we have the displacement current $\partial_t \mathbf{D} = 0$. However, the picture changes if we include the drift velocity as known in plasma physics.

Crossed $\mathbf{E}$ and $\mathbf{H}$ fields cause a drift velocity in a plasma (this applies to both kinds of charges):

$$\mathbf{v}_D = \frac{\mathbf{E} \times \mathbf{B}}{\mathbf{B}^2} = \frac{\left[-vH_x(tJ + H_y(0)),\ v(H_x^2 + H_z^2),\ -vH_z(tJ + H_y(0))\right]}{H_x^2 + ((tJ + H_y(0))^2 + H_z^2}. \tag{12}$$

For $t = 0$:

$$\mathbf{v}_D = \frac{\left[-vH_xH_y(0),\ v(H_x^2 + H_z^2),\ -vH_zH_y(0)\right]}{H_x^2 + (H_y(0))^2 + H_z^2}. \tag{13}$$

Hence for $H_y(0) = 0$ we have $\mathbf{v}_D = \mathbf{v}$. More generally we have $\mathbf{v}_D = \mathbf{v}$ for $tJ = -H_y(0)$. In view of the x- and z-dependence this does usually not happen at all places at the same time. With increasing time $\mathbf{v}_D$ will ultimately vanish. The drift velocity is maximal when $\mathbf{E}$ and $\mathbf{B}$ are perpendicular and vanishes when they are parallel. If we include the drift velocity then the displacement current does not vanish, but it remains at best a term of the order of c^{-2}. Still for very fast changes it may become comparable and even exceed the MHD current.

As the drift velocity affects the free charges only the effect on a conducting liquid will be small. However, the effect on a strongly ionized plasma will be large as long as the field component in the direction of the motion is small. Once this component dominates (thus $\mathbf{E}$ nearly parallel to $\mathbf{B}$) $\mathbf{v}_D$ becomes smaller and smaller. This is surprising since the other components nor the motion change. The various cases may be put to the test in experiments. Clearly the switching on of a magnetic field in a moving fluid or the putting in motion of a fluid in an existing magnetic field needs special investment. This suggest the need to investigate the effect of resistivity as well as the time dependence of the driving velocity which must decrease by creating a magnetic field, unless some energy source is present.

4. CONCLUSION

We gave the relevant solution of the ideal MHD evolution equation in Cartesian coordinates in the case that $\mathbf{v}$ has only one component, depending on the 'other' coordinates. Previous results in spherical and cylindrical coordinates are confirmed: the growth of the magnetic field is in the direction of the velocity and linear in time and proportional to a Jacobian involving the velocity and the 'other' components of the field. Two simple case are considered to give a clear insight.

Next we calculated the remaining electromagnetic quantities and checked the consistency of the MHD approach. Moreover the charge turns out to be negligible (of order $\epsilon\mu$ as realized by Alfvén). The displacement even vanishes in this case, making the MHD approach seemingly perfect. However, we calculated the plasma drift velocity and found that it vanishes for large time and that it exactly equals the driving velocity when the component of the magnetic field in the direction of the velocity vanishes. Including the drift velocity leads to a contribution of the displacement current of order c^{-2}, but for some rapid changes it may become relevant. This investigation is of interest to the study of non-quasi-neutral plasmas or chasmas too.

REFERENCES

1. Callebaut, D. K. and V. I. Makarov, "Generation of sunspots and polar faculae from a kinematic dynamo," *Proc. IX Pulkovo International Conference on Solar Physics: Solar Activity as a Factor of Cosmic Weather*, 379–388, Pulkovo, Main Astronomical Observatory, 196140 St. Petersburg, Russia, July 4–9, 2005.
2. Callebaut, D. K. and V. I. Makarov, "Generation of sunspots and polar faculae from a kinematic dynamo," CDRom *11 Eur. Solar Phys. Meeting — "The Dynamic Sun: Challenges for Theory and Observations"*, p1–call1, September 11–16, 2005 (ESA SP-596, December 2005).

3. Callebaut, D. K. and A. H. Khater, "Generation of sunspot and polar faculae butterflies using bipolar and quadripolar seed field," *Proc. IAU Symposium 233, Solar Activity and Its Magnetic Origin*, 9–16, Cairo, Egypt, March 31–April 4, 2006.
4. Callebaut, D. K., "Proposal for generation of magnetic field between rotating cylindersin," *Proc. Second Catania Workshop 'MHD Laboratory Experiments for Geophysics and Astrophysics'*, Catania, Sicily, 1–3 October 2007, to be published.
5. Callebaut, D. K. and A. H. Khater, "Post-magneto-hydrodynamics," *PIERS Online*, Vol. 3, No. 4, 508–512, Beijing, China, March 26–30, 2007.
6. Callebaut, D. K., G. K. Karugila, and A. H. Khater, "Chasma perturbations," *PIERS Online*, Vol. 1, No. 6, 720–723, Hangzhou, China, August 22–26, 2005.
7. Callebaut, D. K. and H. Kikuchi, "Debye shielding in chasmas," *Proc. PIERS 2007*, Prague, August 27–30, 2007.
8. Callebaut, D. K. and H. Kikuchi, "Extending the concept of Debye length in chasmas," *Proc. PIERS 2008*, Hangzhou, March 24–28, 2008, these proceedings.
9. Kikuchi, H. (ed.), *Relation between Laboratory and Space Plasmas*, Reidel Publ. Co., Dordrecht, The Netherlands, 1981.
10. Kikuchi, H., *Handbook of Atmospheric Electrodynamics*, Vol. 1, 167, CRC Press, Bonn, 1995.
11. Kikuchi, H., *Electrodynamics in Dusty and Dirty Plasmas (Gravito-Electrodynamics and EHD)*, Vol. 258, ASSL, Kluwer Acad. Press, Dordrecht, 2001.

Non-quasi-neutral Plasmas or Chasmas

D. K. Callebaut

Physics Department, CGB, University of Antwerp, Antwerp, B-2020, Belgium

Abstract— Chasmas are a generalization of plasmas, i.e., the condition of quasi-neutrality is dropped. In an accompagnying paper [13] the chasma equilibria were investigated and a shielding length for chasmas was introduced which generalizes the Debye length: $\lambda_D^2 = \epsilon k_B T/(n_- + n_+)e^2$. However, two other relevant quantities occurred in the equilibrium analysis: the chasma angular frequency: $\omega_c^2 = |n_- - n_+|e^2/\epsilon m_-$, and the associated length representing the distance over which a chasma may roughly have constant densities: $\lambda_c^2 = 2\epsilon k_B T/|n_- - n_+|e^2$, where the difference of the densities occurs in contrast to the plasma frequency and the Debye length, but satisfying $(\lambda_D\omega_p)^2 \approx (\lambda_c\omega_c)^2 \approx k_B T/m_-$. Now the perturbation of these chasma equilibria is investigated. In the case that the densities are roughly constant we obtain the dispersion relation which generalizes the one for plasmas, involving the plasma and the chasma frequencies. Previous studies for the stability of a steady state chasma gave not the same results so that further improvement of both cases is desirable.

1. INTRODUCTION

In a plasma one has by definition quasi-neutrality in volumes which have dimensions larger than the Debye length or at least a few times the Debye length. In some situations one has no quasi-neutrality over many times the Debye length. e.g., in certain discharges or in the multipactor effect [1–3] (secondary electron resonance discharge) in the cavities of linear accelerators. It is expected that chasmas occur too in certain extended double layers with currents and magnetic fields playing a role in the solar convective zone and solar atmosphere or in the terrestrial atmosphere. Callebaut and Knuyt [4–7] investigated theoretically the steady state, using an approach based on a singular integro-differential equation, [8–10]. (See [11–18] too.) They showed in 1-D, 2-D and 3-D that the ion density is constant provided the producing beam density is constant. Later Callebaut et al. [11] investigated the situation using the conventional approach involving all the basic equations, which then allowed to do an approximate perturbation analysis in which the steady state motions of the particles were not taken into account. They introduced the so called chasma angular frequency ω_c. (See below).

In an accompanying paper [13], we investigated chasma equilibria, distinct from the chasma steady states considered previously. Moreover, the generalization of the Debye length for chasmas was introduced, together with another characteristic length λ_c over which a chasma equilibrium may have constant densities. λ_c and ω_c have the same structure as the generalized λ_D and ω_p, but use the difference of the densities instead of the sum. The relation $(\lambda_D\omega_p)^2 \approx (\lambda_c\omega_c)^2 \approx k_B T/m_-$ was obtained.

In Section 3 we give first the basic equations and next recall the expressions for the chasma equilibria given in reference [13]. Next we consider perturbations of the chasma equilibria, in particular in the region where the densities and the potential are approximately constant (Section 4), followed by the conclusion (Section 5).

2. BASIC EQUATIONS AND CHASMA EQUILIBRIA

The basic equations here are those of continuity, momentum, Poisson and the polytropic one. They read, using a condensed notation

$$\partial_t n_\mp + \nabla(n_\mp \mathbf{v}_\mp) = \mathbf{0}, \tag{1}$$

$$m_\mp n_\mp(\partial_t \mathbf{v}_\mp + \mathbf{v}_\mp.\nabla\mathbf{v}_\mp) = -\nabla p + e(n_- - n_+)\nabla\varphi, \tag{2}$$

$$\epsilon\Delta\varphi = e(n_- - n_+), \tag{3}$$

$$p = p_- + p_+ = (n_- + n_+)k_B T = K_- n_-^{\Gamma_-} + K_+ n_+^{\Gamma_+}, \tag{4}$$

where the index — stands for quantities related to the electrons (charge $-e$) and + for the ions (once ionized, all of the same mass m_+). The density, velocity, potential and pressure are represented by n, $\mathbf{v}$, p and φ respectively. ϵ is the electric permittivity and equals in vacuum 8.8542×10^{-12} C/Vm; k_B is the Boltzmann constant (1.3807×10^{-23} J/K). The absolute temperature is T. $K_\mp$ are constants

and so are the polytropic exponentials $\Gamma_{\mp}$. If the ions are mono-atomic we have $\Gamma_- = \Gamma_+ = 5/3$ for adiabatic cases. For the equilibrium quantities (still depending on the coordinates) we use the index 0. In equilibrium we have $\mathbf{v} = 0$ and obtain:

$$\begin{aligned} 0 &= -\nabla p_0 + e(n_{-0} - n_{+0})\nabla\varphi_0, && (5)\\ \epsilon\Delta\varphi_0 &= e(n_{-0} - n_{+0}), && (6)\\ p_0 &= p_{-0} + p_{+0} = (n_{-0} + n_{+0})k_B T_0 = K_- n_{-0}^{\Gamma_-} + K_+ n_{+0}^{\Gamma_+}, && (7) \end{aligned}$$

In [13] we obtained in the 1-D case (a plane-parallel chasma, infinite in 2 directions; the x-axis perpendicular to the walls) an asymmetric solution (useful near a boundary)

$$\begin{aligned} \varphi_0 &= \frac{-2k_B T_0}{e}\ln\left(1 + \frac{x}{\lambda_c}\right), \quad n_{-0} = \frac{\lambda_c^2 n_{-0}(0)}{(\lambda_c + x)^2}, \quad n_{+0} = 0, \quad p_0 = n_0(0)k_B T_0,\\ \lambda_c &= \left[\frac{2\epsilon k_B T_0}{(n_{-0}(0) - n_{+0}(0))e^2}\right]^{1/2}, \end{aligned} \tag{8}$$

and a symmetric solution (obtained as a series development)

$$\varphi_0 = \frac{k_B T_0}{e}\sum_{s=1}^{\infty} a_{2s}x^{2s}, \quad a_2 = \lambda_c^{-2}, \quad a_4 = \frac{e^4}{24(\epsilon k_B T_0)^2}\left(n_{-0}(0)^2 - n_{0+}(0)^2\right), \quad a_6 = \ldots, \tag{9}$$

$$n_{\mp 0} = n_{\mp 0}(0)\exp\left(\pm\sum_{s=1}^{\infty} a_{2s}x^{2s}\right), \quad p_0 = (n_{-0} + n_{+0})k_B T_0, \tag{10}$$

where we have adapted the notation of Ref. [13] and replaced λ_- by its more general expression λ_c, taken at the origin. Note that Eq. (5) is satisfied although $n_{\mp}$ and φ were obtained using the Boltzmann equation. However, in the perturbed momentum equation we shall use the two fluid approximation instead of the one fluid one.

3. PERTURBING THE CHASMA EQUILIBRIUM

Writing $X = X_0 + X_1$ for each variable and linearizing yields:

$$\partial_t n_{\mp 1} + \partial_x(n_{\mp 0}v_{\mp 1}) = 0, \tag{11}$$

$$m_{\mp} n_{\mp 0}\partial_t v_{\mp 1} = -\partial_x p_1 \pm e n_{\mp 0}\partial_x\varphi_1 \pm e n_{\mp 1}\partial_x\varphi_0, \tag{12}$$

$$\epsilon\partial_{xx}\varphi_1 = e(n_{-1} - n_{+1}), \tag{13}$$

$$p_1 = \Gamma_- K_- n_{-0}^{\Gamma_- - 1} n_{-1} + \Gamma_+ K_+ n_{+0}^{\Gamma_+ - 1} n_{+1} = \left(\Gamma_- n_{-1} + \Gamma_+ n_{+}\right)k_B T_0 = c_{-s}^2 m_- n_{-1} + c_{+s}^2 m_+ n_{+1}. \tag{14}$$

We eliminate $v_{\mp 1}$:

$$-m_{\mp}\partial_{tt} n_{\mp 1} = -\partial_{xx} p_1 \pm e\partial_x(n_{\mp 0})\partial_x\varphi_1 \pm e n_{\mp 0}\partial_{xx}\varphi_1 \pm e\partial_x(n_{\mp 1})\partial_x\varphi_0 \pm e n_{\mp 1}\partial_{xx}\varphi_0. \tag{15}$$

Now we eliminate $p_1, \partial_{xx}\varphi_1$ and $\partial_{xx}\varphi_0$:

$$\begin{aligned} -m_{\mp}\partial_{tt} n_{\mp 1} = &-m_- c_{-s}^2\partial_{xx} n_{-1} - m_+ c_{+s}^2\partial_{xx} n_{+1} \pm e\partial_x(n_{\mp 0})\partial_x\varphi_1 \pm\\ &\left(e^2 n_{\mp 0}/\epsilon\right)(n_{-1} - n_{+1}) \pm e\partial_x\left(n_{\mp 1}\right)\partial_x\varphi_0 \pm \left(e^2 n_{\mp 1}/\epsilon\right)(n_{-0} - n_{+0}). \end{aligned} \tag{16}$$

The condition to treat n_{-0}, n_{+0} and φ_0 as a constants is $x << \lambda_c$. In the symmetric case we have usually $\lambda_c >> \lambda_D$ so that the approximation is valid in a reasonable range. With this approximation and putting

$$n_{\mp 1} = C_{\mp}\exp[i(\omega t + kx)], \tag{17}$$

with $C_{\mp}$ arbitrary constants we obtain

$$m_{\mp}\omega^2 n_{\mp 1} = m_- c_{-s}^2 k^2 n_{-1} + m_+ c_{+s}^2 k^2 n_{+1} \pm \left(e^2 n_{\mp 0}(0)/\epsilon\right)(n_{-1} - n_{+1}) \pm \left(e^2 n_{\mp 1}/\epsilon\right)(n_{-0}(0) - n_{+0}(0)). \tag{18}$$

These are 2 relations between n_{-1} and n_{+1}. The compatibility condition yields the dispersion equation:

$$\left[\omega^2 - c_{-s}^2 k^2 - e^2 \left(2n_{-0}(0) - n_{+0}(0)\right)/\epsilon m_-\right]\left[\omega^2 - c_{+s}^2 k^2 - e^2 \left(2n_{+0}(0) - n_{-0}(0)\right)/\epsilon m_+\right] =$$
$$= \left[c_{-s}^2 k^2 - e^2 n_{+0}(0)/\epsilon m_-\right]\left[c_{+s}^2 k^2 - e^2 n_{-0}(0)/\epsilon m_+\right]. \quad (19)$$

This is a bi-quadratic equation in ω and a quadratic one in k. In the usual case that $m_+ >> m_-$ it simplifies to

$$\left[\omega^2 - c_{-s}^2 k^2 - e^2(2n_{-0}(0) - n_{+0}(0))/\epsilon m_-\right]\omega^2 = \left[c_{-s}^2 k^2 - e^2 n_{+0}(0)/\epsilon m_-\right]\left[c_{+s}^2 k^2 - e^2 n_{-0}(0)/\epsilon m_+\right] =$$
$$= (m_-/m_+)\left[c_{-s}^2 k^2 - e^2 n_{+0}(0)/\epsilon m_-\right]\left[c_{-s}^2 k^2 - e^2 n_{-0}(0)/\epsilon m_-\right] \approx 0. \quad (20)$$

In the plasma case we recover the well known dispersion relation $\omega^2 = c_{-s}^2 k^2 + e^2 n_0/\epsilon m_-$. In the case of an electron-positron chasma we obtain:

$$= \left[c_s^2 k^2 - e^2 n_{+0}(0)/\epsilon m_-\right]\left[c_s^2 k^2 - e^2 n_{-0}(0)/\epsilon m_-\right]. \quad (21)$$

For the electron-positron plasma this reduces to $\omega^2 - 2c_s k^2 - 2e^2 n_0/\epsilon m_-$ and $\omega^2 = 0$.

Recalling that the chasma frequency ω_c was defined in previous work [11] as

$$\omega_c = \sqrt{\frac{e^2(n_{-0}(0) - n_{+0}(0))}{\epsilon m_-}}, \quad (22)$$

(which satisfies the relation: $(\lambda_c \omega_c)^2 = 2k_B T/m_-$) we may rewrite the relevant part of the dispersion relation (20) as

$$\omega^2 = c_{-s}^2 k^2 + e^2(2n_{-0}(0) - n_{+0}(0))/\epsilon m_- = c_{-s}^2 k^2 + \omega_{-e}^2 + \omega_c^2. \quad (23)$$

(We omit subtleties related to $n_{+0} - n_{-0}$ being positive or not.) Often $\omega_{-e} >> \omega_c$ and the change with the plasma dispersion relation is small. However, in an experiment a small change in ω can have drastic effects e.g. when resonance effects are involved as in the case of h.f. heating.

We may neglect the pressure terms in Eq. (23) provided

$$k^2 << e^2(2n_{-0}(0) - n_{+0}(0))/\epsilon \Gamma_- k_B T_0. \quad (24)$$

On the other hand we must have $k\lambda_c > 1$ in order to have several wavelengths in the region where the densities are constant. Multiplying both sides of Eq. (24) with λ_c^2 leads to the conditions to neglect the pressure terms in the dispersion relation:

$$1 < (k\lambda_c)^2 << \frac{2(2n_{-0}(0) - n_{+0}(0))}{\Gamma_-(n_{-0}(0) - n_{+0}(0))}. \quad (25)$$

both conditions may be satisfied if $n_{-0}(0) \approx n_{+0}(0)$. e.g., in the plasma case there is no problem.

4. CONCLUSION

We have studied the perturbation of a chasma equilibrium considered in paper [13] In the region where we may consider the densities and the potential to be approximately constant we obtained a dispersion equation which is bi-quadratic in ω, but quadratic in k. If $m_+ >> m_-$, it simplifies considerably. However, the chasma dispersion relation contains new combinations of frequencies based on differences between the densities of electrons and ions. Even in the approximation $m_+ >> m_-$, ω_p occurs in combination with ω_c, the so-called chasma frequency.

Most of all we have to realize that the previous analysis applies to the region where the densities are approximately constant, i.e; over a region of the order of λ_c, which is large if the densities do not differ much. However, the neighboring regions, where the densities are not constant, may not only alter the oscillation frequency, but may even have a destabilizing effect on the whole chasma. In fact in an analysis of a steady state chasma, (as occurring in the multipactor effect in accelerators) we obtained nearly always instability. However, it must be mentioned that the analysis there was approximate because the steady state motions were neglected in the perturbation theory. Clearly still many generalizations are possible.

REFERENCES

1. Callebaut, D. K., *Physica*, Vol. 29, 784–802, 1963.
2. Callebaut, D. K., *Physica*, Vol. 31, 1177–1192, 1965.
3. Callebaut, D. K., *Physica*, Vol. 32, 1151–1158, 1966.
4. Callebaut, D. K. and G. K. Knuyt, *Phys. Lett.*, Vol. 29A, 72, 1969.
5. Callebaut, D. K. and G. K. Knuyt, *Plasma Phys.*, Vol. 20, 511–523, 1978.
6. Callebaut, D. K. and G. K. Knuyt, *Plasma Phys.*, Vol. 20, 524–550, 1978.
7. Callebaut, D. K. and G. K. Knuyt, *Relation between Laboratory and Space Plasmas*, (Ed. H. Kikuchi), 207–230, Reidel Publ. Co., Dordrecht, The Netherlands, 1981.
8. Callebaut, D. K. and G. J. Bex, *Workshop on Nonlinear Stability*, (Ed. D. K. Callebaut), 154, 161, UIA Press, University of Antwerp, Antwerp, Belgium, 1989.
9. Callebaut, D. K., *Second URSI-ICPIG-RIKEN Symposium on Critical Problems of Discharge and Plasma Physics and Electro-Plasmo-Meteorology*, 351-01, 40–41, Riken, The Institute of Physical and Chemical Research, Wako, Saitama, Japan, 1995.
10. Callebaut, D. K. and A. H. Khater, *International Symposium on Nonconventional Plasmas*, (Eds. H. Kikuchi and D. K. Callebaut), 1–2, Niigata, Japan, 1997.
11. Callebaut, D. K., G. K. Karugila, and A. H. Khater, "Chasma perturbations," *Progress In Electromagnetic Research Symposium*, 720–723, Hangzhou, China, August 22–26, 2005.
12. Callebaut, D. K. and H. Kikuchi, "Debye shielding in chasmas," *Progress In Electromagnetic Research Symposium*, Prague, 27–30 August, 2007.
13. Callebaut, D. K. and H. Kikuchi, "Extending the concept of Debye length for chasmas," *Progress In Electromagnetic Research Symposium*, Hangzhou, 27–30 March, 2008 (These proceedings, refered to as the accompanying paper.)
14. Kikuchi, H., (ed.), *Relation between Laboratory and Space Plasmas*, Reidel Publ. Co., Dordrecht, The Netherlands, 1981.
15. Kikuchi, H., *Handbook of Atmospheric Electrodynamics*, H. Volland (ed.), Vol. 1, 167, CRC Press, Bonn, 1995.
16. Kikuchi, H., *Electrodynamics in Dusty and Dirty Plasmas (Gravito-Electrodynamics and EHD), ASSL*, Vol. 258, Kluwer Acad. Press, Dordrecht, 2001.
17. Schrittwieser, R. W. and G. Eder, (eds.), *Proc. Second Symposium on Plasma Double Layers and Related Topics*, Institut fr Theoretische Physik, Univ. of Innsbruck, A-6020 Innsbruck, Auqstria, 1984.
18. Schrittwieser, R. W., (ed.), *Fourth Symposium on Double Layers and other Nonlinear Potential Structures in Plasmas*, Innsbruck, Austria, World Scientific, Singapore, 1993.

Usefulness of a Universal Electric-cusp Type Plasma Reactor in Basic Studies and a Variety of Applications in Dust Dynamics, Ionization and Discharge Physics Based on Electrohydrodynamics

H. Kikuchi
Institute for Environmental Electromagnetics, 3-8-18, Komagome, Toshima-ku, Tokyo 170, Japan

Abstract— A universal electriac-cusp type plasma reactor designed more than a decades ago by the present author has successfully been in operation for the last coupe of years and has been proved useful for basic studies and a variety of applications in dust dynamics, ionization and discharge physics, including laboratory simulation of universe, atmospheric and space electricity and plasmas, based on 'Electrohydrodynamics (EHD)/Electromagnetohydrodynamics (EMHD)'. This paper aims to present the structure and operation of this plasma reactor and to show how this device is useful for basic studies and applications, citing a number of examples. The new device is a square box with two lead electrodes (15 mm in diameter and 5 cm in interval) inside, suspended $2.75 \sim 5$ cm above a metallic plate. When a tiny object or dust grain, conducting or dielectric, is placed in the cusp center, electric field line merging toward it occurs from the four or two poles, inducing or polarizing electric charges on its surface or in its volume, negative or positive facing positive or negative poles, respectively. Then a catastrophe occurs, namely zero-electric field without the object or dust suddenly tends to sufficiently high electric fields, almost infinity, around it. We are now ready to be advanced to one of entirely different two directions, depending upon the background gas pressure. One is the case of energy transfer from fields to kinetic energy leading to dust dynamics in a pair of electric mirror for the background gas pressure below the breakdown threshold, and the other is the case of energy transfer from fields to ionization resulting in an electric discharge for the background gas pressure beyond the breakdown threshold. First we deal with the former case. When an uncharged dust grain, conducting or dielectric, is placed onto the center of a quadrupole, dust starts moving between conjugating mirror points and is going back and forth undulating the mid-plane. If a dust grain is negatively or positively charged its motion in periodic cusps and mirrors in the midplane of a quadrupole forming an electric mirror. Next we proceed to three dimensional motion of an uncharged or charged dust grain not in the ecliptic plane of a quadrupole, Then the dust grain is going to helical motion due to helicity generation of an electric quadrupole. The second case when the background gas pressure is beyond the breakdown threshold leads to a variety of electric discharge phenomena, laboratory evidence of 'electric cusp-mirror and reconnection model' as well as the first case and provides basic studies of ionization and discharge processes, laboratory simulations of universe, atmospheric and space electric electricity and plasma phenomena, applications to industrial plasmas, including plasma processing, new material production such as diamond, electric precipitator and so on.

1. INTRODUCTION

A universal electriac-cusp type plasma reactor was designed more than a decades ago by the present author [1] The device is a square box with two lead electrodes (15 mm in diameter and 5 cm in interval) inside, suspended $2.75 \sim 5$ cm above a metallic plate as shown in Fig. 1. When a tiny object or dust grain, conducting or dielectric, is placed in the cusp center, electric field line merging toward it occurs from the four or two poles, inducing or polarizing electric charges on its surface or in its volume, negative or positive facing positive or negative poles, respectively. Then a catastrophe occurs, namely zero-electric field without the object or dust suddenly tends to sufficiently high electric fields, almost infinity, around it.

2. OCCURRENCE OF A CATASTROPHE: DUST AND/OR OBJECT-RELATED ELECTRIC RECOMMECTION

Then we are ready to advance to entirely different two situations, depending upon the background gas pressure as illustrated in Table 1. Namely, the casee. of the background gas pressure is below the threshold 1. The other is case of the background gas pressure beyond the threshold as illustrated in the right column of Table 1 that leads to 'electric discharge' (main discharge or return stroke) through a sequence of those processes.

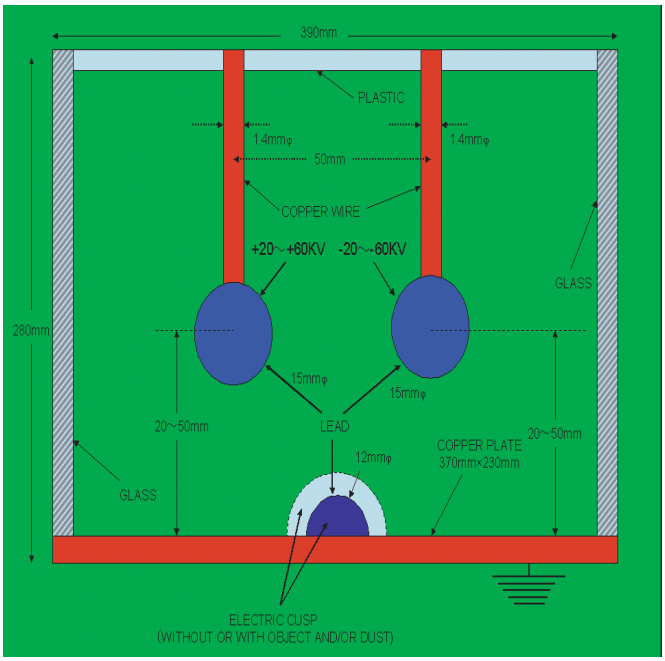


Figure 1: Universal electric-cusp type plasma reactor. Object at the cusp center.

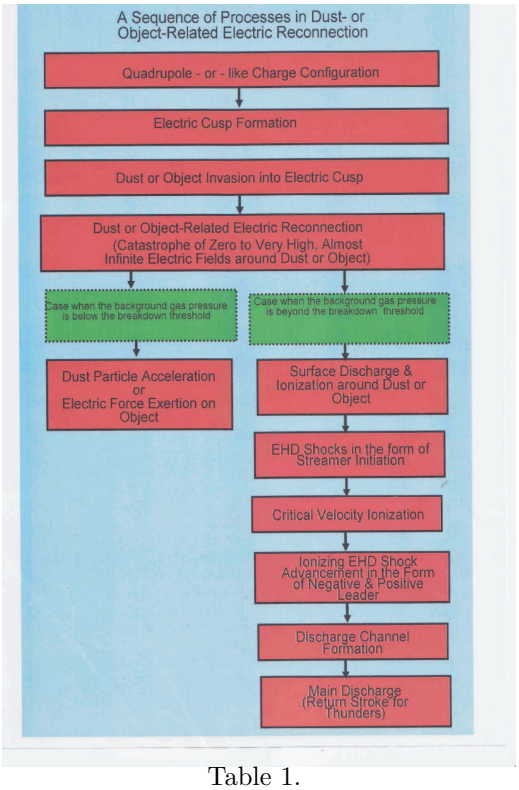


Table 1.

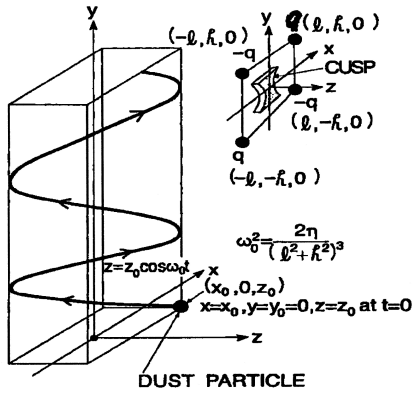


Figure 2: Helical motion of an uncharged dust grain not in the ecliptic plane of a quadrupole.

Figure 3: Electric discharge from positive and negative electrodes to a mesa-type lead placed at a cusp center on a copper plate. Above bright light is a real discharge, while lower weak light is its image. The voltages of electrodes are ±21 KV both and discharge current is around 590 μA. The mesa diameter is 5 mm. It is interesting to compare flat discharge to mesa-type object with sharp discharge to hemisphere in Fig. 4. Also, neither direct discharge between two electrodes nor electrodes to copper ground occurred because of more electric field merging from both electrodes to object.

3. LABORATORY EVIDENCE OF ELECTRIC RECONNECTION MODEL BY DUST-AND/OR OBJECT-RELATED ELECTRIC RECONNECTION BY DUST DYNAMICS IN AN ELECTRIC QUADRUPOLE AND DUST-AND/OR OBJECT-RELATED ELECTRIC DISCHARGE AND ITS VARIETY OF APPLICATIONS INCLUDING ELECTRIC DISCHARGE PHENOMENA, BASIC STUDIES OF IONIZATION AND DISCHATGE PROCESSES, LABORATORY SIMULATIONS OF UNIVERSE, ATMOSPHERIC AND SPACE ELECTRIC ELECTRICITY AND PLASMA PHENOMENA, INDUSTRIAL PLASMA PROCESSING, NEW MATERIAL PRODUCTION SUCH AS DIAMOND, ELECTRIC PRECIPITATOR AND SO ON.

A variety of above dust dynamics and electric discharge observations by a universal electric-cusp type plasma reactor are typically shown in Figs. 2–4.

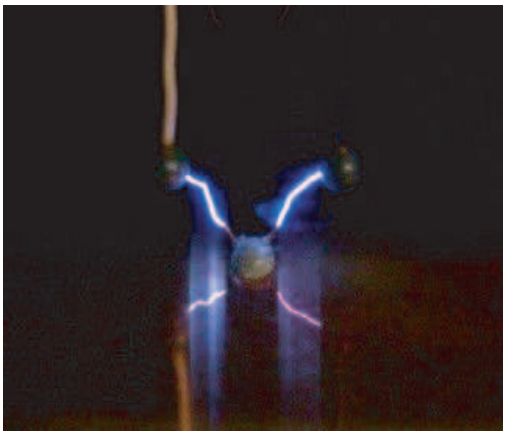


Figure 4: Electric discharge ftpm positive and negative electrodes to a hemispherical lead with double ion needles electroplated and directed to electrodes which placed at a cusp center on a copper plate. Above bright light is a real discharge, while lower weak light is its image. The voltage of both electrodes are ± 21 KV and discharge current is around 500 µA. The length of needle is 5.5 mm. It may be interesting to place object not at a cusp center on a copper plane and to how the discharge behaviour changes. So we typically choose a couple of cases: (i) Object is placed not a cusp center but still inside a cusp region on a copper plane; (ii) Object is placed not a cusp center but inside a cusp region on a copper plane; (iii) Object is placed not a cusp center but outside a cusp region on a copper plane on a copper plane; (iv) Object is placed not a cusp center but just below the positive electrode on a copper plane.

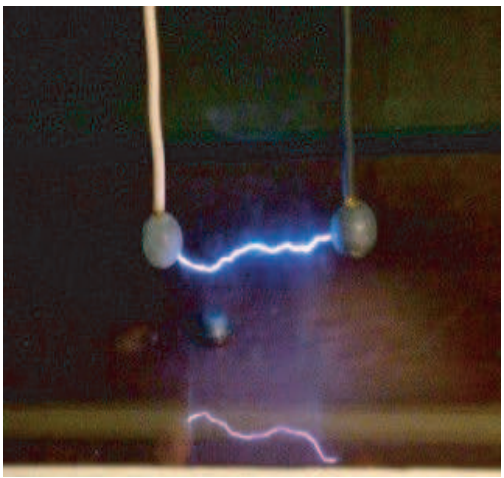


Figure 5: Electric discharge from positive and negative electrodes to a mesa-type lead shifted 1.5 cm left from a cusp center. Above bright light is a real discharge, while lower light is its image. The voltages of both electrodes are ± 21 KV and discharge current is around 560 µA. Mesa diameter is 5 mm. Note that object is still inside a cusp region.

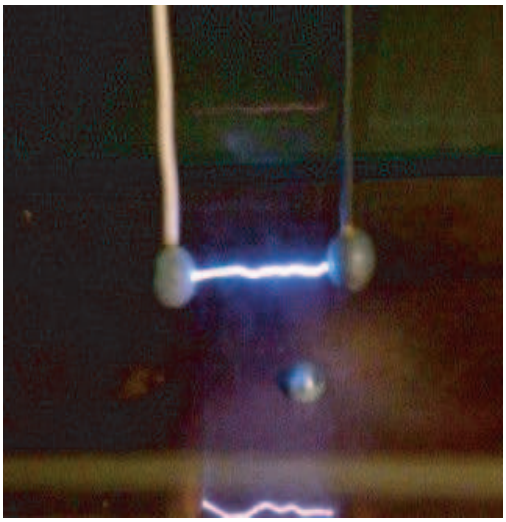


Figure 6: Electric discharge between two positive and negative electrodes for a hemispherical lead shifted 4 cm right from a cusp center. Above bright light is a real discharge, while lower weak light is its image. The voltages of both electrodes are ± 21 KV and discharge current is around 320 μA. Note that discharge from positive and negative electrodes to hemispherical lead no longer occurs because object is outside a cusp region.

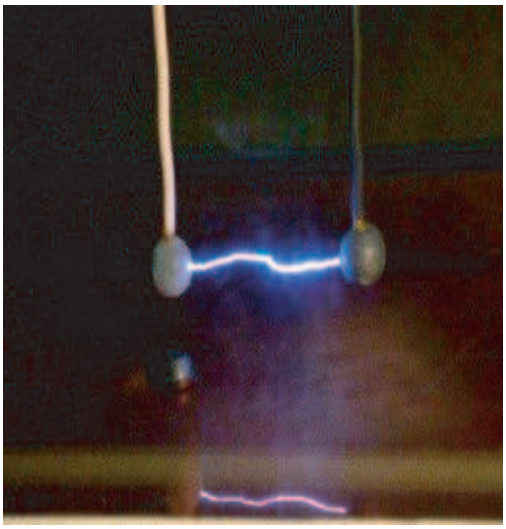


Figure 7: Electric discharge between two electrodes for hemispherical lead placed just under positive electrode. Above bright light is a real discharge, while lower weak light is its image. The voltages of both electrodes are ± 21 KV and discharge current is around 400 μA. Discharge from positive and negative electrodes to object no longer occurs because object is outside a cusp region.

4. CONCLUSION

A universal electriac-cusp type plasma reactor designed more than a decades ago by the present author has successfully been in operation for the last coupe of years and has been proved useful for basic studies and a variety of applications in dust dynamics, ionization and discharge physics, including laboratory simulation of universe, atmospheric and space electricity and plasmas, based on 'Electrohydrodynamics (EHD)/Electromagnetohydrodynamics (EMHD)'. In particular, structure and operation of the novel universal electric-cusp plasma reactor are described in detail.

Specific usefulness are cited below:

(1) Basic studies of dust dynamics and its applications;
(2) Basic studies of ionization and discharge physics;
(3) Laboratory simulation of universe, atmospheric and space electricity and plasmas;

(4) A variety of industrial applications, including plasma processing, new material production such as diamond, and electric precipitator.

REFERENCES

1. Kikuchi, H., *Electrohydrodynamics in Dusty and Dirty Plasmas*, Vol. 258, Astrophysics and Space Science Library, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2001.
2. Kikuchi, H., *Physics and Chemistry of the Earth*, Vol. 21, 549–557, 1996.
3. Kikuchi, H., *Dusty and Dirty Plasmas, Noise and Chaos in Space and in the Laboratory*, Plenum, New York, 1994.

Magnetoplasmons in Graphene Structures

Oleg L. Berman[1], **Godfrey Gumbs**[2], **and Yurii E. Lozovik**[3]

[1]Physics Department, New York City College of Technology, City University of New York, USA
[2]Department of Physics, Hunter College, City University of New York, USA
[3]Institute of Spectroscopy, Russian Academy of Sciences, Russia

Abstract— The dispersion relation for magnetoplasmon excitations in a single layer and a pair of parallel layers of graphene is calculated in the random-phase approximation. Our formalism is valid for integer filling factor. The graphene layers are embedded in a dielectric medium and a strong perpendicular quantizing ambient magnetic foeld is applied. We present numerical results for the for magnetplasmon excitation energies as a function of the in-plane wave vector. Our numerical calculations show that for a graphene layer, there is a collective mode which is not Landau damped by the particle-hole excitations. This magnetoplasmon has a negative as well as a positive group velocity for different ranges of wave vector at fixed magnetic field. For bilayer graphene, the magnetoplasmon mode is split by the inter-layer Coulomb interaction into a symmetric (in-phase) and antisymmetric (out-of-phase) oscillatory mode. A discussion of the plasmon instabilities associated with these layered structures is presented.

1. INTRODUCTION

Recent advances in fabrication techniques have made it possible to produce graphene, which is a two-dimensional (2D) honeycomb lattice of carbon atoms forming the basic planar structure in graphite [1, 2]. Graphene has stimulated considerable theoretical interest as a semi-metal whose electron effective mass may be described by an unusual massless Dirac-fermion band structure. Several novel many-body interactions in graphene have been investigated [3]. In recent experiments, the integer quantum Hall effect (IQHE) has been reported in graphene [4–6]. The quantum Hall ferromagnetism in graphene has been investigated from a theoretical point of view [7]. Graphene has a number of interesting properties as a result of its unusual band structure which is linear near two inequivalent points (K and K') in the Brillouin zone. The single-electron quantum states near K and K' are described by a Dirac-type equation, where the wave functions are spinors because of the two-point basis of the honeycomb lattice. In the presence of a magnetic field, the graphene structure shifts both the Shubnikovde Haas oscillations [8] as well as the step pattern of the IQHE [9]. Both these effects have recently been reported experimentally [4, 5]. The spectrum of plasmon excitations in a single graphene layer embedded in a material with effective dielectric constant ε_b in the absence of an external magnetic field ($B = 0$) was calculated in [10].

In this paper, we calculate the nonlocal dielectric function and the dispersion equation for magnetoplasmons in a single graphene layer and for a pair of spatially separated graphene layers. The charge carriers may be either electrons or holes. A strong perpendicular magnetic field $\mathbf{B}$ is apllied. The calculations are carried out in the random-phase approximation (RPA).

2. MODEL FOR GRAPHENE LAYER AND DISPERSION EQUATION FOR MAGNETOPLASMONS

Let us consider electrons in a single graphene layer in the the xy plane in a perpendicular magnetic field $\mathbf{B}$ parallel to the positive z axis. The effective-mass Hamiltonian for noninteracting electrons in one valley in graphene in the absence of scatterers is given by the following equation. Here, we neglect the Zeeman splitting and assume valley energy degeneracy, describing the eigenstates by two pseudospins [9, 10]. We have

$$\hat{H}_{(0)} = v_F \begin{pmatrix} 0 & \hat{\pi}_x - i\hat{\pi}_y \\ \hat{\pi}_x + i\hat{\pi}_y & 0 \end{pmatrix}, \tag{1}$$

where $\hat{\pi} = -i\hbar\nabla + e\mathbf{A}$, $-e$ is the electron charge, $\mathbf{A}$ is the vector potential, $v_F = \sqrt{3}at/(2\hbar)$ is the Fermi velocity with $a = 2.566$Å denoting the lattice constant, and $t \approx 2.71\,\mathrm{eV}$ is the overlap integral between nearest-neighbor carbon atoms [9].

In the Landau gauge, $\mathbf{A} = (0, Bx, 0)$, the eigenfunctions $\psi_\alpha(\mathbf{r})$ of the Hamiltonian $\hat{H}_{(0)}$ in Eq. (1) are labeled by the set of quantum numbers $\alpha = \{k_y, n, s(n)\}$, where $n = 0, 1, 2, \ldots$ is the Landau level index, k_y is the electron wave vector in the y direction, and $s(n)$ which is defined by

$$s(n) = \begin{cases} 0 & (n = 0), \\ \pm 1 & (n > 0), \end{cases} \tag{2}$$

labels the conduction $(+1)$ and valence $(-1$ and $0)$ band, respectively. The two-component eigenfunction $\psi_\alpha(\mathbf{r})$ is calculated and given by [9]

$$\psi_\alpha(x, y) = \frac{C_n}{\sqrt{L_y}} \exp(ik_y y) \begin{pmatrix} s(n) i^{n-1} \Phi_{n-1}(x + r_B^2 k_y) \\ i^n \Phi_n(x + r_B^2 k_y) \end{pmatrix}, \tag{3}$$

where $r_B = \sqrt{\hbar/eB}$ is the magnetic length, and L_x and L_y are normalization lengths in the x and y directions, respectively. We have

$$C_n = \begin{cases} 1 & (n = 0), \\ 1/\sqrt{2} & (n > 0), \end{cases} \tag{4}$$

and

$$\Phi_n(x) = \left(2^n n! \sqrt{\pi} r_B\right)^{-1/2} \exp\left[-\frac{1}{2}\left(\frac{x}{r_B}\right)^2\right] H_n\left(\frac{x}{r_B}\right), \tag{5}$$

where $H_n(x)$ is a Hermite polynomial. The eigenenergies depend on the quantum numbers n and s only and are given by

$$\epsilon_\alpha = s(n)\epsilon_n = s(n)\frac{\hbar v_F}{r_B}\sqrt{2n}, \tag{6}$$

for which successive levels are not equally spaced. We also note that the ratio of the Zeeman term $\Delta E_Z(B)$ to the separation between adjacent Landau levels $\Delta E_L(B)$ is negligible at High magnetic field. For $B = 10\,\mathrm{T}$, we have $\Delta E_Z(B)/\Delta E_L(B) \approx \mu_B B/(\sqrt{2}\hbar v_F r_B^{-1}) \approx 5 \times 10^{-3}$. Here, $\mu_B = e\hbar/(2m_e)$ is the Bohr magneton with m_e denoting the free electron mass. Therefore, the contributions to the single-electron Hamiltonian from the Zeeman splitting and very small pseudospin splitting caused by two valleys in graphene may be neglected, as done in [9, 11, 12]. We assume energy degeneracy for the two possible spin projections and two graphene valleys described by pseudospins.

We now derive an expression for the dielectric function of a graphene layer in the presence of a perpendicular magnetic field. We use RPA [13] following the procedure described in [14]. The response of the system to a weak external perturbing potential $\phi_{\mathrm{ext}}(\mathbf{r}, t)$ acting at the space-time point $(\mathbf{r}, t)$ can be obtained by calculating the density matrix from its equation of motion $i\hbar\partial\hat{\rho}/\partial t = [\hat{H}, \hat{\rho}]$, where $\hat{H}$ is the total Hamiltonian which includes electron-electron interactions. For a weak perturbation, we write $\hat{\rho} = \hat{\rho}_0 + \hat{\rho}_1$ and $\hat{H} = \hat{H}_0 - e\phi$, where $\hat{\rho}_0$ is the equilibrium density matrix and $\hat{\rho}_1$ its perturbation, $\hat{H}_0$ is the unperturbed Hamiltonian and $\phi(\mathbf{r}, t)$ is the total scalar potential given by $\phi(\mathbf{r}, t) = \phi_{\mathrm{ext}}(\mathbf{r}, t) + \phi_{in}(\mathbf{r}, t)$. In this notation, $\phi_{in}(\mathbf{r}, t)$ is the induced electric potential due to the external perturbation $\phi_{\mathrm{ext}}(\mathbf{r}, t)$. In the basis set given by Eq. (3), we have

$$< \alpha|\hat{\rho}_0|\alpha >= g_s g_v f_0(\epsilon_\alpha)\delta_{\alpha\alpha'}, \qquad < \alpha|H_0|\alpha >= \epsilon_\alpha \delta_{\alpha\alpha'}, \tag{7}$$

where $f^{(0)}(\epsilon_\alpha)$ is the occupation factor of the state $|\alpha\rangle$ determined by the Fermi-Dirac distribution function. We have $f_0(\epsilon_\alpha) = (\exp[\epsilon_\alpha/(k_B T)] + 1)^{-1}$ if the energy is measured from the Fermi level, k_B is the Boltzmann constant and T is temperature. Also, $g_s = 2$ and $g_v = 2$ are the spin and graphene valley degeneracies. Substituting Eq. (7) into the equation of motion for the density matrix, and solving the resulting equation in lowest order, we obtain

$$< \alpha|\hat{\rho}_1|\alpha' >= g_s g_v e\, \frac{f_0(\epsilon_\alpha) - f_0(\epsilon_{\alpha'})}{\hbar\omega + \epsilon_\alpha - \epsilon_{\alpha'}} < \alpha|\phi|\alpha' >, \tag{8}$$

where ω is the frequency.

We take account of screening due to the background medium by assuming that the graphene layer is embedded in a material with effective dielectric constant ε_b and take $\varepsilon_b = 2.5$ in our calculations assuming that the medium is SiO_2 on one side and air on the other. Poisson's equation is

$$\nabla^2\phi_{in} = \frac{4\pi e}{\varepsilon_s}\mathrm{Tr}\left[\hat{\rho}_1\right]\delta(z), \tag{9}$$

where $\varepsilon_s = 4\pi\epsilon_0\varepsilon_b$ and

$$\mathrm{Tr}\left[\hat{\rho}_1\right] = \sum_{\alpha\alpha'}|\alpha><\alpha|\hat{\rho}_1|\alpha'><\alpha'|. \tag{10}$$

Substituting Eq. (8) into Eq. (10), and Eq. (10) into Eq. (9), and Fourier transforming Eq. (9), and introducing the in-plane wave vector q, we have $\phi(\mathbf{q},\omega) = \phi_{\mathrm{ext}}(\mathbf{q},\omega)/\varepsilon(q,\omega)$, where $\varepsilon(q,\omega)$ is the dynamic dielectric function in RPA [13] given by $\varepsilon(q,\omega) = 1 - V_c(q)\Pi(q,\omega)$, where $V_c(q) = 2\pi e^2/(\varepsilon_s q)$ is the 2D Coulomb interaction and $\Pi(q,\omega)$ is the 2D polarization function given by

$$\Pi(q,\omega) = \frac{g_s g_v}{2\pi r_B^2}\sum_{n=0}^{\infty}\sum_{n'=0}^{\infty}\sum_{s(n),s'(n')}\frac{f_{s(n)n} - f_{s'(n')n'}}{\hbar\omega + \epsilon_{s(n)n} - \epsilon_{s'(n')n'}}F_{s(n)s'(n')}(n,n',q), \tag{11}$$

where $F_{ss'}(n,\, n')$ is a form factor arising from the overlap of eigenstates and is given by

$$F_{ss'}(n,n',q) = C_{n_1}^2 C_{n_2}^2\left[-\frac{q^2 r_B^2}{2}\right]^{n_1-n_2}\frac{1}{|(n_1-n_2)!|^2}\times\left(s_1(n_1)s_2(n_2)\left|\frac{(n_1-1)!}{(n_2-1)!}\right| + \left|\frac{n_1!}{n_2!}\right|\right). \tag{12}$$

At $T = 0\,\mathrm{K}$, there will be N_L filled Landau levels in the valence band in Eq. (11) with $f_{n-} = 1$ for $1 \le n \le N_L$ and $s(n) = -1$. For the highest filled Landau level, we have $n = 0$, $f_{0-} = \nu$, where ν is the filling factor of that Landau level. Since the wavefunction of an electron in graphene depends on the guiding center $k_y r_B^2$, the number N_L of completely filled Landau levels, except the highest level ($n = 0$, and ν are given by $2\pi r_B^2 n_{2\mathrm{D}} = N_L + \nu$, where $n_{2\mathrm{D}}$ is the 2D electron density, N_L is an integer, and $0 \le \nu < 1$. We consider the $n = 0$ Landau level as belonging to the valence band, and we set $f_{0+} = 0$ at all temperatures. We note, that at $T = 0\,\mathrm{K}$, the occupation of the Landau levels in the empty conduction band is given by $f_{n+} = 0$.

At high temperatures or weak magnetic field, when the separation between Landau levels is small, and $k_B T \gg \hbar v_F/r_B$, the occupation of the Landau levels is given by the Fermi-Dirac distribution function $f_{n,\,s(n)} = (\exp[\epsilon_{n,s(n)}/(k_B T)] + 1)^{-1}$ for $n > 0$ if the energy is measured from the Fermi level. To simplify our numerical calculations, we assume that the highest Landau level in the valence band is completely full and all other valence levels are empty. Furthermore, we will assume that the magnetic field is sufficiently high that the Landau level separation is large. Consequently, we will take only a few terms in the polarization sums since transions to higher levels have smaller oscillator strengths [10]. Furthermore, since the separation between Landau levels decreases as n increases and is $\propto 1/\sqrt{n}$, it is reasonable to include the Landau levels with $1 < n \le 3$ in our calculations.

The magnetoplasmon dispersion relation for a single graphene layer may now be calculated numerically by seeking the zeros of the dynamical dielectric function, i.e., $\varepsilon(q,\, \omega) = 0$ where $\varepsilon(q,\, \omega)$ is defined above and the polarization function $\Pi(q,\, \omega)$ is obtained from Eq. (11). The magnetoplasmon dispersion relation fora single graphene layer with the highest valence band full and all others empty at $T = 0\,\mathrm{K}$ is presented in Fig. 1.

Our numerical calculations show that there is a negative group velocity for $qr_B > 1$ which is caused by magnetic field. The magnetic field also causes the negative group velocity of magnetoplasmons in a two-dimensional electron gas (2DEG) [17]. The horizontal straight line corresponds to the particle-hole mode excitation energies. There is no Landau damping of the magnetoplasmon branch outside this particle-hole mode region. The effect of the Coulomb interaction is to split off a single-partcle excitation into a collective mode. The dependence of $V_c(q)$ on wave vector leads to the variation of the magnetoplasmon energy with wave vector. For large q, $v_c(q) \to 0$, leading to little effect of the Coulomb interaction on the single-particle excitation energy for $qr_B \gg 1$ in Fig. 1.

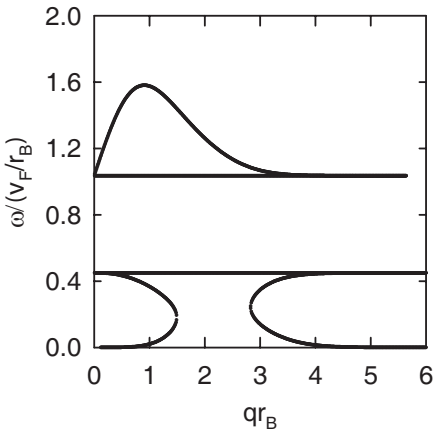


Figure 1: The normalized magnetoplasmon excitation energy as a function of wave vector, in units of r_B^{-1}, in a single graphene layer. We chose $T = 0\,\mathrm{K}$, only one valence band is full and all others are empty, and transitions to the two lowest conduction bands are included in the sums for the polarization function.

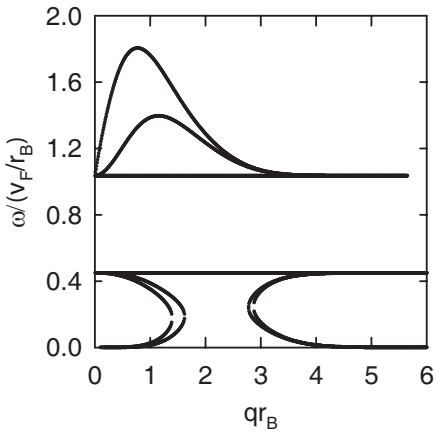


Figure 2: The normalized magnetoplasmon excitation energy at $T = 0\,\mathrm{K}$ as a function of the in-plane wave vector q measured in units of r_B^{-1} for bilayer graphene with separation $D = r_B$. Only the highest Landau level in the valence is occupied and completely full.

3. MAGNETOPLASMONS IN BILAYER GRAPHENE

The possibility of having electrons and holes that are spatially separated in two parallel graphene layers separated by an insulating slab of SiO_2 in the presence of a perpendicular magnetic field has been proposed in [18]. The graphene sheets are surrounded by insulators of dielectric constants ε_1 and ε_2 above and below the graphene layers, respectively. The insulator between the layers has dielectric constant ε_3. The charge carriers, which may be either electrons or holes, could be controlled as they are in coupled quantum wells (CQWs) [19-22] by laser pumping (far infrared in graphene).

The spatial separation of charge carriers in different graphene layers can be achieved by applying an external electric field. Furthermore, the spatially separated carriers can be created by varying the chemical potential, using a bias voltage between a pair of graphene layers, or between gates located near the corresponding graphene sheets. The ratio of the external voltage V_{ext} to the interlayer separation D required to create spatially separated electrons and holes in graphene layers with the 2D density $n_{2\mathrm{D}} = 10^{11}\,\mathrm{cm}^{-2}$ is given by $V_{\mathrm{ext}}/D = 4\pi e n_{2\mathrm{D}} D/\epsilon_s = 4.021 \times 10^4\,\mathrm{V/cm}$. Here, $\epsilon_b = 4.5$ is the dielectric constant of SiO_2. Since the critical electric field E_{cr} for dielectric breakdown in SiO_2 is $E_{\mathrm{cr}} \approx 10^6\,\mathrm{V/cm}$, we conclude that the external electric field for spatially separated electrons and holes is less than the critical electric field for dielectric breakdown in SiO_2.

We may calculate the dispersion equation in the RPA for magnetoplasmons in bilayer graphene when there is no inter-layer hoping using the same methods employed for semiconducting systems employed by the authors in [15, 23–25]. The layers are in the planes $z = 0$ and $z = D$. The spectrum of magnetoplasmons is obtained in RPA by solving the following equation [15, 23]:

$$\left(2V_c(q)\Pi_{11}(q,\omega) - \frac{\varepsilon_1}{\varepsilon_b} - \coth(qD)\right)\left(2V_c(q)\Pi_{22}(q,\omega) - \frac{\varepsilon_2}{\varepsilon_b} - \coth(qD)\right) = \frac{1}{\sinh^2(qD)}, \qquad (13)$$

where $\Pi_{jj}(q,\,\omega)$ is the polarization function of the charge carriers on the first $j = 1$ or the second $j = 2$ graphene layer defined by Eq. (11) and $V_c(q) = 2\pi e^2/(\varepsilon_s q)$. We observe that in the limit $qD \gg 1$, Eq. (13) reduces to the dispersion equation for magnetoplasmons in a single graphene layer. Clearly, the sign of the charge carriers within the plasma has no effect on the dispersion equation of the magnetoplasmons of a two-component system [23]. If $\Pi_{11}(q,\omega) = \Pi_{22}(q,\omega) = \Pi(q,\omega)$, and $\varepsilon_1 = \varepsilon_2 = \varepsilon_b$, then we get from Eq. (13):

$$[(2V_c(q)\Pi(q,\,\omega) - 1)(1 - \exp(-2qD)) - (1 + \exp(-2qD))] = \pm 2\exp(-qD). \qquad (14)$$

The spectrum of magnetoplasmons in two spatially separated graphene layers when only the highest valence band is occupied and $D = r_B$ is presented on Fig. 2. The results of our numerical

calculations for bilayer graphene show there is a symmetric and antisymmetric magnetoplasmon mode for bilayer graphene. The originally degenerate magnetoplasmon mode in each layer of an isolated single graphene layer is shifted from their value by the inter-layer Coulomb interaction. The more energetic mode above the particle-hole mode region is the antisymmetric mode in which the charge density fluctuations in each layer oscillate out of phase with each other. The less energetic mode is the symmetric mode for which the density fluctuations are in phase. For a range of wave vectors the group velocity is negative due to magnetic field analogous to the magnetoplasmon modes in a single graphene layer. In the theory for charged particle energy loss, there is transfer of energy to the collective modes only when the charged particle velocity has the same value as the phase velocity of the collective mode. Plasma instabilities associated with these layered structures will be explored as a source of electro-magnetic radiation.

ACKNOWLEDGMENT

This work is supported by contract FA9453-07-C-0207 of AFRL.

REFERENCES

1. Novoselov, K. S., et al., *Science*, Vol. 306, 666, 2004.
2. Zhang, Y., et al., *Phys. Rev. Lett.*, Vol. 94, 176803, 2005.
3. Sarma, S. D., E. H. Hwang, and W.-K. Tse, *Phys. Rev. B*, Vol. 75, 121406(R), 2007.
4. Novoselov, K. S., et al., *Nature*, Vol. 438, 197, London, 2005.
5. Zhang, Y. B., et al., *Nature*, Vol. 438, 201, London, 2005.
6. Zhang, Y., et al., *Phys. Rev. Lett.*, Vol. 96, 136806, 2006.
7. Nomura, K. and A. H. MacDonald, *Phys. Rev. Lett.*, Vol. 96, 256602, 2006.
8. Mikitik, G. P. and Y. V. Sharlai, *Phys. Rev. Lett.*, Vol. 82, 2147, 1999.
9. Zheng, Y. and T. Ando, *Phys. Rev. B*, Vol. 65, 245420, 2002.
10. Hwang, E. H. and S. D. Sarma, *Phys. Rev. B*, Vol. 75, 205418, 2007.
11. Tőke, C., P. E. Lammert, V. H. Crespi, and J. K. Jain, *Phys. Rev. B*, Vol. 74, 235417, 2006.
12. Iyengar, A., J. Wang, H. A. Fertig, and L. Brey, *Phys. Rev. B*, Vol. 75, 125430, 2007.
13. Pines, D., *Elementary Excitations in Solids*, Benjamin, New York, 1963.
14. Sarma, S. D. and J. J. Quinn, *Phys. Rev. B*, Vol. 25, 7603, 1982.
15. Eguiluz, A., T. K. Lee, J. J. Quinn, and K. W. Chiu, *Phys. Rev. B*, Vol. 11, 4989, 1975.
16. Persson, B. N. J., *Solid State Commun.*, Vol. 52, 811, 1984.
17. Chiu, K. W. and J. J. Quinn, *Phys. Rev. B*, Vol. 9, 4724, 1974.
18. Berman, O. L., Y. E. Lozovik, and G. Gumbs, cond-mat/0706.0244.
19. Snoke, D. W., *Science*, Vol. 298, 1368, 2002.
20. Butov, L. V., *J. Phys.: Condens. Matter*, Vol. 16, R1577, 2004.
21. Timofeev, V. B. and A. V. Gorbunov, *J. Appl. Phys.*, Vol. 101, 081708, 2007.
22. Eisenstein, J. P. and A. H. MacDonald, *Nature*, Vol. 432, 691, 2004.
23. Sarma, S. D. and A. Madhukar, *Phys. Rev. B*, Vol. 23, 805, 1981.
24. Aĭzin, G. R. and G. Gumbs, *Phys. Rev. B*, Vol. 54, 2049, 1996.
25. Sarma, S. D. and E. H. Hwang, *Phys. Rev. Lett.*, Vol. 81, 4216, 1998.

Diagonalization of Translation Operators for Elastic Wave Equations

B. He[1] and W. C. Chew[1,2]

[1]Center for Computational Electromagnetics and Electromagnetics Laboratory
Department of ECE, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[2]University of Hong Kong, Hong Kong, SAR, China

Abstract— It shows that the dyadic Green's function of elastic wave equations can be factorized in terms of vector translators. The diagonalization of translation operators for elastic wave equations is presented. This result is important in applying the fast multipole method to elastic propagation in complex media.

1. INTRODUCTION

Elastic wave equations are difficult to solve because they are composed of longitudinal waves and transverse waves propagating at different speeds. Furthermore, for some boundary conditions, longitudinal waves produce transverse waves on reflection, and vice verse [1]. This difficulty has been overcome by solving the integral formulation of the elastic waves [2]. This integral formulation also facilitates the use of the fast multipole method (FMM) [3–5] to solve the problems of elastic wave scattering.

The crucial step in the FMM for solving the wave equations is the diagonalization of the translation operators [3–7]. Recently, the diagonalizations of the translation operators have been extended to vector and general tensor cases [8, 9]. It can be shown that the longitudinal waves and transverse waves are translated separately [9, 10]. Thus, the diagonalization of the translation operators of the longitudinal waves and transverse waves can be treated separately. However, since Hansen (vector) spherical harmonics and multipole fields are complete sets, it will be shown that the translation operators of the longitudinal waves and transverse waves can be diagonalized simultaneously.

In Section 2, we will introduce compact notations for vector Hansen spherical harmonics and multipole fields, which one can express the vector addition theorem for elastic waves in terms of. In Section 3, we will show that the dyadic Green's function of elastic wave equations can be factorized in terms of vector translators. Then, in Section 4, the derivation of diagonalization of translation operators for elastic wave equations is presented.

2. VECTOR ADDITION THEOREM FOR ELASTIC WAVES

The addition theorem transforms the wave functions from one coordinate system into another. One can refer to [11] for the references and details on addition theorem. In this Section, we shall write the vector addition theorem in terms of Hansen multipole fields in compact notations.

2.1. Hansen Spherical Harmonics and Multipole Fields

Hansen (vector) spherical harmonics can be defined as [1, 12]

$$\mathbf{P}_{JM}(\hat{k}) = -i\mathbf{e}_k Y_{lm}(\hat{k}), \mathbf{B}_{JM}(\hat{k}) = i\frac{\nabla_\alpha Y_{lm}(\hat{k})}{\sqrt{l(l+1)}}, \mathbf{C}_{JM}(\hat{k}) = i\mathbf{e}_k \times \mathbf{B}_{JM}(\hat{k}), \tag{1}$$

where ∇_α is defined by $\nabla_\alpha = \mathbf{e}_\theta \frac{\partial}{\partial\theta} + \frac{\mathbf{e}_\varphi}{sin\theta}\frac{\partial}{\partial\varphi}$. One can refer to [11] for the physical meanings of subscripts J, M, l, m in the viewpoint of angular momentum.

Let $\mathbb{V}_{JM}(\hat{k}) = [\mathbf{C}_{JM}(\hat{k})\ \mathbf{B}_{JM}(\hat{k})\ \mathbf{P}_{JM}(\hat{k})]$. It can be shown [9] that Hansen spherical harmonics satisfy the completeness

$$\sum_{JM} \mathbb{V}_{JM}(\hat{k})\mathbb{V}^T_{JM}(\hat{k}') = \sum_\nu \mathbf{e}^{[1]}_\nu \mathbf{e}^{*[1]}_\nu \delta(\hat{k} - \hat{k}'), \tag{2}$$

and the orthogonality

$$\int d\hat{k}\mathbb{V}^T_{J'M'}(\hat{k}) \cdot \mathbb{V}_{JM}(\hat{k}) = [I]_{3\times3}\delta_{JJ'}\delta_{MM'}, \tag{3}$$

where $[I]_{3\times 3}$ is a 3×3 identity matrix. In the above, $\mathbf{e}_\nu^{[1]}$ are the irreducible unit tensors for vector fields (eigenvectors of the spin 1) [13], which satisfy the orthogonality relationship

$$\mathbf{e}_\nu^{*[1]} \cdot \mathbf{e}_{\nu'}^{[1]} = \delta_{\nu\nu'}, \tag{4}$$

where the values of ν or ν' are integers from -1 to 1.

In order to describe the divergence and curl properties of the vector fields, Hansen (vector) multipole fields are introduced by [1]

$$\begin{aligned}
\mathbf{M}_{JM}(\mathbf{r},k) &= \frac{1}{\sqrt{J(J+1)}}\nabla\times\left[\mathbf{r}h_l^{(1)}(kr)Y_{lm}(\hat{r})\right],\\
\mathbf{N}_{JM}(\mathbf{r},k) &= \frac{1}{k}\nabla\times\mathbf{M}_{JM}(\mathbf{r},k),\\
\mathbf{L}_{JM}(\mathbf{r},k) &= \frac{1}{k}\nabla\left[h_l^{(1)}(kr)Y_{lm}(\hat{r})\right].
\end{aligned} \tag{5}$$

Hansen multipole fields $\mathbf{M}_{JM}(\mathbf{r},k)$ and $\mathbf{N}_{JM}(\mathbf{r},k)$ describe solenoidal waves and $\mathbf{L}_{JM}(\mathbf{r},k)$ describes longitudinal waves.

2.2. Vector Addition Theorem

Since we apply the vector addition theorem to elastic wave equations, which include the longitudinal and transverse waves, we shall write them in one setting.

Let $\mathbb{U}_{JM}(\mathbf{r},k_s,k_c) = [\mathbf{M}_{JM}(\mathbf{r},k_s)\ \mathbf{N}_{JM}(\mathbf{r},k_s)\ \mathbf{L}_{JM}(\mathbf{r},k_c)]$, and

$$[\Psi_{l''m''}(\mathbf{R},k_s,k_c)] = \begin{pmatrix} \Psi_{l''m''}(\mathbf{R},k_s) & 0 & 0\\ 0 & \Psi_{l''m''}(\mathbf{R},k_s) & 0\\ 0 & 0 & \Psi_{l''m''}(\mathbf{R},k_c)\end{pmatrix}, \tag{6}$$

where

$$\Psi_{l''m''}(\mathbf{R},k) = h_{l''}^{(1)}(kR)Y_{l''m''}(\hat{R}), \tag{7}$$

and k_s and k_c are the wave numbers of the transverse and longitudinal waves, respectively. Assuming $\mathbf{r} = \mathbf{r}' + \mathbf{R}$, the vector addition theorem [2] for elastic wave equations reads

$$\mathbb{U}_{JM}(\mathbf{r},k_s,k_c) = \begin{cases} \sum_{J'M'}\mathbb{U}_{J'M'}(\mathbf{r}',k_s,k_c)[\beta_{J'M',JM}(\mathbf{R},k_s,k_c)], & r'>R,\\ \sum_{J'M'}\Re_g\mathbb{U}_{J'M'}(\mathbf{r}',k_s,k_c)[\alpha_{J'M',JM}(\mathbf{R},k_s,k_c)], & r'<R. \end{cases}, \tag{8}$$

with the translators

$$[\alpha_{J'M',JM}(\mathbf{R},k_s,k_c)] = \sum_{l''m''} 4\pi i^{J'+l''-J}[G(J'M'|JM|l''m'')][\Psi_{l''m''}(\mathbf{R},k_s,k_c)], \tag{9}$$

$$[\beta_{J'M',JM}(\mathbf{R},k_s,k_c)] = \sum_{l''m''} 4\pi i^{J'+l''-J}[G(J'M'|JM|l''m'')]\Re_g[\Psi_{l''m''}(\mathbf{R},k_s,k_c)]. \tag{10}$$

The $\Re_g$ operator implies taking the regular part of the function, that is, replacing $h_l^{(1)}$ by j_l. $[G(J'M'|JM|l''m'')]$ is a 3×3 matrix defined by

$$[G(J'M'|JM|l''m'')] = \int d\hat{k}\mathbb{V}_{J'M'}^T(\hat{k})\cdot\mathbb{V}_{JM}(\hat{k})Y_{l''m''}^*(\hat{k}). \tag{11}$$

The matrix $[G(J'M'|JM|l''m'')]$ can be considered a generalization of the Gaunt coefficient with the elements

$$[G(J'M'|JM|l''m'')] = \begin{bmatrix} G_{11} & G_{12} & G_{13}\\ G_{21} & G_{22} & G_{23}\\ G_{31} & G_{32} & G_{33}\end{bmatrix} \tag{12}$$

[1] In this paper, the normalizations and phase factors for Hansen spherical harmonics and multipole fields are different from [9].

[2] In this paper, we limit our discussion to the spherical Hankel function $h_l^{(1)}$ and spherical Bessel function j_l. It is easy to extend vector addition theorem to include the spherical Hankel function $h_l^{(2)}$ and spherical Neuman function y_l [9].

One can refer to [9] for the evaluations of the elements of the matrix $[G(J'M'|JM|l''m'')]$. Since

$$\mathbf{P}^*_{J'M'}(\hat{k}) \cdot \mathbf{B}_{JM}(\hat{k}) = 0, \mathbf{P}^*_{J'M'}(\hat{k}) \cdot \mathbf{C}_{JM}(\hat{k}) = 0 \tag{13}$$

holds for *all* $J'M', JM$, we have

$$G_{13} = G_{31} = G_{23} = G_{32} = 0. \tag{14}$$

The property (14) means that the solenoidal waves ($\mathbf{M}_{JM}(\mathbf{r})$, $\mathbf{N}_{JM}(\mathbf{r})$) and longitudinal waves ($\mathbf{L}_{JM}(\mathbf{r})$) are translated separately [3], so the wave numbers can be different for the solenoidal waves and longitudinal waves in elastodynamics.

3. DYADIC GREEN'S FUNCTION OF ELASTIC WAVE EQUATIONS

In this section, we shall express the dyadic Green's function of elastic wave equations in terms of Hansen multipole fields. Then we factorize it in terms of the vector field translator.

Assuming $r < r'$, we can express the scalar Green's function $G(\mathbf{r}, \mathbf{r}', k)$ in terms of spherical harmonics [14, 15]

$$G(\mathbf{r}, \mathbf{r}', k) = ik \sum_{lm} j_l(kr) Y_{lm}(\hat{r}) h_l^{(1)}(kr') Y^*_{lm}(\hat{r}'). \tag{15}$$

Coupling $\mathbf{e}^{[1]}$ and $\mathbf{e}^{*[1]}$ to both sides of (15) gives tensorial Green's function in terms of Hansen multipole fields as [9]

$$\begin{aligned} \bar{\mathbf{G}}(\mathbf{r}, \mathbf{r}', k) \; = \; & ik \sum_{JM} \left[\Re_g \mathbf{M}_{JM}(\mathbf{r}, k) \mathbf{M}^*_{JM}(\mathbf{r}', k) + \Re_g \mathbf{N}_{JM}(\mathbf{r}, k) \mathbf{N}^*_{JM}(\mathbf{r}', k) \right. \\ & \left. + \Re_g \mathbf{L}_{JM}(\mathbf{r}, k) \mathbf{L}^*_{JM}(\mathbf{r}', k) \right]. \end{aligned} \tag{16}$$

The dyadic Green's function of elastic wave equations can be defined as

$$\tilde{\mathbf{G}}(\mathbf{r}, \mathbf{r}', k_s, k_c) = \frac{1}{\mu} \left[\bar{\mathbf{G}}(\mathbf{r}, \mathbf{r}', k_s) - \frac{\nabla' \nabla}{k^2} G(\mathbf{r}, \mathbf{r}', k_s) \right] + \frac{1}{2\mu + \lambda} \frac{\nabla' \nabla}{k^2} G(\mathbf{r}, \mathbf{r}', k_c), \tag{17}$$

where μ and λ are the Lamé constants. Substituting (15) and (16) into (17) and using the definition of $\mathbf{L}_{JM}(\mathbf{r}, k)$ (5), we have

$$\begin{aligned} \tilde{\mathbf{G}}(\mathbf{r}, \mathbf{r}', k_s, k_c) \; = \; & \sum_{JM} \left\{ \frac{ik_s}{\mu} \left[\Re_g \mathbf{M}_{JM}(\mathbf{r}, k_s) \mathbf{M}^*_{JM}(\mathbf{r}', k_s) + \Re_g \mathbf{N}_{JM}(\mathbf{r}, k_s) \mathbf{N}^*_{JM}(\mathbf{r}', k_s) \right] \right. \\ & \left. + \frac{ik_c}{2\mu + \lambda} \Re_g \mathbf{L}_{JM}(\mathbf{r}, k_c) \mathbf{L}^*_{JM}(\mathbf{r}', k_c) \right\}. \end{aligned} \tag{18}$$

Let

$$[\chi] = \begin{pmatrix} \frac{ik_s}{\mu} & 0 & 0 \\ 0 & \frac{ik_s}{\mu} & 0 \\ 0 & 0 & \frac{ik_c}{2\mu + \lambda} \end{pmatrix}. \tag{19}$$

Equation (18) can be written in a compact form

$$\tilde{\mathbf{G}}(\mathbf{r}, \mathbf{r}', k_s, k_c) \; = \; \sum_{JM} \Re_g \mathbb{U}_{JM}(\mathbf{r}, k_s, k_c) [\chi] \mathbb{U}^T_{JM}(\mathbf{r}', k_s, k_c). \tag{20}$$

Assuming $\mathbf{r}' = \mathbf{r}'' + \mathbf{R}, r'' < R$, and applying (8), we can factorize the dyadic Green's function (20) in terms of the vector field translator

$$\begin{aligned} & \tilde{\mathbf{G}}(\mathbf{r}, \mathbf{r}', k_s, k_c) \\ = \; & \sum_{JM} \Re_g \mathbb{U}_{JM}(\mathbf{r}, k_s, k_c) [\chi] \sum_{J'M'} [\alpha_{J'M',JM}(\mathbf{R}, k_s, k_c)]^T \Re_g \mathbb{U}^T_{J'M'}(\mathbf{r}'', k_s, k_c). \end{aligned} \tag{21}$$

[3] This fact has been discussed in [10], but here our discussion based on the generalization of the Gaunt coefficient is more rigorous and fundamental.

4. DIAGONALIZATION OF THE VECTOR ADDITION THEOREM FOR ELASTIC WAVE EQUATIONS

The diagonalization of the translation operators is the crucial step in the fast multipole method for solving the wave equations [3–5]. We shall derive the diagonal form of the translation operators of the longitudinal waves and transverse waves simultaneously.

Letting

$$[e(\mathbf{R},k_s,k_c)]=\begin{pmatrix} e^{ik_s\hat{k}\cdot\mathbf{R}} & 0 & 0\\ 0 & e^{ik_s\hat{k}\cdot\mathbf{R}} & 0\\ 0 & 0 & e^{ik_c\hat{k}\cdot\mathbf{R}}\end{pmatrix}, \tag{22}$$

and using the Rayleigh plane-wave expansion formula

$$e^{i\mathbf{k}\cdot\mathbf{r}}=\sum_{lm}4\pi i^l Y_{lm}(\hat{r})j_l(kr)Y^*_{lm}(\hat{k}), \tag{23}$$

one can write the translator as

$$[\beta_{J'M',JM}(\mathbf{R},k_s,k_c)]=i^{J'-J}\int d\hat{k}\mathbb{V}^T_{J'M'}(\hat{k})\cdot\mathbb{V}_{JM}(\hat{k})[e(\mathbf{R},k_s,k_c)]. \tag{24}$$

In a fast multipole algorithm, the translator can be factorized by

$$\begin{aligned}&[\alpha_{JM,J'M'}(\mathbf{R}_{ij},k_s,k_c)]=\\ &\sum_{J_1M_1,J_2M_2}[\beta_{JM,J_1M_1}(\mathbf{R}_{ip},k_s,k_c)][\alpha_{J_1M_1,J_2M_2}(\mathbf{R}_{pp'},k_s,k_c)][\beta_{J_2M_2,J'M'}(\mathbf{R}_{p'j},k_s,k_c)]\end{aligned} \tag{25}$$

Using (24) and (9), we compute (25) as

$$\begin{aligned}[\alpha_{JM,J'M'}(\mathbf{R}_{ij},k_s,k_c)] \;=\; &\sum_{J_1M_1,J_2M_2}i^{J-J_1}\int d\hat{k}\mathbb{V}^T_{JM}(\hat{k})\cdot\mathbb{V}_{J_1M_1}(\hat{k})[e(\mathbf{R}_{ip},k_s,k_c)]\\ &\times\sum_{l''m''}4\pi i^{J_1+l''-J_2}[\Psi_{l''m''}(\mathbf{R}_{pp'},k_s,k_c)]\int d\hat{k}'\mathbb{V}^T_{J_1M_1}(\hat{k}')\cdot\mathbb{V}_{J_2M_2}(\hat{k}')Y^*_{l''m''}(\hat{k}')\\ &\times i^{J_2-J'}\int d\hat{k}''\mathbb{V}^T_{J_2M_2}(\hat{k}'')\cdot\mathbb{V}_{J'M'}(\hat{k}'')[e(\mathbf{R}_{p'j},k_s,k_c)].\end{aligned} \tag{26}$$

Using (2) and (4), we have the following formula

$$\sum_{J_1M_1}\mathbb{V}_{J_1M_1}(\hat{k})\mathbb{V}^T_{J_1M_1}(\hat{k}')\cdot\sum_{J_2M_2}\mathbb{V}_{J_2M_2}(\hat{k}')\mathbb{V}^T_{J_2M_2}(\hat{k}'')=\sum_{\nu}\mathbf{e}^{[1]}_{\nu}\mathbf{e}^{*[1]}_{\nu}\delta(\hat{k}-\hat{k}')(\hat{k}'-\hat{k}''). \tag{27}$$

Plugging (27) into (26), we arrives at the diagonalization of the translators

$$\begin{aligned}[\alpha_{JM,J'M'}(\mathbf{R}_{ij},k_s,k_c)] \;=\; &\int d\hat{k}\sum_{\nu}i^J[e(\mathbf{R}_{ip},k_s,k_c)][\tilde{\alpha}(\mathbf{R}_{pp'},k_s,k_c)]\mathbb{V}^T_{JM}(\hat{k})\cdot\mathbf{e}^{[1]}_{\nu}\\ &\mathbf{e}^{*[1]}_{\nu}\cdot\mathbb{V}_{J'M'}(\hat{k})i^{-J'}[e(\mathbf{R}_{p'j},k_s,k_c)]\end{aligned} \tag{28}$$

with

$$[\tilde{\alpha}(\mathbf{R}_{pp'},k_s,k_c)]=\sum_{l''m''}4\pi i^{l''}[\Psi_{l''m''}(\mathbf{R}_{pp'},k_s,k_c)]Y^*_{l''m''}(\hat{k}). \tag{29}$$

Note that $[\alpha(\mathbf{R}_{pp'},k_s,k_c)]$ is a diagonal 3×3 matrix, whose nonzero entries are same as that in the diagonalization of the scalar addition theorem.

5. CONCLUSION

Compact notations for vector Hansen spherical harmonics and multipole fields are introduced to simplify the expression of the vector addition theorem. We have shown that the dyadic Green's function of elastic wave equations can be factorized in terms of vector translators. The diagonalization of translation operators for elastic wave equations is obtained.

There has been growing interest in elastic wave propagation in complex media, e.g., periodic composite media [16]. The diagonalization of the translation operators facilitates in applying FMM to study elastic propagation in complex media, in particular for some cases where large number of unknowns are needed.

ACKNOWLEDGMENT

Support by the Y.T. Lo Endowed Chair grant and AFOSR MURI grant FA9550-04-1-0326WC are gratefully acknowledged.

REFERENCES

1. Morse, P. M. and H. Feshbach, *Methods of the Theoretical Physics*, McGRAW-HILL Book Company, Inc., New York, 1953.
2. Pao, Y. and V. Varatharajulu, "Huygens' principle, radiation conditions, and the integral formulas for the scattering of elastic waves," *J. Acoust. Soc. Am.*, Vol. 59, No. 6, 1361–1371, June 1976.
3. Rokhlin, V., "Rapid solution of integral equations of scattering theory in two dimensions," *J. Comput. Phys.*, Vol. 86, No. 2, 414–439, 1990.
4. Coifman, R., V. Rokhlin, and S. Wandzura, "The fast multipole method for the wave equation: A pedestrian prescription," *IEEE Ant. Propag. Mag.*, Vol. 35, No. 3, 7–12, 1993.
5. Chew, W. C., J. M. Jin, E. Michielssen, and J. M. Song, *Fast and Efficient Algorithms in Computational Electromagnetics*, Artech House, Boston, 2001.
6. Rokhlin, V., "Diagonal forms of translation operators for the Helmholtz equation in three dimensions," *Appl. Comp. Harmon. Anal.*, Vol. 1, 82–93, 1993.
7. Epton, M. and B. Dembart, "Multipole translation theory for three-dimensional Laplace and Helmholtz equations," *SIAM J. Sci. Comput.*, Vol. 16, No. 4, 865–897, 1995.
8. Chew, W. C., "Vector addition theorem and its diagonalization," *Comm. Comput. Phys.*, Vol. 3, 330–341, 2008.
9. He, B. and W. C. Chew, "Diagonalizations of vector and tensor addition theorems," Submitted to *Comm. Comput. Phys.*.
10. Boström, A., "Multiple scattering of elastic waves by bounded obstacles," *J. Acoust. Soc. Am.*, Vol. 67, No. 2, 399–413, February 1980.
11. He, B. and W. C. Chew, "Addition theorem," *Modeling and Computations in Electromagnetics*, 203–226, Ammari, Habib (Ed.), Springer, 2007.
12. Dassios, G. and Z. Rigou, "Elastic Herglotz functions," *SIAM J. Appl. Math.*, Vol. 55, No. 5, 1345–1361, October 1995.
13. Edmonds, A. R., *Angular Momentum in Quantum Mechanics*, Princeton Univ. Press, Princeton, 1960.
14. Gonis, A. and W. H. Butler, *Multiple Scattering in Solids*, Springer, New York, 2000.
15. Biedenharn, L. C. and J. D. Louck, *Angular Momentum in Quantum Physics*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1981.
16. Liu, Z., C. T. Chan, P. Sheng, A. L. Goertzen, and J. H. Page, "Elastic wave scattering by periodic structures of spherical objects: Theory and experiment," *Phys. Rev. B*, Vol. 62, 2446–2457, July 2000.

Finite Element Method Simulation of Photoinductive Imaging for Cracks

Cheng-Chi Tai and Yen-Lin Pan

Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan

Abstract— In this paper, the numerical simulations of photoinductive imaging (PI) method have been performed using the finite element method (FEM) with the 2D transient to characterize corner cracks at the edge of a bolt hole. The PI imaging results have higher spatial resolution in the area of the defect in 2D models as compared with the conventional eddy current (EC) images. The FEM simulation results of 0.5-mm rectangular defects are showed and analyzed. The dependencies of PI signals on EC frequencies and temperature of the thermal spot are also examined. The results demonstrate that the PI method is applicable to examine the geometric shape of corner cracks.

1. INTRODUCTION

Crack detection and sizing is a critical issue in quantitative nondestructive evaluation (NDE). The ultrasonic method is used predominantly to detect subsurface discontinuities, while the EC method is effective for surface cracks. However, one of the main disadvantages of conventional eddy current method is the low spatial resolution, which is constrained by the size of eddy current probes. The PI method is a hybrid NDE technique that combines eddy current and laser-based thermal wave methods. The use of a focused laser beam provides the method with a microscopic resolution while using eddy current pickup sensors.

Moulder et al. [1] showed that this new technique dramatically increased image resolutions, and could be used to calibrate and characterize eddy current probes [2–4]. The method experimentally showed the high-resolution capability inherent in this technique by adapting a photoinductive sensor developed for a fiber optic probe to an existing photoacoustic microscope [2]. The same method will work equally well to characterize cracks on thick metals. Determining the crack dimensions is the interesting detection of a corner crack on the surface surrounding a bolt hole, such as depth and length [5].

In this article, we use the finite element method to simulate the photoinductive imaging (PI) technique for bolt-hole cracks inspection. Based on the simulation results, we also discuss the effects of EC frequencies and temperature of the thermal spot, and compare the PI results with EC images for a 0.5-mm triangular notch.

2. THE PHOTOINDUCTIVE IMAGING METHOD

Photoinductive mapping of eddy current fields interacting with cracks is a newly devised technique that is similar to photothermal (PT) imaging. The physical principles underlying it are illustrated in Fig. 1, which shows the coil of an eddy current probe carrying a current placed in close proximity to the specimen surface. A focused laser beam generates a localized hot spot on the specimen surface from above. The temperature fluctuation causes a local change in the electrical conductivity, which in turn induces a change in the impedance of the eddy current probe. The electrical conductivity of specimen is given by the expression:

$$\sigma = \frac{1}{[\rho_0(1 + \alpha(T - T_0))]} \tag{1}$$

where ρ_0 is the resistivity at temperature T_0, and α is the temperature coefficient of the resistivity. T_0 is the temperature 293 K, and T is the actual temperature in the specimen sub-domain.

The PI effect can be calculated as follows. The dependent variable in this application mode is the azimuthal component of the magnetic vector potential, $\mathbf{A}$, which conforms to the following relation:

$$(j\omega\sigma - \omega^2\varepsilon)\mathbf{A} + \nabla \times (\mu^{-1}\nabla \times \mathbf{A}) = \frac{\sigma V_{loop}}{L}, \tag{2}$$

where ω denotes the angular frequency, σ the conductivity, μ the permeability, ε the permittivity, L the length, and V_{loop} the voltage applied to the coil. The conductivity outside the coil is zero. According to the constitutive relation ($C.R.$), the current density ($\mathbf{J}^e$) can be calculated as follows.

$$\mathbf{J}^e = \sigma\mathbf{E} = -\sigma\left(\nabla V + \frac{\partial \mathbf{A}}{\partial t}\right), \tag{3}$$

where $\mathbf{E}$ is the electric field intensity. The electric potential (V) is obtained from Faraday's law. The defining equation for the magnetic vector potential $\mathbf{A}$ is a direct consequence of the magnetic Gauss'law. The induced current (I) in the coil is calculated by the integration of current density in the cross-sectional area (s) of the coil:

$$\int_S \mathbf{J}^e \cdot \mathrm{d}s = I. \tag{4}$$

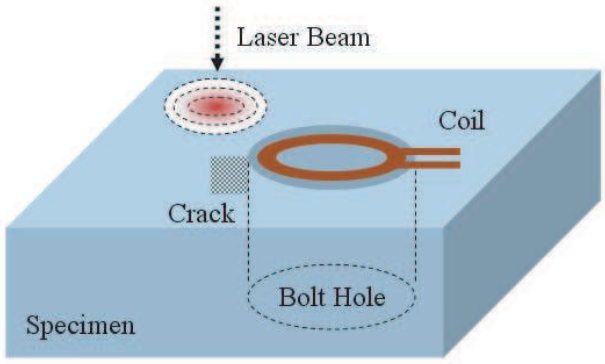


Figure 1: Inspection geometry of the photoinductive field measurement technique.

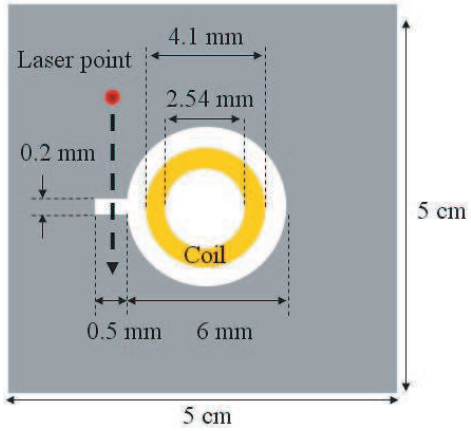


Figure 2: The actual size geometry of the 2D model for PI imaging method. (the bold dotted line is the laser point path).

3. SPECIMEN AND SIMULATION

The specimens used for this study are titanium blocks (Ti-6Al-4V) with 6-mm bolt holes. The notch is 0.5 mm in both length and depth and 0.2 mm in width. The coil probe (inner diameter = 2.54 mm, outer diameter = 4.1 mm, and length = 0.76 mm) was inserted in the bolt hole with the coil firmly positioned flush with the edge of the bolt hole. The actual size geometry of the 2D model for PI imaging method is shown in Fig. 2. The probe was operated at a range of frequencies

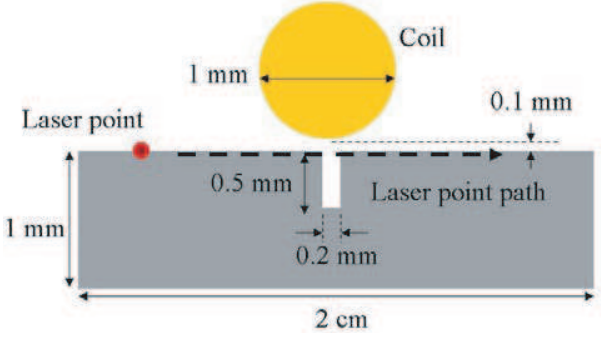


Figure 3: The simplified 2D model for comparing the characteristics of PI imaging method and EC imaging method.

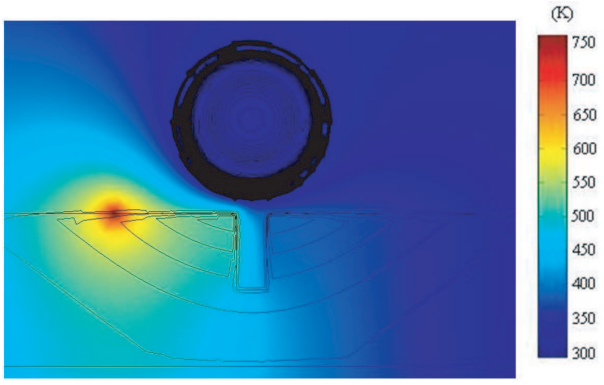


Figure 4: The diffusion of heat from laser beam and eddy current density distributions. EC frequency, 600 kHz; laser temperature, 500 °C (773 K).

from 200 kHz to 1 MHz. The temperature produced by laser beam was at a range of from 100 °C to 500 °C. The simulations were implemented using the COMSOL Multiphysics^{TM} software. In this work, we use the simplified 2D model for comparing the characteristics of PI imaging method and EC imaging method (Fig. 3). We designed the coil that with a height equal to or greater than the depth of the notch. This is to make sure that the eddy currents surround the whole notch and so that the depth information can be revealed. The distance between the specimen and the coil (diameter = 1 mm, length = 0.76 mm) was 0.1 mm. In the case of EC scan that without the laser point, the coil is moving along the x-axis and cross the notch on specimen. In the PI scan, the coil is fixed right on the center of the notch and the laser is moving in the same direction as the previous case. The uniform scan plan with closely spaced scan lines so that flaw orientation and scan spacing would not affect the outcome was assumed. To calculate the impedance edance ($\Delta Z = V_{loop}/I$) of the coil in the simulations, the total induced eddy current of the coil can be obtained by carrying out sub-domain integration of the total current density for the cross-section of the excited coil.

4. RESULTS AND DISCUSSION

The simulation results using the PI imaging method and the conventional EC imaging method are presented and compared in this section. The simulation results of the signals interaction with various temperature and frequency will be first presented. The effects of eddy current were compared by varying the coil excitation frequency from 200 kHz to 1 MHz and the laser beam temperature from 100 °C to 500 °C. The diffusion of heat from laser beam and the eddy current density distribution around the crack are shown in Fig. 4. The temperature fluctuation causes a local change in the electrical conductivity of the specimen and the current density of the specimen. The lines indicate the contour of induced current density on the coil and the specimen. Figs. 5 and 6 show the signal of coil impedances with EC method and PI method, respectively. The center point of the rectangular notch is 0 mm in x-axis, as shown in Figs. 5 and 6. Fig. 5 is the EC image signals of a 0.5-mm rectangular notch at 600-kHz EC frequency, without laser beam. Fig. 6 is the PI image signals of a 0.5-mm rectangular notch at the same EC frequency and 300 °C laser beam temperature. Because the length of the flaw is less than the diameter of the probe, the flaw scan produces a double-peaked response [6]. As shown in the Figs. 5 and 6, both figures conform to this phenomenon. Comparison of flaw impedance measured with two detection method for rectangular notch, the resolution of PI signal is higher than the EC signal. There is a higher sharp edge in PI signal than in EC signal.

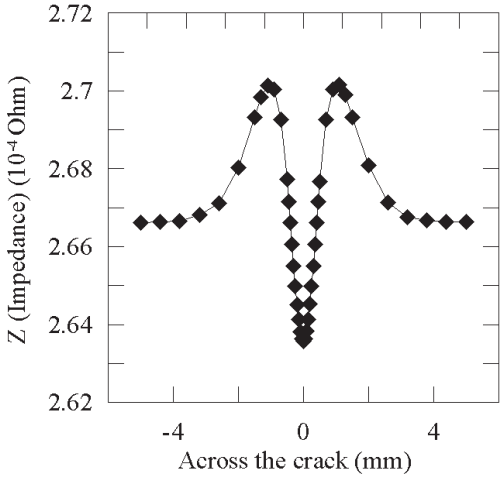


Figure 5: The flaw impedances of EC signal for a 0.5-mm rectangular notch in Ti-6A1-4V. EC frequency, 600 kHz.

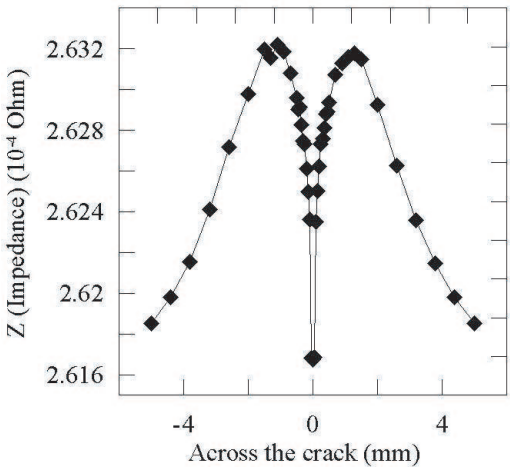


Figure 6: The flaw impedances of PI signal for a 0.5-mm rectangular notch in Ti-6A1-4V. EC frequency, 600 kHz; laser temperature, 300 °C.

The effects of EC frequency on the PI imaging signals and EC imaging signals for transverse scans across a 0.5-mm long and 0.2-mm wide notch are shown in Figs. 7 and 8, respectively. In order to clearly exhibit the crack's effect, the impedance difference between signals with crack and without crack is reported. As shown in Fig. 7, the eddy currents around the crack are more uniform at lower frequencies. But higher EC frequencies generate a stronger PI signal. Fig. 8 illustrates

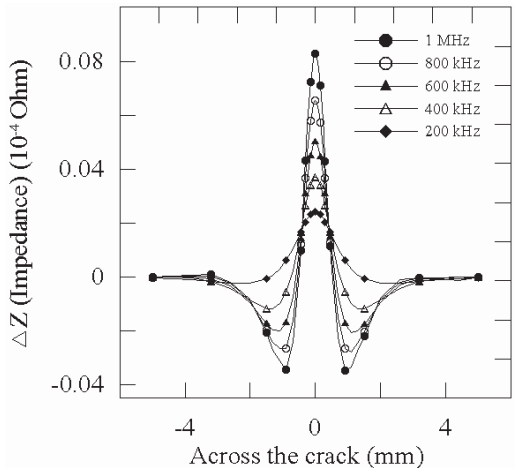


Figure 7: Images with different EC frequency. (EC scan without the laser beam).

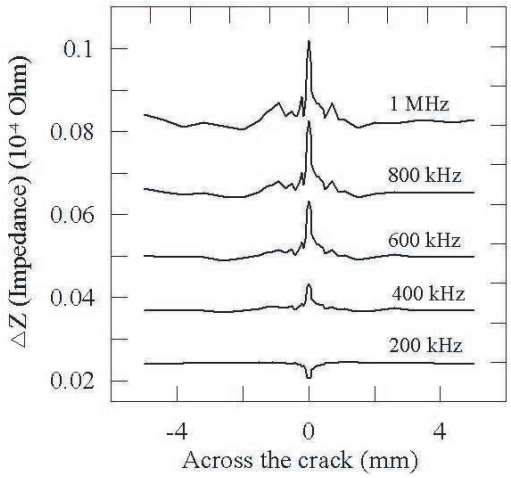


Figure 8: PI imaging signal with different frequency. Laser temperature, 300 °C.

the signal amplitude is increased when higher eddy current frequencies are applied, and therefore the better crack images are obtained by increasing the eddy current frequency. Furthermore, the impedance difference for 200-kHz case is reversed on the notch area. That may due to deeper skin depth and lower current density on the surface of specimens. For the rectangular notches in this titanium alloy, eddy current frequencies above 200 kHz are more suitable for imaging the cracks.

Figure 9 shows that the peak amplitude of PI imaging signals varies with laser point temperature for transverse scans across a 0.5 mm long and 0.2 mm wide notch. There is the same process to clearly display the crack's shape. When lower laser beam temperatures are applied, the peak amplitude of signal is decreased. Because reducing the temperature will generates higher current density and deeper penetration on the surface of this specimen. That makes the eddy currents around the crack are more uniform at lower laser temperature.

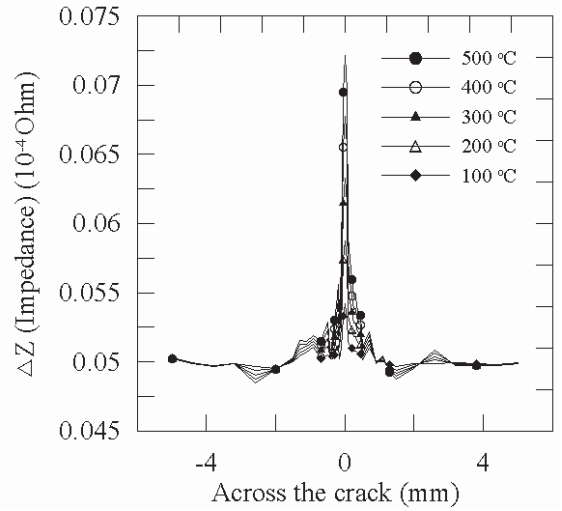


Figure 9: PI imaging signal with different laser point temperature. EC frequency, 600 kHz.

5. CONCLUSIONS

The FEM simulation results demonstrate the feasibility of photoinductive imaging method when applied to the detection of corner cracks. The EC frequency and laser beam temperatures affect PI signal amplitude and resolution. The PI images have higher spatial resolution in the area of the defect in 2D models when compared with the conventional EC images. The higher PI signal amplitude can be obtained by increasing the laser beam temperature.

ACKNOWLEDGMENT

This research was supported by the grant from National Science Council, Taiwan (NSC 96-2628-E-006-256-MY3). Also, this work made use of Shared Facilities supported by the Program of Top 100 Universities Advancement, Ministry of Education, Taiwan.

REFERENCES

1. Moulder, J. C., N. Nakagawa, K. S. No, Y. P. Lee, and J. F. McClelland, "Photoinductive imaging: a new NDE technique," *Review of Progress in Quantitative NDE*, edited by D. O. Thompson and D. E. Chimenti, Plenum Press, Vol. 8A, 599, New York, 1989.
2. Moulder, J. C. and N. Nakagawa, "Characterizing the performance of eddy current probes using photoinductive field-mapping," *Research in Nondestructive Evaluation*, Vol. 4, 221–236, 1992.
3. Hughes, M. S., J. C. Moulder, M. W. Kubovich, and B. A. Auld, "Mapping eddy current probe fields using the photoinductive effect," *NDT & E International*, Vol. 28, 251, 1995.
4. Nakagawa, N. and J. C. Moulder, "Eddy current probe calibration via the photoinductive effect," *NDT and E International*, Vol. 28, 245–246, 1995.
5. Tai, C.-C. and J. C. Moulder, "Bolt-hole corner crack inspection using the photoinductive imaging method," *Journal of Nondestructive Evaluation*, Vol. 19, 81–93, 2000.
6. Auld, B. A., S. R. Jefferies, and J. C. Moulder, "Eddy-current signal analysis and inversion for semielliptical surface cracks," *Journal of Nondestructive Evaluation*, Vol. 7, 79–94, 1988.

Staggered-grid Pseudospectral Time Domain (PSTD) Method Using Real Fourier Transform for 2.5D Electromagnetic Wave Propagation

Lanbo Liu[1,2], **Benjamin Barrowes**[2], **and Zhao Zhao**[1]

[1]Department of Civil & Environmental Engineering, University of Connecticut, USA
[2]Cold Regions Research and Engineering Laboratory, USA

Abstract— The staggered grid algorithm was originally invented for achieving better stability and efficiency in the finite difference time domain (FDTD) method for modeling the electromagnetic wave propagation. Seismologists extended the staggered grid approach to the pseudospectral time domain (PSTD) scheme to model seismic wave propagation. However, no detailed formulation of the staggered grid approach for electromagnetic (EM) simulations has been explicitly discussed. We present the staggered grid PSTD for EM simulations by shifting the spatial derivatives halfway between 2 adjacent nodes and making the Nyquist wave number a non-zero pure real value of $-\pi/\Delta x$. By doing this, the Nyquist information of the original spatial function is preserved, and the differentiation operator is more stable. In the Fourier domain, adding trigonometric factors in the classic Fourier coefficients is equivalent to the staggered grid approach in the original space domain. A staggered grid PSTD algorithm makes the time marching more stable, and numerical dispersion is suppressed for models with sharp contrasts in material properties. In this paper, we have applied the staggered grid PSTD method to 2.5D electromagnetic wave propagation simulations using the real Fourier transform. We discuss this method and apply it to model a Ground Penetrating Radar (GPR) system.

1. INTRODUCTION

The staggered grid scheme was first proposed by Yee [1] for achieving better accuracy and higher stability in modeling electromagnetic (EM) wave propagation using the finite difference time domain (FDTD) method. It is a robust and well studied algorithm and has been widely used in the EM research community. Meanwhile, in the pseudospectral time domain (PSTD) method, which theoretically only needs two grids per wavelength [2] for EM wave simulations, the field values and material properties [2] are collocated at the same nodes (non-staggered) for simplicity. The great success of the staggered grid scheme in FDTD made a few researchers [3–8] turn their attention to trying to apply the staggered grid scheme for the PSTD method as applied to seismic wave propagations. It has been demonstrated that using a staggered grid in the PSTD scheme made the time marching more stable and suppressed numerical dispersion for models with large contrast in material properties. To the best of the author's knowledge, there is no detailed discussion on the application of the staggered grid PSTD to model the electromagnetic wave propagations in the published literature. This paper discusses staggered grid PSTD for EM for the 2.5D problem [9, 10] and the unsplit stretched coordinate perfectly matched layer (PML) absorption boundary conditions [11].

The staggered grid scheme is realized by shifting the spatial derivatives of the electric or magnetic field halfway between 2 nodes and makes the Nyquist wave number a non-zero pure real value of $-\pi/\Delta x$ [7]. By doing this the Nyquist information of the original spatial function is preserved and the differentiation operator is more stable. In the Fourier domain, adding trigonometric factors in the classic Fourier coefficients is equivalent to the staggered grid approach in the original space domain [7] with minimal increase of computational cost.

The staggered grid PSTD algorithm development presented in this paper was associated with the use of the unsplit stretched coordinate PML [11] which is based on the PML scheme [12]. The 3D spreading decay of the spherical wave excited by a point source was precisely preserved for the case of medium Physical properties is invariant in one of the horizontal directions, by adapting a 2.5D simulation scenario.

Finally, the efficacy of the staggered grid PSTD was illustrated by modeling a subsurface discrimination system like GPR using simple geometries.

2. THE 2.5D MAXWELL'S EQUATIONS WITH A PML FORMULATION

The electromagnetic (EM) wave propagation in isotropic, inhomogeneous media with the current density source $\mathbf{J}$ is governed by the Maxwell equations

$$\begin{aligned} \nabla \times \mathbf{H} &= \varepsilon \frac{\partial \mathbf{E}}{\partial t} + \sigma \mathbf{E} + \mathbf{J}_s \\ \nabla \times \mathbf{E} &= -\mu \frac{\partial \mathbf{H}}{\partial t} \end{aligned} \tag{1}$$

where the dielectric permittivity $\varepsilon(\mathbf{r})$, magnetic permeability $\mu(\mathbf{r})$, and electric conductivity $\sigma(\mathbf{r})$ are all spatially dependent. $\mathbf{J}_s$ is the source current. With the introduction of the stretched coordinate perfectly matched layer (PML) technique as the absorption boundary condition [12] used to truncate the model domain is

$$\widehat{\nabla} = \frac{1}{e_x}\frac{\partial}{\partial x}\hat{x} + \frac{1}{e_y}\frac{\partial}{\partial y}\hat{y} + \frac{1}{e_z}\frac{\partial}{\partial z}\hat{z}$$

and

$$e_\eta = a_\eta + i\frac{\omega_\eta}{\omega}, \qquad (\eta = x,\, y,\, z)$$

Thus, the full 3D Maxwell's Equations (1) in a Cartesian coordinate system can be written as

$$\begin{aligned} a_\eta \varepsilon \frac{\partial \mathbf{E}}{\partial t} + (a_\eta \sigma + \omega_\eta \varepsilon)\mathbf{E} + \omega_\eta \sigma \int_{-\infty}^{t} \mathbf{E} dt &= \frac{\partial}{\partial \eta}(\hat{\eta} \times \mathbf{H}) - \mathbf{J} \\ a_\eta \mu \frac{\partial \mathbf{H}}{\partial t} + \mu \omega_\eta \mathbf{H} &= -\frac{\partial}{\partial \eta}(\hat{\eta} \times \mathbf{E}) \end{aligned} \tag{2}$$

In the following analysis we adapt the un-split stretched coordinate PML approach [11] as the absorption boundary condition. For a 2.5D problem, i.e., the 3D spherical wave generated by a point source propagating through a 2D medium, proper mathematical treatment of the numerical simulation problem leads to significant reduction in demands of computational resources such as memory and computational time. In the following discussion and development, we assume the medium is inhomogeneous in the xz-plane and homogeneous in one of the horizontal directions, say the y-direction; i.e., the dielectric permittivity $\varepsilon(\mathbf{r})$, magnetic permeability $\mu(\mathbf{r})$, and electric conductivity $\sigma(\mathbf{r})$ are functions of only x and z.

$$\varepsilon = \varepsilon(\mathbf{r}) = \varepsilon(x,\, z), \quad \mu = \mu(\mathbf{r}) = \mu(x,\, z), \quad \sigma = \sigma(\mathbf{r}) = \sigma(x,\, z)$$

Furthermore, we assume that the imposed point current sources are all in the $y = 0$ plane. Depending on the source polarization, the current sources and the corresponding field components may be an even or odd function about the $y = 0$ plane [9]. For example, the total electric field E can be expressed as $\mathbf{E} = \mathbf{E}_e + \mathbf{E}_o$, with $\mathbf{E}_e$ the even function and $\mathbf{E}_o$ the odd function, which are defined as

$$\begin{aligned} \widehat{\mathbf{E}}_e(x,\, k_y,\, z,\, t) &= \int_{-\infty}^{\infty} \mathbf{E}(x,\, y,\, z,\, t)\cos(k_y y) dy \\ \widehat{\mathbf{E}}_o(x,\, k_y,\, z,\, t) &= \int_{-\infty}^{\infty} \mathbf{E}(x,\, y,\, z,\, t)\sin(k_y y) dy \\ \widehat{\mathbf{H}}_e(x,\, k_y,\, z,\, t) &= \int_{-\infty}^{\infty} \mathbf{H}(x,\, y,\, z,\, t)\cos(k_y y) dy \\ \widehat{\mathbf{H}}_o(x,\, k_y,\, z,\, t) &= \int_{-\infty}^{\infty} \mathbf{H}(x,\, y,\, z,\, t)\sin(k_y y) dy \end{aligned} \tag{3}$$

using the cosine and sine functions in the real Fourier transform. Here k_y is the Fourier integration parameter equivalent to the wave number in y-direction. For simplicity in the following discussion we focus on the case of the electromagnetic fields generated by a current dipole source in y-direction and in the $y = 0$ plane ($J_y = J_{ye}$). It can be shown that this current source excites only the field

components of (E_{xo}, E_{ye}, E_{zo}, H_{xe}, H_{yo}, H_{ze}). Using the unsplit stretched coordinate PML [11], the 2.5D Maxwell equations in components form are

$$\begin{aligned}
&\varepsilon[a_x a_z \frac{\partial \widehat{E}_{xo}}{\partial t} + (a_z\omega_x + a_x\omega_z)\widehat{E}_{xo} + \omega_z\omega_x \widehat{E}^*_{xo}] + \sigma(a_x \widehat{E}_{xo} + \omega_x \widehat{E}^*_{xo}) \\
= \;& a_z a_x k_y \widehat{H}_{ze} + (a_z\omega_x + a_x\omega_z)k_y \widehat{H}^*_{ze} - a_x \frac{\partial \widehat{H}_{yo}}{\partial z} \\
&\varepsilon[a_x a_z \frac{\partial \widehat{E}_{ye}}{\partial t} + (a_z\omega_x + a_x\omega_z)\widehat{E}_{ye} + \omega_z\omega_x \widehat{E}^*_{ye}] + \sigma \widehat{E}_{ye} \\
= \;& a_x \frac{\partial \widehat{H}_{xe}}{\partial z} + \omega_x \frac{\partial \widehat{H}^*_{xe}}{\partial z} - a_z \frac{\partial \widehat{H}_{ze}}{\partial x} - \omega_z \frac{\partial \widehat{H}^*_{ze}}{\partial x} - J_{ye} \\
&\varepsilon[a_x a_z \frac{\partial \widehat{E}_{zo}}{\partial t} + (a_z\omega_x + a_x\omega_z)\widehat{E}_{zo} + \omega_z\omega_x \widehat{E}^*_{zo}] + \sigma(a_z \widehat{E}_{zo} + \omega_z \widehat{E}^*_{zo}) \\
= \;& a_z \frac{\partial \widehat{H}_{yo}}{\partial x} + \omega_z \frac{\partial \widehat{H}^*_{yo}}{\partial x} - a_z a_x k_y \widehat{H}_{xe} - (a_z\omega_x + a_x\omega_z)k_y \widehat{H}^*_{xe}
\end{aligned} \tag{4}$$

for the equations originated from the Ampere's law and

$$\begin{aligned}
\mu[a_z \frac{\partial \widehat{H}_{xe}}{\partial t} + \omega_z \widehat{H}_{xe}] &= \frac{\partial \widehat{E}_{ye}}{\partial z} - k_y(a_z \widehat{E}_{zo} + \omega_z \widehat{E}^*_{zo}) \\
\mu[a_z a_x \frac{\partial \widehat{H}_{yo}}{\partial t} + (a_x\omega_z + a_z\omega_x)\widehat{H}_{yo} + \omega_z\omega_x \widehat{H}^*_{yo}] &= a_z \frac{\partial \widehat{E}_{zo}}{\partial x} + \omega_z \frac{\partial \widehat{E}^*_{zo}}{\partial x} - a_x \frac{\partial \widehat{E}_{xo}}{\partial z} - \omega_x \frac{\partial \widehat{E}^*_{xo}}{\partial z} \\
\mu[a_x \frac{\partial \widehat{H}_{ze}}{\partial t} + \omega_x \widehat{H}_{ze}] &= k_y(a_x \widehat{E}_{xo} + \omega_x \widehat{E}^*_{xo}) - \frac{\partial \widehat{E}_{ye}}{\partial x}
\end{aligned} \tag{5}$$

for the equations derived from the Faraday's law. The asterisk '$*$' symbol in Equations (4) and (5) denotes integration of the corresponding EM component with respect to time. Note that the transformed variables, i.e., the EM field components with a '$\widehat{\ }$' hat are the Fourier transform of the corresponding component and are generally complex.

In actual calculations, the Fourier integration parameter k_y must be discretized. This discretization can be achieved by assuming the source-medium configuration to be periodic along the y-direction. A repetitive length of L_y along y-direction gives $k_y = n\Delta k_y = 2\pi n/L_y$ ($n = 1, 2, 3, \ldots, N$). By now it is clear that the Fourier integration parameter k_y is nothing more than the wave number in y-direction. Using a generic parameter of U_η, we can get the EM components in the real physical domain as [9]

$$U_\eta(x, y_d, z, t) = \frac{1}{L_y} \sum_{k_y=-N\Delta k_y}^{N\Delta k_y} \widehat{U}_\eta(x, k_y, z, t) e^{-ik_y y_d} \tag{6}$$

The truncation of the wave number summation is controlled by the shortest wavelength or the largest wave number of the EM wave simulated, i.e., we have

$$k_{\max} = N\Delta k_y \geq 1.1 \frac{2\pi f_{\max}}{v_{\min}} \tag{7}$$

3. STAGGERED GRID PSTD ALGORITHM USING THE REAL FOURIER TRANSFORM

In the vertical-xz plane, the spatial derivative of an EM component on a traditional collocated grid system with PSTD calculations was realized by the Fourier transform and inverse Fourier transform pair in the wave number domain along one particular direction. For example, the spatial derivative of the y-component of the electric field with respect to x-direction can be expressed as

$$\frac{\partial E_y}{\partial x} = FFT_x^{-1}(ik_x FFT_x(E_y)) = \int_0^\infty ik_x \left(\int_{-\infty}^\infty E_y(x, y, z) e^{-ik_x x} dx \right) e^{ik_x x} dk_x \tag{8}$$

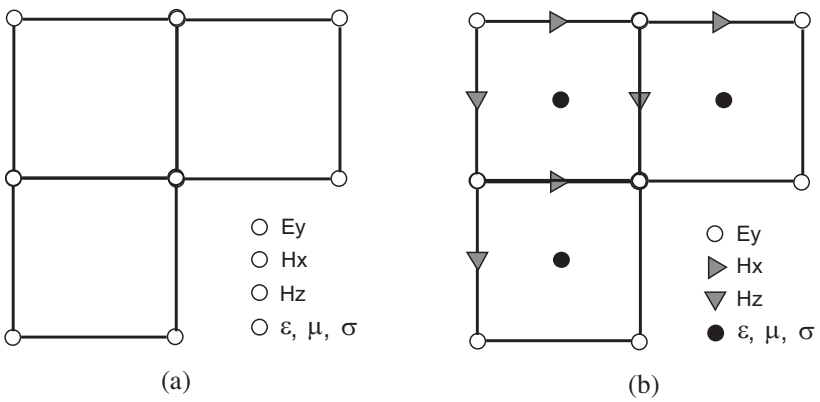


Figure 1: The non-staggered grid (a) and staggered grid for a 2D example applied to the EM fields.

However, on a staggered grid system, the spatial derivative of the electric field with respect to x-direction should be evaluated halfway between two grid points at $\triangle x/2$ to match the location of the magnetic components, exactly as used in the FDTD scheme as shown in Figure 1.

Thereafter, for the same spatial derivative using the PSTD approach the differentiation can be written as

$$\begin{aligned}\frac{\partial E_z}{\partial x} &= \int_0^{\infty} ik_x e^{ik_x \frac{\Delta x}{2}} \left(\int_{-\infty}^{\infty} E_z(x,y,z) e^{-ik_x x} dx \right) e^{ik_x x} dk_x \\ &= \int_0^{\infty} ik_x e^{ik_x \frac{\Delta x}{2}} \widehat{E}_z(k_x,y,z) \cdot e^{ik_x x} dk_x \approx \sum_{k_x=0}^{k_{xN}} ik_x e^{ik_x \frac{\Delta x}{2}} \widehat{E}_z e^{ik_x x} dk_x \end{aligned} \tag{9}$$

where K_{xN} is the Nyquist wave number, and the sum affects only the inverse Fourier transform. Here, the spectral differential operator (ik_x) used in non-staggered grid computations was replaced by $ik_x \exp(ik_x(\Delta x/2))$ for staggered grid computations [5]. No extra memory storage is required in the staggered grid computation. Using the real discrete Fourier transform Equation (8) can be expressed as [7]

$$\frac{\partial}{\partial x}[f((n+\frac{1}{2})\Delta x)] = \sum_{l=0}^{\frac{N}{2}} A_1(l\Delta k)\cos(\frac{2\pi nl}{N}) + \sum_{l=0}^{\frac{N}{2}-1} B_1(l\Delta k)\sin(\frac{2\pi nl}{N}) \tag{10}$$

where

$$\begin{aligned} A_1(l\Delta k) &= (l\Delta k)\cos(\tfrac{\pi nl}{N})B(l\Delta k) \mp \sin(\tfrac{\pi nl}{N})A(l\Delta k) \\ B_1(l\Delta k) &= (-l\Delta k)\cos(\tfrac{\pi nl}{N})A(l\Delta k) \pm \sin(\tfrac{\pi nl}{N})B(l\Delta k) \end{aligned}$$

and *A(lΔk)* and *B(lΔk)* are the original Fourier coefficients defined as

$$\begin{aligned} A(l\Delta k) &= \tfrac{1}{N}\sum_{n=0}^{N-1} f(n\Delta x)\cos(\tfrac{2\pi nl}{N}), \text{ for } \quad l = 0,\ \tfrac{N}{2} \\ A(l\Delta k) &= \tfrac{2}{N}\sum_{n=0}^{N-1} f(n\Delta x)\cos(\tfrac{2\pi nl}{N}), \text{ for } \quad l = 1,\ 2,\ 3, \ldots,\ \tfrac{N}{2}-1 \\ B(l\Delta k) &= \tfrac{2}{N}\sum_{n=0}^{N-1} f(n\Delta x)\sin(\tfrac{2\pi nl}{N}), \text{ for } \quad l = 1,\ 2,\ 3, \ldots,\ \tfrac{N}{2}-1 \end{aligned}$$

4. NUMERICAL EXAMPLE

The effectiveness of the staggered grid approach on suppressing the numerical dispersion is illustrated by the following 2 numerical realizations. Figure 2(a) shows the spatial first derivative of a one-point impulse using non-staggered grid and staggered grid PSTD with 256 sampling points. Obviously, the first spatial derivative using staggered grid PSTD greatly suppresses the ringing caused by numerical dispersion shown in the derivative using non-staggered grid approach. For a 5-point Gaussian impulse, as shown in Figure 2(b), the ringing shown in the non-staggered grid

derivative is much less than its 1-point impulse counterpart. And the derivative using staggered grid PSTD is extremely adequate and no observable ringing exits at all. This comparison can also be extrapolated to infer that the staggered grid approach are particularly useful for model domains containing high contrast boundaries that act as secondary diffraction sources corresponding to a short duration impulse.

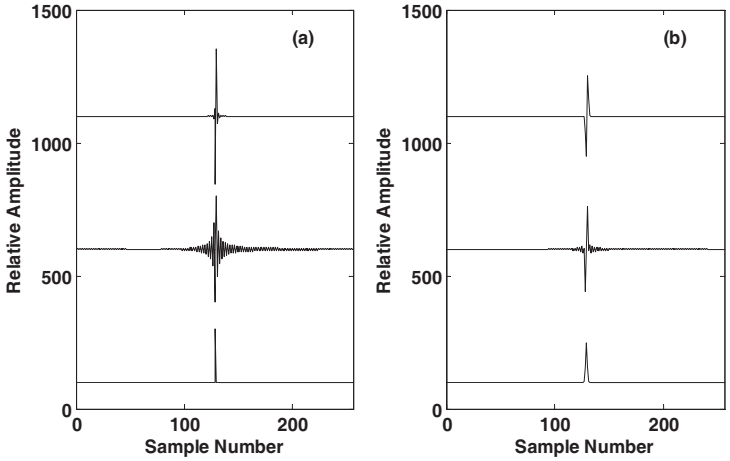


Figure 2: The 1-point impulse (bottom curve), its first derivative using non-staggered grid PSTD (middle curve), and the non-staggered grid and the first derivative using staggered grid PSTD (top curve) for the 1-point impulse (a); and the similar test using a 5-point Gaussian impulse (b).

5. CONCLUSION

A numerical algorithm for the simulation of EM wave propagation using the staggered grid PSTD calculation in a 2.5D model domain has been derived. The model domain was truncated using the unsplit stretched coordinate PML as the absorption boundary condition. To improve the numerical efficiency, the real FFT was employed. The approach was illustrated by an example in TE mode that most ground penetrating radar surveys commonly use, which is consistent with the case of y-direction symmetric situation that the components of E_y, H_x, and H_z are even functions and E_x, E_z, and H_y are odd function about the $y = 0$ plane. The 2.5D approach preserved the fidelity of the EM geometric spreading.

ACKNOWLEDGMENT

This research was supported by USACE Engineering Research and Development Center (ERDC).

REFERENCES

1. Yee, K. S., "Numerical solution of initial boundary problems involving Maxwell's equations in isotropic media," *IEEE Trans. Antennas Propag.*, Vol. 14, 302–307, 1966.
2. Liu, Q. H., "The PSTD algorithm: A time-domain method requiring only two cells per wavelength," *Microwave Opt. Technol. Lett.*, Vol. 15, No. 3, 158–165, 1997.
3. Fornberg, B., "High-order-finite-difference and the pseudospectral method on staggered-grids," *SIAM J. Num. Anal.*, Vol. 27, 904–918, 1990.
4. Witte, D. C., "The pseudospectral method for simulating wave propagation," PhD dissertation, Columbia University, 1989.
5. Chen, H. W., "Staggered-grid pseudospectral viscoacoustic wave field simulation in two-dimensional media," *J. Acoust. Soc. Am.*, Vol. 34, No. 10, 1064–1076, 1996.
6. Ozdenvar, T. and G. McMechann, "Algorithms for staggered-grid computations for poroelastic, elastic, acoustic, and scalar wave equations," *Geophys. Prospecting*, Vol. 45, 403–420, 1997.
7. Zhao, Z., J. Xu, and S. Horiuchi, "Differentiation operation in the wave equation for the pseudospectral method," *Earth Planets Space*, Vol. 53, 327–332, 2001.
8. Zhao, Z., Z. Zhao, J. Xu, R. Kubota, and L. Liu, "Strong ground motion simulation for seismic hazard assessment in an urban area," *Journal of Geophysics and Engineering*, 308–316, 2007.
9. Furumura, T. and H. Takenaka, "2.5-D modelling of elastic waves using the pseudospectral method," *Geophys. J. Int.*, Vol. 124, 820–832, 1996.

10. Tang, Z. B. and Q. Liu, "The 2.5-D FDTD and fourier PSTD methods and applications," *Microwave Opt. Technol. Lett.*, Vol. 36, No. 6, 430–436, 2003.
11. Zeng, Y. Q. and Q. Liu, "A multidomain PSTD method for 3D elastic wave equations," *Bull. Seism. Soc. Am.*, Vol. 94, No. 3, 1002–1015, 2004.
12. Chew, W. C. and W. H. Weddon, "A 3-D perfectly matched medium from modified Maxwell's equations with stretched corrdinates," *Micro. Opt. Tech. Lett.*, Vol. 7, No. 13, 599–604, 1994.

Microscopic Biological Cell Level Model Using Modified Finite-difference Time-domain at Mobile Radio Frequencies

C. H. See, R. A. Abd-Alhameed, P. S. Excell, and D. Zhou
Mobile and Satellite Communications Research Centre, University of Bradford
Richmond Road, Bradford, West Yorkshire, BD7 1DP, UK

Abstract— The paper demonstrates the modelling of the interaction mechanism between the biological tissues and electromagnetic field at mobile communication frequency ranges. The implementing of modified FDTD numerical method using frequency scaled FDTD with Floquet periodic boundary conditions and modified PMLs, the microdosimetric modelling of bioelectromagnetic interactions at cellular level, is established. In order to include the membrane effect on the biological tissues model in the analysis, the LE-FDTD is exploited to embed the lumped element cell-membrane model on the surface of the proposed tissue model in the FDTD computational domain. A new different structures of biological tissues are modeled and discussed, this includes a cluster of cylindrical cells. In order to imitate the effect of periodic replication of assemblages, Floquet periodic boundary conditions are imposed on the proposed model. Thus, the analysis of a large structure of cells is made more computationally efficient than the modeling of the entire structure. The total field distributions were shown in the context.

1. INTRODUCTIO

The interest in diagnostic and therapeutic applications of RF/microwaves in Medicine and in the assessment of possible health hazards due to EM radiation have stimulated the development of research streams in both modelling and experiments for evaluating EM power deposition in the interior of the human body or biological system. In order to establish precisely the required safety standard for regulating human exposure to EM waves, different aspects of studying the problem such as tissue level [1–6], cell level [7] and ionic level [8] have been carried out theoretically and experimentally. In this particular research area, the Finite-Difference Time-Domain (FDTD) method has been a overwhelming majority of the numerical techniques to solve various of different electromagnetic problem due to its simplicity and capable to handle complex geometry.

This paper is devoted to investigate the EM field distribution over the new proposed cylindrical cells equivalent tissue model by using modified FDTD numerical technique. Different EM approaches have been used to analyze this problem, in particular, the lumped-element FDTD has been implemented to model the cell's membrane. This was achieved with the use of Hodgkin-Huxley (HH) model and the Floquet theorem, in order to mimic the infinite model of the tissue and in turn to reduce the computational time. Due to the analysed structure under consideration is considerably smaller than the wavelength of mobile communication frequency GSM900/GSM1800 and also the time steps required for GSM900/GSM1800 frequency involves some millions of iterations, therefore, quasi-static FDTD is exploited to perform the computation of the analysis. The electric field distribution along the centre of the analysed various structures are discussed in which the computational results are found in well agreement with the previous published results [4, 6].

2. SUMMARY OF THE METHOD

Figure 1 shows a simple geometry for the elucidation of three-dimensional periodic boundary implementation. As can be observed, the periodic boundaries are imposed on the x- and y-sides of the structures, while modified Absorbing Boundary Conditions (ABCs) [6] are applied to truncate the space lattice along z-axis. The coordinate points (i_o, j_o, k_o) and (i_N, j_M, k_p) denote a space point in a uniform rectangular lattice, where i_o, j_o, k_o are the smallest lattice grid number in x, y, z direction respectively and i_N, j_M, k_p are the largest lattice gird number in x, y, z direction respectively.

Consider the problem space is filled with the lossless medium ($\sigma = 0$, $\varepsilon_o \varepsilon_r$ and $\mu = \mu_o$) and normal incident plane wave is propagating along z axis. The tangential electric fields distribution on plane i_0, i_N, j_0 and j_M are illustrated in Fig. 2. As can be seen, the red arrows are representing the tangential electric field components which are located on edge of the surface plane, while the black arrows are indicated as the rest of the tangential electric field components which are located on the surface plane. It should be noted that the explanation of the implementation method of the

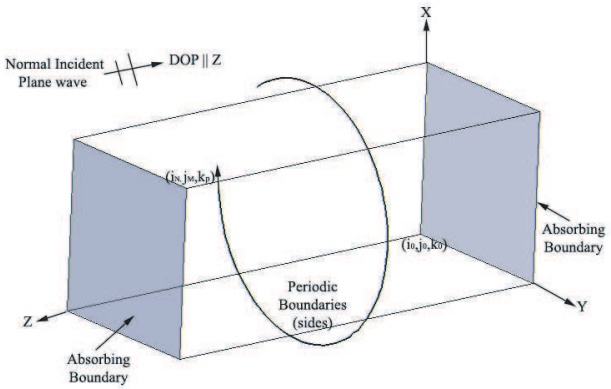


Figure 1: Geometry used in the analysis of 3-D infinite periodic structure illuminated by a normal incident plane wave.

periodic boundary condition into FDTD computation domain in the following context, are based on the normal incidence methods [9, 10], therefore, it is only applicable when the normal incidence plane wave is used. For the sake of simplicity and consistency of explanation of updating equations for the periodic boundary condition on the surfaces of the geometry shown in Fig. 1, the updating equations of the tangential electric components (E_x, E_z and E_y) which are not on the edge of the surface will be firstly to be discussed, and subsequently, the updating equations of the edged tangential electric component (E_z) are demonstrated.

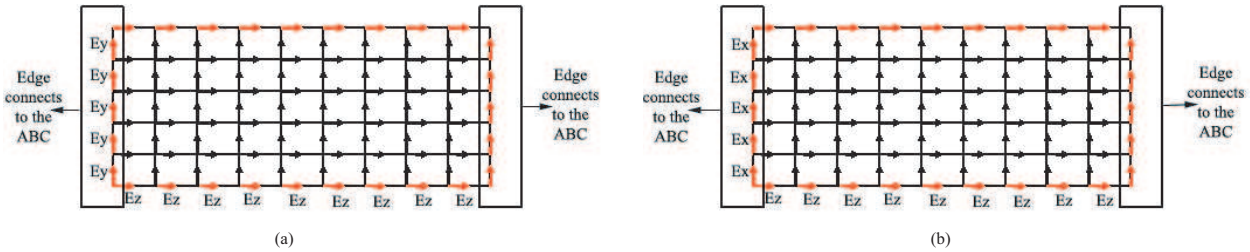


Figure 2: (a) Location of $E_z(\rightarrow)$ and $E_y(\uparrow)$ components in plane $i = i_0$ and i_N, (b) Location of $E_z(\rightarrow)$ and $E_x(\uparrow)$ components in plane $j = j_0$ and j_M.

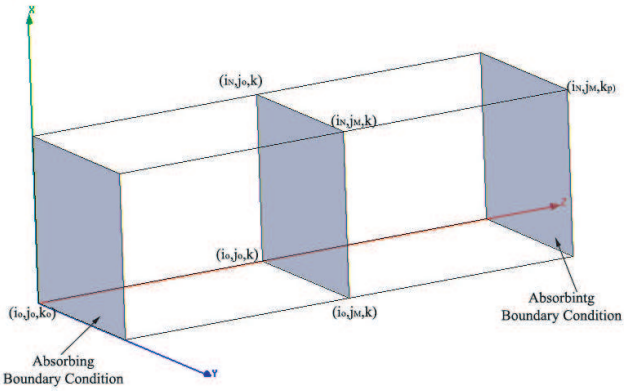


Figure 3: Location of the edged tangential component (E_z).

Consider the four surfaces planes at $i = i_0$, i_N and $j = j_0$, j_M, in which the floquet periodic boundary condition should be applied as shown in Fig. 1. From Fig. 2(a), the 3D updating equations for the tangential electric components (E_y and E_z) which are not located on the edge of surface

plane $i = i_0$ and i_N, can be stated in the following forms:

$$E_y^{n+1}(i,j,k) = E_y^n(i,j,k) + \frac{\Delta t}{\varepsilon\Delta}\left(\begin{array}{l} H_x^{n+\frac{1}{2}}(i,j,k+1/2) - H_x^{n+\frac{1}{2}}(i,j,k_p-1/2) \\ +H_z^{n+\frac{1}{2}}(i-1/2,j,k) - H_z^{n+\frac{1}{2}}(i+1/2,j,k) \end{array} \right) \quad (1)$$

$$E_y^{n+1}(i_N,j,k) = E_y^{n+1}(i,j,k) \quad (2)$$

$$E_z^{n+1}(i,j,k) = E_z^n(i,j,k) + \frac{\Delta t}{\varepsilon\Delta}\left(\begin{array}{l} H_y^{n+\frac{1}{2}}(i+1/2,j,k) - H_y^{n+1/2}(i_N-1/2,j,k) \\ +H_x^{n+\frac{1}{2}}(i,j-1/2,k) - H_x^{n+\frac{1}{2}}(i,j+1/2,k) \end{array} \right) \quad (3)$$

$$E_z^{n+1}(i_N,j,k) = E_z^{n+1}(i,j,k) \quad (4)$$

where Δt is the time increment and Δ is the space lattice increment.

As for the 3-D updating equations for the tangential components (E_x and E_z) which are not located on the edge of the surface at plane $j = j_0$ and j_M as shown in Fig. 2(b), can be derived as follows:

$$E_x^{n+1}(i,j,k) = E_x^n(i,j,k) + \frac{\Delta t}{\varepsilon\Delta}\left(\begin{array}{l} H_z^{n+\frac{1}{2}}(i,j+1/2,k) - H_z^{n+\frac{1}{2}}(i,j_M-1/2,k) \\ +H_y^{n+\frac{1}{2}}(i,j,k-1/2) - H_y^{n+\frac{1}{2}}(i,j,k+1/2) \end{array} \right) \quad (5)$$

$$E_x^{n+1}(i,j_M,k) = E_x^{n+1}(i,j,k) \quad (6)$$

$$E_z^{n+1}(i,j,k) = E_z^n(i,j,k) + \frac{\Delta t}{\varepsilon\Delta}\left(\begin{array}{l} H_y^{n+\frac{1}{2}}(i+1/2,j,k) - H_y^{n+1/2}(i-1/2,j,k) \\ +H_x^{n+\frac{1}{2}}(i,j_M-1/2,k) - H_x^{n+\frac{1}{2}}(i,j+1/2,k) \\ \end{array} \right) \quad (7)$$

$$E_z^{n+1}(i,j_M,k) = E_z^{n+1}(i,j,k) \quad (8)$$

As for the case of the edged tangential electric components (E_x, E_y, E_z) on plane $i = i_0$, i_N and $j = j_0$, j_M, and due to the E_x and E_y components are on the edge of the absorbing boundary conditions as shown in Fig. 2(a) and (b), therefore, they are assumed to be updated by the ABC updating equations. It should be noted that the only edged E_z tangential components will be considered here for the periodic boundary condition. According to Fig. 3, due to the presence of normal incidence plane wave, hence, four equations of E_z tangential components can be updated simultaneously. This can be simply done by using the following modified updated equations:

$$E_z^{n+1}(i,j,k) = E_z^n(i,j,k) + \frac{\Delta t}{\varepsilon\Delta}\left(\begin{array}{l} H_y^{n+\frac{1}{2}}(i+1/2,j,k) - H_y^{n+1/2}(i_N-1/2,j,k) \\ +H_x^{n+\frac{1}{2}}(i,j_M-1/2,k) - H_x^{n+\frac{1}{2}}(i,j+1/2,k) \end{array} \right) \quad (9)$$

$$E_z^{n+1}(i_N,j_M,k) = E_z^{n+1}(i,j,k) \quad (10)$$

$$E_z^{n+1}(i_N,j_M,k) = E_z^{n+1}(i,j,k) \quad (11)$$

$$E_z^{n+1}(i_0,j_M,k) = E_z^{n+1}(i,j,k) \quad (12)$$

$$(13)$$

3. SIMULATION AND RESULTS

A stack of ten cylindrical cells was investigated, as shown in Fig. 4. The radius and height of the each cell were considered to be 10 µm and 20 µm respectively. A plane wave of 100 V/m, propagating in the z-direction and polarized in the x-direction was used as excitation source. Note that the incident plane wave excitation was applied on a plane lying between the PML region and the outer limit of the FDTD grid. The PML, shown in Fig. 4, was 6 FDTD elements wide, the grading factor g was 10.1383 and the grid structure was effectively extended to infinity in the x- and y-directions, by imposing the Floquet boundary condition along the x and y axes. The Floquet periodic boundary condition plays an important role to mimic the presence of an extended 3-dimensional structure of biological cells, simulating connected tissue. The FDTD problem space was $220 \times 20 \times 20$ FDTD elements of size 1 µm while a discretization time step δt of 1.3 femtoseconds was chosen to drive the FDTD computation, to meet the requirements of the Courant stability criterion.

The shape of the living cells can be so diverse. In order to have better understanding on the effects of EM field interaction with different geometries of biological tissues, a cluster of such

cylindrical cells model of the tissue is proposed for the present work. This analysis is performed in which the material properties unchanged as in the case of spherical and cubical cell structures, as in [6]. Fig. 5 describes the 2D view of electric filed distribution of the proposed cylindrical-cell tissue at 10 GHz, while the electric field distribution along the centre of the analysed structure is shown in Fig. 6, where the Fig. 7 are the amplified version of the Fig. 6.

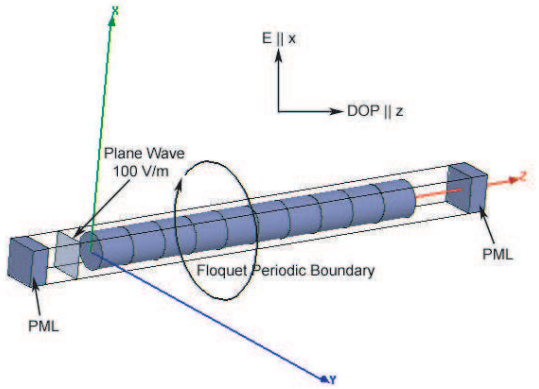


Figure 4: The three-dimensional view of the simulated cubical structures in FDTD computational domain.

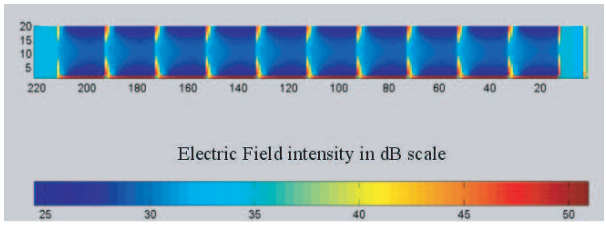


Figure 5: Modulus of the electric field on xz-plane at intermediate frequency 10 GHz (logarithmic scale).

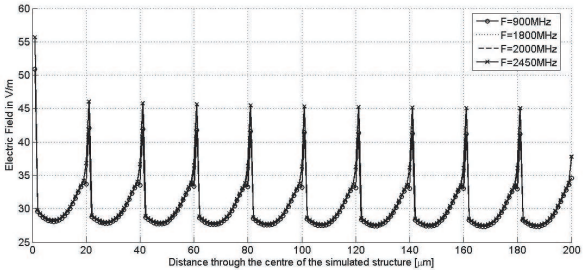


Figure 6: Penetration of electric field along z axis, through the centre of the simulated structure.

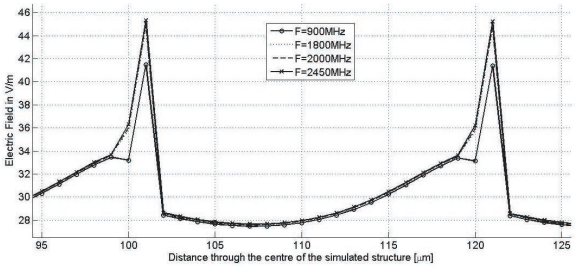


Figure 7: Penetration of electric field (enlargement of Fig. 5).

The loading effect of the HH model into the cylindrical-cell tissue has also been studied. Fig. 8 demonstrates the difference of 15% in field magnitude with and without the presence of HH model in the proposed simulated structure. The results show consistent difference with the previous spherical- and cubical-cells tissue simulated structures [6].

The comparison of the field distribution of spherical-, cubical- and cylindrical- cells tissue model, through the centre of the analysed structure, is elucidated in Fig. 9. The peak field on the membrane of the cylindrical structure is found to be about 1.7 times higher than in the cytoplasm, which is

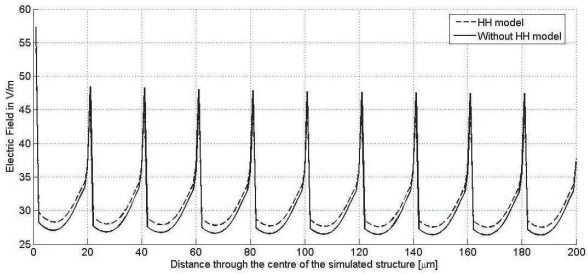


Figure 8: Electric field distribution along z-axis, through the centre of the simulated cubical-cell structure in Fig. 3, incorporating Hodgkin-Huxley model and driven at 2400 MHz.

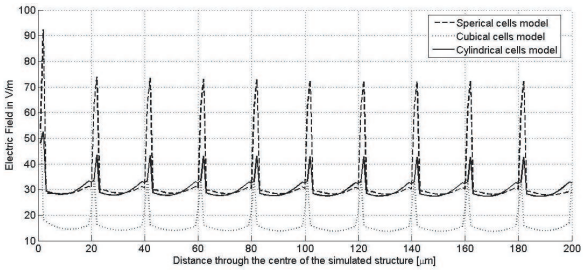


Figure 9: Comparison of field values of three different simulated structures at 900 MHz.

distinct with the previous two models [6]. As can be noticed, the peak field value of this cylindrical model is higher than the cubical model and lower than the spherical model, whereas the peak field value on the cytoplasm is the about the same as in the spherical model and double the value found in the cubical structure.

4. CONCLUSIONS

The combination of quasi-static FDTD, an arbitrarily-oriented lumped element membrane model, the modified Berenger's absorbing boundary condition and the Floquet periodic boundary condition represent a significant advance in verisimilitude of biological cell modeling, is demonstrated. By adopting the proposed methodology, it permits the computationally-efficient FDTD method to model small cell size object with reasonable computing time and accuracy.

REFERENCES

1. Gandhi, O. P. and J. Chen, "Numerical dosimetry at power-line frequencies using anatomically based models," *Bioelectromagnetics*, Vol. 13, 43–60, 1992.
2. Gandhi, O. P., G. Lazzi, and C. M. Furse, "Electromagnetic absorption in the human head and neck for mobile telephones at 835 and 1900 MHz," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 44, 1884–1896, 1996.
3. Kotnik, T. and D. Miklavcic, "Theorectical evaluation of the distributed power dissipation in biological cells exposed to electric fields," *Bioelectromagnetics*, Vol. 21, 385–394, 2000.
4. Emili, G., A. Schiavoni, F. L. Roselli, and R. Sorrentino, "Computation of electromagnetic field inside a tissue at mobile communications frequencies," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 51, 178–186, 2003.
5. Hadjem, A., D. Lautru, C. Dale, M. F. Wong, V. F. Hanna, and J. Wiart, "Study of Specific Absoption Rate (SAR) induced in two child head models and in adult heads using mobile phones," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 53, 4–11, 2005.
6. See, C. H., R. A. Abd-Alhameed, and P. S. Excell, "Computation of electromagnetic fields in assemblages of biological cells using a modified finite difference time domain scheme," *IEEE Transactions on Microwave Theory and Techniques*, No. 55, 1986–1994, Sept 2007.
7. Martin, S. M. S., J. L. Sebastian, M. Sancho, and J. M. Miranda, "A study of the electric field distribution in erythrocyte and rod shape cells from direct RF exposure," *Phys. Med. Biol.*, Vol. 48, 1649–1659, 2003.
8. Apollonio, F., M. Liberti, and G. D'Inzeo, "Theoretical evaluation of GSM/UMTS electromagnetic fields on neuronal network response," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, 3029–3035, 2002.
9. Taflove, A. and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-domain Method*, 2nd ed, Artech House, Inc, 2000.
10. Alexanian, A., N. J. Kolias, R. C. Compton, and R. A. York, "Three-dimensional FDTD analysis of quasi-optical arrays using Floquet boundary conditions and Berenger's PML," *IEEE Microwave and Guided Wave Letters*, No. 6, 138–140, 1996.

A Hybrid of Genetic Algorithm and Particle Swarm Optimization for Antenna Design

W. T. Li, L. Xu, and X. W. Shi

National Key Laboratory of Antenna and Microwave Technology, Xidian University
Xi'an 710071, China

Abstract— In this paper, a new effective optimization algorithm called PGHA is presented, which combines in the most effective way the properties of two of the most popular evolutionary optimization approaches now in use for the optimization of electromagnetic structures, the Particle Swarm Optimization (PSO) and Genetic Algorithms (GA). Some improved genetic mechanisms based on non-linear ranking selection, competition and selection among several crossover offspring and adaptive change of mutation scaling are adopted in the paper to overcome the drawbacks of standard genetic algorithm. Furthermore, the proposed algorithm is successfully applied to design a linear array with ten elements and a circular array with thirty one elements and obtain the desired beam forms. We try to use a modified Bernstern polynomial to reduce the number of variables when calculating the circular array, and simulation results show the abroad foreground of PGHA in the antenna array design.

1. INTRODUCTION

The global synthesis of antenna arrays that generate a desired radiation pattern is a highly nonlinear optimization problem. Pattern synthesis is known as the process of choosing the parameters of an antenna array to produce desired radiation characteristics. Many analytical methods have been proposed for its solution. Examples of analytical techniques include the well-known Taylor method and the Chebishev method [1]. However, analytical or calculus-based methods are generally unable to optimize problem with precedence constraint. To this end, stochastic methods are necessary [2, 3] in order to efficiently deal with large nonlinear search spaces and to extend the analysis.

Advantages of evolutionary computation are the capability to find a global optimum, without being trapped in local optima, and the possibility to face nonlinear and discontinuous problems, with great numbers of variables. Genetic Algorithms (GA) [4] have proven to be a useful method of optimization for difficult and discontinuous multidimensional engineering problems. A new method of optimization, Particle Swarm optimization (PSO) [5], is able to accomplish the same goal as GA optimization in a new and faster way. Since PSO and GA both work with a population of solutions, combining the searching abilities of both methods seems to be a good approach. The purpose of this paper is to investigate the foundations and performance of the two algorithms when applied to the design of two antenna array designs.

2. A HYBRID OF GENETIC ALGORITHM AND PARTICLE SWARM OPTIMIZATION

Genetic Algorithm [6] is an iterative stochastic optimizer that works on the concept of survival of the fittest motivated by Darwin, using methods based on the mechanics of natural genetics and natural selection to construct search and optimization procedures that best satisfies a predefined goal. Floating-point GA uses floating-point number representation for the real variables and thus is free from binary encoding and decoding. It takes less memory space and works faster than binary GA. To overcome the drawbacks of standard genetic algorithm such as prematurity and easily trapping in local optimum, some improved genetic mechanisms are adopted in the paper, such as non-linear ranking selection [7], different from the conventional algorithm in which two parents only produce two offspring, the new heuristic crossover operators defined below:

First, it produces three chromosomes from two parents I^A and I^Baccording to the following mechanisms:

$$I_1 = rI^A + (1-r)\, I^B \tag{1}$$

$$I_2 = (1-r)\, I^A + rI^B \tag{2}$$

$$I_3 = \frac{I^A + I^B}{2} \tag{3}$$

where r is uniform random number in $[0, 1]$. Then, among I_1 to I_3, the two with the largest fitness value are used as the offspring of the crossover operation.

The mutation operator is defined as follows: For a parent p, if variable p_k was selected at random for this mutation, the result is: $\bar{p} = (p_1 \cdots \bar{p}_k \cdots p_n)$ where

$$\bar{p}_k \in \left\{ \max\left(p_k - \mu \frac{p_k^{\max} - p_k^{\min}}{2}, p_k^{\min} \right), \min\left(p_k + \mu \frac{p_k^{\max} - p_k^{\min}}{2}, p_k^{\max} \right) \right\} \tag{4}$$

where $p_k^{\max}$, $p_k^{\min}$ are upper and lower bounds of p_k respectively, μ decreased with the increase of iterations.

$$\mu(\tau) = 1 - r^{[1-(\tau/T)]^b} \tag{5}$$

where r is uniform random number in $[0, 1]$, T is the maximum number of iterations, τ is the current iteration number, b is the shape parameter.

Particle Swarm Optimization is one of the more recently developed evolutionary technique, and it is based on a suitable model of social interaction between independent agents (particles) and it uses social interaction knowledge (also called swarm intelligence) in order to find the global maximum or minimum of a genetic function [8]. While for the GA the improvement in the population fitness is assured by pseudobiological operators, such as selection, crossover and mutation, the main PSO operator is velocity update:

$$\vec{v}_i(\tau+1) = w\vec{v}_i(\tau) + c_1\phi_1(\vec{p}_i(\tau) - \vec{x}_i(\tau)) + c_2\phi_2(\vec{p}_g(\tau) - \vec{x}_i(\tau)) \tag{6}$$

$$\vec{x}_i(\tau+1) = \vec{x}_i(\tau) + \vec{v}_i(\tau+1) \tag{7}$$

where

$\vec{v}_i(\tau) =$ particle velocity

$\vec{x}_i(\tau) =$ particle variables

ϕ_1, $\phi_2 =$ independent uniform random numbers

c_1, $c_2 =$ learning factors

$\vec{p}_i =$ local best solution

$\vec{p}_g =$ best global solution

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{T} \cdot \tau$$

where $w_{\max}$ and $w_{\min}$ is maximum and minimum value of the weighting factor respectively. T is the maximum number of iterations, τ is the current iteration number.

The PSO algorithm updates the velocity vector for each particle then adds that velocity to the particle position or values. Velocity updates are influenced by both the best global solution and the best local solution in the present population. The advantages of PSO are that it is easy to implement and there are few parameters to adjust. Our PGHA consists essentially in a strong co-operation of the two evolutionary algorithms described above, since it maintains the integration of the two techniques for the entire run.

In each iteration the population is divided into two parts and they are evolved with the two techniques respectively. They are then recombined in the updated population, that is again divided into two parts in the next iteration for another run of genetic or particle swarm operators.

3. PROBLEM FORMULATION

In amplitude-phase synthesis of the linear array, the far field array factor of this array can be written as

$$F(\theta) = \sum_{n=1}^{N} a_n \exp(j(n\frac{2\pi}{\lambda} d \sin\theta + \beta_n)) \tag{8}$$

where n the element number, λ the wavelength, β_n the excitation current phases of the elements, a_n the excitation current amplitudes of the elements, j the imaginary, d is the inter-element spacing, and θ is the polar angle of far-field measured from broadside (-90° to $+90^\circ$).

We try to use a modified Bernstein polynomial [9] to reduce the number of variables when calculating the circular array. The modified Bernstein polynomial is

$$F(U)=\begin{cases} B_1+\dfrac{1-B_1}{A^{MA}(1-A)^{M(1-A)}}U^{MA}(1-U)^{M(1-A)}, & 0\le U\le A \\ B_2+\dfrac{1-B_2}{A^{MA}(1-A)^{M(1-A)}}U^{MA}(1-U)^{M(1-A)}, & A\le U\le 1 \end{cases} \tag{9}$$

where B_1, B_2, M, A are parameters in the polynomial, B_1 and B_2 specify the left and right endpoints $F(0)$ and $F(1)$ respectively, while increasing M sharpens and narrows the peak of $F(U)$. For $\theta=90^\circ$, the far field array factor of the circular array is

$$E(90^\circ,\phi)=\sum_{n=1}^{N}F(U)_n\exp\left(j\frac{2\pi}{\lambda}r\left(\cos(\phi-\beta_n)-\cos(\beta_n)\right)\right) \tag{10}$$

where $F(U)_n$ is the nth excitation amplitude, being the nth equal sampling point value of $F(U)$. β_n is the excitation current phases of the elements, and r is the circle radius of the array.

A key point of optimization is the construction of the target function. In this paper, the fitness function to be maximized for array optimization problem can be expressed as follows:

$$Fitness=\frac{1}{\alpha\times|\mathrm{MSLL}-\mathrm{SLVL}|+\gamma\times|F_o(\theta)-F_d(\theta)|} \tag{11}$$

where MSLL the highest sidelobe level, SLVL the desired sidelobe level, and $F_o(\theta)$ and $F_d(\theta)$ are, respectively, the pattern obtained by using PGHA and the desired pattern. The values of α and γ should be selected by experience such that the fitness function is capable of guiding potential solutions to obtain satisfactory array pattern performance with desired properties.

4. NUMERICAL RESULTS

With the aim to validate the effectiveness of the developed technique, two examples are considered here. Firstly, we consider a uniform linear array of 10 isotropic elements spaced 0.5λ along x-axis

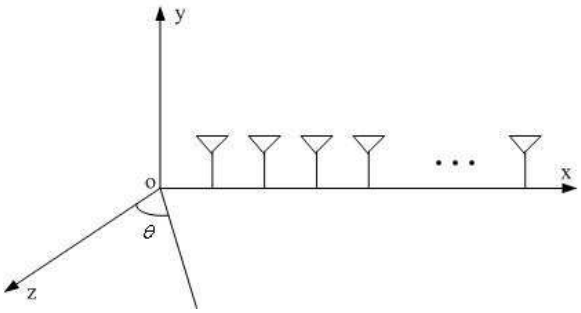


Figure 1: Linear array geometry.

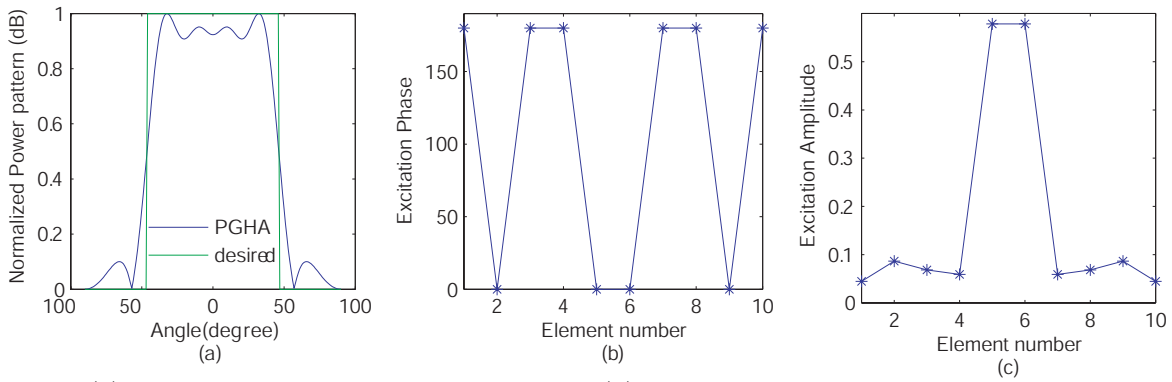


Figure 2: (a) Radiation pattern for the linear array, (b) Excitation amplitude distribution for the linear array, (c) Excitation phases in degree for the linear array.

in order to generate a sector beam as illustrated in Fig. 1. In the amplitude-phase synthesis, the phase is limited to 0 or π. Fig. 2(a) shows the normalized absolute power patterns in dB. Fig. 2(b) shows common amplitude distribution and Fig. 2(c) shows the phase distributions in degree.

Then, we consider a circular array of 31 isotropic radiators spaced 0.5λ apart along a circle of radius 6 wavelengths as illustrated in Fig. 3. To reduce the number of variables, we try to use a modified Bernstein polynomial according to Eq. (9). Fig. 4(a) shows the normalized absolute power patterns in dB. Fig. 4(b) shows the optimized distribution.

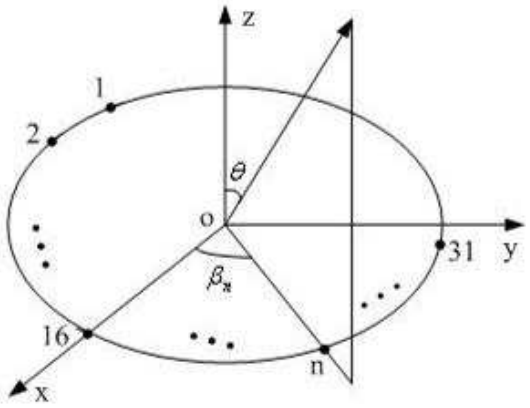


Figure 3: Circular array geometry.

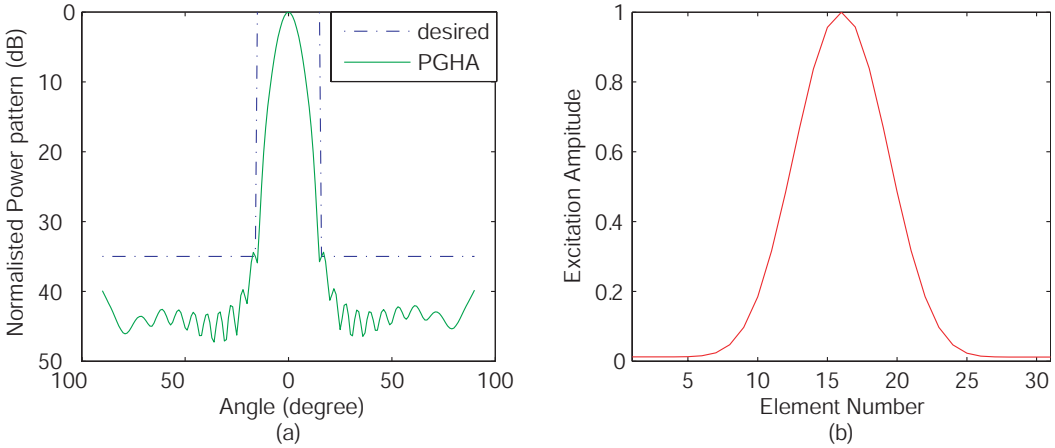


Figure 4: (a) Radiation pattern for the circular array, (b) Excitation amplitude distributions for the circular array.

5. CONCLUSIONS

An optimization method for the synthesis of linear array pattern and circular array pattern functions has been proposed and assessed. In order to take advantage of the peculiarities of these two methods, the proposed algorithm integrates the main features of GA and PSO into the optimization process. Results clearly show a very good agreement between the desired and synthesized specifications for the two cases. Since the algorithm proposed in this paper is reliable and effective and this feature makes it suitable for a wider application in electromagnetics.

ACKNOWLEDGMENT

The authors would like to thank the financial support form national natural science fund of P. R. China (No. 60571057).

REFERENCES

1. Balanis, C., *Antenna Theory Analysis and Design*, Wiley-Interscience, New York, 1997.

2. Shu, L. S., S. Y. Ho, and S. J. Ho, "A novel orthogonal simulated annealing algorithm for optimization of electromagnetic problems," *IEEE Transactions on Magnetics*, Vol. 40, No. 4, 1791–1795, 2004.
3. Lee, K. C. and J. Y. Jhang, "Application of particle swarm algorithm to the optimization of unequally spaced antenna arrays," *Journal of Electromagnetic Waves and Applications*, Vol. 20, No. 14, 2001–2012, 2006.
4. Hoorfar, A., "Evolutionary programming in electromagnetic optimization: A review," *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 3, 523–537, 2007.
5. Jin, N. and Y. Rahmat-Samii, "Advances in particle swarm optimization for antenna designs: Real-number, binary, single-objective and multiobjective implementations," *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 3, 556–567, 2007.
6. Zhai, Y. W., X. W. Shi, and Y. J. Zhao, "Optimized design of ideal and actual transformer based on improved micro-genetic algorithm," *Journal of Electromagnetic Waves and Applications*, Vol. 21, No. 13, 1761–1771, 2007.
7. Mahanti, G. K. and A. Chakrabarty, "Phase-only and amplitude-phase synthesie of dual-pattern linear antenna arrays using floating-point genetic algorithms," *Progress in Electromagnetics Research*, PIER 68, 247–259, 2007.
8. Clerc, M. and J. Kennedy, "The particle swarm-explosion, stability, and convergence in a multidimensional complex space," *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 1, 58–73, 2002.
9. Boeringer, D. W. and D. H. Werner, "Efficiency-constrained particle swarm optimization of a modified bernstein polynomial for conformal array excitation amplitude synthesis," *IEEE Transactions on Antennas and Propagation*, Vol. 53, No. 8, 2662–2673, 2005.

Fabrication and Analysis of Valve-less Micro Pumps

Nan-Chyuan Tsai, **Wei-Ming Huang**, **Chao-Wen Chiang**, **and Rong-Mao Lee**
Department of Mechanical Engineering, National Cheng Kung University
No. 1, University Road, Tainan 70101, Taiwan

Abstract— Due to dramatic progress in microfluidic devices and their applications, the development of microfluidic systems become more and more pertinent in recent years, such as fluidic activation devices, micro-sensors and micro-pumps. In this research, micro-electro-mechanical fabrication technologies are employed to fabricate an innovative valve-less micro-pump under magnetic control. The micro-coil is fabricated by electroforming and utilized to generate control force. *Prior to* fabrication of micro-pump, commercial software *CFD-RC* and *ANSOFT* are used first to simulate the performance of the designed micro-pump. *CFD-RC* is to observe the flow field and pressure in the micro-channel, and *ANSOFT* to evaluate if the magnetic force is sufficient enough to activate the membrane of the micro-pump. Finally, the micro-pump is equipped with a personal computer, in which *dSPACE* simulation module *DS1104* is embedded, to realize the closed-loop micro-system.

1. INTRODUCTION

The micro-pump plays an important role in the microfluidic devices because it can provide considerably precise flow rate. Therefore, it becomes more popular in medical applications, e.g., micro total analysis system ([1, 2]) and Lab-on-a-chip [3]. In recent years, a wide variety of valve-less micro-pumps have been developed. Berg proposed a two-stage peristaltic micro-pump for applications in microfluidics [4]. It mainly consists of an air blister, air tubes and a chamber. However, Berg's micro-pump needs a certain of additionally-added gas as the driving force. Xu proposed a micro-pump actuated by NiTi/Si composite diaphragm [5]. The actuation diaphragm performs reciprocating motion by applying the elastic force of NiTi thin-film. An oxygen micro-pump, whose actuation power is generated *via* the chemical production of oxygen gas, was proposed by Choi [6]. The oxygen gas is produced by decomposition of hydrogen peroxide under aid of catalyst. Nevertheless, the generation of oxygen is hardly controlled quantitatively.

In this report, an innovative valve-less micro-pump actuated by magnetic force is presented. The magnetic force is induced *via* the micro-coil and interacts with a stationary magnetic field generated by the permanent magnet to control the vibration of membrane. The magnetic force is controlled by the coil current. Due to the feedback control strategy, the pumping rate can be actively regulated to adjust the flow rate.

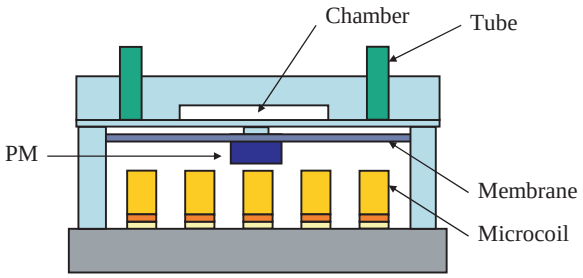


Figure 1: Sketch of the proposed micro-pump.

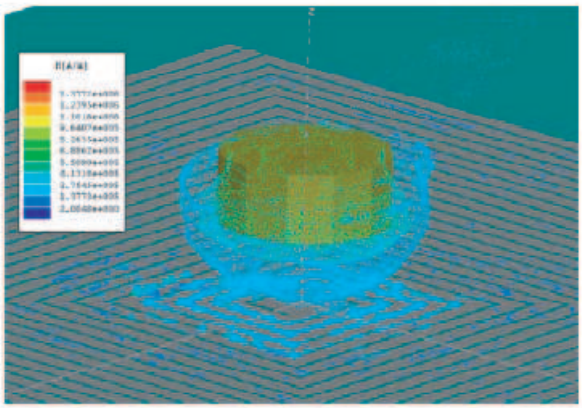


Figure 2: Magnetic field between PM and micro-coil.

2. FABRICATION PROCESS

The micro-pump under an actuation mechanism driven by magnetic force is manufactured *via* micro-electro-mechanical fabrication technologies. The micro-pump consists of a circular chamber

with nozzle/diffuser elements, a membrane and a set of micro-coil, as depicted in Fig. 1. A NdFeB-type permanent magnet is tightly bonded with the membrane so that the motion of the membrane can be controlled by the magnetic field (see Fig. 2). The strength (i.e., the flux density) of the magnetic field is adjusted by the current in the electroformed micro-coil fabricated on the other wafer (i.e., the bottom chip of Fig. 1). In order to avoid overheating of the micro-coil and ensure it can generate sufficient magnetic force enough to activate the membrane, the commercial software *ANSOFT* is employed to verify the dimensions of micro-coil and estimate the maximum magnetic force induced by coil current. The induced magnetic flux intensity is shown in Fig. 2.

2.1. Microfluidic Channel

The microfluidic channel is made of Polydimethylsiloxane (PDMS). The fabrication process is illustrated in Fig. 3. The process is started with coating a thin SU-8 about 50 µm thick, on a glass wafer, as shown in Fig. 3(a). The circular chamber with nozzle/diffuser is defined and developed as shown in Fig. 3(b). The PDMS is uniformly spin-coated and then baked at 70 °C for 4 hours, as illustrated in Fig. 3(c). After natural cooling under room temperature, the micro-channel is formed by molding, as shown in Fig. 3(d). Eventually, tubes are inserted as the inlet and the outlet of the microfluid, as shown in Fig. 3(e).

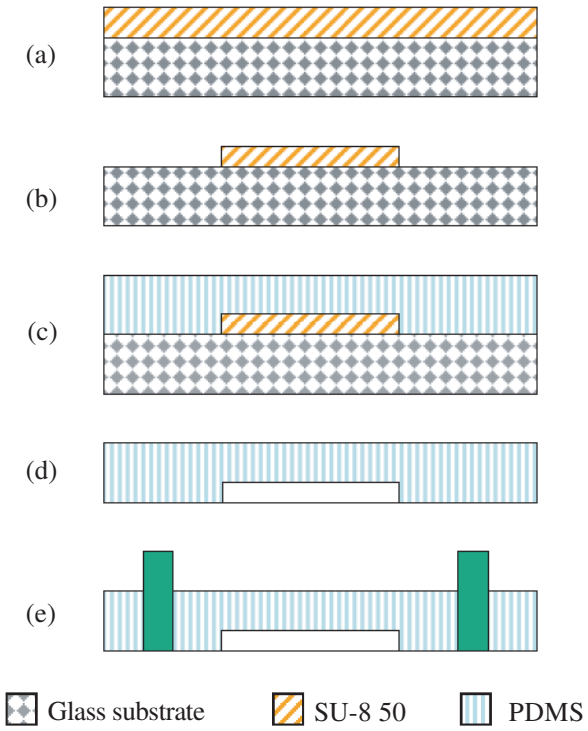


Figure 3: Outlines of the fabrication process of microfluidic mechanism.

2.2. Mechanical Membrane

The membrane is also made of PDMS. A sensing element, in which a permanent magnet is embedded, is attached with the membrane for detecting the vibration of the membrane, as shown in Fig. 4.

2.3. Micro-electromegnetic Actuation

The micro-coil is fabricated by electroforming technology. Photoresist SU-8 is employed to define the patterns of micro-coil. *Prior to* electroform process, a bed layer whose material is superior for electric conduction is required. The bed layer utilized in the research is the alloy of chromium (Cr) and gold (Au), whose width are 200 Å and 1000 Å respectively. Because Au can hardly be electroformed on the glass substrate intimately, a very thin layer of Cr is coated by electroforming on the glass substrate at first as the interface between Au and glass substrate. Au possesses excellent conductivity so that Au is chosen as the main material of bed layer. In this study, the KMPR is adopted as the thick-film photoresist. The N-methyl pyrrolidone (NMP) is used to remove KMPR

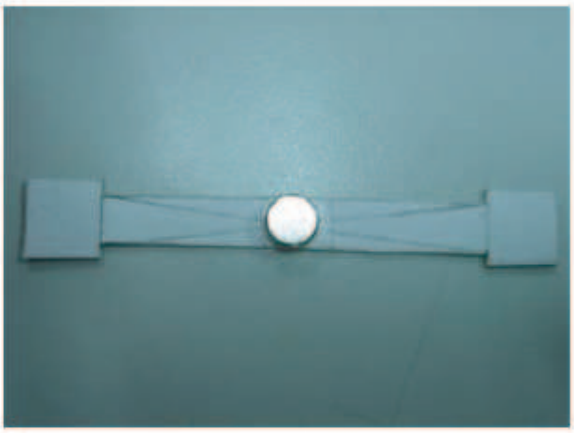


Figure 4: Diagram of the sensing element.

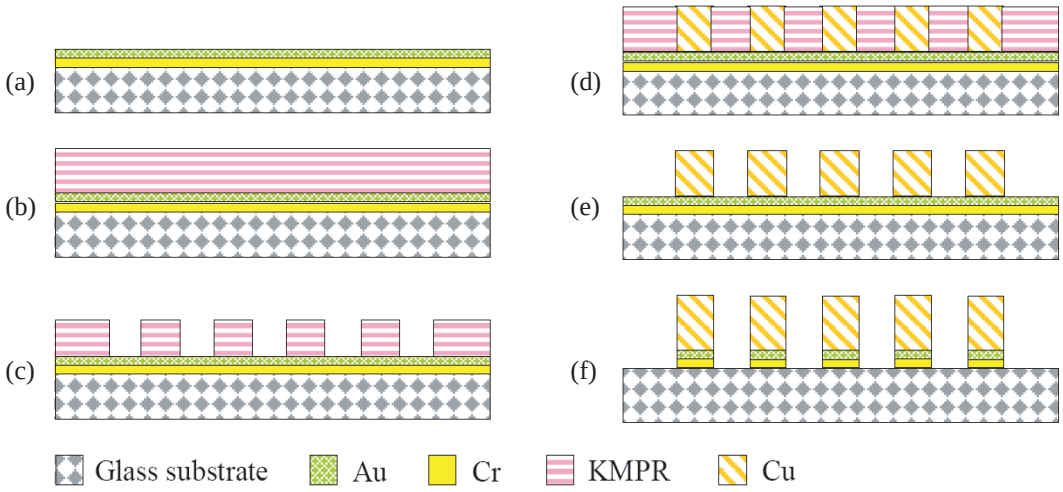


Figure 5: Outlines of the fabrication process of micro-coil.

after electroforming. It is noted that removing the photoresist after electroforming can improve the dissipation of heat generated by micro-coil.

The fabrication process of micro-coil is illustrated in Fig. 5. Cr and Au are electroformed on the glass substrate. KMPR is coated on the wafer by spin coater, as illustrated in Fig. 5(a) and Fig. 5(b). After soft bake, the wafer is under exposure about 90 seconds. Afterwards the wafer is soaked in TMAH about 5 minutes for photo development, as illustrated in Fig. 5(c), and then cleaned by ionic water. Before photolithography, the wafer has to be under hard bake at 150 °C for 3 minutes to enhance the adhesion of photoresist. The micro-coil, which is made of copper, is electroformed on the wafer, as illustrated in Fig. 5(d). In order to remove the undesired photoresist after electroforming, the wafer is soaked in NMP about 3 minutes, as illustrated in Fig. 5(e). Eventually, the wafer is soaked in etching solution to remove the undesired alloys of Au and Cr, as illustrated in Fig. 5(f). The micrograph of micro-coil is shown in Fig. 6. By stacking the micro-coil, the membrane and the microfluidic channel in sequence, the micro-pump is accomplished, as illustrated in Fig. 7.

3. EXPERIMENTAL RESULTS

The micro-pump is equipped with a personal computer to realize the closed-loop control system. An eddy current type gap sensor is employed to measure the vibration of membrane. The PID controller is designed and implemented with *dSPACE* simulation module *DS1104* which consists of a set of 16-bit A/D and D/A convertors with conversion speed of 2 µs and 10 µs respectively. Fig. 8 is the block diagram of closed-loop PID control system. Fig. 9 and Fig. 10 are the forced response of the closed-loop micro-pump. The command is a sinusoidal signal. From Fig. 9 and Fig. 10, the

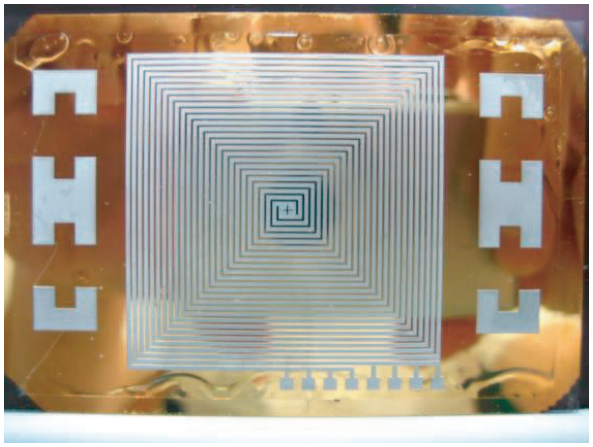


Figure 6: Micrograph of micro-coil.

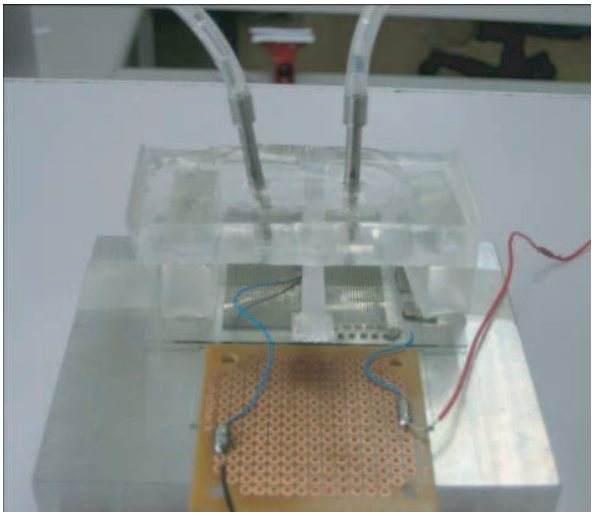


Figure 7: Picture of the proposed micro-pump.

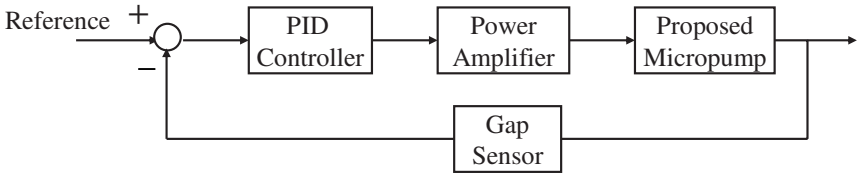


Figure 8: Block diagram of closed-loop PID control system.

amplitude of membrane can be successfully controlled by tuning the amplitude of the commands. In another words, the flow rate can be precisely adjusted by the feedback control loop by giving an appropriate reference.

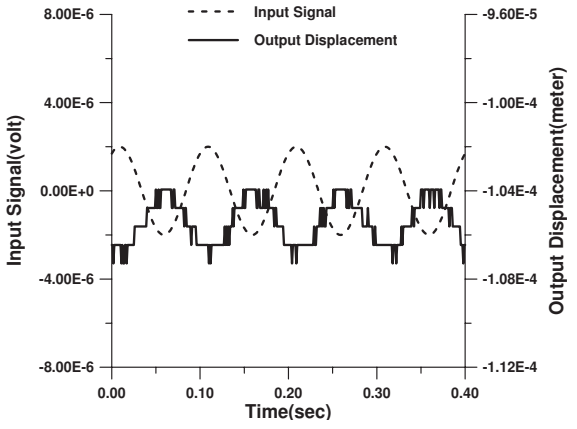


Figure 9: Forced response under sinusoidal command with amplitude 2 μV.

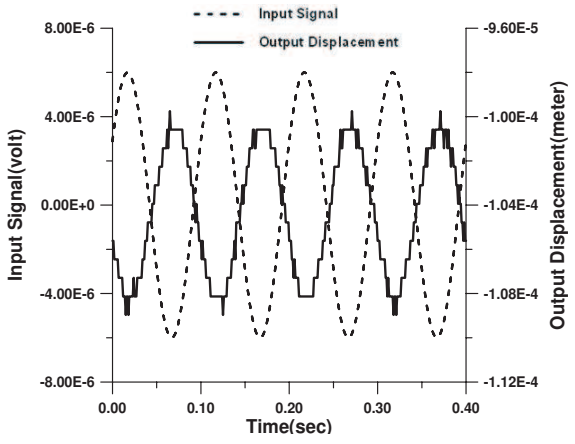


Figure 10: Forced response under sinusoidal command with amplitude 6 μV.

4. CONCLUSION

A valve-less micro-pump has been successfully fabricated by applying electroforming and MEMS technologies. The pump mainly consists of a microfluidic channel, a membrane and a set of micro-coil. The microfluidic and membrane are both made of PDMS. The outstanding property of PDMS is that it possesses low Young's modulus, and therefore PDMS is popularly used as the material of the vibratory membrane. For the fabrication of micro-coil, the electroforming technology is

adopted. Commercial software *ANSOFT* is employed to simulate the performance of the designed micro-pump. The PID controller is implemented *via dSPACE* simulation module *DS1104*. The efficacy of micro-pump under closed-loop control by magnetic force is verified. The amplitude of vibratory membrane can be successfully controlled by tuning the amplitude of the commands so that the flow rate can be precisely adjusted.

ACKNOWLEDGMENT

The authors would like to thank the Center for Micro/Nano Technology Research, National Cheng Kung University, Tainan, Taiwan, and National Nano Devices Laboratory (NDL) for equipment access and technical support. This research was partially supported by National Science Council (Taiwan) with Grant NSC 96-2221-E-006-255.

REFERENCES

1. Seibel, K., L. Scholar, M. Walder, H. Schafer, A. Schafer, T. Pletzer, R. Puschl, M. Waidelich, H. Ihmels, D. Ehrhardt, and M. Bohm, "A novel technology to create monolithic instruments for micro total analysis systems," *Materials Research Society Symposium Proceedings*, Vol. 869, *Materials, Integration and Technology for Monolithic Instruments*, 119–124, 2005.
2. Jen, C. P. and Y. C. Lin, "Design and simulation of bi-directional microfluid driving systems," *Journal of Micromechanics and Microengineering*, Vol. 12, No. 2, 115–121, 2002.
3. Kim, E. G., J. G. Oh, and B. Choi, "A study on the development of a continuous peristaltic micropump using magnetic fluids," *Sensors and Actuators, A: Physical*, Vol. 128, No. 1, 43–51, 2006.
4. Berg, J. M., R. Anderson, M. Anaya, B. Lahlouh, M. Holtz, and T. Dallas, "A two-stage discrete peristaltic micropump," *Sensors and Actuators, A: Physical*, Vol. 104, No. 1, 6–10, 2003.
5. Xu, D., L. Wang, G. Ding, Y. Zhou, A. Yu, and B. Cai, "Characteristics and fabrication of NiTi/Si diaphragm micropump," *Sensors and Actuators, A: Physical*, Vol. 93, No. 1, 87–92, 2001.
6. Choi, Y. H., S. U. Son, and S. S. Lee, "A micropump operating with chemically produced oxygen gas," *Sensors and Actuators, A: Physical*, Vol. 111, No. 1, 8–13, 2004.

RF MEMS Extended Tuning Range Varactor and Varactor Based True Time Delay Line Design

Yaping Liang[1], **C. W. Domier**[2], **and N. C. Luhmann, Jr.**[2]

[1]Department of Electronics and Information Engineering, Hangzhou Dianzi University, China
[2]Department of Applied Science, University of California, Davis, USA

Abstract— MEMS varactors are one of the important passive MEMS devices. Their applications include use in VCOs, tunable impedance matching networks, tunable filters, phase shifters, and true time delay lines. The shunt capacitive structure has been employed in most of the conventional MEMS varactor designs because of its simplicity. However, the capacitance ratio of this conventional shunt capacitive MEMS varactor is limited to 1.5 because of the MEMS Pull-In effect, which happens when the deflection between the MEMS top and bottom metal plates increase beyond 1/3 of the airgap between the two metal plates. At that time, the top metal plate will quickly snap down. This effect is the major limitation in MEMS varactor designs and can cause nonlinearity and mechanically instability. In order to eliminate this Pull-In effect, the author employed the so-called MEMS extended tuning range structure. This structure utilizes a variable height top metal beam with separate actuation parts. The airgap between the center part of the top beam and the bottom plate has been designed to be less than 1/3 of the airgap between the top beam and the bottom actuation pads. When DC bias is applied to the actuation parts, the entire top beam will move down together. Consequently, before the Pull-In effect happens at the actuation parts, the center part has already traveled through its entire tuning range, which means that the capacitive ratio of this kind of MEMS varactor can go to infinity.

A fabrication process employing a GaAs substrate has been designed based on surface micromachining technology. The maximum capacitance ratio of the designed MEMS extended tuning range varactor is 5.39 with a C_{max} value of 167 fF. Based on this MEMS varactor design, a Ka-band MEMS varactor based distributed true time delay line has been designed. This distributed true time delay line includes a high impedance CPW transmission line with 70 Ω unloaded impedance at 28 GHz and eight MEMS extended tuning range varactors based on the varactor design periodically loaded on the CPW line. The testing results show that a 56° phase delay variation has been achieved at 28 GHz. The measured insertion loss at 28 GHz is −1.07 dB at the up-state and −2.36 dB at the down-state. The measured return losses, S_{11} and S_{22}, are both below −15 dB at 28 GHz and below −10 dB over the entire tested frequency range of 5 GHz to 40 GHz.

1. INTRODUCTION

MEMS varactors are one of the important passive MEMS devices. They have considerable advantages compared with other semiconductor devices, including low loss, very high Q at mm-wave frequencies, high power handling capability, low power consumption, and high IIP3. The RF MEMS varactor can be employed in a phase shifter or true time delay line design to replace the GaAs Schottky varactor diode for low-loss, broadband, and high frequency applications in modern communication, automotive and defense applications. It can also be used in low loss tunable circuits including matching networks, tunable filters, and low noise oscillators.

2. RF MEMS EXTENDED TUNING RANGE VARACTOR

Conventional RF-MEMS varactors usually employ a shunt parallel plate capacitor whose capacitance is determined by the spacing between a fixed bottom plate and a movable suspended top plate. Electrostatic actuation occurs when an electrostatic force is created by applying a DC voltage between the capacitor plates, thereby displacing the movable plate toward the fixed plate. However, this shunt capacitance MEMS varactor structure suffers from the so-called Pull-In effect [1]. It happens when the displacement between the two metal plates exceeds 1/3 of the entire airgap. At that moment, the electrostatic attraction force loses balance with the mechanical restoring force and that causes the two metal plates to quickly snap into contact. The Pull-In effect is the major limitation in MEMS varactor designs. It will cause nonlinearity and mechanical instability of the MEMS varactors. In order to avoid the snap down, the designed capacitance ratio of the conventional MEMS capacitive varactor is usually set to 1.2 to 1.5 [2].

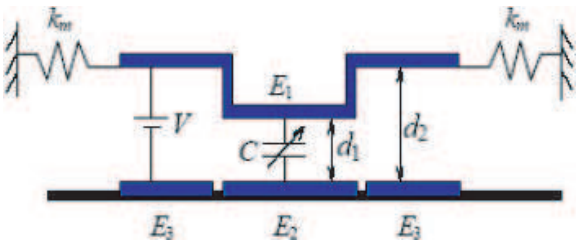


Figure 1: Model of a MEMS extended tuning range varactor structure.

In order to eliminate this Pull-In effect, one approach is to employ the so-called MEMS extended tuning range structure [3]. This structure, as shown in Figure 1, utilizes a variable height top metal beam E_1 with separate actuation parts E_3. The airgap between the center part of the top beam E_1 and the bottom plate E_2 has been designed to be less than 1/3 of the airgap between the top beam E_1 and the bottom actuation pads E_3. When DC bias is applied to the actuation parts, the entire top beam E_1 will move down together. Consequently, before the Pull-In effect happens at the actuation parts, the center part has already traveled through its entire tuning range, which means that the capacitive ratio of this kind of MEMS varactor can theoretically approach infinity.

A MEMS extended tuning range varactor has been designed at 28 GHz on a GaAs substrate by using the Ansoft HFSS and Agilent ADS simulation tools. Figure 2 shows the designed five-mask fabrication process. The most important and difficult step in building this extended tuning range structure is to form the variable height top metal beam E_1. Here, it has been realized by spinning two layers of photoresist continuously with different masks to pattern, see Figure 2(c) and (d). The

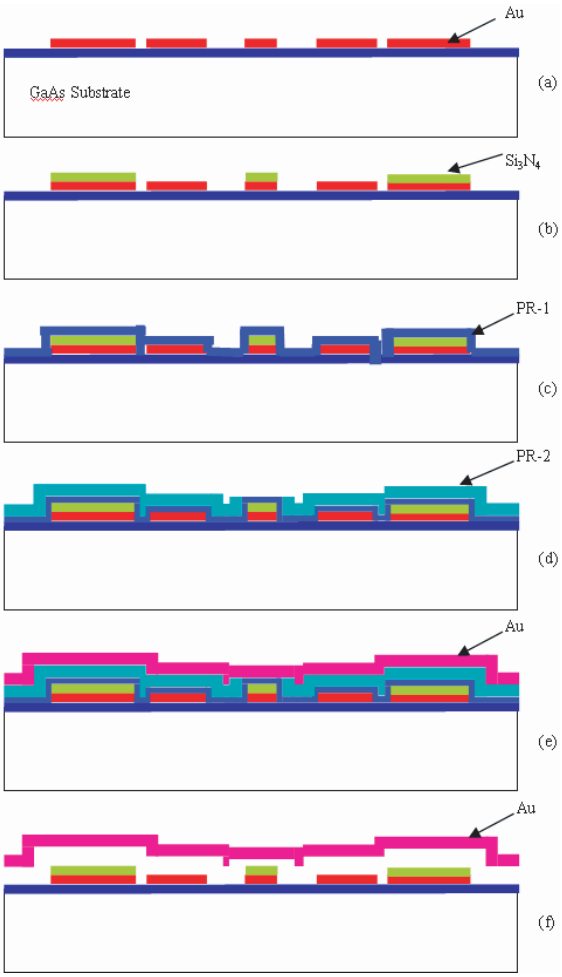


Figure 2: Fabrication process steps.

first step (Figure 2(a)) is to evaporate 0.7 µm of gold to form the signal lines and actuation pads using a gold lift-off process. The second step (Figure 2(b)) is to use PECVD to deposit 3000 Å of Si_3N_4 and to use dry Reactive Ion Etching (RIE) to form the dielectric layer between the bottom and top metal beams. The third step (Figure 2(c)) is to spin 1 µm thick photoresist and pattern the anchor points. The fourth step (Figure 2(d)) is to spin another 2 µm thick photoresist layer and pattern the anchor points and the center lower beam E_2. The fifth step (Figure 2(e)) is to electroplate 2 µm of gold and use photolithography to form the upper beam E_1. The final step (Figure 2(f)) is to use a dry etch to remove the sacrificial layer and release the whole structure.

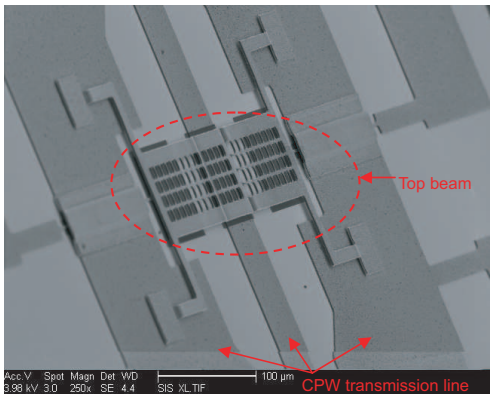


Figure 3: SEM picture of a MEMS extended tuning range varactor.

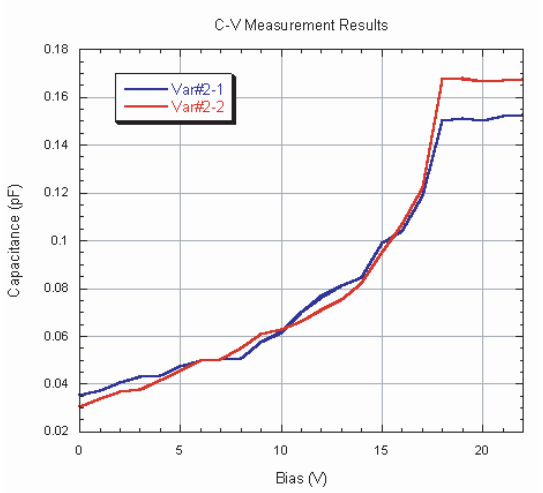


Figure 4: On-wafer C-V testing results.

Figure 3 shows an SEM picture of one of the fabricated MEMS varactors. On-wafer measurements by using an HP 4279A C-V meter have been employed and the results show that the maximum capacitance ratio is 5.39 with a C_{max} value of 167 fF (see Figure 4).

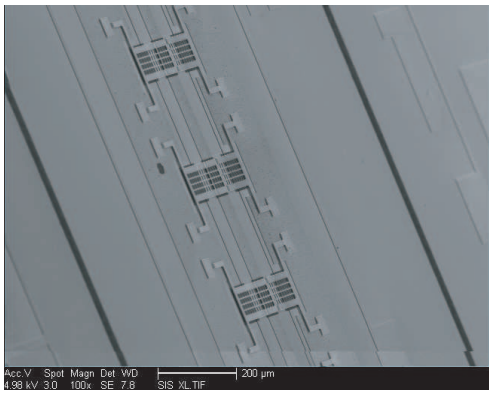


Figure 5: SEM picture of a MEMS varactor based true time delay line.

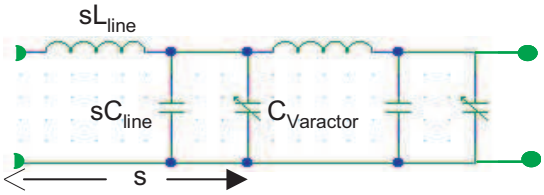


Figure 6: Equivalent circuit of unit section LC ladder network.

3. RF MEMS VARACTOR BASED TRUE TIME DELAY LINE

RF MEMS varactor based true time delay line technology employs a distributed LC ladder structure by parallel loading the MEMS varactors on high impedance coplanar waveguide (CPW) transmission lines. Figure 5 shows an SEM picture of a portion of one of the fabricated MEMS extended tuning range varactor based true time delay lines. The unit section equivalent circuit of the distributed LC ladder network is shown in Figure 6. When the operation frequency is far below the Bragg cutoff frequency of the LC ladder network, the group velocity remains essentially constant as the frequency is varied [4].

This MEMS varactor based true time delay line comprises 8 MEMS extended tuning range varactors loaded on a 70 Ω CPW transmission line operated at 28 GHz. The on-wafer testing results show that the insertion loss at 28 GHz is −2.36 dB in the down-state; the return loss, S_{11} and S_{22}, are both below −15 dB at 28 GHz (see Figure 7). The measurement phase delay is 56° at 28 GHz (see Figure 8).

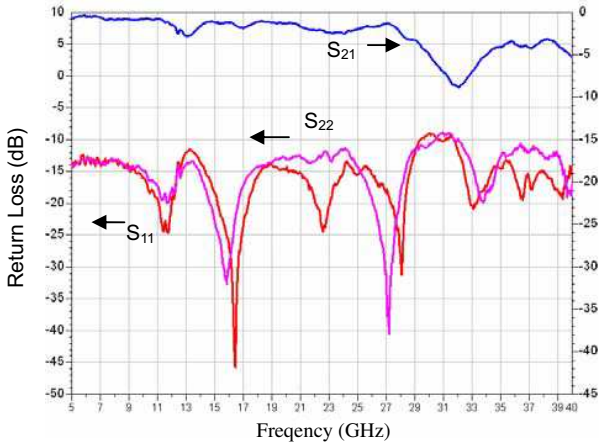


Figure 7: Measured down-state S-parameters of the MEMS varactor based true time delay line.

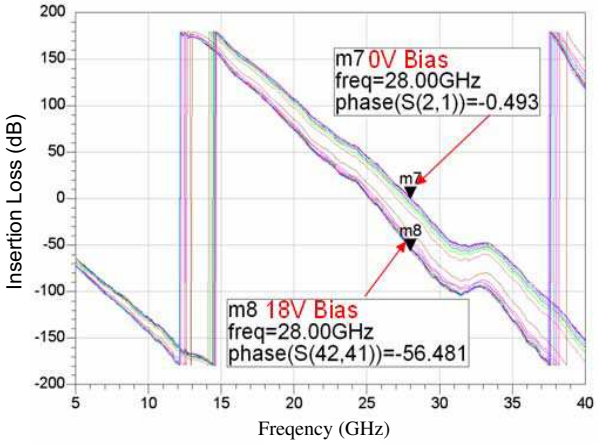


Figure 8: S_{21} phase delay.

4. CONCLUSIONS

A novel RF MEMS extended tuning range varactor structure has been employed to eliminate the Pull-In effect of the conventional MEMS varactor designs. On-wafer measurement results show that the maximum capacitance ratio is 5.39 for the extended tuning range MEMS varactors. A 28 GHz proof-of-principle MEMS varactor based true time delay line design employed the MEMS extended tuning range varactor structure. The maximum phase delay is 56° with a usable range extending from 5 to 40 GHz over which the line has demonstrated both low insertion loss and high return loss.

ACKNOWLEDGMENT

The authors are grateful to Miao Lu, Xiaodong Hu, and Yongjun Yan, of the Hebei Semiconductor Research Institute in China, for fabricating our MEMS varactors and delay lines. The authors would also like to thank Mehmet Ozgur and Michael Huff of the MEMS and Nanotechnology Exchange, funded by DARPA, for fabricating additional MEMS devices and circuits. This work was supported in part by the U.S. Department of Defense under Grant No. NBCH1050014 and by the U.S. Department of Energy under Grant No. DE-FG02-99ER54531.

REFERENCES

1. Senturia, S. D., *Microsystem Design*, Kluwer Academic Publishers, Boston, MA, 2001.
2. Rebeiz, G. M., *RF MEMS Theory, Design, and Technology*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2003.
3. Zou, J., C. Liu, et al., "Development of a wide tuning range MEMS tunable capacitor for wireless communication system," *Technical Digest of International Electron Devices Meeting*, 403–406, 2000.
4. Hsia, R. P., "Nonlinear transmission lines and applications," PhD Dissertation, UC Davis, 1997.

Conductance and Wave Impedance of Electrons

Raphael Tsu[1] **and Timir Datta**[2]

[1]University of North Carolina at Charlotte, Charlotte, NC 28223, USA
[2]University of South Carolina, Columbia, SC 29208, USA

Abstract— In circuit theory, impedance is used to characterize input/output. For waves, in addition to impedance with lump entities, an additional parameter, the wave impedance or characteristic impedance of the transmission medium which involves the propagation constant, needs to be considered. Electromagnetic waves are considered as 'true waves' because the energy-momentum relation is linear allowing a wave packet to maintain at all time, and in any frame of reference. However, for electrons, a Gaussian packet spreads because of the non-linear E-k dispersion, as well as the fact that k, the de Broglie wave vector, depends on the frame of reference. In spite of these concerns that electron, perhaps, is not really a wave, we present the concept of wave impedance of electrons. Within a finite mean free path, the propagation of an electron in a quantum wire of quantum waveguide in general can be characterized by the wave impedance.

1. INTRODUCTION

Before we touch on the need to introduce the concept of wave impedance, let us review what is involved with electromagnetic waves. In circuit theory, we deal with lump elements, such as resistors, capacitors and inductors. For a given input/output, the totality of the system, even involving operational amplifier, for example, may be represented by an equivalent circuit. If a transmission line is placed between input and output terminals, propagation down the transmission line represented by a wave-impedance is often all is necessary to describe the reflection and transmission from the input to the output. Let us ask what are the meaningful measurements needed to characterize the input/output. For example, the current flow in a diode as a detector, or the voltage at the control gate of a MOSFET can be determined with just these parameters mentioned. When the lump elements consist of tunnel junction or a quantum dot, obviously the wave nature of electron should be involved, particularly if something more than a single tunnel junction such as a superlattice or multiple quantum wells or simply a long Quantum Wire, QWire. It is then obvious we need to introduce the concept of wave impedance for electrons [1], to clarify or to expand on various concepts, a universal conductance defined by Landauer [2], contact conductance defined by S. Datta [3], and the wave conductance defined by Tsu [4], and by T. Datta and Tsu [5]. Basically, we need to distinguish the difference between an open and closed system. A transmission line is an open system, with the conductance per spin given by $G_0 = 2e^2/h$, while for a close system, the conductance per spin is given by $G_c = e^2/h$. Extending the transmission medium from 1D to 3D, the wave impedance Z_0 involves even fractions somewhat reminiscent to fractional quantum Hall constant. Which one of these definitions is most applicable to a given measurement will be examined.

2. LANDAUER CONDUCTANCE FORMULA

It was noted that conductance consists of discrete components depending on the number of longitudinal modes, as well as transverse degree of freedom [1]. In the Tsu-Esaki expression for the resonant tunneling [6], integration was performed over the transverse degree of freedom first, noting that the 2D-DOS for unbounded case is simply $m^*/\pi\hbar^2$. Let us instead integrate over the longitudinal direction first, dk_ℓ or dE_ℓ. Defining the function $F(E) \equiv 2\sum_t [1 + \exp(E + E_t)/k_BT]$, then the net tunneling current between two contacts given by Mitin-Kochelap-Stroscio 1999 [7] becomes

$$I = \frac{2e}{L}\sum_t \sum_{k_\ell} \frac{1}{\hbar}\frac{dE_\ell}{dk_\ell}\{F(E + eV - E_F) - F(E - E_F)\}. \tag{1}$$

And with $T \to 0$, and $V \to 0$, $F(E+eV-E_F)-F(E-E_F) \to eV\partial(E_F-E)$, then the conductance $G = \partial I/\partial V$, from (1) becomes the Landauer's conductance formula

$$G = 2\,G_0 \sum_t |T|^2(E_F,\, E_t), \tag{2}$$

in which the conductance per spin, $G_0 = e^2/h = 38.6\,\mu\text{S}$, referred to by Landauer as the universal conductance [2, 8], with its inverse $Z_0 = 25.9\,\text{k}\Omega$, and by S. Datta 1995 [3], as contact conductance. Both will be discussed more in detail after the section on the wave impedance of electron.

3. ELECTRON QUANTUM WAVEGUIDE (EQW)

Let us take the electron quantum waveguide shown in Fig. 1, the propagating wave-vector k_z including the potential energy eV is given by [9]

$$k_z^2 = \frac{2m_e}{\hbar^2}(E + eV) - k_{t\,nm}^2. \tag{3}$$

The transverse momentum vector k_{tmn} of the mode (m, n) is given by

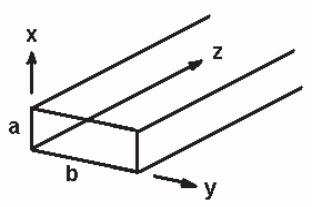


Figure 1: Section of an electron waveguide embedded in an infinite potential barrier on all sides. The applied potential is 0 for $z < 0$, and V for $z > 0$.

$$k_{tnm}^2 = (\frac{m\pi}{a})^2 + (\frac{n\pi}{b})^2 \equiv k_c^2, \tag{4}$$

and $E_{t,\,mn} = (\hbar^2 k_c^2)/2m$, we see that the transverse energy is nothing but the energy at the cutoff propagation vector, k_c as in usual waveguide case. The density of states, DOS for a cross-sectional area A is

$$DOS = \frac{1}{A}\sum_{n,\,m}\int_0^{k_{z\,M}} dk_z, \tag{5}$$

where k_{zM} is the maximum value for a given set (m, n), E and eV. Below cutoff, k_z is purely imaginary, propagation is not possible. At $T = 0$, $E = E_F$, above cutoff, the current density with two spins is given by

$$I = \frac{2e}{2\pi}\sum_{n,\,m}\int_0^{k_{z\,M}} \frac{1}{\hbar}\frac{\partial E}{\partial k_z}dk_z = \frac{2e}{h}\sum_{n,\,m}(E_F + eV) - E_{t,\,mn}. \tag{6}$$

At

$$m = n = 0,\; I_{00} = 2\frac{e}{h}(E_F + eV),$$

and for

$$\text{e}V_{00} + E_F > 0,\; G_{00} = \frac{\partial I_{00}}{\partial V} = 2\frac{e^2}{h} \equiv 2G_0;$$

and at

$$m = 1,\; n = 0,\; I_{10} = 2\frac{e}{h}(E_F + eV) - E_{t,\,10},$$

and for

$$eV_{10} + E_F > E_{t,\,10},\; G_{10} = \frac{\partial I_{10}}{\partial V} = 2\frac{e^2}{h} \equiv 2G_0;$$

etc. resulting in $G = G_{00} + G_{10}$, continuing to the general case of (m, n), $G = \frac{\partial I}{\partial V}$ is

$$G = \sum_{n,\,m} 2G_0\theta(E_F + eV - \hbar^2 k_{tnm}^2/2m), \tag{7}$$

in which θ is the unit step function, having a series of steps depending on how many modes, (m, n) are included. With a negative sign for e, $+eV$ becomes $-eV$ in (7). The factor of 2 in front of

G_0 is for the two spins. Thus for spin polarized case, there should be $G_0(+1/2)$ and $G_0(-1/2)$ without the factor 2. It is important to recognize that for single mode operation, only one step in G appears. The origin of these extra modes is due to the inclusion of modes (m, n) coming from the incident electrons having transverse energy. In free space, we simply take $a = b = \infty$, then $G = G_0\theta(E_F - eV)$, which ensures that the potential of the waveguide is below that of E_F. Otherwise no transmission is possible from the source of electrons at the energy range of $E = 0$ to E_F. It is important to recognize that an incident electron with transverse energy can enter the waveguide, without transverse-longitudinal scattering, only the longitudinal energy contributes to conductance. What is this conductance? From the derivation, clearly it is an input conductance, what Datta 1995 referred to as contact conductance. What happens to the output impedance? The setup of the problem implicitly assumes the output end is terminated in its own characteristic impedance. What happens to the sending end? Since all transverse modes forming the allowed modes entering the EQW are assumed to be uncoupled, reflection coefficient of each mode is zero for the planar boundary conditions.

The quantized conductance steps in units of $2e^2/h$ for the transport of electrons in constricted geometry was first pointed out by Landauer and misleadingly referred to as ballistic behavior or even some mechanism of energy loss in these systems [2, 8]. We shall show that these conductance steps are entirely a wave phenomenon similar to the propagation of an E&M wave through free space with a wave impedance of $377\,\Omega$. Unlike photons where it is quite simple to direct the wave in a given direction, however, electrons on a Fermi surface are in all directions. Thus, multimode operation dominates over a single mode operation in general. *It should be apparent that this sum of modes is totally controllable by the potential V as in tunable filters.*

4. WAVE IMPEDANCE OF ELECTRON

Shortly after the publication on the electron waveguide [4], T. Datta was asked by Tsu to work jointly on a generalized formulation of the electron wave impedance based on the original definition by Schelkunoff for electromagnetic waves [10]. The primary reason was to dispel the issue addressed by Landauer about the real conductance implying losses. Principally, in electromagnetics, the wave impedance in free space, $\eta = (\mu/\varepsilon)^{1/2} = 36.6\,\Omega$. In reality, the concept of wave impedance manifests in the transport of energy whether to a load with reflection and dissipative losses, or simply flowing on. Although electrons are not pure waves, with entanglement due to Pauli's exclusion principle, the differential equation, at least represented by the Schrödinger equation, is a typical wave equation, thus, should have similar wave impedance, even for sound. As long as we are dealing with a lossless case, the conductance is purely real. Reflection and transmission, consequences of boundary conditions in sections where the differential equation is piece-wise analytic, are handled through the use of impedance matching at these boundaries identical to the role of impedances in circuit theory. Although wave impedance shares some common notion with impedance in a lump network, the two are definitely not the same.

Following Datta-Tsu 2003 [1, 5], the wave impedance is defined via a plane wave normalized in a volume AL, with the area A transverse to the direction z, of length L. However, in solids, L must be less than the mean free path Λ, otherwise waves have no meaning. Then

$$E = \int_0^L dz \int_0^b dy \int_0^a dx \psi^* H \psi = \frac{\hbar^2 k_0^2}{2m}, \tag{8}$$

and

$$I = e \int_0^b dy \int_0^a dx \frac{\hbar}{i2m}(\psi^* \psi_z - \psi \psi_z^*) = e\frac{\hbar k_z}{mL}, \tag{9}$$

where H is the kinetic energy operator. To get an expression for impedance in units of ohm, we need to divide E in Eq. (8) by the charge to obtain the potential V. The impedance $Z = V/I$, so that

$$Z = \frac{\hbar}{2e^2}\frac{k_0^2 L}{k_z} = \frac{\hbar}{2e^2}\frac{k_0 L}{\sqrt{1 - k_c^2/k_0^2}}, \tag{10}$$

in which k_c is given by Eq. (4). Furthermore, along the direction of propagation, unlike the sine dependences in the x and y directions, the wave function dependence on z is $\exp(ik_z z)$, a

propagating function, therefore, periodic boundary conditions must be applied, i.e., $k_z L = 2\ell\pi$, with ℓ being any integer. Then Eq. (10) becomes

$$Z = Z_0 \ell \left[1 - k_c^2/k_0^2\right]^{-1}, \quad \text{with} \quad Z_0 = h/2e^2. \tag{11}$$

The factor $[1-k_c^2/k_0^2]^{-1}$, is not present in G for the EQW. However, in deriving the conductance of EQW, we allow electrons with transverse energy to enter the waveguide, although only the longitudinal energy contributes to the conductance. Here, an incident electron having transverse degree of freedom applies to an electron incident at an angle from the impedance along the z-axis. Since our formulation does not allow the transverse energy to be channel into the z-direction, for comparison with the derivation for EQW, we should have taken an electron having energy only in the z-direction. Then $k_z = k_0$ and the factor $[1 - k_c^2/k_0^2]^{-1}$ does not appear. However with $Z_0 = h/2e^2$, $Z_0^{-1} = 2G_0$, which is a factor of 2 larger than the derivation for EQW. There must be a satisfactory explanation for the factor of 2. Let us invoke the difference between open and close systems. An open system, (a) consists of an input to a tunnel junction or quantum well on the left, and an output to a similar thing to the right, with a long transmission line or quantum wire between the two; and a closed system, (b) consists of two contacts connecting a resonant tunneling diode or quantum well. Thus we recognize that (a) is represented by the solution of the wave impedance, or EQW. Therefore, $Z_0 = h/2e^2$, or $G_0 = 2e^2/h$ per spin, which is a factor of two larger than the Landauer conductance. Now, as we recognize that (b) is nothing but putting two Z_0 in series, or $Z_C = 2Z_0$. Then $G_c = G_0/2$. To summarize, the conductance per spin for the wave impedance treatment is twice the value for that treated by Landauer, which was originally obtained by Tsu and Esaki for tunneling between two contacts held at their respective Fermi levels. Considering as a closed system, it was pointed out simply that the potential seen by the tunnel junction $eV = (\mu_R - \mu_L)/2$, therefore $Z_0 = h/e^2$ [1]. Let us conclude by pointing out that the correct way to treat the case where a signal is applied to a transmission line, or in our case, the quantum wire or quantum wave guide is discussed in this work. The treatment by using the resonant tunneling approach as first shown by Tsu and Esaki [6], or Landauer [2], really only applies to lump circuit as in circuit theory, treating the junction as a lump element. However, it is important to recognize that whatever treatment is considered, it is the wave nature of electrons gives rise to this conductance per spin given by e^2/h per spin. The so-called universal conductance is really a wave conductance for a closed system. We may call it an input conductance [1], or a contact conductance [3], or a universal conductance [2], what it matters is the fact that it is a wave impedance entirely similar to the wave impedance of an E&M wave, which is not at all similar to the impedances of lump elements such as a resistance and capacitance.

Let us discuss another important point with the electron wave impedance. In a solid, L must be less than the mean free path, or $L \leq \Lambda$, with Λ the mean free path or the coherent length of electron. Otherwise, the wave nature loses its meaning. Let us define a minimum $k_{zm} \equiv 2\ell\pi/\Lambda$ for $\ell = 1$, corresponding to the lowest energy of the electron. The longer is the mean free path Λ, the lower is the allowed lowest energy and its corresponding lowest k-vector. What happens to the conductance for $\ell = 2, 3, \ldots$ The higher orders correspond to higher energy states which has a larger phase space for scattering leading to higher losses, or in less pronounced conductance oscillations. Or simply the transmission at resonance is lower. Only in free space with infinitely long coherence length, the conductance corresponding to higher $\ell's$ are allowed. Note that in electromagnetic case, the set of indices (ℓ, m, n) are always canceling each other because the k-vector and the photon energy are linear. Therefore, the wave impedance for E&M is always isotropic in free space. However, it is not always so in wave impedance for electrons.

5. WAVE IMPEDANCE IN 3D

In arbitrary direction of propagation, and using periodic boundary conditions for all three directions, the wave impedance in an unbounded free space is given by

$$Z_{\ell,m,n} = Z_0 \Xi_{\ell,m,n},$$

where

$$\Xi_{\ell,m,n} = \left[\frac{\dfrac{\ell^2}{L^2} + \dfrac{m^2}{a^2} + \dfrac{n^2}{b^2}}{\dfrac{\ell}{L^2} + \dfrac{m}{a^2} + \dfrac{n}{b^2}} \right]. \tag{12}$$

Even with $L = a = b$, the function $\Xi_{\ell,m,n}$ listed below consists of fractions except in one dimensional case, reminding the fractional quantum numbers in the fractional Hall effects [11], Note that there are degeneracy in 3D case. Suppose there is only one electron traversing the space, one can always pick one of the axes of the cube to align with the direction of propagation, then the wave impedance will be given by the fourth column marked 1-D. Now, a second electron is propagating in a direction not collinear with the first. Since we cannot align the coordinates with both, the complicated impedances will appear at the detector. Therefore in principle, these fractional terms will enter into the observed. For interacting electrons, the simplest may be just the Coulomb term e^2/r_{ij}, the situation is obviously very complex. Quantum number dependence of $\Xi_{\ell,m,n}$ in 1, 2 & 3 dimensions is listed in Table 1. Note that many fractions are involved. As in the wave impedance of E&M waves, when various photons are coupled, for example, in a photonic crystals, it is debatable what simplification, if any, the concept of wave impedance may provide. Under what circumstances the concept of wave impedance may help to identify processes is remain to be discovered.

Table 1: Quantum number dependence of $\Xi_{\ell,m,n}$ in 1, 2 & 3 dimensions.

ℓ	m	n	1D	2D	3D
1	1	1	1	1	1
2	1	1	-	5/3	3/2
2	2	1	-	-	9/5
2	2	2	2	2	2
3	1	1	-	5/2	11/5
3	2	1	-	-	7/3
3	2	2	-	13/5	17/7
3	3	1	-	-	19/7
3	3	2	-	-	11/4
3	3	3	3	3	3
4	1	1	-	17/5	3
4	2	1	-	-	3
4	2	2	-	10/3	3
4	3	1	-	-	13/4
4	3	2	-	-	29/9
4	3	3	-	25/7	17/5
4	4	1	-	-	11/3
4	4	2	-	-	18/5
4	4	3	-	-	41/11
4	4	4	4	4	4

6. CONCLUSIONS

The pre-factor for the conductance per spin $G_0 = e^2/h$ for a closed system is double for an open system, with $2G_0$. Whenever tunneling is involved, conductance is given by the wave impedance or wave conductance. The so-called universal conductance, input conductance, or contact conductance, describe the very same wave conductance of a closed system having right and left contacts. What led Landauer to assume that the contact is reflection-less? Generally contacts are nearly reflection-less at resonance but the details are accounted for by the transmission term rather than the 'pre-factor' G_0. Why then is the tunneling at resonance seems to be related to G_0? This is because tunneling as in Fabry-Perrot interferometer, at resonance, the transmission coefficient is nearly unity in a lossless system. Experimental data gives unity for the transmission is because at low temperatures without phonons, different modes from different transverse degree of freedoms are nearly independent. As soon as longitudinal and transverse momenta are mixed, these equal

steps of conductance are smeared. But why in the case of Si-QD that Nicollian and Tsu worked on showed conductance jumps in units of G_0 even at room temperatures? The answer lies in the fact that, for size $\sim$ few nm, the quantized energies are so far apart, being almost unaffected by phonons, a primary contributor for mixing of modes [12].

ACKNOWLEDGMENT

One of us, Ray Tsu, would like to thank Lee Casperson for some stimulating discussions regarding the wave nature of electrons.

REFERENCES

1. Tsu, R., *Superlattice to Nanoelectronics*, Chapter 11, Elsevier, 2005.
2. Landauer, R., *Phil. Mag.*, Vol. 21, 863, 1970.
3. Datta, S., *Electronic Transport in Mesoscopic Systems*, Cambridge Univ. Press, 1995.
4. Tsu, R., *Advanced Semiconductor Heterostructures*, 221, Eds. M. Dutta & M. Stroscio, World Sci., Singapore, 2003.
5. Datta, T. and R. Tsu, QWI_LANE2.19.Nov. 2003, http://arXiv.org/cond-mat/0311479, 2003.
6. Tsu, R. and L. Esaki, *Appl. Phys. Lett.*, Vol. 22, 562, 1973.
7. Mitin, V. V., V. A. Kochelap, and M. A. Stroscio, *Quantum Heterostructures*, Cambridge University Press, Cambridge, 1999.
8. Landauer, R., *IBM J. Res. Devlop.*, Vol. 1, 223, 1957.
9. Tsu, R., *Advanced Semiconductor Heterostructures*, 221, Eds. M. Dutta & M. Stroscio, World Sci., Singapore, 2003.
10. Schelkunoff, S. A., *B. S. T. J.*, Vol. 17, 17, 1938.
11. Chakraborty, T. and P. Pietilainen, *The Fractional Hall Effects*, Springer, 1988.
12. Tsu, R., *Microelectronic J.*, MEJ_2184, 2007.

A New Definition of Capacitance of Few Electron Systems

Tim LaFave Jr.[1] **and Raphael Tsu**[2]

[1]Department of Physics and Astronomy, University of Iowa, Iowa, IA 52242, USA
[2]Department of Computer and Electrical Engineering
University of North Carolina at Charlotte, Charlotte, NC 28223, USA

Abstract— Due to the discreteness of electronic charges in a nanoscale system, capacitance is defined in terms of the total interaction energy of N-electrons confined in a dielectric sphere. Specifically, the distribution of N-electrons is obtained from minimization of the total interaction energy including Coulomb, polarization and the self-polarization terms. And recently, by including the formation energy, the work done on the system, the capacitance agrees with the case when $N = 1$ and $N = \infty$. While Schrödinger equation is crucial in dealing with kinetic energy, it is the Poisson equation that allows properly accounting for spatial symmetry properties resulting from the discrete nature of electrons.

1. INTRODUCTION

When considering nanoscale systems the spatial symmetry properties of electrons play a role in the characterization of its electronic properties. These symmetry properties are most important for systems consisting of few electrons since the resulting potential landscape shapes its electrostatic boundary conditions. While the quantum mechanical Schrödinger equation provides an accurate evaluation of kinetic energy, here a classical electrostatic interactions model is demonstrated to provide the total electrostatic potential energy using only the Poisson equation to properly account for spatial symmetry properties resulting from the discrete nature of electrons. Within this model, the N-dependent expression of capacitance is written

$$C(N) \equiv \frac{Q}{V_{TOT}(r_i)} = N\frac{Q}{V(N)} = \frac{(Ne)^2}{E(N)} \tag{1}$$

with the N-dependent energy given by this interactions model,

$$E(N) = e\left[\phi_C(N) + \phi_P(N) + \phi_S(N)\right] = E_C(N) + E_P(N) + E_S(N) + W(N) \tag{2}$$

where subscripts C, P and S correspond to electrostatic Coulomb, polarization and self-polarization interactions, respectively. The total interaction energy may then be written,

$$E_I(N) = E_C(N) + E_P(N) + E_S(N)$$

and $W(N)$, the total work done on the dielectric system, is the total formation energy. Eq. (1) agrees with the conventional capacitance when $N = 1$ and in the conducting limit when $N \to \infty$, as the dielectric constant also tends to infinity.

The confinement of electrons to nanoscale systems is crucial to the development of current and near future materials and devices. We have previously demonstrated variation of the classical electrostatic capacitance of dielectric spheres consisting of few electrons [1], a resulting equilibrium configuration consistent with the classical single shell Thomson model [2–7], and derived a new time-independent, electrostatic capacitance expression concurrent with the fundamental definition, $C \equiv Q/V$ [2, 7], based on spatial symmetry of electrons. The validity of the discrete charge dielectric (DCD) model is substantiated by its agreement with existing models in necessary limits and a correspondence to electron orbital symmetry differences inherent to natural atomic systems [7].

The DCD model is based on the interaction properties of charges and charge distributions within a dielectric system. The work needed to perform this redistribution was previously shown [2] as being precisely equal in magnitude to the self-polarization interaction energy of all electrons in the system. In more common terms, this energy is formation energy. An evaluation of the formation energy is necessary to obtain the true spatial distribution of electrons at equilibrium from which the absolute ground state energy of the system is obtained.

The Coulomb interaction acts to maintain separation of electrons, pushing them as far apart as the "box" in which they reside will allow. The "box" in which the electrons of our model reside

is an electrostatic potential well provided solely by the formation of induced sheets of charge at the dielectric interface of a uniform dielectric sphere. For the purpose of simplicity the dielectric function of the media inside and outside the sphere are constant. Inside the sphere, $\varepsilon = 12$, while outside $\varepsilon' = 4$. In practical terms these values may represent a uniform silicon sphereembedded in a uniform silicon dioxide medium.

2. THE FORMATION ENERGY

In the absence of any excess (or "free") electrons, the medium is electrically neutral. However, with the insertion of an electron, the inherent dipoles of the medium become polarized such that there will appear a net charge at the dielectric interface. The medium presents an electric resistance to the electron just as the electron presents an electric resistance to the medium. Once the electron and medium come to equilibrium, the electron is located at the origin of the sphere forming a spherically symmetry equipotential landscape. The total energy in the system derives from the interaction of the electron with the potential presented by the charged sphere and the interaction resulting from the polarization of dipoles in the once-neutral dielectric material. The self-polarization energy may be written simply as

$$E_S(r_i) = e\phi(r_i) - W$$

where e is the charge of an electron, r_i, is its location and W is the formation energy associated with the surface charge. The formation energy is obtained by integration over the potential function with respect to the charge,

$$W = \int_0^\infty \phi(r_i)dq = \frac{e^2}{8\pi\varepsilon_0\varepsilon a}\sum_{l=0}^{\infty}\frac{(\varepsilon-\varepsilon')(l+1)}{[\varepsilon'+l(\varepsilon+\varepsilon')]}\left(\frac{r_i}{a}\right)^{2l} = \frac{e\phi(r_i)}{2}. \tag{3}$$

This is justified if $dq \to 0$, assuming the formation of a net surface charge from the polarization of numerous dipole elements. Hence, the formation energy results from one-half of the total interaction, $e\phi(r_i)$, which may be obtained by solution of the electrostatic Poisson equation [8]. The total potential energy of the single electron system is then

$$E(r_i,\, N=1) = e\phi(r_i) = W + E_S(r_i). \tag{4}$$

Upon minimization of $E(r_i,\, N=1)$, $r_i = 0$, and the ground state energy is obtained,

$$E_0(N=1) = e\phi(0).$$

Half of the total ground state energy is associated with the electron, the other half with the induced surface charge. Had the minimization of energy been performed with respect only to the self-polarization interaction energy the electron would still be located at the origin.

In general, the total energy in the system is given by

$$E(N) = E_I(N) + W(N)$$

with $W(N)$ being the summation of the work done by all N electrons in the system as individually evaluated as in Eq. (3)

$$W(N) = \sum_{i=1}^{N} W_i.$$

Note that in Table 1 for $N = 1$, minimization is unnecessary as the electron is at the origin. The left side shows minimization without the total formation energy $W(N)$, while the right side gives minimization of the total interaction energy with the formation energy term. Although inclusion of $W(N)$ in the minimization scheme is for consistency when N is large so that the result should converge to that of the conventional expression. The value for $E_0(N)$ is less than $E_0^1(N)$ indicating even if we did not use the correct way by physical considerations the lower minimized value would have led to the correct formalism.

A system consisting of two electrons will be in equilibrium when the electrons are on either side of the origin, $r_i = r_j = b$. A sphere radius of $a = 10\,\text{nm}$ is chosen to arguably exclude magnetic moment interaction energy [2, 7]. Minimization exclusively with respect to self-polarization yields $b = 0.6321a$ and a total system energy of $E_I(N) = 0.11145\,\text{eV}$. Minimization with respect to the

Table 1: Left-Data for minimization with respect to the interaction energy, $E_I(N)$, with $E_0^1(N) = [E_I(N)]_{\rm MIN} + W(N)$ and Right-Data for minimization to the total system energy $E_0(N) = [E_I(N) + W(N)]_{\rm MIN}$.

	Minimization with respect to $E_I(N)$				Minimization with respect to $E_I(N)$+$W(N)$			
N	b	$E_I(N)$	$W(N)$	$E_0^1(N)$	b	$E_I(N)$	$W(N)$	$E_0(N)$
	(a)	(eV)	(eV)	(eV)	(a)	(eV)	(eV)	(eV)
1	0.0000	0.01200	0.01200	0.02400	0.0000	0.01200	0.01200	0.02400
2	0.6321	0.08157	0.02988	0.11145	0.5270	0.08253	0.02750	0.11003
3	0.7271	0.20794	0.05032	0.25825	0.6313	0.21019	0.04478	0.25497
4	0.7740	0.38988	0.07282	0.46269	0.6884	0.39369	0.06359	0.45727
5	0.8030	0.63195	0.09703	0.72898	0.7248	0.63752	0.08357	0.72108
6	0.8237	0.92871	0.12294	1.05164	0.7518	0.93623	0.10478	1.04101
7	0.8387	1.28777	0.15005	1.43782	0.7714	1.29743	0.12678	1.42421
8	0.8507	1.70248	0.17856	1.88104	0.7875	1.71443	0.14981	1.86424
9	0.8604	2.17603	0.20825	2.38428	0.8007	2.19038	0.17370	2.36408
10	0.8684	2.70848	0.23899	2.94747	0.8116	2.72541	0.19829	2.92371

total system energy yields $b = 0.5270a$, and a total ground state energy of $E_0(N) = 0.11003\,\text{eV}$. Data for $1 \leq N \leq 10$ are provided in Table 1. The ground state energy for $N = 10$, for example, obtained by minimizing with respect to the total energy of the system is 24 meV lower than having minimized with respect to the interaction energy alone.

3. CAPACITANCE MEASUREMENTS

The familiar approach of measuring the capacitance of a system is to measure a potential difference between two metal plates directly in contact with the dielectric medium. However, in such a system, the metal plates predominantly define the potential well within which electrons are confined. In this case, the symmetry of excess electrons is quenched by the shape and size of the metal plates, and capacitance is a constant.

Another approach of measuring capacitance is with an electrometer that measures the electric potential some distance away from the device itself. An example of this is evident in Ref. 9. Through induction then, the influence of the electrometer on the symmetry of excess electrons in the system may be minimized. The nanoscale device of interest might be a semiconducting dielectric sphere embedded in an insulating medium that may be charged by electrons by means of tunneling. The electrometer may detect an electron's presence at such a "dot" when a sharp rise in potential is observed. This potential results from unique sources of charge within the device itself. In a dot charged with a single electron, two unique charge sources are available: the electron and the surface charge. With present technology, the electrometer is unable to distinguish between the electron and the surface charge. Therefore, if the electron moves away from the origin of the sphere, the electrometer is unable to distinguish if it is the electron's motion or the change in the magnitude of the surface charge that results in a change in potential at its location. In a classical sense, however, such a change would not be predicted as the conventional capacitance model assumes that all charge is spherically symmetric within the system — either a single charge at the origin or an infinitely divisible surface charge. The conventional capacitance expression,

$$C_{CONV} = \int_0^Q \phi dq = \frac{Q^2}{2E}$$

represents the total charge in a many electron system ($Q = Ne$) while the surface charge is implicitly included with the electrostatic potential appearing in the energy, E. The factor of $1/2$ appears as a consequence of assuming that electrons are infinitesimally small and do not repel each other

(hence, the integration over q). However, the factor of $1/2$ in the conventional expression is argued to derive from an average over the total potential of a many-electron system in the presence of metallic contacts. For the case of few electrons, however, the expression is misleading. In particular, we demonstrate that the total energy in a single electron sphere is E when it is connected and $2E$ when it is isolated, or unconnected.

The capacitance of a system must be defined with respect to an equipotential. The symmetry of a connected single electron device is relatively simple as the equipotential landscape is well defined by the metal surfaces. The difference between the capacitance of a connected and unconnected single electron system has been shown [10] to be a factor of $1/2$. In particular, a connected single electron system has the capacitance,

$$C_C = \frac{e^2}{2E}$$

as in the conventional expression. This is not surprising, as the conventional capacitance for spheres has been supported by numerous experimental measurements, particularly for many-electron systems, but also by single-electron Coulomb blockade studies [11]. For the single electron case, the equipotential symmetry is the same as though the charge were uniformly distributed at the sphere's surface. However, for an unconnected system, free from metallic contacts, the capacitance can be shown to be [10]

$$C_U = \frac{e^2}{E} = 2C_C.$$

To rectify this difference, it is worth noting that a single electron in a dielectric sphere involves a self-polarization interaction and an interaction of surface charge formation of equal magnitude. Together, these two energy components constitute the total system energy. In previous work we've shown that the single electron case agrees with the conventional capacitance if only half of the total energy of the system (the self-polarization interaction energy) is included. The model considered here is an unconnected sphere, and consequently agrees with the unconnected capacitance expression as it includes both interaction and formation energy components.

4. MONOPHASIC CAPACITANCE

In cases of more than one electron, symmetry must play a central role in a definition of capacitance as the equipotential landscape is defined by the spatial configuration of electrons. A new expression for capacitance has emerged from the DCD model, coined the *monophasic* capacitance [2, 7] to distinguish it as an electrostatic capacitance resulting from unique spatial symmetries of electrons at *equilibrium* defined within a given phase of the system. In previous work [2], the monophasic capacitance was shown to agree with the classical model in a conducting limit as well as for the single electron case (in both conducting and dielectric cases).

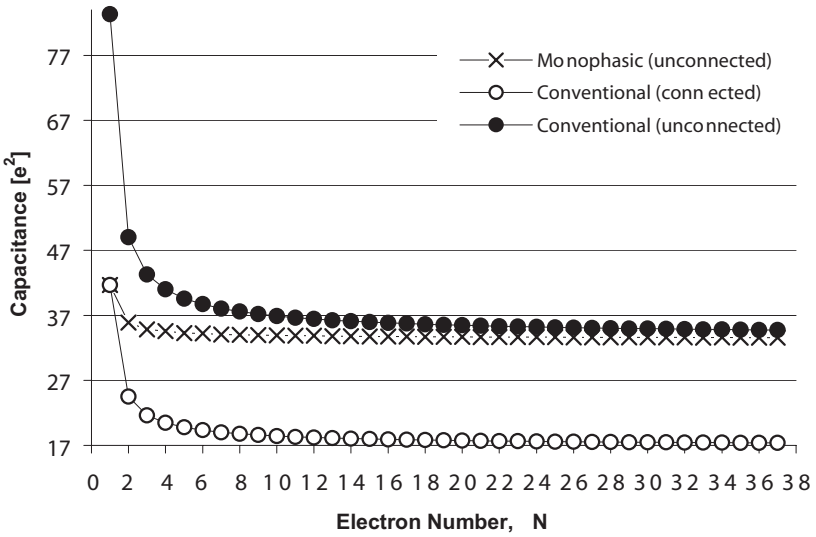


Figure 1: The monophasic capacitance, connected and unconnected conventional capacitance.

The monophasic capacitance expression,

$$C(N) = \frac{(Ne)^2}{E_I(N) + W(N)} = \frac{Q^2}{E_0(N)} \tag{5}$$

differs from the conventional expression only by a factor of 1/2. In the single electron case, the difference between a connected and unconnected system is attributed to the inclusion of the formation energy in the monophasic definition as a consequence of having derived the capacitance from the average total electrostatic potential experienced by each electron which forces an evaluation with respect to the total energy in the system rather than merely the interaction energy associated with electrons themselves. Since $E_0(N) = 2E_{CONV}$ for a single electron system, then $E_I(N) = 2E_{CONV}$ indicating that the conventional expression, as written, includes only an evaluation of the interaction energy when the system is connected. When connected, $2E_{CONV}$ implies the inclusion of the formation energy in the single electron case (the only spherically symmetric possibility). In the case of many electrons, the factor of 1/2 is also consistent depending upon whether the system is connected or not.

The monophasic capacitance is inherently more detailed than the conventional capacitance expression, having explicitly included the formation energy, $W(N)$. In Figure 1, the conventional capacitance of both connected and unconnected systems is shown using only the interaction energy $E_I(N)$ derived from the DCD model. The monophasic capacitance, including both interaction and formation energy, is also shown. In the single electron case, the monophasic capacitance for an unconnected dielectric sphere agrees with the connected conventional capacitance indicating an error in the connected conventional expression unless the denominator $(2E)$ is taken as the total energy of the system including formation energy (as required for an unconnected system). As well, the unconnected conventional capacitance for the single electron case demonstrates the complication of merely changing the conventional expression by a factor of 1/2. However, as N becomes large, the monophasic capacitance tends toward the unconnected conventional capacitance, as the formation energy becomes a smaller fraction of the total energy in the system as shown in Figure 2.

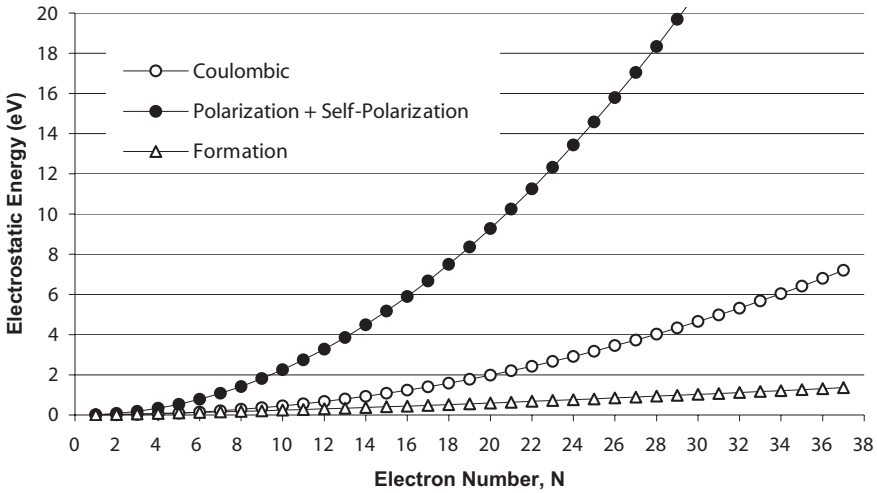


Figure 2: The confining interactions of polarization and self-polarization give rise to a significantly greater energy component than the repulsive Coulombic interaction. The formation energy, however, represents a very small fraction of the total energy in the system, becoming negligible as $N \to \infty$.

5. DISCUSSION

As a system, the confining interactions give rise to polarization and self-polarization energy components. The interactions that counter confinement are Coulombic in nature. The formation energy represents an interaction similar to the Coulombic interaction as it results from the polarization of numerous dipoles within the medium. In regions where the medium cannot balance this interaction, a net charge appears — at the interface. As the number of excess electrons increases, the energy needed to form a charged dielectric structure significant enough to confine all the electrons becomes a smaller fraction of the total energy in the system as shown in Figure 2. On the other hand, the

energy resulting from interactions that counter the repulsive Coulomb interaction is greater than the Coulombic interaction energy itself (Figure 2).

The formation energy may arguably be related to a potential gate or switch that must be raised or lowered to allow electrons to flow in and out of the system. This energy may then be unrecoverable as it is inherent to the system. Knowledge of the formation energy is necessary to determine the true equilibrium configuration of electrons in the system.

In an unconnected system free of metallic contacts the equipotential landscape is a complex three-dimensional surface dependent only upon the spatial symmetry of the electrons themselves. Such is the case with biological, chemical and atomic systems. This inherently makes for a difficult task in defining capacitance. However, the monophasic capacitance definition reveals the significance of the underlying principles of electronic charge interactions, the crucial role of formation energy, and the imperative exploitation of the symmetry properties of discrete charges. The monophasic definition resorts to an average electrostatic potential to circumvent complications of evaluating a complex three-dimensional equipotential surface.

Here we have demonstrated the value of the DCD model by illustrating the significance of the electrostatic formation energy. As nanoscale devices are fabricated and utilized, the spatial symmetry properties of electrons becomes central to an understanding of such devices. Unfortunately, the conventional capacitance expression is misleading, as it does not bear out these symmetry properties.

While metallic contacts are universally applied to provide well-defined equipotential surfaces in order to rigidly define input and output, it is not generally recognized that they also serve to minimize effects due to the discreteness of electronic charge. Indeed, metallic contacts effectively diminish the potential of fully exploiting symmetries afforded by the discrete nature of electrons. As electronic devices shrink to the nanoscale, it is doubtful that metallic contacts will remain in their traditional role. An example is the impossibility of applying coaxial shielding of transmission lines. Perhaps the answer lies in replacing metallic contacts by super-ionic dielectrics in nanoscale devices. However, it is remarkable that living organisms function exceptionally well in the absence of metallic contacts. Perhaps the motivation of using an average electrostatic potential in our definition of capacitance may be universally adopted within studies of living organisms.

REFERENCES

1. Zhu, J., T. LaFave Jr., and R. Tsu, "Classical capacitance of few-electron dielectric spheres," *Microelectron. J.*, Vol. 37, 1293–1296, 2006.
2. LaFave Jr. T. and R. Tsu, "Capacitance: A property of nanoscale materials based on spatial symmetry of discrete electrons," *Microelectron. J.*, doi:10.1016/j.mejo.2007.07.105, 2007.
3. Thomson, J. J., "The structure of the atom," *Philos. Mag.*, Vol. 7, 237, 1904.
4. Thomson, J. J., "The electron in chemistry," Franklin Institute, Philadelphia, 1923.
5. Erber, T. and G. M. Hockney, "Equilibrium configurations of N equal charges on a sphere," *J. Phys. A*, Vol. 24, L1369–L1377, 1991.
6. Altschuler, E. L., T. J. Williams, E. R. Ratner, F. Dowla, and R. Wooten, "Method of constrained global optimization," *Phys. Rev. Lett.*, Vol. 74, 2671–2674, 1995.
7. LaFave Jr., T., "The classical electrostatic periodic table, capacitance of few electron dielectric spheres and a novel treatment of one- and two-electron finite quantum wells," Ph.D. Dissertation, University of North Carolina, Charlotte, 2006.
8. Böttcher, C. J. F., *Theory of Electric Polarization, Vol. I: Dielectrics in Static Fields*, Elsevier, 1973.
9. Osborn, K. D., M. W. Keller, and R. P. Mirin, "Single-electron transistor spectroscopy of InGaAs self-assembled quantum dots," *Physica E*, Vol. 21, 501–505, 2004.
10. Tsu, R., *Superlattice to Nanoelectronics*, Ch. 11, Elsevier, 2005.
11. Grabert H. and M. H. Devoret, *Single Charge Tunneling*, Plenum Press, New York, 1992.

Quantum Wires and Field Theory

B. Bellazzini[1], **M. Mintchev**[2], **and P. Sorba**[3]

[1]INFN and Scuola Normale Superiore, Piazza dei Cavalieri 7, Pisa 56127, Italy
[2]INFN and Dipartimento di Fisica, Università di Pisa, Largo Pontecorvo 3, Pisa 56127, Italy
[3]LAPTH, 9, Chemin de Bellevue, BP 110, F-74941 Annecy-le-Vieux cedex, France

Abstract— Quantum graphs are networks of one-dimensional wires connected at nodes. The interest for such structures increased these recent times with the development of nano-scale technology. We focus our attention on star graphs made of n edges with one junction. The related bosonic fields propagate in the bulk, either freely or submitted to a four-fermion interaction, and interact at the vertex, which can be considered as a defect. Hereafter, a quantum field theoretical framework is developed and applied to the computation of physical quantities, such that the electric and spin conductance. More precisely, our approach combines results from the spectral theory of the Schrodinger operator on quantum graphs with an algebraic technique for dealing with quantum fields with defects (impurities). At the vertex, all possible interactions preserving unitarity are taken into account, but special attention is given to scale-invariant ones, which lead to the critical properties of the system. Then bosonisation and vertex operators on quantum graphs are investigated to solve exactly, for scale invariant boundary conditions, the four-fermion bulk interaction (Tomonaga-Luttinger model). At this point, we are in position to derive the charge and spin transport, and establish a simple relationship among them.

1. INTRODUCTION

First applied for describing electron transport in organic molecules, quantum wires appeared more recently in the study of one-dimensional electron gas. These networks of junctions and wires are very thin — of the order of few nanometers. For this reason such devices, which are now created and tested in laboratory, are fairly well modeled by graphs. Among others, the most interesting physical problems concern the charge and the spin transport. The fact that these problems can be investigated in depth by means of quantum field theory (QFT) is our basic motivation for the construction and the study of quantum fields on graphs, such a program being also interesting from the mathematical point of view.

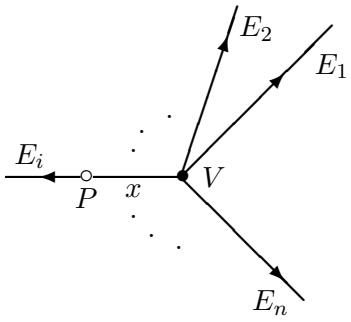


Figure 1: A star graph Γ with n edges.

For simplicity, we consider below graphs of the type shown in Fig. 1. They are called star graphs and represent the building blocks for generic graphs. Each point P in the bulk $\Gamma \setminus V$ of the graph Γ belongs to some edge E_i and can be parametrized by the pair of coordinates (x, i), where $x > 0$ is the distance of P from the vertex (junction) V along E_i. The embedding of Γ and the relative position of its edges in the ambient space are irrelevant in what follows.

Hereafter, we review some previous works [1], [2], [3] on QFT on star graphs. The paper is basically constituted by two parts. We start in Section 2 by setting the framework, describing some specific properties of quantum fields on graphs related to symmetries and boundary conditions. Special attention is devoted to the scale invariant boundary conditions, which define in the physical

terminology the so called critical points of the system. Our construction is based on the point-like character of the boundary interactions at the vertex, the theory of self-adjoint extension of Hermitian operators on graphs [4] and the algebraic technique [5] for dealing with defects in QFT. Then this approach is used in Section 3 for solving the Tomonaga-Luttinger model on Γ and answering some basic questions left open in the pioneering work by Nayak and collaborators [6]. We derive here both the charge and spin conductance and establish a simple relationship among them. We conclude briefly mentioning some perspectives before limiting as much as possible the set of references, due to the lack of allowed space (the interested reader will find a more complete list in [3]).

2. THE FRAMEWORK

2.1. The Field φ and Its Dual $\widetilde{\varphi}$

The massless scalar field φ on Γ is defined by the equation of motion

$$\left(\partial_t^2 - \partial_x^2\right)\varphi(t,\, x,\, i) = 0, \qquad x > 0,\ i = 1,\, \ldots,\, n, \tag{1}$$

the initial conditions (canonical commutation relations)

$$[\varphi(0,\, x_1,\, i_1),\, \varphi(0,\, x_2,\, i_2)] \,=\, 0, \tag{2}$$

$$[(\partial_t\varphi)(0,\, x_1,\, i_1),\, \varphi(0,\, x_2,\, i_2)] \,=\, -\mathrm{i}\delta_{i_1 i_2}\delta(x_1 - x_2), \tag{3}$$

and the vertex boundary condition

$$\sum_{j=1}^{n} \left[A_{ij}\varphi(t,\, 0,\, j) + B_{ij}(\partial_x\varphi)(t,\, 0,\, j)\right] = 0, \qquad \forall t \in \mathbb{R},\ i = 1,\, \ldots,\, n, \tag{4}$$

A and B being in general two $n \times n$ complex matrices. Clearly the pairs $\{A,\, B\}$ and $\{CA,\, CB\}$, where C is any invertible matrix, define equivalent boundary conditions. The results of [4] for the operator $-\partial_x^2$ on Γ imply that the Hamiltonian of the system is self-adjoint, provided that

$$AB^* - BA^* = 0, \tag{5}$$

where we denote by * the Hermitian conjugation, and the composite matrix $(A,\, B)$ has rank n. The general solution of these constraints on $\{A,\, B\}$ is

$$A = C(\mathbb{I} - U) \ \text{ and } \ B = -\mathrm{i}C(\mathbb{I} + U) \tag{6}$$

where C is an arbitrary invertible matrix and U is any unitary matrix.

The problem of quantizing (1) with the initial conditions (2, 3) and the boundary condition (4) has a unique solution. It can be written in the form

$$\varphi(t,\, x,\, i) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi\sqrt{2|k|}} \left[a_i^*(k)\mathrm{e}^{\mathrm{i}(|k|t - kx)} + a_i(k)\mathrm{e}^{-\mathrm{i}(|k|t - kx)}\right], \tag{7}$$

where $\{a_i(k),\, a_i^*(k) \,:\, k \in \mathbb{R}\}$ generate the reflection-transmission (boundary) algebra [5] corresponding to the boundary condition (4). This is an associative algebra $\mathcal{A}$ with identity element $\mathbf{1}$, whose generators $\{a_i(k),\, a^{*i}(k) \,:\, k \in \mathbb{R}\}$ satisfy the commutation relations

$$a_{i_1}(k_1)\, a_{i_2}(k_2) - a_{i_2}(k_2)\, a_{i_1}(k_1) = 0, \tag{8}$$

$$a_{i_1}^*(k_1)\, a_{i_2}^*(k_2) - a_{i_2}^*(k_2)\, a_{i_1}^*(k_1) = 0, \tag{9}$$

$$a_{i_1}(k_1)\, a_{i_2}^*(k_2) - a_{i_2}^*(k_2)\, a_{i_1}(k_1) = 2\pi\left[\delta_{i_1 i_2}\delta(k_1 - k_2) + S_{i_1 i_2}(k_1)\delta(k_1 + k_2)\right]\mathbf{1}, \tag{10}$$

and the constraints

$$a_i(k) = \sum_{j=1}^{n} S_{ij}(k)a_j(-k), \qquad a_i^*(k) = \sum_{j=1}^{n} a_j^*(-k)S_{ji}(-k). \tag{11}$$

Let us mention that two kinds of representations of $\mathcal{A}$ have been constructed and applied to physical situations: the Fock representation $\mathcal{F}(A)$ and the Gibbs representation $\mathcal{G}_\beta(A)$ at inverse

temperature $\beta \sim 1/T$. This last one has been used to determine the Casimir energy density on Γ and the correction to the Stefan-Boltzmann law, due to the interaction at the vertex V [1].

Now, the S-matrix in (10, 11) equals [4]

$$S(k) = -[(\mathbb{I} - U) + k(\mathbb{I} + U)]^{-1}[(\mathbb{I} - U) - k(\mathbb{I} + U)] \tag{12}$$

and has the following simple physical interpretation: the diagonal element $S_{ii}(k)$ is the reflection amplitude on the edge E_i, whereas $S_{ij}(k)$ with $i \neq j$ is the transmission amplitude from E_i to E_j. Being fully characterized by the S-matrix (12), the vertex V can be viewed as a sort of point-like defect. It is not surprising therefore that the algebra $\mathcal{A}$, appearing in the context of QFT with boundaries or defects [5], applies in the present case as well. $\mathcal{A}$ provides a simple algebraic description of the boundary value problem at hand and defines convenient coordinates in field space.

We assume in what follows that U is such that

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}k}{2\pi} \mathrm{e}^{\mathrm{i}kx} S_{ij}(k) = 0, \qquad x > 0. \tag{13}$$

This condition guarantees the absence of bound states and implies the canonical commutation relations (2, 3).

By construction (12) is unitary $S(k)^* = S(k)^{-1}$ and satisfies Hermitian analyticity: $S(k)^* = S(-k)$.

The dual field $\widetilde{\varphi}$ is defined in terms of φ by the relations

$$\partial_t \widetilde{\varphi}(t,\, x,\, i) = -\partial_x \varphi(t,\, x,\, i), \quad \partial_x \widetilde{\varphi}(t,\, x,\, i) = -\partial_t \varphi(t,\, x,\, i), \quad x > 0,\ i = 1,\, \ldots,\, n, \tag{14}$$

which imply that

$$\left(\partial_t^2 - \partial_x^2\right) \widetilde{\varphi}(t,\, x,\, i) = 0, \qquad x > 0,\ i = 1,\, \ldots,\, n \tag{15}$$

as well. The solution is

$$\widetilde{\varphi}(t,\, x,\, i) = \int_{-\infty}^{\infty} \frac{\mathrm{d}k\, \varepsilon(k)}{2\pi\sqrt{2|k|}} \left[a_i^*(k) \mathrm{e}^{\mathrm{i}(|k|t - kx)} + a_i(k) \mathrm{e}^{-\mathrm{i}(|k|t - kx)}\right], \tag{16}$$

where $\varepsilon(k)$ is the sign function. Both φ and $\widetilde{\varphi}$ are local fields, but we stress that they are not relatively local. This feature is fundamental for bosonization.

The invariance of the equations of motion under time translation implies the conservation of the energy momentum tensor $\{\theta_{tt}(t,\, x,\, i),\, \theta_{tx}(t,\, x,\, i)\}$ and the associated Kirchhoff's rule

$$\sum_{i=1}^{n} \theta_{tx}(t,\, 0,\, i) = 0 \tag{17}$$

is satisfied by construction. This equation guarantees energy conservation (no dissipation)and represents a meeting point between boundary conformal field theories and the concept of scale invariance on a star graph with n edges, n bigger than 1.

The equations of motion (1, 14) are also invariant under the transformations

$$\varphi(t,\, x,\, i) \longmapsto \varphi(t,\, x,\, i) + c, \quad \widetilde{\varphi}(t,\, x,\, i) \longmapsto \widetilde{\varphi}(t,\, x,\, i) + \widetilde{c}, \quad c,\, \widetilde{c} \in \mathbb{R}, \tag{18}$$

which implies the conservation of the currents

$$j_\nu(t,\, x,\, i) = \partial_\nu \varphi(t,\, x,\, i), \qquad \widetilde{j}_\nu(t,\, x,\, i) = \partial_\nu \widetilde{\varphi}(t,\, x,\, i), \qquad \nu = t,\, x. \tag{19}$$

These currents have a deep physical meaning. In fact, in the framework of bosonization j_ν and $\widetilde{j}_\nu$ control the charge and spin transport respectively. It is therefore crucial to check the relative Kirchhoff's rules. Using the solution (7) and the constraints (11) one finds that j_x satisfies Kirchhoff's rule if and only if

$$\sum_{j=1}^{n} S_{ij}(k) = 1, \qquad \forall\, i = 1,\, \ldots,\, n,\ k \in \mathbb{R}. \tag{20}$$

Analogously, for $\widetilde{j}_x$ one finds

$$\sum_{j=1}^{n} S_{ij}(k) = -1, \qquad \forall\, i = 1, \ldots, n,\ k \in \mathbb{R}. \tag{21}$$

From (20) and (21) one draws the important conclusion that Kirchhoff's rule cannot be satisfied for both j_ν and $\widetilde{j}_\nu$ simultaneously.

2.2. Scale Invariance

Our system is scale invariant if and only if

$$S(k) = S(\rho^{-1}k) \quad \forall k,\, \rho \in \mathbb{R}, \tag{22}$$

which, combined with (12), implies that S is k-independent. Then

$$S^* = S^{-1} \qquad S^* = S \qquad S^t = S \tag{23}$$

The classification of these S-matrices (critical points) is now a simple matter. Indeed, one easily deduces from above that the eigenvalues of S are ± 1. Let us denote by p the number of eigenvalues -1. The values $p = 0$ and $p = n$ correspond to the familiar Neumann ($S_N = \mathbb{I}$) and Dirichlet ($S_D = -\mathbb{I}$) boundary conditions respectively. For $0 < p < n$ the scale-invariant S-matrices depend on $p(n-p) \geq 1$ parameters.

The conditions (20) and (21) have a simple impact on the scale invariant S-matrices. Requiring (20), one eliminates the Dirichlet point ($p = n$) and is left only with $p(n-p-1)$ parameters for $0 < p < n$.

The case $n = 3$ is an instructive example. Indeed, imposing (20) one is led for $p = 2$ to the isolated critical point

$$S_2 = \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}, \tag{24}$$

which is invariant under edge permutations. For $p = 1$, one gets instead one-parameter family

$$S_1(\alpha) = \frac{1}{1+\alpha+\alpha^2}\begin{pmatrix} -\alpha & \alpha(\alpha+1) & 1+\alpha \\ \alpha(\alpha+1) & \alpha+1 & -\alpha \\ \alpha+1 & -\alpha & \alpha(\alpha+1) \end{pmatrix}, \qquad \alpha \in \mathbb{R}, \tag{25}$$

which is not invariant under edge permutations for generic α.

Analogous considerations apply to the case when, instead of the Kirchhoff's rule (20), one imposes (21). In conclusion, we see that the currents j_ν and $\widetilde{j}_\nu$ nicely illustrate the obstructions which may appear when symmetries on $\mathbb{R}$ are lifted to Γ.

2.3. Vertex Operators

We introduce first the right and left chiral fields

$$\varphi_{i,\,R}(t-x) = \varphi(t,\,x,\,i) + \widetilde{\varphi}(t,\,x,\,i), \qquad \varphi_{i,\,L}(t+x) = \varphi(t,\,x,\,i) - \widetilde{\varphi}(t,\,x,\,i). \tag{26}$$

Defining the chiral charges as

$$Q_{i,\,Z} = \frac{1}{4}\int_{-\infty}^{\infty} \mathrm{d}\xi\, \partial_\xi \varphi_{i,\,Z}(\xi), \qquad Z = R,\, L. \tag{27}$$

we introduce a family of vertex operators parametrized by $\zeta = (\sigma,\,\tau) \in \mathbb{R}^2$ and defined by

$$v(t,\,x,\,i;\,\zeta) = z_i\, q(i;\,\zeta) : \exp\left\{\mathrm{i}\sqrt{\pi}\left[\sigma\varphi_{i,\,R}(t-x) + \tau\varphi_{i,\,L}(t+x)\right]\right\} :, \tag{28}$$

where the value of the normalization constant $z_i \in \mathbb{R}$ will be fixed later on,

$$q(i;\,\zeta) = \exp\left[\mathrm{i}\sqrt{\pi}\,(\sigma Q_{i,\,R} - \tau Q_{i,\,L})\right], \tag{29}$$

and $: \cdots :$ denotes the normal product in the algebra $\mathcal{A}$. For later use we take any $\zeta = (\sigma > 0, \tau)$ with $\sigma \neq \pm\tau$ and set

$$\zeta' = (\tau, \sigma). \tag{30}$$

Then we define

$$\mathcal{V}(t,\,x,\,i;\,\zeta) = \eta_i v(t,\,x,\,i;\,\zeta), \qquad \mathcal{V}(t,\,x,\,i;\,\zeta') = \eta'_i v(t,\,x,\,i;\,\zeta'), \tag{31}$$

where $\{\eta_i,\,\eta'_i\}$ are the so called Klein factors.

3. THE FOUR-FERMION INTERACTION ON Γ

In this section we introduce and investigate non-trivial bulk interactions on Γ. We focus on the Tomonaga-Luttinger (TL) model which captures the universal features of a wide class of one-dimensional quantum many-body systems called Luttinger liquids. More precisely, it will be enough for our purposes to analyze the (TL) model for special values of the coupling constants and Fermi velocity (namely $g_+ = -g_- \equiv g\pi > 0$ and $v_F = 1$). The classical equations of motion of this system, known as a Thirring model, can be written in the following matrix form

$$i(\gamma_t\partial_t - \gamma_x\partial_x)\psi(t,\, x,\, i) = 2\pi g[\gamma_t J_t(t,\, x,\, i) - \gamma_x J_x(t,\, x,\, i)]\psi(t,\, x,\, i), \tag{32}$$

where

$$\psi(t,\, x,\, i) = \begin{pmatrix}\psi_1(t,\, x,\, i)\\ \psi_2(t,\, x,\, i)\end{pmatrix}, \quad \gamma_t = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \gamma_x = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \tag{33}$$

and

$$J_\nu(t,\, x,\, i) = \overline{\psi}(t,\, x,\, i)\gamma_\nu\psi(t,\, x,\, i), \qquad \overline{\psi} \equiv \psi^*\gamma_t. \tag{34}$$

is the conserved vector current. The number of edges n is arbitrary.

Let us at this point mention that our approach is different from the one of [**?**], where using instanton gas expansion and strong-weak coupling duality on a star graph with $n = 3$ edges, Nayak and collaborators established the existence of a critical point in which the electric conductance G is enhanced[1] with respect to that on the line, namely

$$G = \frac{4}{3}G_{\text{line}}. \tag{35}$$

This inspiring result, which has been confirmed more recently by different authors, raises some interesting open problems:

(i) existence of other critical points and their classification;

(ii) behavior under edge permutations;

(iii) is enhancement of the conductance universal or reduction is possible as well?

(iv) what is the law governing the spin transport?

These questions are answered in the rest of the paper. The main idea behind our analysis is to modify the boundary conditions in such a way that they become linear *after* bosonization. We show that this is indeed possible and, applying the results of Section 2, in our case the model can be solved exactly. All critical points can be classified and their characteristic features are easily investigated.

3.1. The Thirring Model on Γ and Its Solution

For quantizing the above defined Thirring model, we set $\sigma > 0$ and

$$\psi_1(t,\, x,\, i) = \frac{1}{\sqrt{2\pi}}\mathcal{V}(t,\, x,\, i;\, \zeta), \qquad \psi_2(t,\, x,\, i) = \frac{1}{\sqrt{2\pi}}\mathcal{V}(t,\, x;\, \zeta'). \tag{36}$$

In order to have canonical fermions we require

$$\sigma^2 - \tau^2 = 1, \tag{37}$$

implying $\sigma \neq \pm\tau$.

The quantum current J_ν is constructed by point-splitting and satisfies:

$$J_\nu(t,\, x,\, i) = -\frac{1}{(\sigma+\tau)\sqrt{\pi}}\partial_\nu\varphi(t,\, x,\, i) = -\frac{1}{(\sigma+\tau)\sqrt{\pi}}j_\nu(t,\, x,\, i), \tag{38}$$

Because of (38) the quantum equation of motion takes the form

$$i(\gamma_t\partial_t - \gamma_x\partial_x)\psi(t,\, x,\, i) = -\frac{2g\sqrt{\pi}}{(\sigma+\tau)} : (\gamma_t\partial_t\varphi - \gamma_x\partial_x\varphi)\,\psi : (t,\, x,\, i). \tag{39}$$

[1] A possible explanation [**?**] of this result is based on the so-called Andreev reflection.

Now, using the vertex realization (36) of ψ, one easily verifies that (39) is satisfied provided that

$$\tau(\sigma+\tau)=g. \tag{40}$$

Combining (37) and (40) one determines σ and τ in terms of the coupling constant:

$$\sigma=\frac{1+g}{\sqrt{1+2g}}>0, \qquad \tau=\frac{g}{\sqrt{1+2g}}. \tag{41}$$

We generated above a solution of the Thirring model on Γ satisfying the boundary conditions

$$J_x(t,\,0,\,i)=-\sum_{k=1}^{n}S_{ik}J_x(t,\,0,\,k), \tag{42}$$

This boundary condition (42) describes the dissipationless splitting of the charge current at the junction $x=0$. It is *quadratic* in the fermion field ψ_α, but *linear* in φ. This fundamental difference with respect to the boundary conditions adopted in [**?**] is the crucial novelty allowing to solve the model exactly.

Let us discuss now the spin-transport. It is well known that there is no true spin in one space dimension (and therefore on Γ) because there are no space-rotations. Nevertheless, one can associate a "spin" $-\frac{1}{2}$ to ψ_1 and $\frac{1}{2}$ to ψ_2. This assignment is not only formal, because there exists a conserved current describing the transport of this quantum number. In fact, one can also construct an axial current such that

$$\widetilde{J}_\nu(t,\,x,\,i)=-\frac{1}{2(\sigma-\tau)\sqrt{\pi}}\partial_\nu\widetilde{\varphi}(t,\,x,\,i)=-\frac{1}{2(\sigma-\tau)\sqrt{\pi}}\widetilde{j}_\nu(t,\,x,\,i), \tag{43}$$

which is indeed conserved. Moreover,

$$\Big[\widetilde{J}_t(t,\,x,\,i),\psi_\alpha(t,\,y,\,j)\Big]=\begin{cases}-\frac{1}{2}\delta(x-y)\delta_{ij}\psi_1(t,\,y,\,j), & \alpha=1,\\ \frac{1}{2}\delta(x-y)\delta_{ij}\psi_2(t,\,y,\,j), & \alpha=2.\end{cases} \tag{44}$$

We turn finally to the critical points. As already discussed in Section 2.2, the critical points are fully classified. If one requires in addition the conservation of the charge associated with the electric current (38), one must impose the Kirchhoff's rule (20). In this way, besides the Neumann point, the remaining critical points for $n=3$ are S_2 and $S_1(\alpha)$, defined by (24, 25). We will see in the next section that S_2 is precisely the critical point discovered by Nayak and collaborators in [**?**]. The one-parameter family $S_1(\alpha)$ is new and is not invariant under edge permutations for generic α. This statement clarifies points (i) and (ii) at the beginning of this section.

3.2. Charge and Spin Transport

In order to derive the electric and spin conductance, we couple the system to a *classical* external field $A_\nu(t,\,x,\,i)$ by means of the substitution

$$\partial_\nu\longmapsto\partial_\nu+\mathrm{i}A_\nu(t,\,x,\,i) \tag{45}$$

in Eq. (32). The resulting Hamiltonian is time dependent. The conductance can be extracted from the expectation value $\langle J_x(t,\,x,\,i)\rangle_{A_\nu}$, more precisely, from the linear term of the expansion of $\langle J_x(t,\,x,\,i)\rangle_{A_\nu}$ in terms of A_ν. This term can be computed by linear response theory which gives for the conductance tensor

$$G_{ij}=G_{\mathrm{line}}\left(\delta_{ij}-S_{ij}\right), \tag{46}$$

where

$$G_{\mathrm{line}}=\frac{1}{2\pi(\sigma+\tau)^2}=\frac{1}{2\pi(1+2g)} \tag{47}$$

is the conductance on the line $\mathbb{R}$.

The simple expression (46), describing the electric conductance of the Thirring model at a critical point, has a number of remarkable properties. As expected, it satisfies the Kirchhoff's rule

$$\sum_{j=1}^{n}G_{ij}=0, \qquad i=1,\,\ldots,\,n, \tag{48}$$

provided that (20) holds, i.e., the electric charge is conserved. The conductance G_{ii} of the edge E_i is enhanced with respect to $G_{\rm line}$ if $S_{ii} < 0$ and reduced if $S_{ii} > 0$. In particular, at the critical point (24) one reproduces the result (35) of [**?**]. Note that enhanced conductance is associated with attractive Casimir force and vice versa.

The properties (23) of the S-matrix imply $|S_i^i| \leq 1$, leading to the simple bound

$$0 \leq G_{ii} \leq 2G_{\rm line}, \tag{49}$$

where we have used that $G_{\rm line}$ is positive. Another constraint on the diagonal elements of G is the sum rule

$$\mathrm{Tr}G = 2pG_{\rm line}, \tag{50}$$

p being the number of eigenvalues -1 of S.

Inserting the new family of critical points (25) in (46) we conclude that both enhancement and reduction of the conductance with respect to the line are possible. This fact is illustrated by the plots in Fig. 1 and answers the question (iii) from the beginning of this section.

Away of criticality the conductance takes the form

$$G_{ij}(\omega) = G_{\rm line}\left[\delta_{ij} - S_{ij}(\omega)\right], \tag{51}$$

where S is given by (12) and ω is the frequency of the Fourier transform $\widehat{A}_x(\omega, i)$ of the external field $A_x(t,\, i)$ applied to the system. Since $S_{ij}(\omega)$ are in general complex, away from criticality the system admits non-trivial inductance as well. Let us compare for instance a star graph with $n = 2$ edges to an impedance $Z\,(\omega)$, characterized by the condition $Z\,(1) = R + \mathrm{i}L$. The ingoing electric currents $I\,(\omega, i) = \langle J_x(\omega, i)\rangle_{A_\nu}$ can be expressed in the form

$$\begin{pmatrix} I\,(\omega,1) \\ I\,(\omega,2) \end{pmatrix} = \begin{pmatrix} Z^{-1}\,(\omega) & -Z^{-1}\,(\omega) \\ -Z^{-1}\,(\omega) & Z^{-1}\,(\omega) \end{pmatrix} \begin{pmatrix} \widehat{A}_x(\omega,1) \\ \widehat{A}_x(\omega,2) \end{pmatrix}. \tag{52}$$

Confronting this relation with (51) and using the explicit form (12) of $S\,(\omega)$ one finds

$$Z\,(\omega) = R + \mathrm{i}\omega L, \tag{53}$$

so that the most general impedance, which can be obtained for $n = 2$, has a real component $R = G_{\rm line}^{-1}$ and a generic inductive component $\mathrm{i}\omega L$ depending on boundary conditions. Following the above procedure, one can analyze all the admittances characterizing star graphs with $n > 2$ away of criticality.

Finally, to answer the last question (iv), we investigate the spin transport governed by the dual current (43). In this case one should evaluate the expectation value $\langle \widetilde{J}_x(t,\, x,\, i)\rangle_{A_\nu}$. The result for the spin conductance tensor is

$$\widetilde{G}_{ij} = \widetilde{G}_{\rm line}\left(\delta_{ij} + S_{ij}\right), \tag{54}$$

where

$$\widetilde{G}_{\rm line} = \frac{1}{4\pi(\sigma^2 - \tau^2)} = \frac{1}{4\pi}. \tag{55}$$

We see that the spin conductance differs from the electric one. In particular, the conservation of the electric charge (20) spoils the Kirchhoff's rule for $\widetilde{G}_{ij}$. Therefore, the spin is not a conserved quantum number in this case. Alternatively, one can impose the Kirchhoff's rule (21), which guarantees spin conservation but breaks down the charge conservation.

Comparing (46) and (54), one discovers a simple but deep interplay between charge and spin transport: enhancement of the electric conductance corresponds to reduction of the spin conductance and vice versa. This feature clearly shows an effective separation in the dynamics of the charge and spin degrees of freedom of the model.

4. OUTLOOK AND PERSPECTIVES

The content of the above investigation can be generalized in several directions, which can certainly help a better understanding of the physics of quantum wires. Among others, we have in mind the study of generic graphs, integrable and more involved bulk interactions, finite temperature systems and dissipative phenomena at the vertex.

REFERENCES

1. Bellazzini, B. and M. Mintchev, "Quantum fields on star graphs," *J. Phys. A*, Vol. 39, 11101–11117, [arXiv:hep-th/0605036], 2006.
2. Bellazzini, B., M. Mintchev, and P. Sorba, "Bosonization and scale invariance on quantum wires," *J. Phys. A*, Vol. 40, 2485–2507, [arXiv:hep-th/0611090], 2007.
3. Bellazzini, B., M. Burrello, M. Mintchev, and P. Sorba, "Quantum field theory on star graphs," to appear in *Proceedings of Symposia in Pure Mathematics*, [arXiv:hep-th/0801.2852], 2008.
4. Kostrykin, V. and R. Schrader, "Kirchhoff's rule for quantum wires," *J. Phys. A*, Vol. 32, 595–630, 1999.
5. Mintchev, M., E. Ragoucy, and P. Sorba, "Reflection-transmission algebras," *J. Phys. A*, Vol. 36, 10407–10429, [arXiv:hep-th/0303187], 2003.
6. Nayak, C., M. P. A. Fisher, A. W. W. Ludwig, and H. H. Lin, "Resonant multilead point-contact tunneling," *Phys. Rev. B*, Vol. 59, 15 694–15 704, 1999.

Noise Associated with Microwave Intensity Modulation of Semiconductor Lasers

Moustafa Ahmed[1] and Minoru Yamada[2]

[1]Department of Physics, Faculty of Science, Minia University, El-Minia 61519, Egypt
[2]Division of Electrical Engineering and Computer Science
Graduate School on Natural Science and Technology, Kanazawa University
Kakuma-machi, Kanazawa 920-1192, Japan

Abstract— Intensity and frequency noises associated with microwave intensity modulation of semiconductor lasers are theoretically investigated in both time and frequency domains. The modulation dynamics is classified into six distinct types. The noise spectra are higher when the laser emits pulses than when its signal varies continuously with time. Under low-frequency strong modulation, the low-frequency noise is most enhances as the laser emits irregular spike-like pulses.

1. INTRODUCTION

The behavior of semiconductor lasers under high-speed modulation is attractive for both study and application. Small-signal sinusoidal GHz modulation is employed to reduce optical feedback noise in optical disc systems [1]. Under large-signal modulation, laser pulses as short as a few picoseconds can be obtained [2] for use in time resolved dynamical studies and for carrying information at high bit rates in optical communication systems [3]. The noise performance of semiconductor lasers is an important criterion of the performance of such applications. Intensity and frequency noises of semiconductor lasers stem from intrinsic intensity and phase fluctuations associated with quantum transitions of electrons between the valence and conduction bands [4].

Evaluation of relative intensity noise (RIN) and frequency noise (FN) under modulation is based on solving the laser rate equations augmented by a sinusoidal current source and Langevin noise sources. The nonlinear nature of the equations happens to induce inconsistence of the time variation between the optical signal and the exciting electrical signal. Forms of inconsistence include irregular signals, period multiplication, or even chaos [5, 6], which may deteriorate laser noise. Several groups investigated regular and irregular dynamics and determined the corresponding modulation conditions [7–9]. Kao and Lin [9] showed that including the noise sources in the rate equations act as a virtual Hopf precursor for period doubling. Determining the noise properties of the induced regular and irregular dynamics is necessary to help laser-based system designers to choose operation conditions with optimum noise performance. Yamada and Higashi [1] applied a small-signal model to predict RIN and its variation with modulation parameters. However, such a model is unsuitable to investigate laser dynamics under large-signal modulation.

In this paper, we investigate semiconductor laser dynamics under sinusoidal intensity modulation and characterize the corresponding frequency spectra of RIN and FN. We classify these dynamics into six types with distinct dynamic characteristics in both time and frequency domains. These types are "continuous periodic signal (CPS)", "continuous periodic signal with relaxation oscillation (CPSRO)", "continuous periodic signal with period doubling (CPSPD)", "periodic pulse (PP)", "periodic pulse with relaxation oscillation (PPRO)", and "periodic pulse with period doubling (PPPD)". The classification is achieved in terms of the time characteristics of the pulse, phase portrait and frequency spectra of RIN and FN. Variations of the low-frequency (LF) RIN and FN with both modulation frequency and index are also presented. We show that the GHz regime of both RIN and FN is higher when the signal is pulsed than when it varies continuously with time. The LF noise is most enhanced when the laser emits irregular spike-like pulses under low-frequency strong modulation.

In the next section, the theoretical model of laser modulation dynamics is given. The results of classifying laser dynamics and evaluation of noise are given in Section 3. The conclusions appear in Section 4.

2. THEORETICAL MODEL OF LASER MODULATION DYNAMICS

Analog modulation of semiconductor lasers is described by the following stochastic rate equations of the photon number $S(t)$, optical phase $\theta(t)$ and injected electron number $N(t)$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = (A - BS - G_{th})\,S + \frac{a\xi}{V}N + F_S(t) \tag{1}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = 2\pi\Delta v(t) = \frac{\alpha a\xi}{2V}(N - \overline{N})S + F_\theta(t) \tag{2}$$

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \frac{1}{e}I(t) - AS - \frac{N}{\tau_e} + F_N(t) \tag{3}$$

where $\Delta v(t)$ is the shift of the lasing frequency, G_{th} is threshold gain, and $\overline{N}$ is the time average of $N(t)$. A and B are coefficients of linear and nonlinear gain, respectively, and are given in terms of N as

$$A = \frac{a\xi}{V}\,(N - N_g) \tag{4}$$

$$B = \frac{9}{2}\frac{\pi\hbar c}{\varepsilon_0 n_a^2 V\lambda}\left(\frac{\xi\tau_{in}}{\hbar}\right)^2 aR_{cv}^2(N - N_S) \tag{5}$$

a is the tangential gain coefficient, N_g is the electron number at transparency, N_s is an electron number characterizing B, ξ is the field confinement factor in the active layer whose volume is V and refractive index is n_a respectively, α is the linewidth enhancement factor, τ_e is the electron lifetime, R_{cv} is the dipole moment and λ is the emission wavelength. ε_0 and c are the dielectric constant and speed of light in free space, respectively.

The modulating injection current $I(t)$ is assumed sinusoidal consisting of bias and modulation components,

$$I(t) = I_b + I_m\cos(2\pi f_m t) \tag{6}$$

where I_b and I_m are the bias and modulation currents; they define the modulation index $m = I_m/I_b$. f_m is the modulation frequency. The terms $F_S(t)$, $F_\theta(t)$ and $F_N(t)$ are Langevin noise sources describing the intrinsic fluctuations on $S(t)$, $\theta(t)$ and $N(t)$, respectively. These noise sources have Gaussian statistics and are δ-correlated. Techniques of instantaneous generation of these noise sources can be found in [4].

3. NUMERICAL RESULTS AND DISCUSSION

Rate Equations (1)–(3) are solved numerically by the fourth-order Runge-Kutta method. The laser is assumed to be biased above the threshold level I_{th}, $I_b = 3I_{th}$. The integration is achieved over 512 cycles of period $T = 1/f_m$ using an integration step of $\Delta t \sim 5$ ps. Both RIN and FN are calculated as the fast Fourier transforms of fluctuations on $S(t)$ and $\Delta v(t)$, respectively [4], employing the longer half of their temporal trajectories. This ascertains that the laser transients are discarded and the output is stabilized. In the calculations, Fabry-Perot 1.55 µm-InGaAsP lasers are considered whose parameters have the following typical values. $a = 7.85\times10^{-12}\,\mathrm{m^3s^{-1}}$, $\xi = 0.2$, $\alpha = 4$, $V = 60\,\mu\mathrm{m}^3$, $n_a = 3.56$, $N_g = 5.31\times10^7$, $N_s = 4.05\times10^7$, $\tau_e = 2.83$ ns. $R_{cv}^2 = 8.575\times10^{-57}$ $\mathrm{C^2m^2}$, and $G_{th} = 7.817\times10^{10}\,\mathrm{s^{-1}}$. The corresponding threshold current is $I_{th} = 3.17$ mA. The relaxation frequency of the unmodulated laser is $f_{r0} = 5.14$ GHz, which lies in the microwave region of the electromagnetic spectrum.

3.1. Dynamic Types of Laser Modulated Signal

Characteristics of the investigated dynamic types of modulation are illustrated in Fig. 1. The figure illustrates the time variation of $S(t)$, ($S(t)$ vs $N(t)$)-phase portrait, and spectra of RIN and FN. Both $S(t)$ and $N(t)$ are normalized by their biased values $S_b = 5\times10^5$ and $Nb = 5.61\times10^7$, respectively. Figs. 1(a)–(d) correspond to the CPS type which occupies the lower range of m at a given frequency f_m. Fig. 1(a) shows that $S(t)$ varies continuously and regularly, and the portrait 1(b) indicates a single loop whose thick border is manifest of fluctuations. The spectra of RIN and FN in Figs. 1(c) and (d), respectively, are characterized by the frequency $f_m = 1.2f_{r0}$ with a much lower noise peak at f_{r0}. The LF-part of noise is flat with LF-RIN = −151 dB/Hz and LF-FN = 1.1 MHz.

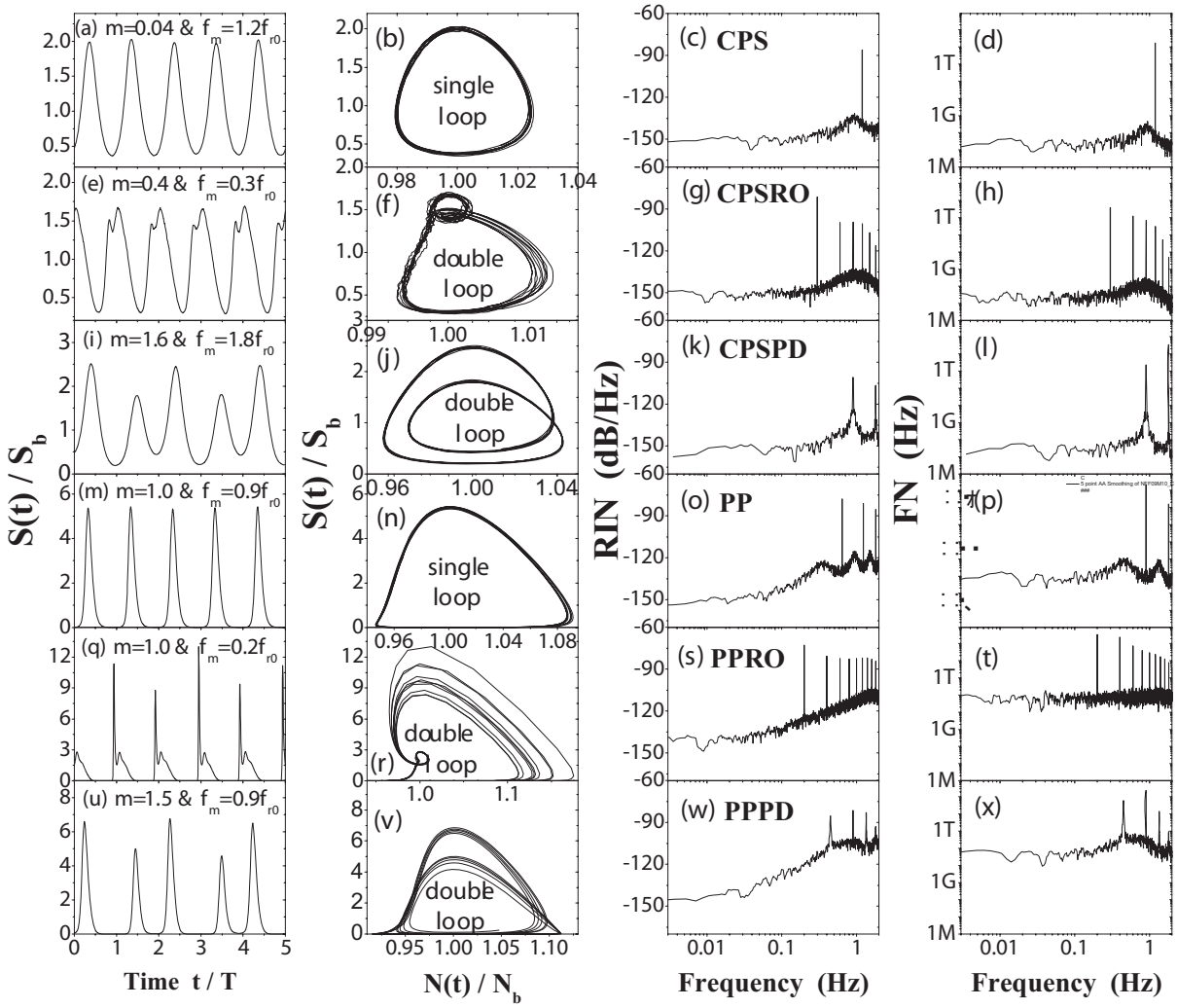


Figure 1: Characterization of the modulation types: (a)–(d) CPS, (e)–(h) CPSRO, (i)–(l) CPSPD, (m)–(p) PP, (q)–(t) PPRO, and (u)–(x) PPP, by time variation of $S(t)$, ($S(t)$–$N(t)$) portrait, and spectra of RIN and FN, respectively.

When f_m is low, the CPS type is developed to CPSRO with the increase of m. This brings the gain under G_{th} in the lower cycles of current exciting relaxation oscillation over the long period T. This type is characterized in Figs. 1(e)–(h). The time trajectory of $S(t)$ is superposed by a relaxation sub-peak, and the phase portrait indicates a main loop attached with a smaller one referring to the superposing sub-peak. The noise spectra of Figs. 1(g) and (h) are characterized by both f_m and f_{r0}. LF-RIN and LF-FN are little higher than those of the CPS type.

The third type characterizing continuous signals is the CPSPD type. It dominates the operating regime of high-frequency ($f_m > 1.3f_{r0}$) and strong ($m \geq 1.2$) modulation. As Fig. 1(i) shows, the signal exhibits period doubling, which is identified in the phase portrait of Fig. 1(j) with the double loop. The spectra of RIN and FN in Figs. 1(k) and (l), respectively, are characterized by f_m and its half-harmonic. The low-frequency noise components are almost in the same level of the CPS and CPSRO types.

The laser intensity happens to discontinue giving rise to regular and irregular pulses. The PP type is preferred to generate short pulses. It appears when $0.7f_{r0} \leq f_m \leq 1.1f_{r0}$ with $m > 0.5$. The latter induces a switching effect associated with turn-on delay of $S(t)$, and the former corresponds to modulation period T comparable to or shorter than the setting time of the relaxation oscillation. This combined effect generates the discontinuous pulses. Fig. 1(m) shows the periodic pulses of $S(t)$. The portrait 1(n) refers to a single loop; it is broad around $S(t) \approx 0$ and gets narrower with the increase of $S(t)$. As shown in Figs. 1(o) and (p), the GHz regime of the noise is higher than those of continuous signals. On the other hand, the LF-RIN is less than $-150\,\mathrm{dB/Hz}$ at $0.002f_{r0}$,

and increases with frequency. The LF-FN is flat with a level as high as 1 GHz indicating worse coherency.

The noisiest type is PPRO which follows the CPSRO type with the increase of m beyond 0.6. It is characterized by periodic and discontinuous relaxation oscillations. This is shown in Fig. 1(q), which demonstrates spike-like pulsation. The phase portrait 1(s) shows a big loop representing the spikes and a child loop referring to the sub-relaxation oscillation peaks. The influence of noise sources appears in Fig. 1(q) as unequal spikes and in Fig. 1(s) as different paths (attractors) of the big loop. This type corresponds to the spike generation predicted by Lee et al. [10]. The GHz regime of noises in Figs. 1(s) and (t) are highest showing peaks at the modulation frequency $f_m = 0.2f_{r0}$ and its multiples. The LF-RIN increases with frequency from RIN< -140 dB at $0.002f_{r0}$. FN is flat with a level as high as 100 GHz indicating deterioration of laser coherency.

Figures 1(u)–(x) characterize the PPPD type. It happens to appear under high-frequency ($f_m \geq f_{r0}$) and strong ($m > 1$) modulation. As given in Fig. 1(u), the pulses of $S(t)$ has period doubling. The phase portrait 1(v) shows a double-loop referring to the two unequal peaks of $S(t)$. The multiple paths of each loop are manifest of laser fluctuations. Both noise spectra in Figs. 1(w) and (x) are characterized by $f_m = 0.9f_{r0}$ as well as its half-harmonic. The RIN spectrum is high in the GHz regime, but its LF-components increase with frequency from less than -140 dB at $0.002f_{r0}$. Similar to the PPRO type, the LF-part of FN is flat with levels about 100 GHz. The simulated period doubling dynamics were observed in experiment by Henery et al. [8].

3.2. Variation of Low-frequency Noise with Modulation Index and Frequency

Variations of the LF-components of both RIN and FN with m are plotted in Figs. 2(a)–(c), when the modulation frequency is much lower than f_{r0} ($f_m = 0.1f_{r0}$), comparable to f_{r0} ($f_m = 0.9f_{r0}$), and higher than f_{r0} ($f_m = 1.6f_{r0}$), respectively. The shown noise data are averaged over frequency components less than $0.08f_{r0}$. Fig. 2(a) shows that under low-frequency modulation, the noises are lowest under both continuous wave (CW) (at very weak modulation) and the CPS type, increase when the CPSRO type is excited, and are enhanced when $m > 0.65$ under the PPRO type. Both LF-RIN and LF-FN are pronounced when $m > 1.4$ because the pulse exhibits the first overshot (spike) of relaxation oscillations accompanied with smaller discontinuous shots. That is, the noise performance and laser coherence are worst in this regime of modulation. When $f_m = 0.9f_{r0}$, Fig. 2(b) shows that both LF-RIN and LF-FN are higher under the PP type than under the CPS type. When $m = 1.2$, the periodic pulses are irregular with unequal heights which pronounces the noise levels as shown in the figure. The noise levels in Fig. 2(c) are comparable to those of continuous signals in Figs. 2(a) and (b). It is clear that the levels of RIN and FN in the regime of PPPD are higher than those in the regimes of CPS and CPSRO.

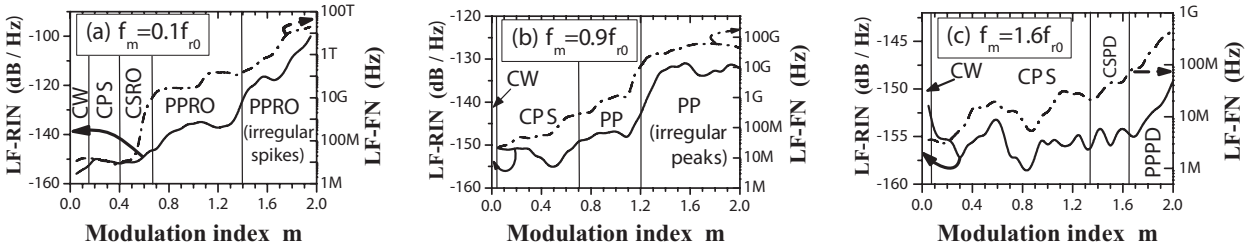


Figure 2: Variations of the average LF-RIN and LF-FN with m when (a) $f_m = 0.1f_{r0}$, (b) $f_m = 0.9f_{r0}$, and (c) $f_m = 1.6f_{r0}$.

4. CONCLUSIONS

RIN and FN of semiconductor lasers are characterized under microwave sinusoidal intensity modulation. The noise spectra are higher when the laser emits pulses than when its signal varies continuously with time. The LF-components of FN are almost flat. The LF-RIN is also flat when the signal is continuous. When the signal is pulsed, the LF-RIN increases with frequency. The LF-noise is lowest when the signal is continuous and periodic under weak and moderate modulation. The LF-noise is most enhanced and the laser coherency deteriorates under low-frequency ($f_m < 0.4f_{r0}$) strong ($m > 0.6$) modulation where the laser emits irregular spike-like pulses.

REFERENCES

1. Yamada, M. and T. Higashi, "Mechanism of the noise reduction method by suppression of high-frequency current for semiconductor injection laser," *IEEE J. Quantum Electron.*, Vol. 27, No. 3, 380–388, 1991.
2. Bimberg, D., K. Ketterer, H. E. Scholl, and H. P. Vollmer, "Generation of 4ps light pulses from directly modulated V-groove lasers," *Electron. Lett.*, Vol. 20, No. 8, 343–345, 1984.
3. Makeyev, O. N., Y. A. Zarkevitch, and V. I. Smirnov, "Analysis of the basic characteristics of high-speed fiber-optic transmission systems," *Telecom. Radio Eng.*, Vol. 45, 41–44, 1990
4. Ahmed, M., M. Yamada, and M. Saito, "Numerical modeling of intensity and phase noises in semiconductor lasers," *IEEE J. Quantum Electron.*, Vol. 37, No. 12, 1600–1610, 2001.
5. Winful, H. G., Y. C. Chen, and J. M. Liu, "Frequency locking, quasiperiodicity, and chaos in modulated self-pulsing semiconductor lasers," *Appl. Phys. Lett.*, Vol. 48, No. 10, 616–618, 1986.
6. Hori, Y., H. Serizawa, and H. Sato, "Chaos in a directly modulated semiconductor laser," *J. Opt. Soc. Am. B*, Vol. 5, No. 5, 1128–1133, 1988.
7. Tang, M. and S. Wang, "Simulation studies of bifurcation and chaos in semiconductor lasers," *Appl. Phys. Lett.*, Vol. 48, No. 14, 900–902, 1986.
8. Henery, E., L. Chusseau, and J M. Lourtioz, "Dynamic behavior of semiconductor laser under strong sinusoidal modulation: modeling and experiments at 1.3 µm", *IEEE J. Quantum Electron.*, Vol. 26, No. 4, 633–41, 1990.
9. Kao, Y. H. and H. T. Lin, "Persistent properties of period doubling in directly modulated semiconductor lasers," *Phys. Rev. A*, Vol. 48, No. 3, 2292–2298, 1993.
10. Lee, C. H., T. H. Yoon, and S. Y. Shin, "Period doubling and chaos in a directly modulated laser diode," *Appl. Phys. Lett.*, Vol. 46, No. 1, 95–97, 1984.

Temperature Effects Induced a Radically Different Behavior on the Transport Properties of Si Delta-doped GaAs Quantum Wells

L. M. Gaggero-Sager[1] and I. Rodriguez-Vargas[2]

[1]Universidad Autónoma del Estado de Morelos, Ave. Universidad 1001, 62209 Cuernavaca, México
[2]Unidad Académica de Física, Universidad Autónoma de Zacatecas
Calzada Solidaridad Esquina con Paseo la Bufa S/N, 98060 Zacatecas, ZAC., México

Abstract— The electronic structure of a delta-doped quantum well of Si in GaAs matrix is studied for different temperatures. The calculation is carried out self-consistently in the framework of the Hartree approximation. The energy levels and the mobility trends are reported for various impurity densities. The apparent contradictory temperature dependence of the mobility reported between Zheng and Gurtovoi et al. [1, 2] can be explain by means of the temperature variation of the electronic structure. Values of 1.9 and 0.9 are obtained for the ratio of the mobility at 300 K and 77 K corresponding to impurity densities of $5 \times 10^{12}\,\mathrm{cm}^{-2}$ and $3 \times 10^{12}\,\mathrm{cm}^{-2}$, respectively, in excellent agreement with the experimental ones, 2.1 and 0.85.

1. INTRODUCTION

There are several experimental works in which an enhancement in the transport properties of double and triple delta-doped quantum wells in different materials is reported [1–5]. Moreover, there are apparent contradictory reports in Si delta-doped GaAs wells [1, 2]. In one case, Zheng et al. [1] measured the mobility of single and double delta-doped wells as a function of temperature finding an increase in the mobility by a factor of 2.5 from 77 K to 300 K. This result was quite surprising, since in this kind of systems is expected a broad mobility peak near the liquid nitrogen temperature [6–8]. Not only is amazing the increase in the mobility, it is also amazing the factor, 2.5. Motivated for the unexpected results of Zheng and collaborators [1], Gurtovoi and coworkers [2] accomplished transport measurements also in single and double delta-doped systems, however they never found the trend observed by Zheng et al. It is important to mention that the only difference between both experiments was the doping level. So, a natural question arises, what could be the explanation for such so contradictory results?, taking into account that a change in the doping level only modify the electronic structure.

On the other hand, from the theoretical point of view there are few works dealing with temperature effects and transport properties in delta-doped systems [9–14]. However, none of these reports is aimed to explain the apparent contradictory results found by the abovementioned experimental groups [1, 2].

The aim of the present paper is to explain both experiments, showing that the changes in the electronic structure are the responsibles of the apparentcontradiction.

2. METHODOLOGY

The calculation is performed in the framework of the Envelope Function Approximation (EFA). Thus,

$$-\frac{\hbar^2}{2m^*}\frac{\mathrm{d}^2}{\mathrm{d}z^2}F_i(z) + V(z)F_i(z) = E_iF_i(z). \tag{1}$$

$F_i(z)$ is the z-dependent envelope function, E_i is the i-th energy level and m^* are the corresponding effective mass.

The spatial band bending is described by the Poisson equation. The energy origin is taken at $z = 0$ ($V(0) = 0$) and the electric field tends to zero when z tends to infinity. Integrating twice and using the Fubini-Lebesgue theorem to change the order of integration, the Poisson equation adopts a simple form, i.e.,

$$V(z) = \frac{4\pi e}{\epsilon_r}\int_0^z (z-\chi)\rho_e(\chi)\mathrm{d}\chi + \frac{2\pi e^2}{\epsilon_r}n_{2D}|z|, \tag{2}$$

n_{2D} is the bidimensional impurity concentration ($\rho_{imp} = n_{2D}\delta(z)$), ρ_e is the free charge density in the well region and ϵ_r is the dielectric constant. The charge density, $\rho_e(z)$, is given by

$$\rho_e(z) = -\frac{em^*k_BT}{\pi\hbar^2}\sum_i \ln\left[1+\exp\left(\frac{E_F-E_i}{k_BT}\right)\right]|F_i(z)|^2, \tag{3}$$

where E_F is the Fermi level and n is the occupation number. For non vanishing temperatures, all levels are, in principle, occupied. Nevertheless, if the system is locally neutral, the continuous part of the spectrum may be considered empty. Then, the sum in (3) is extended only up to the value of the occupation number in the band. At the same time, the charge neutrality requires that $en_{2D} = -\int_{-\infty}^{\infty} \rho_e(\chi)\mathrm{d}\chi$. Thus, the Fermi level is given implicitly by:

$$en_{2D} = \frac{em^* k_B T}{\pi\hbar^2} \sum_{i=1}^{n} \ln\left[1 + \exp\left(\frac{E_F - E_i}{k_B T}\right)\right] \tag{4}$$

These calculations, which consist of the simultaneous solution of the Schrödinger and Poisson equations, tend to be rather awkward and furthermore, on occasion present divergence. They are carried out in an iterative manner, that is, using in each step an input potential, the charge density is calculated and with this, an output potential is calculated. The input potential of the following iteration will be a mixture of the input and output potentials from the step of the previous iteration. One of the key issues in obtaining the quickest convergence, and of course, non-divergence, is this mixture. In this section, we will discuss a method which we have developed for the mixture after some experience in self-consistent calculations. The ideal self-consistent calculation is that in which the input and output potentials are identical. Let us imagine that we construct a bidimensional space formed by input potential at point z on the x-axis and the output potential at the same point z on the y-axis, the self-consistent solution would be when the point were exactly above the line with slope 1. If $V^{m\,in}(z)$ ($V^{m\,out}(z)$) is the input (output) potential in step m of self-consistency, in general the point in the space will not be above the line with slope 1, we will construct a straight line with the two points from the previous steps and with the intersection of these points on the said line, we will obtain the following input potential. In what follows, we will suppose that the zero of energies is placed at the origin. The mixture was made in the following manner.

$$V^{(m+1)\,i}(z) = A(z)V^{m\,in}(z) + (1 - A(z))V^{m\,out}(z) \tag{5}$$

where

$$A(z) = \left\{ \begin{array}{ll} \dfrac{V^{m\,out}(z) \;-\; V^{(m-1)\,out}(z)}{V^{m\,out}(z) \;-\; V^{(m-1)\,out}(z) \;-\; V^{m\,in}(z) \;+\; V^{(m-1)\,in}(z)} & \text{if } \Delta V^{\,m}(z)\,\Delta V^{(m-1)}(z) \;\leq 0 \\ \dfrac{\left|V^{m\,in}(z)\right|}{V^{m\,in}_{\max}(\infty)}\beta & \text{in other case} \end{array} \right\} \tag{6}$$

where $V^{m\,in}_{\max}(\infty)$ is the absolute value of the maximum value of the input potential in step m, $\beta = 1 - \frac{m}{n_{\max}}$ where $n_{\max}$ is the maximum number of iterations proposed to be carried out, it changes, but is usually $n_{\max} \sim 200$ and $\Delta V^m(z) = V^{m\,in}(z) - V^{m\,out}(z)(\Delta V^{(m-1)}(z) = V^{(m-1)\,in}(z) - V^{(m-1)\,out}(z))$. The first part of the equation for $A(z)$, demands that both points (that corresponding to step m and that corresponding to step $(m-1)$ be on both sides of the line in order to avoid divergence. The second part of the equation for $A(z)$, at the origin, $V^{(m+1)\,i}(0) = V^{m\,out}(0)$, so that the input potential of step $m+1$ slowly begins to differentiate itself from the output potential from step m as it distances itself from the origin.

With the self-consistent values of the electronic structure we calculated the relative mobility between $T = 77\,\mathrm{K}$ and $T = 300\,\mathrm{K}$ following the ideas stated in [13, 14],

$$\mu_{rel} = \frac{\mu^{T=300\,\mathrm{K}}}{\mu^{T=77\,\mathrm{K}}} = \frac{k_B T^{Ni} \sum_{i=1}^{n_{Ni}} \ln\left[1 + \exp\left(\dfrac{E_F^{Ni} - E_i^{Ni}}{k_B T}\right)\right] \int |F_i^{Ni}(z)|^2 |z| \mathrm{d}z}{k_B T^R \sum_{i=1}^{n_R} \ln\left[1 + \exp\left(\dfrac{E_F^R - E_i^R}{k_B T}\right)\right] \int |F_i^R(z)|^2 |z| \mathrm{d}z}$$

We have used the following values as input parameters: $m^* = 0.067 m_0$; m_0 being the free electron mass, $\epsilon_r = 12.5$, and $n_{2D} = 3 \times 10^{12}\,\mathrm{cm}^{-2}$ and $n_{2D} = 5 \times 10^{12}\,\mathrm{cm}^{-2}$. As a convergence criterion we have taken the depth of the well to vary less than 0.01 meV for two consecutive steps.

Gurtovoi et al. [2] shown that if the mobility of a group of electrons in higher subbandas is increased, the measured concentration will decrease and be lower than the net subband concentration. These authors also shown that the product of the sheet concentration and mobility is a more realistic criterion to evaluate the sample quality. On the other hand, the double well system at a certain

interwell distance behaves like a single well one, so in the specific case of Gurtovoi and collaborators this criteria is reached at 350 Å. However, in single quantum wells with $n_{2D} = 3 \times 10^{12}\,\mathrm{cm}^{-2}$ they observed that the measured concentration is very different for 77 K and 300 K. In the double delta doped quantum well is also observed a little variation of the concentration for interwell distances of 380 Å and 580 Å. In principle, the carrier concentration should remain unchanged with a variation of the interwell distance since the impurity density is always the same. The observed variations can be due to the experimental error, so, we will compared our results to the Gurtovoi data for the above mentioned interwell distances, 380 Å and 580 Å, since seems to be the data with higher accuracy. For both distances the mobility decreases with temperature, with relative values of 0.81 and 0.88, respectively. Taking into account that the electronic structure is the same in both cases, we can averaged these results and consider that the relative mobility for a single delta-doped well with an impurity concentration of $3 \times 10^{12}\,\mathrm{cm}^{-2}$ is 0.85. The relative mobility in our theoretical calculation gives 0.9.

Zheng et al. [1] grown a variety of double and single delta-doped quantum wells, the impurity concentration in all samples was $n_{2D} = 5 \times 10^{12}\,\mathrm{cm}^{-2}$. Despite all data reported by these authors we are interested only on the results presented for the single delta-doped wells. In this system, the Hall mobility was measuredand an increased of 2.5 from 77 K to 300 K was obtained. In the case of a double delta-doped well with an interwell distance of 400 Å the Hall mobility is enhanced by a factor of 1.7. So, using the same average procedure as in the data reported by Gurtovoi et al. [2], with the purpose of taking into account (of some way) the experimental error, we can talk of an increase in the Hall mobility of 2.1. In our theoretical computations this factor is 1.9.

In summary, the Hall mobility in single delta-doped wells presents an enhancement with temperature of 0.9 and 2.1 for impurity concentrations of $n_{2D} = 3 \times 10^{12}\,\mathrm{cm}^{-2}$ and $n_{2D} = 5 \times 10^{12}\,\mathrm{cm}^{-2}$, respectively. Our results give values of 0.85 and 1.9. All this shown that the delta-doped wells present a completely different behavior for different impurity concentrations. So, the study of these systems can not be generalized readily. Indeed, the report of Gurtovoi et al. [2] assumes that the data presented by Zheng and colleagues [1] can be faulty, however as we already demonstrated here both experimental works are correct and the apparent discrepancies can be explain by the change in the electronic structure.

Si delta-doped GaAs QW's have a quite different behavior with respect to B delta-doped Si QW's [11], contrary to what should be supposed. In the case of B delta-doped Si wells the electronic structure changes dramatically with temperature, for example the number of levels passes from 7 to 13 from cero Kelvin to room temperature, while in the system reported here the trend is opposite, i.e., the number of levels remains practically unchanged, Figs. 1 and 2. The Fermi level (measured from the band bottom) goes down about 70 meV from 0 K to 300 K for B delta-doped Si wells, while in Si delta-doped GaAs wells this dropping is only the half, 30 meV. Two other features that present

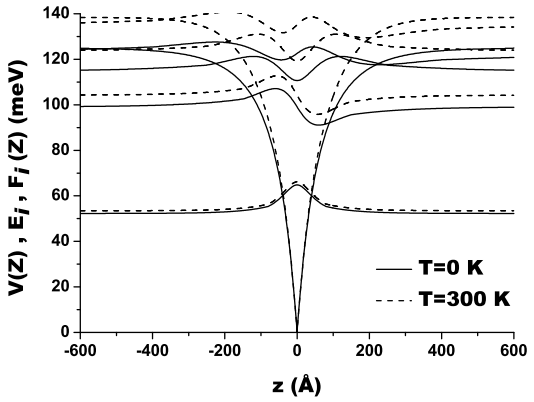


Figure 1: Potential profile, energy levels and square wave functions of Si δ-doped GaAs quantum wells for an impurity density of $3 \times 10^{13}\,\mathrm{cm}^{-2}$. The solid and dashed lines correspond to 0 K and 300 K, respectively.

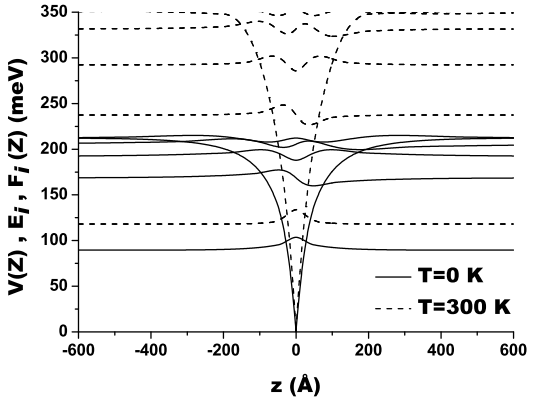


Figure 2: Potential profile, energy levels and square wave functions of Si δ-doped GaAs quantum wells for an impurity density of $5 \times 10^{13}\,\mathrm{cm}^{-2}$. The solid and dashed lines correspond to 0 K and 300 K, respectively.

important changes with temperature, in B delta-doped Si wells, are the energetic distance between levels and the potential depth, while in Si delta-doped GaAs wells the inter-level energy-distance is practically unchanged and the potential depth presents a little increase with temperature, Figs. 1 and 2.

There are also experimental evidence in p-type delta-doped quantum wells in GaAs that supports the behavior observed here. Noh et al. observed that the resistivity behaves quite different for different impurity densities [15, 16]. Obviously, the results obtained in the present paper can not explained the mentioned experimental data, but, these works pointed out that a surprising behavior of the mobility with temperature for n-type delta-doped wells in GaAs is possible.

3. CONCLUSION

In summary, Si-delta-doped GaAs systems present a notably different behavior with temperature as compared to B-delta-doped quantum wells in Si. Despite the similarities in both systems (both due to planar doping) the behavior with temperature is quite different, which argue that it is not possible to extrapolated the trends from one system to another. All seems to indicate, in contradiction with intuition, that each quantum well will attend a particular behavior with temperature.

ACKNOWLEDGMENT

L. M. G-S acknowledges UAEM (State Autonomous University of Morelos, at Cuernavaca, Mexico) for support.

REFERENCES

1. Zheng, X., T. K. Carns, K. L. Wang, and B. Wu, *Appl. Phys. Lett.*, Vol. 62, 504, 1993.
2. Gurtovoi, V. L., V. V. Valyaev, S. Y. Shapoval, and A. N. Pustovit, *Appl. Phys. Lett.*, Vol. 72, 1202, 1998.
3. Wu, C. L., W. C. Hsu, H. M. Shieh, and W. C. Liu, *Appl. Phys. Lett.*, Vol. 64, 307, 1994.
4. Kao, M. J., W. C. Hsu, R. T. Hsu, Y. H. Wu, T. Y. Lin, and C. Y. Chang, *Appl. Phys. Lett.*, Vol. 66, 2505, 1995.
5. Radamson, H. H., M. R. Sardela, Jr., O. Nur, M. Willander, B. E. Sernelius, W.-X. Ni, and G. V. Hansson, *Appl. Phys. Lett.*, Vol. 64, 1842, 1994.
6. Schubert, E. F., J. E. Cunningham, and W. T. Tsang, *Solid State Commun.*, Vol. 63, 591, 1987.
7. Pan, N., J. Carter, G. S. Jackson, H. Hendriks, X. L. Zheng, and H. M. Kim, *Appl. Phys. Lett.*, Vol. 59, 458, 1991.
8. P. M. Koenraad, F. A. P. Blom, C. J. G. M. Langerak, M. R. Leys, J. A. A. J. Perenboom, J. Singleton, S. J. R. M. Spermon, W. C. van der Vleuten, A. P. J. Voncken, and J. H. Wolter, *Semicond. Sci. Technol.*, Vol. 5, 861, 1990.
9. Koenraad, P. M., A. C. L. Heessels, F. A. P. Blom, J. A. A. J. Perenboom, and J. H. Wolter, *Physica B*, Vol. 184, 221, 1993.
10. Hai, G.-Q., N. Studart, and F. M. Peeters, *Phys. Rev. B*, Vol. 52, 11273, 1995.
11. Gaggero-Sager, L. M. and R. Pérez-Alvarez, *Appl. Phys. Lett.*, Vol. 70, 212, 1997.
12. Ozturk, E., Y. Ergun, H. Sari, and I. Sokmen, *Eur. Phys. J. AP*, Vol. 21, 97, 2003.
13. Rodriguez-Vargas, I., L. M. Gaggero-Sager, and V. R. Velasco, *Surf. Sci.*, Vol. 537, 75, 2003.
14. Rodriguez-Vargas, I. and L. M. Gaggero-Sager, *J. Appl. Phys.*, Vol. 99, 033702, 2003.
15. Noh, J. P., F. Shimogishi, and N. Otsuka, *Phys. Rev. B*, 67, 075309, 2003.
16. Noh, J. P., F. Shimogishi, Y. Idutsu, and N. Otsuka, *Phys. Rev. B*, 69, 045321, 2004.

Evaluation of Leakage Losses in Optical Bragg Waveguides

Jie Li and Kin Seng Chiang
Department of Electronic Engineering, City University of Hong Kong
83 Tat Chee Avenue, Hong Kong, China

Abstract— Using a ray-optics model, we analyze with numerical examples the transmission characteristics of a practical optical Bragg waveguide that contains truncated Bragg reflectors. In particular, by treating the truncated Bragg reflector as a perturbation of an infinite periodic structure, we derive an explicit expression for the calculation of the leakage losses of the waveguide.

1. INTRODUCTION

An optical Bragg waveguide, also known as a one-dimensional photonic bandgap waveguide, can be formed by placing a low-index guiding layer between two Bragg reflectors, which are periodic structures consisting of alternating high- and low-index layers [1–5]. A Bragg reflector reflects light over certain optical frequency bands, which are referred to as forbidden bands or bandgaps [1–5]. Light in these bands can be trapped in the guiding layer of the waveguide and thus form the guided modes of the waveguide, regardless of the fact that the guiding layer has a lower refractive index than the surrounding cladding. Optical Bragg waveguides exhibit many unusual properties, such as strong wavelength-dependence transmission, giant dispersion effects, and sharp polarization discrimination, which can find new applications in optical signal transmission and processing. An ideal optical Bragg waveguide consists of a Bragg reflector that is made up of infinite multilayer stacks of alternating refractive indices (an infinite periodic structure). A practical optical Bragg waveguide, however, has a truncated Bragg reflector (a finite periodic structure), which unavoidably introduces leakage losses. It is therefore of fundamental importance to calculate the leakage losses caused by a truncated Bragg reflector for the design and fabrication of such waveguides. In general, the leakage losses are calculated by numerical methods [6, 7], such as the finite-difference time-domain method and the transfer-matrix method, which are time consuming and provide limited physical insight. In this paper, we present an analytical expression for the evaluation of the leakage losses in an optical Bragg waveguide that contains a guiding layer sandwiched between two identical truncated Bragg reflectors. Our method of analysis is based on a ray-optics model [5] together with a perturbation analysis. Using our method, we analyze a number of Bragg waveguides and determine how the leakage losses in these waveguides depend on the number of the periods in the Bragg reflector.

2. METHOD OF ANALYSIS

Figure 1 shows the configuration of a symmetric optical Bragg waveguide, which consists of a guiding layer with refractive index n_g and thickness d_g sandwiched between two identical truncated Bragg reflectors. Each Bragg reflector is made up of N periods of multilayer stacks of alternating refractive indices n_1 and n_2 with corresponding thicknesses d_1 and d_2, where $n_1 > n_2 \geq n_g$. The period of the multilayer structure is $\Lambda = d_1 + d_2$. The refractive index of the outermost medium is assumed to be n_2. The index profile of the waveguide varies only in the x direction, as shown in Fig. 1; it is invariant in the y and z directions. Light is assumed to propagate along the z direction with a propagation constant β. The effective index of the light wave is $n_{eff} = \beta/k_0$, where $k_0 = 2\pi/\lambda$ is the free-space wavenumber with λ being the free-space optical wavelength.

In the ray-optics model [5], the light wave bounces back and forth along the z direction. When the Bragg reflector has an infinite number of periods, light within the bandgaps of the Bragg reflector undergoes complete reflection and can be trapped in the guiding layer. To be a guided wave, the field pattern must be invariant in the z direction, which implies that a standing wave must be set up in the x direction. The total phase shift acquired by the wave traveling over a round trip in the x direction between the two boundaries of the guiding layer must be equal to an even number of π, which is known as the transverse resonance condition. The phase shift acquired by the wave in the guiding layers over a distance of d_g is $k_g d_g$, where $k_g = k_0(n_g^2 - n_{eff}^2)^{1/2}$ is the x-component of the wave vector. The transverse-resonance condition is then expressed as [5].

$$2k_g d_g + 4\phi = 2m\pi, \quad \text{for} \quad m = \text{integers}, \tag{1}$$

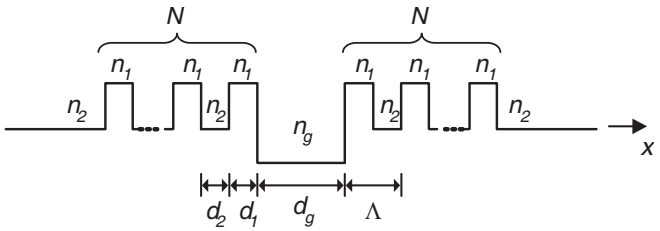


Figure 1: Refractive-index profile of a Bragg waveguide formed with two identical truncated Bragg reflectors.

where 2ϕ is the phase shift acquired at the boundary between the guiding layer and the Bragg reflector. For a semi-infinite Bragg reflector ($N = +\infty$), the reflection phase 2ϕ is a real number and the corresponding reflectivity r is 100% [5]. The range of optical frequencies that satisfy Eq. (1) is governed by the bandgap of the Bragg reflector, which is given by the condition $|\chi| > 1$ with $\chi = \cos(k_1 d_1)\cos(k_2 d_2) - (1/2)(K_1/K_2 + K_2/K_1)\sin(k_1 d_1)\sin(k_2 d_2)$, where $k_i = k_0(n_i^2 - n_{eff}^2)^{1/2}$, and $K_i = k_i$ for the TE polarization and $K_i = k_i/n_i^2$ for the TM polarization (for $i = 1, 2$) [5]. By solving Eq. (1), we can find the effective indices of all the guided waves within the bandgaps [5]. The effective indices are real and the guided modes are lossless.

In the practical situation where the Bragg reflector is finite, the reflectivity r is always smaller than unity in magnitude, which accounts for the leakage loss, and the reflection phase 2ϕ becomes a complex number. To see these, we apply the transfer-matrix method to the multilayer structure [1, 4, 5] and derive the reflectivity r as

$$r = \exp(i2\phi_g)\frac{1 + \Psi\,(1 - K_2/K_g)\exp(-i\phi_g)}{1 + \Psi\,(1 + K_2/K_g)\exp(i\phi_g)]} \tag{2}$$

with

$$\Psi = \frac{\sin K\Lambda}{\sin NK\Lambda}\frac{\exp(iNK\Lambda)}{2T_{21}}\left|\frac{\cos\phi_g}{\cos\phi_0}\right|\exp(i\phi_0), \tag{3}$$

where $\phi_g = \tan^{-1}(K_2/K_g \tan\phi_0)$, $\exp(i2\phi_0) = [\exp(-iK\Lambda) - T_{11}]/T_{12}$, $K_g = k_g$ for the TE polarization and $K_g = k_g/n_g^2$ for the TM polarization, $K = \cos^{-1}(\chi)/\Lambda$ is the Bloch wavenumber of the wave propagating in an infinite Bragg reflector [1, 5], $T_{11} = \exp(-ik_2 d_2)\cos(k_1 d_1) - i(1/2)[(K_1/K_2)+(K_2/K_1)]\sin(k_1 d_1)\}$, and $T_{12} = \exp(ik_2 d_2)\{i(1/2)[(K_2/K_1)-(K_1/K_2)]\sin(k_1 d_1)\}$. The reflection phase 2ϕ, defined as $r = \exp(i2\phi)$ [5], can be solved from Eq. (2). It can be shown from Eq. (2) that 2ϕ is a complex number and, as a result, Eq. (1) admits a complex effective index, denoted as $n_{eff} = n_{effr} + i \cdot n_{effi}$. The real part n_{effr} governs the phase velocity of the wave propagating along the z direction, which can be obtained from the real part of Eq. (1) by setting $n_{effi} = 0$ (since $|n_{effi}| \ll |n_{effr}|$), while the imaginary part n_{effi} gives the leakage loss, which can be obtained from the imaginary part of Eq. (1).

Within the bandgaps (i.e., when $|\chi| > 1$), the Bloch wavenumber is a complex number given by $K = n\pi/\Lambda \pm i\xi/\Lambda$, where $n = 1, 2, 3, \ldots$ is the bandgap order and $\xi = |\mathrm{Im}[\cos^{-1}(\chi)]|$ is the decay rate of the Bloch wave along the infinite periodic structure [5]. Eq. (3) can be written as

$$\Psi = \frac{\chi(1 - 1/\chi^2)^{1/2}}{\exp(2N\xi) - 1}\frac{1}{T_{21}}\left|\frac{\cos\phi_g}{\cos\phi_0}\right|\exp(i\phi_0). \tag{4}$$

In the case $N = +\infty$, we have $\Psi = 0$ from Eq. (4) and recover the lossless result $r = \exp(i2\phi_g)$ [5]. In the case of a finite yet sufficiently large N, we have $\exp(2N\xi) \gg 1$ and hence $\Psi \ll 1$. The finite periodic structure can then be treated as a perturbation of the infinite periodic structure. We thus have $r = \exp[i(2\phi_g + \delta)]$ with

$$\delta \cong \frac{K_2}{K_g}\frac{\cos^2\phi_g}{\cos^2\phi_0}\frac{2(\chi^2 - 1)}{|T_{21}|^2}\exp(-2N\xi)\left[\frac{\mathrm{Im}(T_{11})}{\chi(1 - 1/\chi^2)^{1/2}} + i\right]. \tag{5}$$

By denoting the effective index of the guided mode in the ideal Bragg waveguide as $n_{eff}^{(g)}$ and expanding Eq. (1) with $k_g \cong k_g^{(g)} + [n_{eff} - n_{eff}^{(g)}] \cdot k\prime_g$ and $2\phi \cong 2\phi^{(g)} + [n_{eff} - n_{eff}^{(g)}] \cdot 2\phi\prime_g + i\delta$, where

$k_g^{(g)} = k_g(n_{eff} = n_{eff}^{(g)})$, $\phi^{(g)} = \phi(n_{eff} = n_{eff}^{(g)})$, $k\prime_g = dk_g/dn_{eff}$ and $\phi\prime_g = d\phi_g/dn_{eff}$, we obtain the effective index of the corresponding leaky mode as $n_{eff} = n_{eff}^{(g)} + \delta/(-2\phi\prime_g - k\prime_g d_g)$. By writing $n_{eff} - n_{eff}^{(g)} = \Delta n_{effr} + i \cdot n_{effi}$, we have

$$\Delta n_{effr} = G \exp(-2N\xi) \frac{\mathrm{Im}(T_{11})}{\chi(1 - 1/\chi^2)^{1/2}} \tag{6}$$

and

$$n_{effi} = G \exp(-2N\xi) \tag{7}$$

with

$$G = \frac{K_2}{K_g} \frac{\cos^2\phi_g}{\cos^2\phi_0} \frac{2(\chi^2 - 1)}{|T_{21}|^2} \frac{1}{(-2\phi'_g - k'_g d_g)} \tag{8}$$

$$k\prime_g = -k_g n_{eff}/(n_g^2 - n_{eff}^2) \tag{9}$$

$$\begin{aligned} 2\phi'_g = {} & \frac{n_{eff}}{n_g^2 - n_{eff}^2} \sin 2\phi_g - \frac{4k_g n_{eff}}{n_g^2 - n_{eff}^2} \left\{ \frac{d_1}{2r_{1g}} \frac{1}{1 - (e^{iK\Lambda})^2} \left| \frac{\cos\phi_g}{\cos\phi_1} \right|^2 \left[1 + \frac{1}{k_1 d_1} \sin k_1 d_1 \cos(k_1 d_1 - 2\phi_1) \right] \right. \\ & \left. + \frac{d_2}{2r_{2g}} \frac{\left(e^{iK\Lambda}\right)^2}{1 - (e^{iK\Lambda})^2} \left| \frac{\cos\phi_g}{\cos\phi_0} \right|^2 \left[1 + \frac{1}{k_2 d_2} \sin k_2 d_2 \cos(k_2 d_2 + 2\phi_0) \right] \right\} \end{aligned} \tag{10}$$

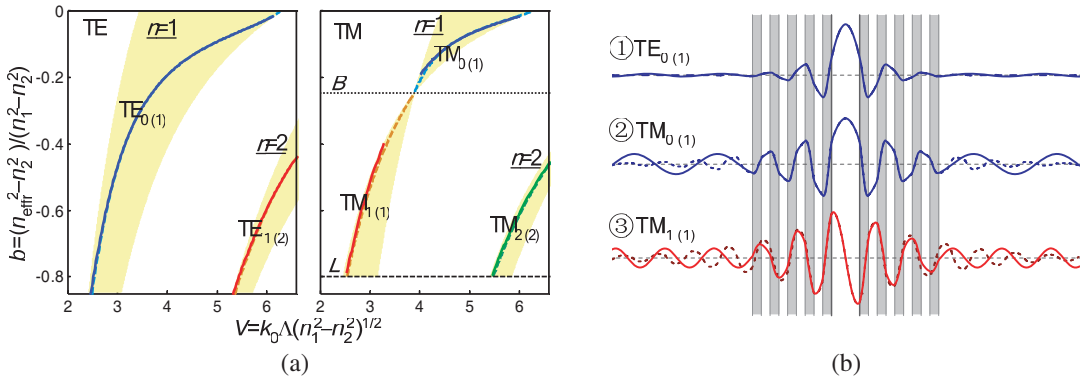
where $\phi_1 = \tan^{-1}(K_2/K_1 \tan\phi_0)$, $r_{1g} = r_{2g} = 1$ for the TE polarization, and $r_{1g} = n_1^2/n_g^2$ and $r_{2g} = n_2^2/n_g^2$ for the TM polarization. The expressions for Δn_{effr} and n_{effi} given above show explicitly how the effective index and the leakage loss change with the period number N. As shown in the expressions, both $|\Delta n_{effr}|$ and n_{effi} decrease exponentially with the value of N. The factor that governs the rate of the exponential decay is 2ξ. When the bandgap edge is approached (i.e., $|\chi| \to 1$ from above), we have $\xi \cong 0$ and the perturbation condition $\exp(2N\xi) \gg 1$ is no longer satisfied. In that case, both Δn_{effr} and n_{effi} should approach 0.

3. NUMERICAL EXAMPLES

We use five parameters to characterize the Bragg waveguide: the dielectric-constant ratio $r_{12} = n_1^2/n_2^2$, the high-index fraction of a period $p = d_1/\Lambda$, the relative guiding-layer index $a = (n_g^2 - n_2^2)/(n_1^2 - n_2^2)$, the relative guiding-layer thickness $q = d_g/\Lambda$, and the period number N. The TE modes are independent of r_{12}. Only the TM modes depend on all these parameters. The propagation constant and the leakage loss of the waveguide can be expressed in the normalized form as $b = (n_{effr}^2 - n_2^2)/(n_1^2 - n_2^2)$ and $\zeta = 2n_{effr} n_{effi}/(n_1^2 - n_2^2)$, respectively, which depend on the operating condition of the waveguide defined by the normalized frequency $V = k_0\Lambda(n_1^2 - n_2^2)^{1/2}$. In our numerical examples, we assume $q = 1.6,\ a = 0,\ p = 0.5$, and $r_{12} = 2.25$.

Figure 2(a) shows the variation of b with V for a truncated structure with $N = 5$ (solid) and the ideal structure with $N = +\infty$ (dashed). The shaded regions are the bandgaps. The modes are labeled as the $\mathrm{TE}_{m(n)}$ or $\mathrm{TM}_{m(n)}$ modes, where m is the mode order from Eq. (1) and n is the bandgap order [5]. The Brewster condition of the TM polarization is shown as the B-line along which $b = b_B = 1/(1 - r_{12}^2)$ and the minimum value of b is shown as the L-line along which $b = b_{\min} = 1/(1 - r_{12})$ [5]. As shown in Fig. 2(a), a truncated Bragg reflector does not affect much the propagation constant. Some mode-field patterns are shown in Fig. 2(b). The field in the truncated structure is oscillatory with finite amplitude in the external medium, which gives rise to a leakage loss. On the other hand, the field in the ideal structure decays away through the periodic structure and suffers from no radiation loss.

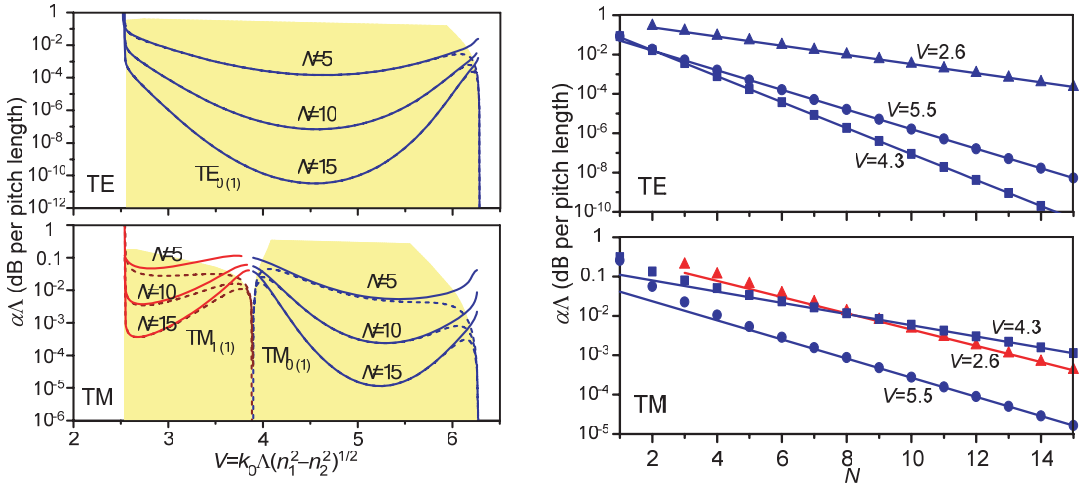
The normalized leakage loss ζ is related to the power attenuation coefficient α in the waveguide by $\alpha = 0.869 k_0 n_{effi}$ in dB or $\alpha\Lambda = 0.434 V\zeta/(b - b_{\min})^{1/2}$ in dB per pitch length (dB/Λ). Fig. 3(a) shows the variation of $\alpha\Lambda$ with V within the first bandgap ($n = 1$) for several values of N, where the results calculated from Eq. (1) (solid) and Eq. (7) (dashed) are compared. As shown in the figure, the leakage loss decreases with an increase in N and becomes large as the band edge is approached. The results calculated by the perturbation formula Eq. (7) are in good agreement with those from Eq. (1), except near the bandgap edges. Fig. 3(b) shows the variation of $\alpha\Lambda$ with



(a) (b)

Figure 2: (a) Dispersion curves and (b) mode-field distributions for a truncated Bragg waveguide with $N = 5$ (solid) with reference to those of the ideal waveguide with $N = +\infty$ (dashed).

N for several values of V. The results obtained from Eq. (1) (points) and Eq. (7) (curves) are in good agreement. As expected, the leakage loss decreases exponentially with an increase in N. The slopes of the straight lines shown in Fig. 3(b) are given by 2ξ. For example, at $V = 2.6$, we have $2\xi = 0.537$ and 0.475 for the TE and TM modes, respectively.



(a) (b)

Figure 3: Variations of the attenuation coefficient $\alpha\Lambda$ with (a) V and (b) N.

4. CONCLUSION

Using the ray-optics model, we analyze the leaky modes of an optical Bragg waveguide that contain two identical truncated Bragg reflectors and derive an explicit expression for the calculation of the leakage losses in the waveguide. With numerical examples, we investigate the effects of the number of the periods in the truncated Bragg reflectors on the propagation constants and the leakage losses of the modes. Our method of analysis can be extended to the analysis of more general Bragg waveguides that contain non-identical Bragg reflectors.

ACKNOWLEDGMENT

This research work was supported by a research grant from the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No. CityU 1248/04E].

REFERENCES

1. Yeh, P., A. Yariv, and C. S. Hong, "Electromagnetic propagation in periodic stratified media. I. General theory," *J. Opt. Soc. Am.*, Vol. 67, No. 4, 423–438, 1977.
2. Yeh, P. and A. Yariv, "Bragg reflection waveguides," *Opt. Commun.*, Vol. 19, No. 3, 427–430, 1976.

3. Dods, S. R. A., "Bragg reflection waveguide," *J. Opt. Soc. Am. A*, Vol. 6, No. 9, 1465–1476, 1989.
4. West, B. R. and A. S. Helmy, "Dispersion tailoring of the quarter-wave Bragg reflection waveguide," *Opt. Express*, Vol. 14, No. 9, 4073–4086, 2006.
5. Li, J. and K. S. Chiang, "Guided modes of one-dimensional photonic bandgap waveguides," *J. Opt. Soc. Am. B*, Vol. 24, No. 8, 1942–1950, 2007.
6. Lenz, G., E. Baruch, and J. Salzman, "Polarization discrimination properties of Bragg-reflection waveguides," *Opt. Lett.*, Vol. 15, No. 22, 1288–1290, 1990.
7. Dasgupta, S., A. Ghatak, and B. P. Pal, "Analysis of Bragg reflection waveguides with finite cladding: An accurate matrix method formulation," *Opt. Commun.*, Vol. 279, No. 1, 83–88, 2007.

Progress in Theoretical Design and Numerical Simulation of High Power Terahertz Backward Wave Oscillator

Hai Zhang[1], **Jianguo Wang**[1,2], **and Changjiang Tong**[2]

[1]School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[2]Northwest Institute of Nuclear Technology, Xi'an 710024, China

Abstract— Results of theoretical analysis and numerical simulation studies of a MW-class, overmoded terahertz oscillator are presented. The device consists of a large diameter cross-section, slow wave structure with a unique profile of wall radius specifically designed to support surface wave and to provide a strong beam-wave coupling at a moderate voltage. Under the condition of 500 kV voltage and 2 kA beam current, the 2.5-D particle-in-cell simulation predicted the output power of 41 MW at the frequency of 0.143 THz. And an efficiency of 4.1% was also obtained with a perfect time plot and fine spectrum characteristic.

1. INTRODUCTION

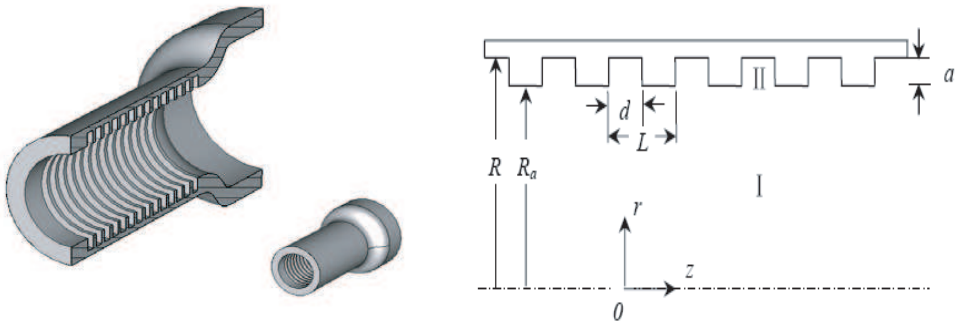
"Terahertz (THz) fields" is a generic term for electromagnetic waves within a spectrum between 0.1 and 10 THz.The interest in this frequency range is fuelled by the fact that this range of frequency is the place where unique physical phenomena with characteristic features are produced [1]. For example, the spectral energy distribution in observable galaxies shows that 50% of the total luminosity are located in the THz frequency range. And THz signals are the information carriers in the ultra wideband communications systems, which are developed now and are expected to become a commercial reality in the next decade.

However, due to the difficulties in generation and detection of the Terahertz signals, they were until recently an almost unexplored area of research. The developments of ultra fast optical techniques, the manufacturing of semi-insulating semiconductors and the micromachining of vacuum electron devices have boosted the terahertz fields as a new research area [2]. This paper presented the recent results of design and simulation of 0.14 THz high power relativistic backward wave oscillator (BWO) in our laboratory.

2. GENERAL CONSIDERATIONS

2.1. Overmoded Slow Wave Structure (SWS)

The main function of SWS in the BWO is to support slow waves with the phase velocity below that of the light, and to ensure strong enough coupling impedance over the frequency range of interest for an electron beam located relatively far from the structure's inner surface. In order to meet these requirements, various axial profiles of wall radius for the periodic structure were analyzed, such as sinusoidal, rectangular, trapezoidal and semicircular. And finally a spatially periodic structure with the rectangular profile was chosen because it can meet the both requirements (slow wave and



(a) (b)

Figure 1: (a) 3-D SWS model with rectangular rippled wall, (b) 2-D axial symmetric cross-section of the SWS

strong coupling) and supply the fabrication convenience in the millimeter dimensional range. The model of the selected structure is illustrated in Fig. 1.

In this structure, the transverse diameter $D(D = 2R)$ is designed to several times the free-space wavelength λ thereby reducing the internal field stress for the same power flow, and in other words it can increase the power-handling capacities of the high power terahertz devices [3]. For a TM_{01}mode propagating in the selected SWS, an approximate relation between the maximum power $P_{\max}$ and the maximum strength of electric field allowable at the wall, $E_{\max}$, can be given by

$$P_{\max} = 8.707 \left(\frac{E_{\max} \cdot \lambda}{511}\right)^2 \frac{\pi^2}{2} \sigma'^4 \sqrt{1 - \left(\frac{\upsilon_{0,1}}{2\pi D'}\right)^2} \frac{1}{\upsilon_{0,1}^2} \tag{1}$$

where $D' = \pi D/\lambda$, λ is the free-space wavelength, $\upsilon_{0,1}$ is the first root of equation: $J_0(x) = 0$, $J_0(x)$ is the Bessel function of order 0. From the formula (1), we can find that the output power $P_{\max}$ is dramatically increased with the enlargement of transverse diameter D. And if $D/\lambda \geq 1.76$, the SWS is defined as an overmoded SWS.

2.2. Surface Wave Operation

Linear beam relativistic BWOs are based on the interaction between an electron beam and the electromagnetic field containing slow-wave components. Such a field can be realized in the spatial periodic structure where the electromagnetic field at an eigenfrequency ω can be expressed as an

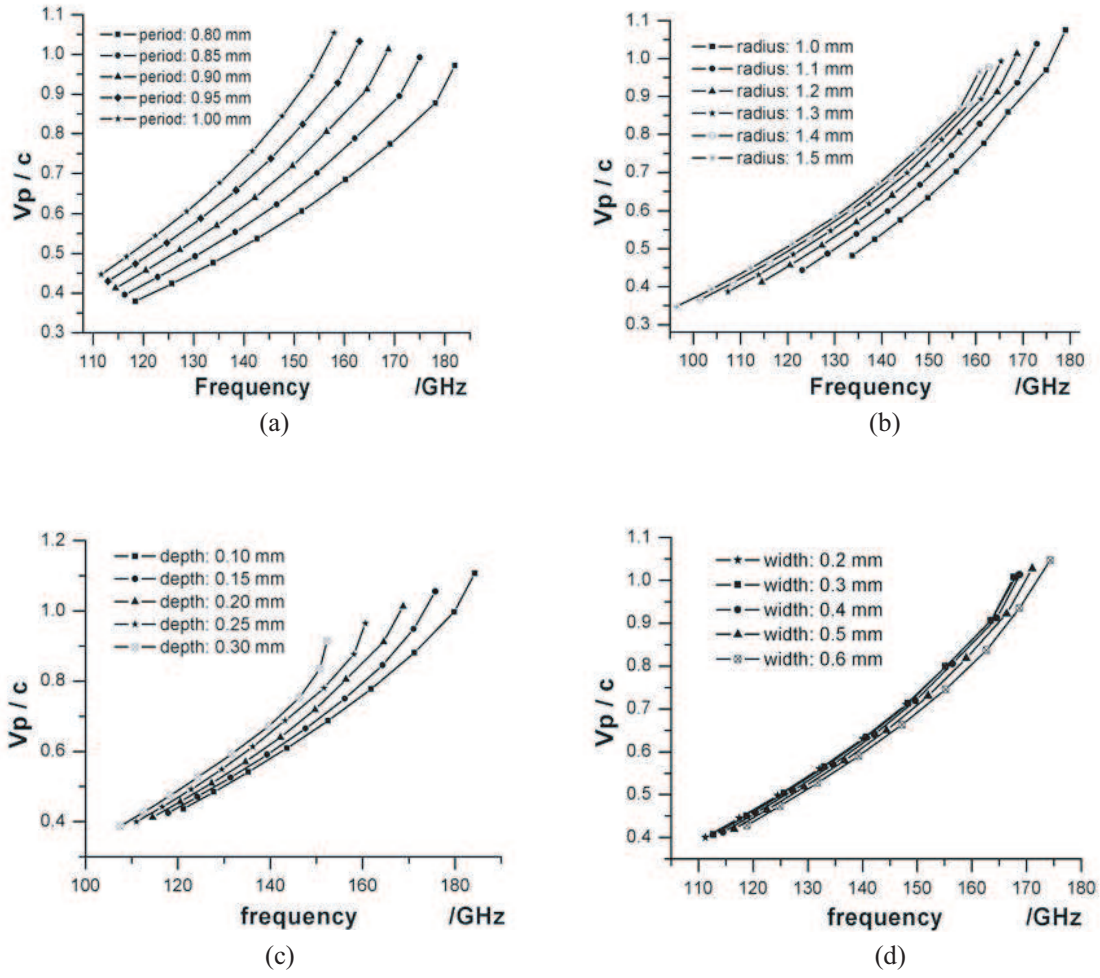


Figure 2: (a) Dispersion diagrams corresponding to different periods of rectangular SWS, (b) dispersion diagrams corresponding to different inner radii of cylindrical waveguide, (c) dispersion diagrams corresponding to different depth of rectangular slots, (d) dispersion diagrams corresponding to different width of rectangular slots.

infinite sum of spatial harmonics:

$$E_z = \sum_{n=-\infty}^{+\infty} a_n G_0\left(k_{c,n} r\right) e^{j k_{z,n} z} e^{-j\omega t} + c.c \tag{2}$$

where

$$k_{z,n} = k_{z,0} + \frac{2\pi n}{L}, \quad n = 0, \pm 1, \pm 2, \cdots$$

L is the spatial period of the SWS, n is the spatial harmonics number. And the transverse and longitude wave numbers satisfy the equation

$$k_{z,n}^2 + k_{c,n}^2 = \left(\frac{\omega}{c}\right)^2 \tag{3}$$

If $|k_{z,n}| < \frac{\omega}{c}$, the spatial harmonic is fast and volumetric since its field profile is described by the ordinary Bessel function, $G_0 = J_0$; However, when $|k_{z,n}| < \frac{\omega}{c}$, the harmonic is slow, and the field profile is described by the modified Bessel function, i.e., $G_0 = I_0$. The field of this harmonic has the characteristic that the field is mainly localized near the surface of SWS, which is defined as surface wave. Our goal is to design this kind of SWS for which the eigenmode in a certain frequency range consist of only slow spatial harmonics, which will play an important role in the realization of high power overmoded THz wave sources.

3. DISPERSION RELATION OF SWS

Dispersion diagrams, to some extent, are the most important characteristic of SWS. From the diagrams, the operation frequency of device can be approximately determined, and the other eignvalues such as coupling impedance and linear growth rate can also be derived as well [4]. We investigated the dispersion relation of SWS with various dimensional parameters. And the dependence of the dispersion characteristics on these parameters is showed in Fig. 2. It is apparent that the frequency of interest (0.14 THz) is included in the frequency range of the SWS. From the calculated diagrams, we can also preliminary determine the values of dimensional parameters of SWS as well as the acceleration voltage of the electron beam in order to ensure the device operate at the frequency of 0.14 THz.

4. PARETICLE-IN-CELL (PIC) SIMULATION AND RESULT ANNALSIS

The analytic theory of the microwave electronics can provide an exact description of the normal modes of the SWS in the absence of a beam. Also, it supplies us with an accurate picture of small-amplitude behavior when the beam is added to the system. However, in the regime of large-amplitude, nonlinear operation as the situation investigated in this paper, the PIC computer simulations should be used in order to examine the behavior of the THz BWO [5]. The simulation model is illustrated in Fig. 3, and the values of parameters are given in Table 1.

Table 1: Value of structural parameters for the PIC simulation.

Parameters	R_a	L	d	a	D_1	D_2	H	N
value	3.0 mm	0.9 mm	0.4 mm	0.2 mm	4.0 mm	4.0 mm	1.0 mm	20

The simulation was carried out with the boundary conditions that the waveguide wall is a perfect conductor, that there was an axial symmetry of z axis, and that the electromagnetic waves were outgoing at the ends of the structure to a good approximation. As the initial condition for the simulation, there were no electromagnetic fields in the SWS, and the electron beam was just incident at the left-end side of the structure. Under the condition of 500 kV voltage, 2 kA beam current and 4T axial guiding magnetic field, the oscillator began to work. Fig. 4 shows the temporal behavior of typical field component at a certain point in the SWS and the corresponding Fourier transform of this signal. We can find that the oscillator steadily operated at the frequency of 0.143 THz. The time plot of the output power is indicated in Fig. 5. And the peak power of 41 MW (efficiency equals to 4.1%) was obtained from the device. Fig. 6 exhibits the distribution of longitude electric field

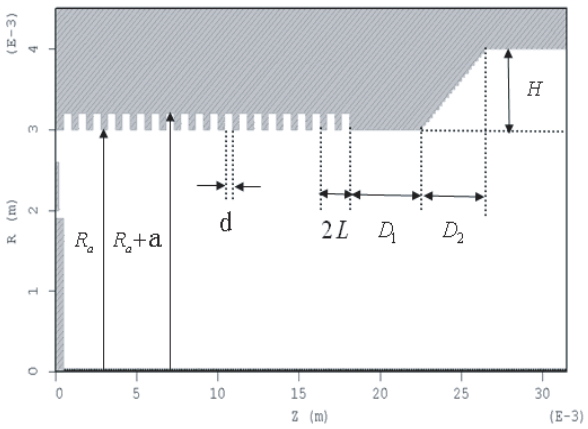


Figure 3: THz BWO model in the PIC simulation.

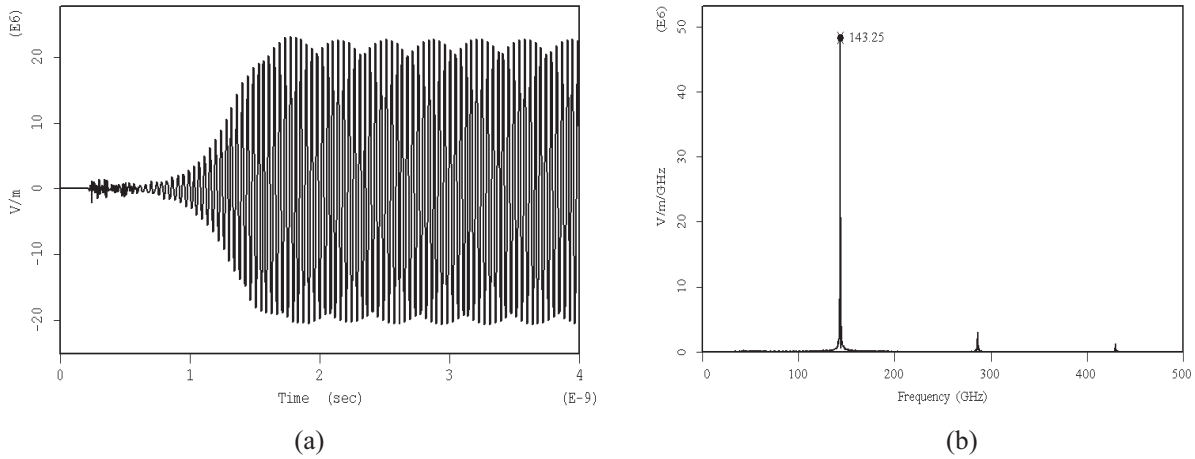


Figure 4: (a) Time plot of the generated signal, (b) spectrum of the generated signal.

E_Z along the radial direction at the coordinate of $z = 5\,\text{mm}$. It is found that the field decreased along the given radius from the inner surface of the SWS to the center axis, that is to say, the field accumulated almost near the inner surface, which confirmed that the system we designed operated as a surface wave oscillator.

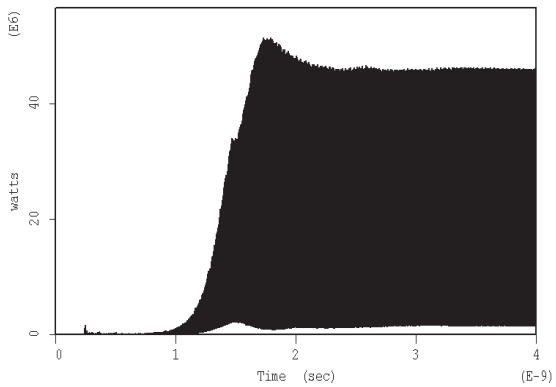


Figure 5: Time plot of the output power.

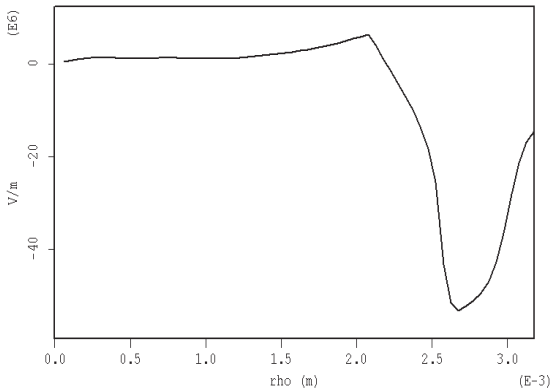


Figure 6: Distribution of E_Z along the radius.

5. CONCLUSIONS

In this paper, we presented the preliminary research progress on the overmoded, MW class terahertz source in our lab. It was found that an overmoded interaction SWS must support surface wave that are synchronous with the electron beam, and simultaneously exhibit large values of the coupling impedance. The dispersion characteristics of the SWS with periodic rectangular rippled-wall were examined. And a set of optimal values of the system parameters were specified. With the 2.5-D particle-in-cell simulation method, the power lever of 41 MW and the efficiency of 4.1% were obtained at the frequency of 0.14 THz. Further innovation of the device structure is required in order to enable the generation of higher wave power at higher frequencies.

REFERENCES

1. Dragoman, D. and M. Dragoman, "Terahertz fields and applications," *Progress in Quantum Electronics*, Vol. 28, No. 1, 1–66, 2004.
2. Taniuchi, T and H. Nakanishi, "Continuously tunable terahertz-wave generation in GaP crystal by collinear difference frequency mixing," *Electronics Letters*, Vol. 40, No. 5, 327–328, 2004.
3. Vlasov, A. N, A. G. Shkvaruntes, J. C. Rodgers, et al., "Overmoded GW-class surface-wave microwave oscillator," *IEEE Trans. on Plasma Sci.*, Vol. 28, No. 3, 550–560, 2000.
4. Guo, H. Z., Y. Carmel, and V. L. Granatstein, "A novel high accurate synthetic technique for determination of the despersive characteristics in periodic slow wave circuits," *IEEE Trans. MTT*, Vol. 40, No. 11, 2086–2094, 1992.
5. Swegle, J. A., J. W. Poukey, and G. T. Leifeste, "Backward wave oscillators with rippled wall resonators: Analytical theory and numerical simulation," *Phys. Fluids*, Vol. 28, No. 9, 2882–2894, 1985.

Modeling of Passively Mode-locked Broadband Dual-gain-media Nd:glass Laser

Song Han and Li Yan
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County, Baltimore, MD 21250, USA

Abstract— Passive mode-locking of Nd:glass laser is studied based on a gain model that includes Stark-split sub-transitions. The full gain model appropriately describes the passive mode-locking characteristics of both a single-medium and a dual-gain-media Nd:glass laser. The simulation results are in good agreement with the experiments.

1. INTRODUCTION

Many broadband gain media used in ultrafast lasers, such as Ti:sapphire, Cr:forsterite and Cr:LiSAF, are homogeneously broadened. There are other gain media, such as neodymium-doped glass and erbiumdoped glass, whose lasing transitions consist of inhomogeneous broadening and multiple Stark-split subtransitions. Careful handling of this type of gain media in certain mode-locking operations is needed. In active mode-locking, the mode-locking strength is usually weak, and the pulse bandwidth is comparable to the CW lasing bandwidth, which is mainly contributed from a single major sub-transition. Thus, it is appropriate to include only the major sub-transition in the gain model [1]. However, for passive modelocking that usually has a strong pulse shortening mechanism, the broadband gain can be efficiently utilized, and it is necessary to include all the Stark-split sub-transitions to account for the overall broadband gain. This is particularly true for modeling a dual-gain-media Nd:glass laser [2].

In this paper, we modeled the gain of Nd:glass laser by considering the homogeneous broadening, the inhomogeneous broadening and the Stark-split sub-transitions. It gives the main features of Nd:glass gain, i.e., having a relatively narrow CW lasing bandwidth under a broad overall gain linewidth. Under passive mode-locking, the full gain model appropriately describes the mode-locking characteristics of Nd:glass laser. For the dual-gain-media Nd:glass laser, the full gain clearly shows how the overall gain is reshaped and broadened, which leads to the generation of shorter pulses. The simulation results are in good agreement with the experiments.

2. EXPERIMENTAL CHARACTERIZATION OF KLM STRENGTH

The schematic of the laser is shown in Figure 1. The transmittance of the output coupler was 1.5%. The reflection of each mirror had a reflection of 99.5%. The loss at each Brewster's surface was estimated to be ∼0.2%. With the unsaturated absorber loss of ∼1%, the total roundtrip internal cavity loss (excluding the loss from the output coupler) was ∼7.2%. The Nd:glass laser was pumped by a CW Ti:sapphire laser with P_{pump} of ∼1 W. About 60% of the pump power was delivered to the Nd:fluorophosphate glass while the Nd:silicate glass received 40% of the total pump power. When the laser was operated at the center of the stability zone, it had the smallest cavity loss δ_0, and generated a maximum free-running power $P_{out,CW_{\max}}$ of 64 mW, with the threshold pump power P_{th} of 60 mW. Femtosecond pulses were generated by Kerr-lens mode-locking (KLM), while the saturable absorber acted to start the mode-locking and help maintain the pulse stability. The pulse bandwidth was adjusted by changing the insertion of the prism into the beam path. Experimentally, when the group-delay dispersion (GDD) was adjusted to be about $-320\,\mathrm{fs}^2$, we obtained 41 nm pulse bandwidth and 38 fs pulses, which are the shortest pulses ever generated from a bulk Nd:glass laser [2].

We described the effect of KLM by a fast saturable absorber model $\delta_A(t) = \delta_{a0}/(1+|\tilde{a}(t)|^2/P_{s,abs})$, where δ_{a0} is the unsaturated absorber loss and $P_{s,abs}$ is the saturation power. We estimated the cavity linear loss according to

$$\frac{P_{out,CW}}{P_{out,CW_{\max}}} = \frac{\frac{P_{pump}}{P_{th,CW_{\max}}}\frac{\delta_0 - \ln R}{\delta_{CW} - \ln R} - 1}{\frac{P_{pump}}{P_{th,CW_{\max}}} - 1} \tag{1}$$

where $P_{out,CW}$ and δ_{CW} are the CW power and the cavity loss at near the edge of the stability zone under KLM. From the measured $P_{out,CW}$ of 35 mW, we obtained $\delta_{CW} = 13.7\%$. The difference between δ_0 and δ_{CW} gave the unsaturated absorber loss $\delta_{a0} = 6.5\%$.

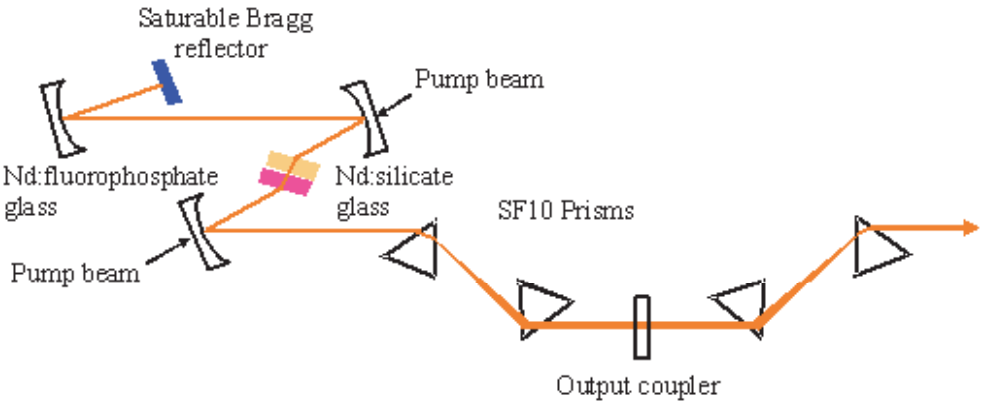


Figure 1: Schematic diagram of the passively mode-locked dual-gain-media Nd:glass laser.

3. MODELING

The observed fluorescence linewidth of various Nd:glasses come actually from a convolution of the homogeneous broadening, the inhomogeneous broadening and twelve Stark-split sub-transitions. Due to Stark-split sub-transitions, the overall gain of the Nd:glass is broad (25–30 nm) and asymmetric. In our model, we set the homogeneous linewidth Δv_h and the inhomogeneous linewidth Δv_{ih} to be 24 cm^{-1} and 96 cm^{-1}, respectively, corresponding to an inhomogeneity $\Delta v_{ih}/\Delta v_h = 4$ [3]. The sub-transition locations were set according to reference [3], and the average Stark-split line-separation is about 45 cm^{-1}. The gain cross section of each Stark-split sub-transition was adjusted individually for both the Nd:silicate glass and Nd:fluorophosphate glass to make their overall gain profiles fit to the measured fluorescence spectra [4], as listed in Table 1. Figure 2(a) shows the small signal gain of the Nd:silicate glass and its CW lasing spectrum. The simulated gain linewidth is ∼26 nm and the CW lasing bandwidth is ∼8 nm, which are in good agreement with the published data [4]. For Nd:glass, the fast thermal relaxations cause cross saturation of populations of different sub-levels. Through this mechanism, the CW lasing bandwidth is mainly determined by the single major sub-transition [5]. If only the major Stark-split subtransition is modeled, it would give to a substantially narrower gain linewidth of 13 nm and a slightly broader CW lasing bandwidth of 11 nm. For the passive mode-locking of the Nd:glass, where the modelocking strength is strong, it is necessary to include all the Stark-split sub-transitions to account for the overall broadband gain.

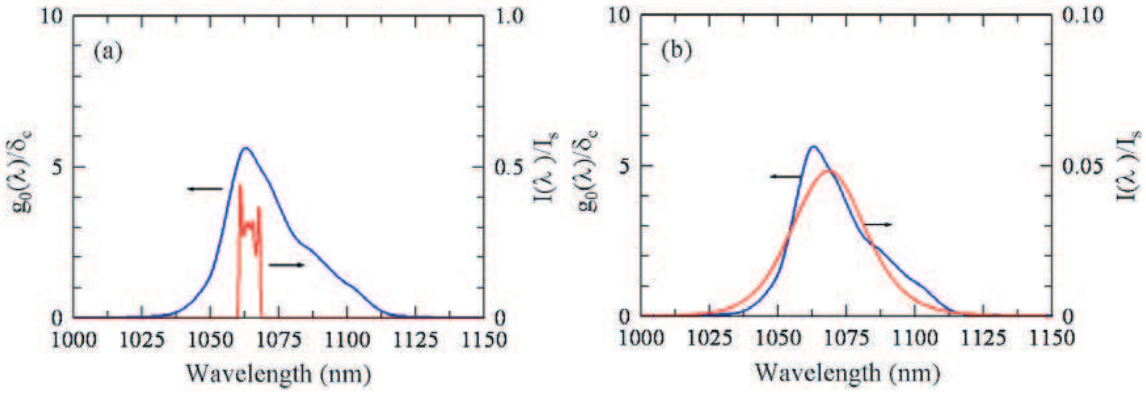


Figure 2: (a) Unsaturated gain profile and CW lasing spectrum; (b) Unsaturated gain profile and pulse spectrum of Nd:silicate glass laser (unsaturated absorber loss $\delta_{a0} = 6.5\%$).

4. MODE-LOCKING PERFORMANCE OF SINGLE-MEDIUM ND:GLASS LASER

We simulated the passive mode-locking of Nd:silicate glass laser based on the gain model described above. Besides the gain, the simulation includes the cavity loss, GDD, self-phase modulation (SPM) and self-amplitude modulation (SAM) by a fast saturable absorber with the saturation level set to

Table 1: Peak cross-section of each Stark-split sub-transition for Nd:glass [3].

	Wavelength (nm)	$\sigma_0(10^{-20}\,\mathrm{cm}^2)$ Nd:glass	$\sigma_0(10^{-20}\,\mathrm{cm}^2)$ Nd:fluorophosphates	$\sigma_0(10^{-20}\,\mathrm{cm}^2)$ Nd:silicate
(A)-(1)	1064	1.0010	1.0010	1.0010
(A)-(2)	1073	0.4390	0.7024	0.6146
(A)-(3)	1086	0.0760	0.0760	0.0760
(A)-(4)	1095	0.1520	0.1824	0.1976
(A)-(5)	1099	0.0100	0.0033	0.0059
(A)-(6)	1104	0.2970	0.0660	0.1485
(B)-(1)	1048	0.4830	0.1725	0.1725
(B)-(2)	1057	0.9660	0.4830	0.5367
(B)-(3)	1070	0.0100	0.0180	0.0160
(B)-(4)	1078	0.2760	0.4140	0.3588
(B)-(5)	1080	0.1200	0.2040	0.1560
(B)-(6)	1088	0.3560	0.4628	0.4628

be about optimal ($P_{ave}/P_{s,abs} = 10^{-1}$). Figure 2(b) shows the pulse spectrum of Nd:silicate glass laser at the soliton stability boundary (due to excessive gain filtering) where the shortest pulse can be generated and stably maintained. The unsaturated absorber loss is set to be 6.5%, same as the estimated δ_{a0} from the experiment of the dual-gain-media Nd:glass laser, for the purpose of comparison. The strong inhomogeneous broadening of Nd:silicate glass introduces less gain filtering to the pulse spectrum and helps create a broad bandwidth of 32 nm. Due to the steep slope at the short-wavelength side of the gain profile, a red shift of the spectral center occurs so that the pulse experiences less gain filtering. Figure 3 shows the pulse bandwidth generated at the soliton stability boundary for different unsaturated absorber loss. A stronger SAM action from the KLM can push the soliton-like pulse into a smaller dispersion region and therefore creates a broader bandwidth. The simulation shows that for Nd:glass laser, an unsaturated absorber loss $\delta_{a0} \geq 5\%$ is needed to generate a pulse bandwidth comparable to the gain linewidth.

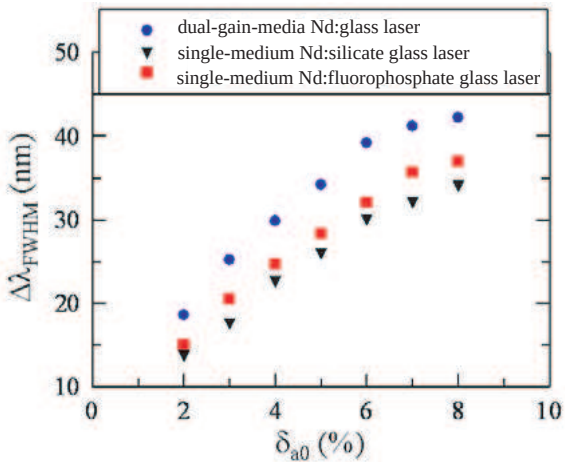


Figure 3: Pulse bandwidth as a function of δ_{a0} for single-medium Nd:glass laser and dual-gain-media Nd:glass laser.

5. MODE-LOCKING PERFORMANCE OF DUAL-GAIN-MEDIA ND:GLASS LASER

For generating short pulses from Nd:glass laser, gain reshaping technique is usually used to broaden the effective gain linewidth, which is otherwise limited by the asymmetry of the gain profile. One approach is to introduce a filter at the short wavelength and reshape the overall gain profile [6].

In our scheme of dual-gain-media Nd:glass laser, the gain reshaping and broadening are achieved through the combination of two gain media into one oscillator, as shown in Figure 1. We simulated the mode-locking of the dual-gain-media laser. Figure 4(a) shows the unsaturated gain profile when the pump ratio of the Nd:fluorophosphate glass and Nd:silicate glass equals to 2. The benefits are two fold. First, it creates a maximum broadening of unsaturated gain by 6.5 nm, which is about 25% increase compared to that of a Nd:silicate glass or Nd:fluorophosphate glass. Second, the gain profile becomes more symmetric. When the same unsaturated absorber loss $\delta_{a0} = 6.5\%$ is applied, with this reshaped gain, the dual-media Nd:glass laser generates a broader pulse bandwidth compared with the single-medium Nd:glass laser, as shown in Figure 4(b). The simulated pulse bandwidth is ∼40 nm at the small GDD boundary ($-330\,\mathrm{fs}^2$) of the stable soliton mode-locking. Both the pulse bandwidth and the GDD are within 10% with respect to that from the experiments. Since the gain becomes more symmetric in the dual-gain-media scheme, the pulse spectrum experiences only ∼2.3 nm red-shift. In comparison, for the single-medium Nd:silicate glass laser with asymmetric gain, the red-shift is ∼4.4 nm, which is noticeably larger. The similar feature was also observed in the experiments, in which the spectral red-shift for the dual-gain-media Nd:glass laser and single-medium Nd:silicate laser were ∼5 nm and ∼8 nm, respectively. The differences in the absolute amounts were likely caused by a large loss of the saturable Bragg absorber at the edge of high reflectivity coating ∼1.032 nm in the experiments.

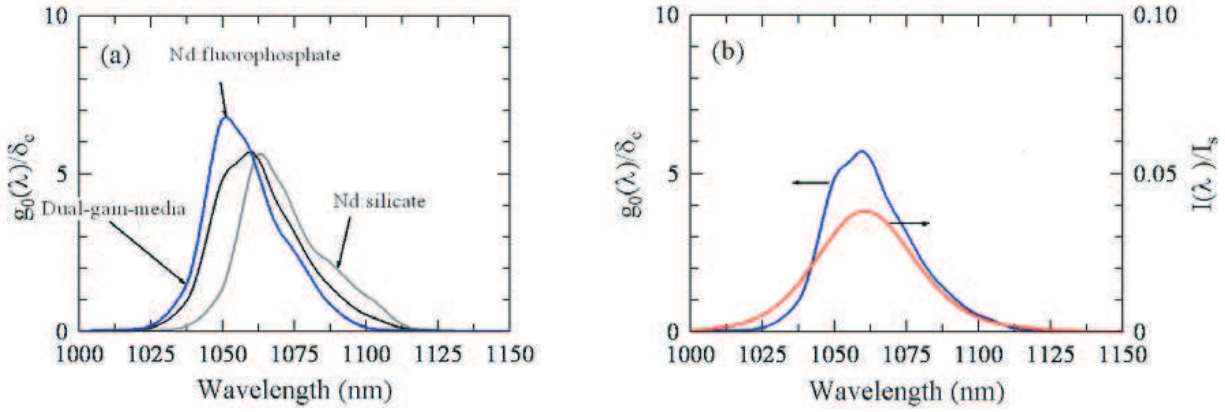


Figure 4: (a) Gain reshaping and broadening, (b) Pulse spectrum broadening of dual-gain-media Nd:glass laser (unsaturated absorber loss $\delta_{a0} = 6.5\%$).

REFERENCES

1. Lu, W., L. Yan, and C. R. Menyuk, "Dispersion effects in an actively mode-locked inhomogeneously broadened laser," *IEEE J. Quantum Electron.*, Vol. 38, 1317–1324, 2002.
2. Han, S., W. Lu, B. Y. Sheh, L. Yan, M. Wraback, H. Shen, J. Pamulapati, and P. G. Newman, "Generation of sub-40 pulses from a mode-locked dual-gain-media Nd:glass lser," *Appl. Phys. B*, Vol. 74, 177–179, 2002.
3. Mann, M. M. and L. G. DeShazer, "Energy level and spectral broadening of neodymium ions in laser glass," *J. Appl. Phys.*, Vol. 41, 2951–2957, 1970.
4. Kopf, D., F. X. Kärtner, and U. Keller, "Diode-pumped mode-locked Nd:glass lasers with an antiresonant fabry-perot saturable absorber," *Opt. Lett.*, Vol. 20, 1169–1171, 1995.
5. Yan, L., P.-T. Ho, C. H. Lee, and G. L. Burdge, "Generation of ultrashort pulses from a neodymium glass laser system," *IEEE J. Quantum Electron.*, Vol. 25, 2431–2440, 1989.
6. Kopf, D., F. X. Kartner, K. J. Weingarten, and U. Keller, "Pulse shortening in a Nd:glass laser by gain reshaping and soliton formation," *Opt. Lett.*, Vol. 19, 2146–2148, 1994.

Metamaterial Structures for Compact Millimeter Wave Antenna Applications

Cuong Tran Manh[1], **Habiba Hafdallah Ouslimani**[1], **Geraldine Guida**[1]
Alain Priou[1], **Herve Teillet**[2], **and J. Y. Daden**[2]

[1]Applied Electromagnetic Group (GEA), Université Paris X, Nanterre
50 rue de Sevres, Ville d'Avray 92410, France
[2]THALES Communications, 160 Boulevard de Valmy, Colombes 92704, France

Abstract— This paper proposes the study of some high impedance surface (HIS) structures for compact antenna applications in the millimeter-wave domain. The millimeter wave domain is now very important for high speed wireless and high bit rate optical (> 40 Gbits/s) communications systems. The HIS structures provide many advantages for antennas as they enhance their performances; HIS structures have capability to block the surface wave, to reduce the coupling effect, to present high real impedance at the resonance frequency ($R_e(Z) \gg 377\Omega$) and to reduce the global thickness of the low profile antenna. Several high impedance surfaces structures are analyzed and their properties compared. We perform this analysis on structures which composed of rectangular lattices patches periodic arrangements, Jerusalem lattices shape and 2LC shape (two LC split loops). For each structure, we are interested in the frequency behavior of the reflection phase to determine the resonance frequency and the band-gap as well as in the losses (joule effect) in the structure. All the dimensions and shapes of the unit cell geometry are optimized in order to use the dielectric substrate available in our laboratory. The high impedance surface is modeled using HFSS (Ansoft) code based on finite element methods. We chose the structure presenting the best performances to design the metamaterial antenna with coaxial feed and finite surface witch is modeled with 7×7 and 9×9 double rang unit cells. In comparison with conventional antenna type, placed above a metal ground plan, the antenna placed above the HIS has smoother radiation profile, less power wasted in the backward direction, better return loss (at least -10 to -15 dB better) and higher gain and directivity (at least $+1$ dB). The layout of the HIS structures circuits (many varieties) are now edited and the manufacturing process in progress. The results of the HFSS simulations will be compared with the experimental free space and coaxial measurements in the millimeter-wave domain.

1. INTRODUCTION

We have known that a patch antenna consist of a metal patch suspended over a ground plane and separate with the ground plane by a dielectric substrate, so is acts as a cavity. We use herein the rectangular shape for radiator patch, and we choose the coaxial feeding method, the feeding point has chosen at the patch's corner in the purpose of impedance matching. Antennas of this type are low-profile but highly resonant.

In the circuit, the ground plane is always finite, and its edges contribute to the radiation pattern. In addition to space waves, the antenna generates surface waves in the ground plane, which then radiate from edges and corners. The combined radiation from the patch and the ground plane edges interfere to form a series of multipart lobes and nulls at various angles. The edges radiate backwards as well as forwards, causing a significant amount of wasted power in the backward hemisphere and ripples in the antenna pattern. This problem is exacerbated, if the substrate is thick, or has a high dielectric constant.

Many authors propose the suppression of the surface waves, by embedding the patch antenna in a highimpedance ground plane as [1–5]. In the following example (Fig. 1), D. Sievenpiper shows a comparison between the radiation pattern of a microstrip patch and HIS ground plane patch antenna.

The measurements (Fig. 1) are at a frequency in which the two antennas have the same return loss. In both the H and E-plane, the patch on the ordinary metal ground plane shows significant radiation in the backward direction, and ripples in the forward direction. The pattern is not rotationally symmetric, and is much thinner in the H-plane than in the E-plane. Conversely, the patch placed over the high-impedance ground plane produces a smooth, symmetric pattern with little backward radiation.

In this paper, after presenting some potential planar microstrip high impedance surfaces and their frequency response, we chose the structure presenting the best performances, to design the

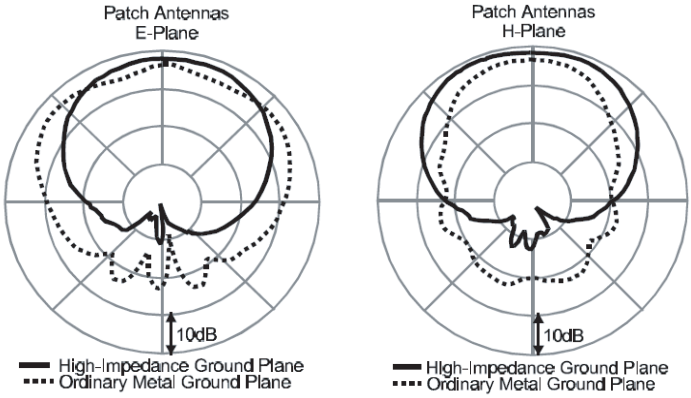


Figure 1: E and H-plane radiation patterns of two patch antennas [1].

metamaterial antenna with coaxial feeding method and finite ground plane surface. The later is modeled with 7×7 double rang unit cells. In comparison with conventional antenna type placed above a metal ground plan, the antenna placed above the HIS has smoother radiation profile, less power wasted in the backward direction, better return loss (at least -4 dB to -8 dB better) and higher gain and directivity (at least $+4$ dB)

2. HIS DEFINITION

High Impedance Surface (HIS) structures, also designed as Artificial Magnetic Conductor (AMC) or Perfect Magnetic Conductor (PMC) [3, 6] may be very useful for antenna applications and in a very large variety of microwave other devices [4, 5]. Electromagnetic band gap structures (EBG) have been widely studied for their behaviour as High Impedance Surface (HIS), since they show a stop band frequencies behaviour. The AMC condition is characterized by a resonance frequency where the phase of the reflection coefficient is zero and its magnitude equal to one. In contrast, an HIS may deviate a little from this condition, sometimes yielding more flexibility in antenna design. In this case, we defined a band gap frequency as a frequency range in which the reflection phase cross from $+90°$ to $-90°$ [9].

Like the proposition of Sievenpiper and co-others [1, 4], planar periodic array of metallic patches with connection via to the ground plane exhibit a high impedance with an exactly zero degree reflection phase at the resonance frequency. Array of patches without via connection to the ground plane exhibit also this property. We have known that, there is a problem with most of proposed HIS structures because they present a shift of the resonant frequency versus the incidence angle [15].

3. SIMULATION MODEL OF HIS STRUCTURES

3.1. HFSS Modelling

The performances of the HIS structures are studied using numerical simulation from a finite element method HFSS codes (FEM-HFSS ANSOFT Version 10.1). We use three structures: the square patches structure, the Jerusalem structure and the two LC boucles "2LC" structure. The Figure 2 (below) gives the studied planar structures called also Uni-planar Compact Photonic Bandgap

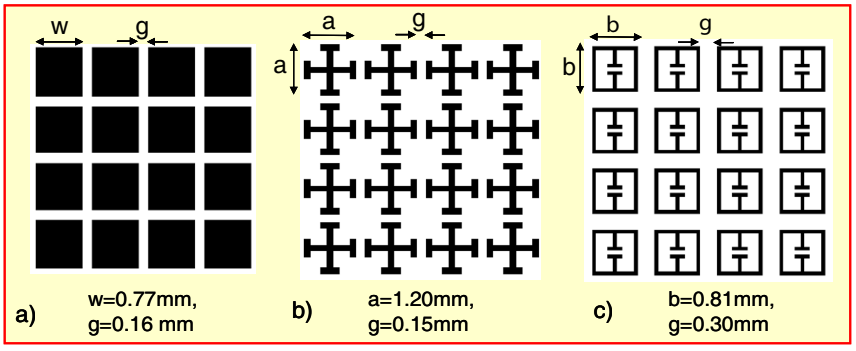


Figure 2: Square patches (a), Jerusalem (b) and the "2LC" (c) structures.

(UCPBG) [10, 11].

Patch antennas are studied with all the three type of HIS structure listed above. The simulation set-up is presented in Figure 3. A plan wave model is established to evaluate the reflection phase of the EBG surface. The plan wave is launched to normally illuminate the EBG structure like the method in [7] and [12]. To model an infinite periodic structure, we used a single unit of EBG structure with periodic boundary condition on the four sides of the sample in simulation.

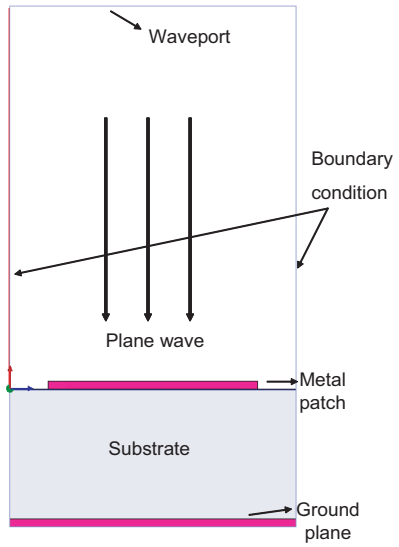


Figure 3: Setup simulation for one cell "UCPBG".

3.2. Numerical Results

The geometries (all dimensions) of the three types of structures are optimized in the millimeter wave domain, corresponding to the 40–45 GHz frequency range. The used substrate has the following characteristics: thickness $h = 0.51\,\text{mm}$, permittivity $\varepsilon = 3.38$ (RO4003 substrate) and $\tan(\delta) = 0$. The metallic cell patches and ground plan are chosen to be perfect electric conductor (no loss in the material).

The optimized dimensions for the unit cell and the array period for the three structures are respectively; $W \times W = 0.7742 \times 0.7742\,\text{mm}^2$, array period $P = 0.931\,\text{mm}$ for the rectangular patch, array period $P = 1.11\,\text{mm}$ for the "2LC" and array period $P = 1.30\,\text{mm}$ for the Jerusalem structure.

The obtained results on the optimized structures are shown in Figure 4. For each type, we can see the diagram of return loss, the reflection phase and the real and imaginary part of the surface impedance.

Broader band gap results in a better control of the antenna backward radiation, thus it is essential to choose the structure which shows the largest band gap. We know also, that the band gap will be seriously reduced for incoming plane wave with a certain incidence angle. In our simulation, the chosen structure showed the best real part of surface impedance approximately $6 \cdot 10^4\,\Omega$. After analysis, we show that the rectangular HIS structure presented the best performances for antenna.

Table 1: For the different cells: the resonance frequency, the bandgap width and the dimension period.

Type of structure	Period(mm)	Frequency(GHz)	Bandwidth(%)
Rectangular	0.93	41.2	41.46
Jerusalem	1.30	40.6	27.16
2LC	2.04	41.5	22.89

At the resonance frequency, the reflection coefficient has zero-phase and the surface impedance is real and maximum (infinite in theory). For all three structures, the minimum return loss occurs

approximately at the resonance frequency. The relative bandwidth of the three structures is defined as:

$$B_{\text{bandwidth}} = \left(f_{[+90^\circ]} - f_{[-90^\circ]}\right) / f_{[0^\circ]}$$

And the results are reported in Table 1.

One can observe that, the rectangular patch HIS presents the best results in term of band gap width, stability and satisfied features for antenna application. It was naturally chosen to simulate the antenna performances. In order to verify the constancy of the resonance frequency, particularly when the unit cell is integrated to the array and when it interact with the other cells, we have simulate new patches structures with lager number of elements (4×4 to 6×6). The results showed that the bandgap region does not shift since the period of structure remains exactly the same (imposed by the unit cell).

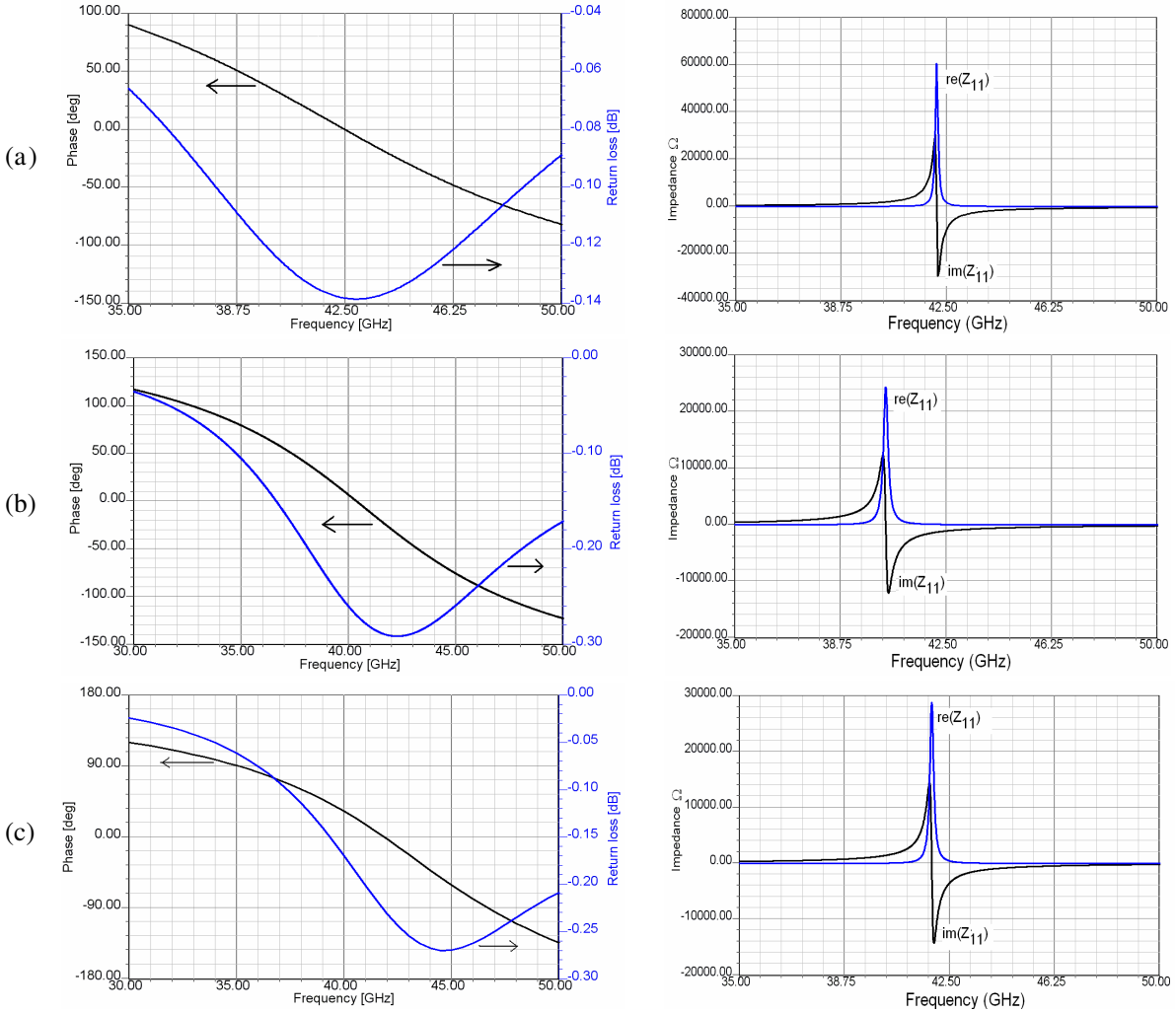


Figure 4: Numerical simulation results: (a) Rectangular, (b) Jerusalem and (c) "2LC" patches. Left: Return loss magnitude (dB) and phase (degrees). Right: Impedance surface $\mathrm{Z_{in}}(f)$.

4. ANTENNA PATCH SIMULATION WITH AND WITHOUT HIS STRUCTURE

Figure 5 shows the antenna topology that we choose to study in this paper. Following the analysis and results obtained in Section 3, we use the high impedance surface, rectangular type to realize our antenna. Using this "composite" ground plan structure, we expect better performances (enhance of the antenna gain, directivity and return loss). All the results presented here compare the performances of the two types of ground plan structures; conventional antenna and HIS antenna (metamaterial antenna).

The radiator element is exactly the same in the two antennas (rectangular patch antenna of 1.75×1.60 mm and 0.03 mm copper thickness). The radiator element is fed using coaxial line and

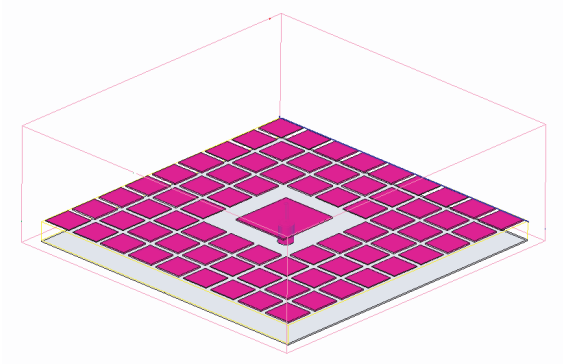


Figure 5: A patch antenna embedded in a high-impedance ground plane.

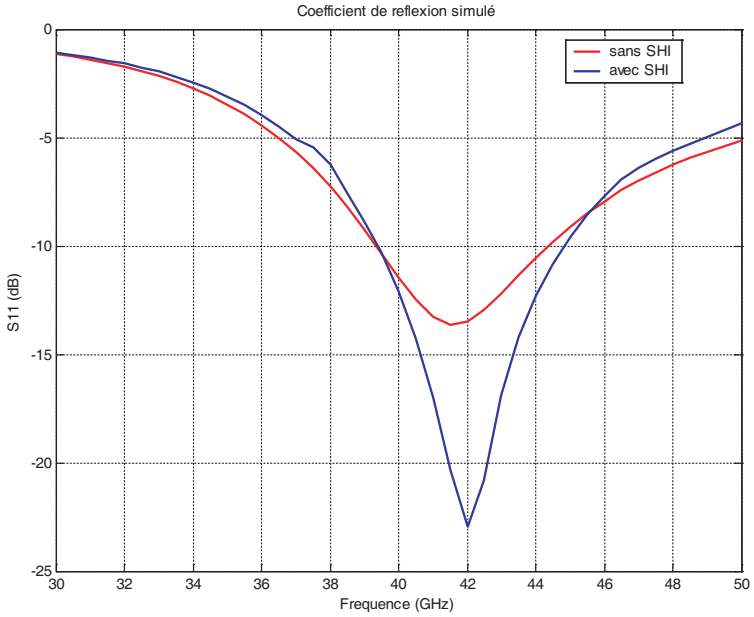


Figure 6: Comparison of the return loss (MS_{11}) between HIS ground plane and conventional patch antennas.

positioned in order to obtain the best impedance matching (0.2 mm at the corner of radiator).

We have show here in the Figures 7 and 8 the simulation results of two antennas at 42 GHz, the point in which both of them have the same value of return loss (see the Figure 6), the comparison in

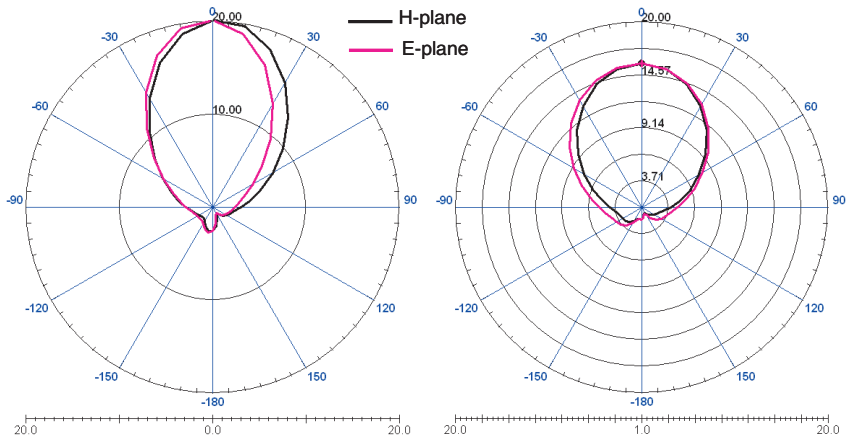


Figure 7: Radiation pattern comparison of two kinds of antenna at 42 GHz.

Figure 8 has shown clearly the advantage of using HIS structure for patch antenna; We receive the augmentation 4.5 dB of gain and/or directivity for patch with HIS in our simulation (see Figure 8) compared with the conventional patch, thus leading to an increase in bandwidth and efficiency of antenna [14]. From the Figure 7, we see that the radiation pattern of the metamaterial antenna is more directive (+20 dB) than the conventional antenna (+15 dB).

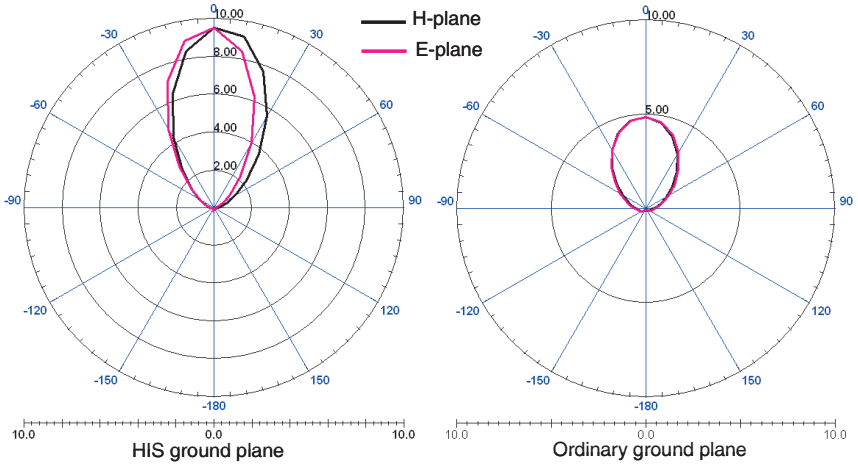


Figure 8: Gain/Directivity comparison of two kinds of antenna at 42 GHz.

5. CONCLUSIONS

We have used finite element method (FEM) to numerically determined characteristics of high impedance surface structures. Numerical results for return loss, reflection phase and input impedance are shown. The structure which shows the best performance is chosen to design a metamterial patch antenna. The results confirm and show the good characteristics of antenna with HIS ground plane. We have seen that, around of the resonant frequency, the return loss of the designed antenna decrease significantly and its gain and directivity are notably improved or enhanced.

The layout of many varieties of HIS structures circuits are now edited and the manufacturing process in progress. The comparison and discussion between the numerical results obtained by the simulation and experimental results will be presented in our future work.

REFERENCES

1. Daniel Sievenpiper Thesis, University California, Los Angeles, 1999.
2. Sievenpiper, D., L. Zhang, R. F. J. Broas, N. G. Alexopoulos, and E. Yablonovitch, "High impedance electromagnetic surfaces with a forbidden frequency band," *IEEE Trans., Micro., Theory Tech.*, Vol. 47, 2059–2074, 1999.
3. Yang, F. and Y. Rahmat-Samii, "Reflection phase characterizations of the EBG ground plane for low profile wire antenna applications," *IEEE Transacations on Antennas and Propagation*, Vol. 51, No. 10, 2003.
4. Gonzalo, R., P. de Maagt, and M. Sorolla, "Enhanced patch antenna performance by suppressing surface waves using photonic-bandgap substrates," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 47, No. 11, 1999.
5. Cheype, C., C. Serier, M. Thèvenot, T. Monédière, A. Reineixn, and B. Jecko, "An electromagnetic bandgap resonator antenna," *IEEE Transactions on Antennas and Propagation*, Vol. 50, No. 9, 2002.
6. Kim, Y., F. Yang, and A. Z. Elsherbeni, "Compact artificial magnetic conductor designs using planar square spiral geometries," *Progress in Electromagnetics Research*, PIER 77, 43–54, 2007.
7. Schurig, D., J. J. Mock, and D. R. Smith, "Electric-field-coupled resonators for negative permittivity metamaterials," *Appl Physics Letter*, Vol. 88, 041109, 2006.
8. Abdelwaheb Ourir Thesis, University Paris X, 2006.
9. Engheta, N. and R. W. Ziolkowski, *Electromagnetic Metamaterials: Physics and Engineering Explorations*, Wiley-IEEE Press, August 2006.

10. Yang, L., M. Fan, and Z. Feng, "A spiral electromagnetic bandgap structure and its application in microstrip antenna arrays," *Microwave Conference Proceedings*, 2005.
11. J. M. Bell, M. F. Iskander, and J. J. Lee, "Ultrawideband hybrid EBG/Ferrite grounf plane for low-profile array antennas," *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 1, 2007.
12. Zheng, Q. R., B. Q. Lin, Y. Q. Fu, and N. C. Yuan, "Characteristics and applications of a novel compact spiral electromagnetic bandgap structure," *J. of Electromagn. Waves and Appl.*, Vol. 21, No. 2, 199–213, 2007.
13. Liu, C. H., Y. G. Lu, X. G. Luo, and C. L. Du, "The numerical simulation of AMC characteristics for UCPBG on grounded dielectric substrate," *J. of Electromagn. Waves and Appl.*, Vol. 21, No. 6, 755–768, 2007.
14. Alù, A., F. Bilotti, N. Engheta, and L. Vegni, "Subwavalength, compact, resonant patch antennas loaded with metamaterials," *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 1, 2007.
15. Hosseini, M., A. Pirhadi, and M. Hakkak, "A novel AMC with little sensitivity to angle of incidence using an optimized jerusalem cross FSS," PIER 64, 43–51, 2006.

Force Generation of Selemion Governed by the Charge Quantity

H. Tamagawa
Department of Human and Information Systems, Faculty of Engineering
Gifu University, Japan

Abstract— It was reported previously that the force generated by Selemion appeared to have a direct correlation to the charge given to it. In this work, a quite simple relationship was observed between charge and force as reported before - linear relationship -, as long as the redox reaction of silver layers on Selemion surfaces was induced. Consequently, it was expected that the control of charge given to Selemion would be reflected as the control of force generated by Selemion. In fact, it was experimentally confirmed that the force Selemion generated could be easily controlled with the control of charge given to Selemion itself.

1. INTRODUCTION

Fabricating an electroactive polymer actuator has been an attractive theme for a number of researchers for the past decades [1–11]. The author of this paper has been involved in the investigation especially on the polymer actuators consisting of ion exchange polymer membranes [12–14]. Since the ion exchange polymer membranes exhibit large bending under a small applied voltage [2–6, 9–11], it has been believed that they can be employed as the bending mode polymer actuator materials.

Precise controllability of bending and force generation behavior of ion exchange polymer membrane actuator (hereafter called IEPM actuator) are fundamental factors for practical IEPM actuators. It was observed previously that a largely dehydrated IEPM called Selemion (Asahi glass, Co. Ltd., Japan), sandwiched between thin silver layers, retained the precise controllability of its bending curvature with the control of charge given to it [14]. Further, it was observed at the same time that the force generated by Selemion could be controlled by the control of charge as well [14].

In order to verify if the force generated by Selemion was truly governed by the charge given to Selemion itself, the further experiments were carried out, and the results are shown in this paper.

2. EXPERIMENTAL

2.1. Specimen Preparation

Selemion (containing -SO_3H group, thickness = 140 μm) was employed as a starting material of specimen. Its top and bottom surfaces were plated with silver with the silver mirror reaction. This silver-plated Selemion was dehydrated in a desiccator with a desiccant. Several hours before performing the experiments, it was cut into the strip shape, where its structure is shown in Fig. 1. This strip shape specimen was left in the air so that it could absorb a minute quantity of water from the air, since previously a minute quantity of water was found to play an essential role for the induction of precisely controllable Selemion bending by the charge control [13].

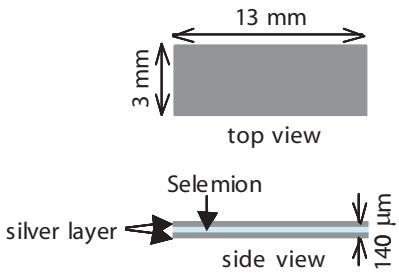


Figure 1: Structure of specimen.

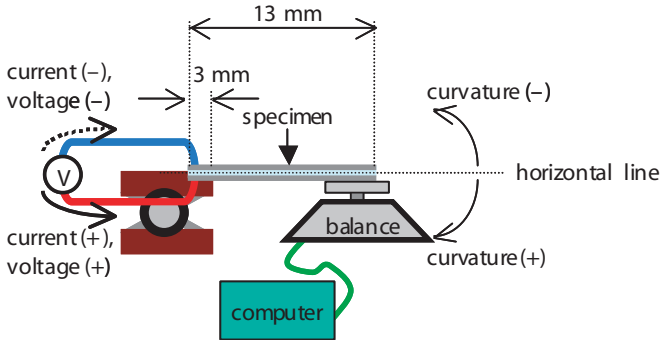


Figure 2: The experimental setup employed for the measurement of Selemion force.

2.2. Force Measurement under the Given Voltage

Force generated by Selemion was measured along with current under the given voltage. A Selemion specimen was horizontally clamped with the electrodes at the 3 mm narrow end, and the other end was placed slightly above the stage surface of balance as depicted in Fig. 2. Imposing a voltage on Selemion caused its bending, where the definitions of positive and negative of voltage, current, and curvature are given in Fig. 2. Downward bending of Selemion induced by the positive applied voltage pressed downward the stage of balance, exerting a force to the stage. The force was detected with the balance, and this force data was acquired in real time with the computer connected to the balance.

2.3. Force Control with Charge

The controllability of force with the control of charge was investigated. Force generated by Selemion was measured by using the same experimental setup as described in the previous section, but the given current was imposed to the Selemion in order to control the charge.

3. RESULTS AND DISCUSSIONS

3.1. Charge vs. Force

According to Ref. [13], the precise bending curvature controllability of Selemion was observed only when the redox reaction of Ag layers on Selemion surfaces was induced. There was a threshold voltage, V_{th}, at least required for the induction of Ag redox reaction [14]. Without controlling the bending curvature of Selemion, it would be impossible to control the force Selemion generates. Therefore it was necessary to know V_{th}, and V_{th} was experimentally determined as follows: Force was measured along with the current under the applied voltage increasing at the rate of $50\,\mathrm{mVs^{-1}}$ from 0 mV to 3000 mV employing the setup described in the section 2.2. Fig. 3 shows V (voltage) vs. I (current). Note that I represents the current of Selemion whose width is 1 mm, namely, I was obtained by the simple calculation that the actual measured current was divided by the specimen width. The abrupt increase of I was observed around at $V = 1300\,\mathrm{mV}$. It suggests the occurrence of redox reaction at $V = 1300\,\mathrm{mV}$. V vs. F (force) is shown in Fig. 4. F represents the force generated by Selemion whose width is 1 mm as well, namely, F was obtained by the calculation that the actual measured force was divided by the specimen width. F exhibited abrupt increase around at $V = 1300\,\mathrm{mV}$ as well. Thus $V = 1300\,\mathrm{mV}$ was defined as V_{th}.

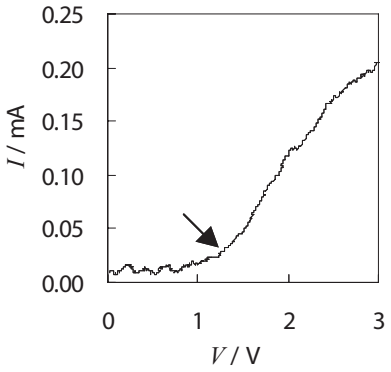


Figure 3: V vs. I. V increased at the rate of $50\,\mathrm{mVs^{-1}}$. I increased discontinuously at $V = 1300\,\mathrm{mV}$ indicated with an arrow.

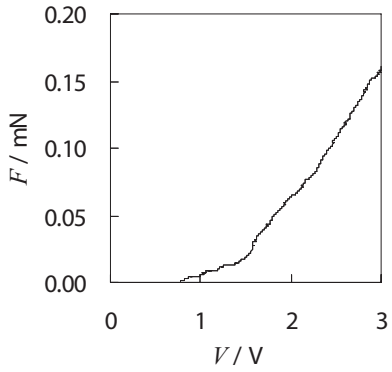


Figure 4: V vs. F. V increased at the rate of $50\,\mathrm{mVs^{-1}}$. F increased discontinuously at $V = 1300\,\mathrm{mV}$ indicated with an arrow.

Q (charge given to Selemion) vs. F under the application of constant $V = 2000\,\mathrm{mV}$ was measured, where $V = 2000\,\mathrm{mV}$ was imposed for 45 s. Here, Q was given by $Q(t) = \int_0^t I(t')\mathrm{d}t'$, where t represents time. Since 2000 mV was higher than V_{th}, the redox reaction of Ag layers on Selemion surfaces occurred. The result is shown in Fig. 5. Q vs. F formed a perfect straight line, that is, a direct correlation between Q and F. It was observed no or unordered F at the early stage in the dotted circled area in Fig. 5. It must have been caused by the technical difficulty in measuring F, when the bending curvature of Selemion was positive but quite small, close to $0\,\mathrm{mm^{-1}}$. As described earlier, the straight shape Selemion was horizontally clamped as depicted in Fig. 2 for

the measurement of F, and one end of it was placed slightly above the stage of balance initially. Such a slight gap between the end of Selemion and the stage of balance made it impossible to detect F induced by the slight bending of Selemion. Another reason of no or unordered F must have been due to the insufficient sensitivity of balance to F.

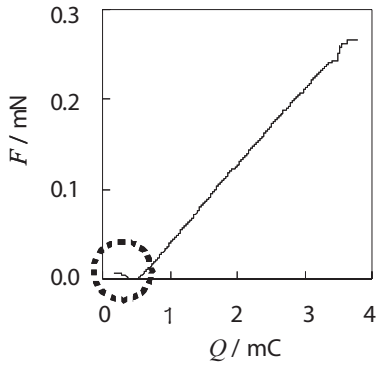


Figure 5: Q vs. F under the constant $V = 2000\,\mathrm{mV}$, where $V = 2000\,\mathrm{mV}$ was imposed for 45 s. See the text about the dotted circled area.

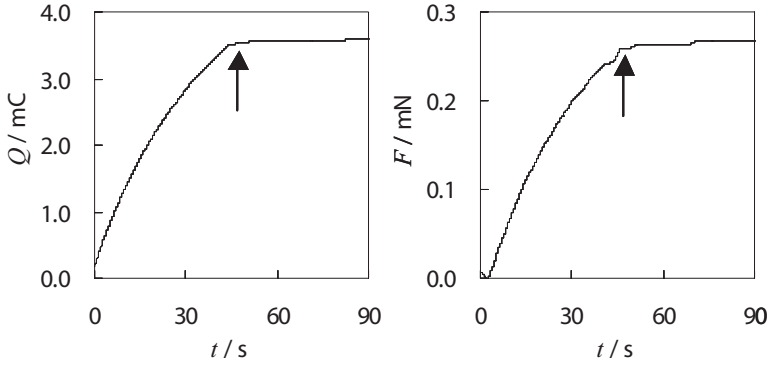


Figure 6: t vs. Q and t vs. F under the constant $V = 2000\,\mathrm{mV}$ followed by the open circuit at $t = 45\,\mathrm{s}$ indicated with the arrows.

In order to further see the relationship between Q and F, the measurement for Q and F was carried out under $V = 2000\,\mathrm{mV}$ for 45 s followed by the open circuit. Fig. 6 shows the results of t vs. Q and t vs. F. At the moment indicated with the arrows in Fig. 6, the circuit was opened. Q and F were maintained constant after $t = 45\,\mathrm{s}$. F did not decay even in the open circuit state. It strongly suggests that F has a direct correlation to Q.

3.2. Force Controllability with the Charge Control

From the experimental results and discussions so far, it was expected that F could be precisely controlled by the control of Q. In order to verify this speculation, the relationship of Q vs. F was investigated by measuring F under the controlled Q as described in the section 2.3. Experimental condition was as follows: a constant rate of Q was given to Selemion initially, and discontinuously the rate of Q was raised at $t = 45\,\mathrm{s}$, where Q was always controlled actually by controlling I imposed on Selemion. Fig. 7 shows t vs. Q and t vs. F. Discontinuous change of F was observed at $t = 45\,\mathrm{s}$ from this diagram along with the discontinuous change of Q at the same moment, both are indicated with the arrows. Fig. 7 was rearranged into Q vs. F as shown in Fig. 8. Q vs. F formed an almost complete straight line, despite the observation of discontinuous change of Q vs. F at $t = 45\,\mathrm{s}$. It might be interpreted as F was governed by Q, and F could be controlled with the control of Q. Therefore F was again measured under the complicatedly controlled Q. The

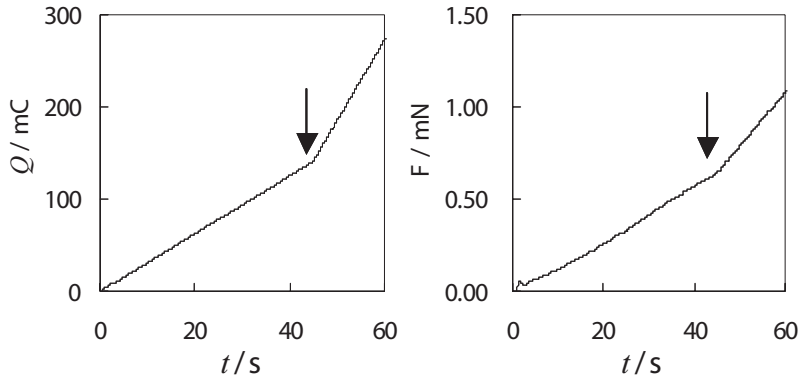


Figure 7: t vs. Q and t vs. F. The rate of Q was discontinuously raised at $t = 45\,\mathrm{s}$, indicated with an arrow. The rate of F also discontinuously increased at $t = 45\,\mathrm{s}$, indicated with an arrow.

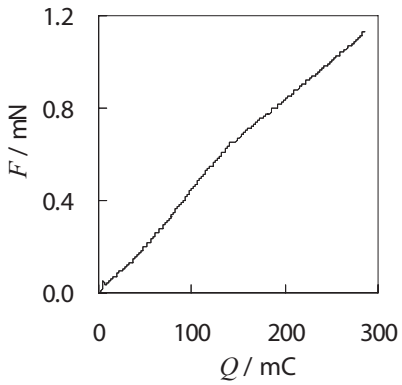


Figure 8: Q vs. F obtained by rearranging the data shown in Fig. 7.

behaviors of Q given to Selemion and F are shown as a function of t in Fig. 9, where the rate of Q from $t = 0\,\mathrm{s}$ to $20\,\mathrm{s}$ was $4.4\,\mathrm{mCs}^{-1}$ and $-4.4\,\mathrm{mCs}^{-1}$ from $t = 20\,\mathrm{s}$ to $22\,\mathrm{s}$, and the same Q cycle was repeated from $t = 22\,\mathrm{s}$ onward. t vs. Q looks quite similar to t vs. F, and these diagrams were rearranged into Q vs. F shown in Fig. 10. Q vs. F was almost on the straight line, despite the repeated significant rate change of Q, where the absolute value of V was always above V_{th}. This result also suggests that F is governed by Q, and the control of Q can leads to the control of F.

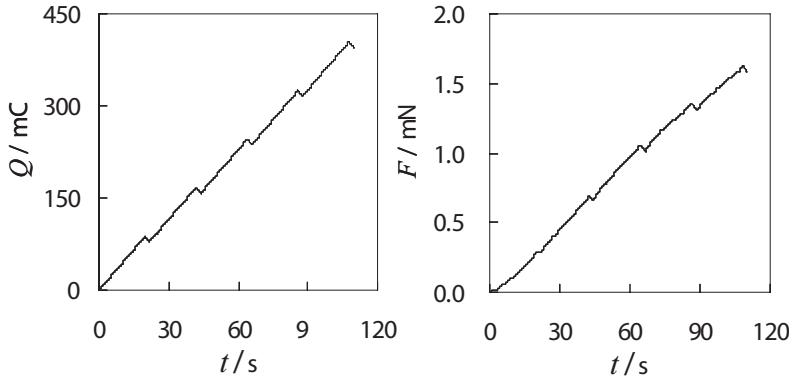


Figure 9: t vs. Q and t vs. F. Q increased constantly at the rate of $4.4\,\mathrm{mCs}^{-1}$ for $20\,\mathrm{s}$ followed by its constant decrease at the rate of $-4.4\,\mathrm{mCs}^{-1}$ for $2\,\mathrm{s}$, and the same cycle was repeated from that time onward.

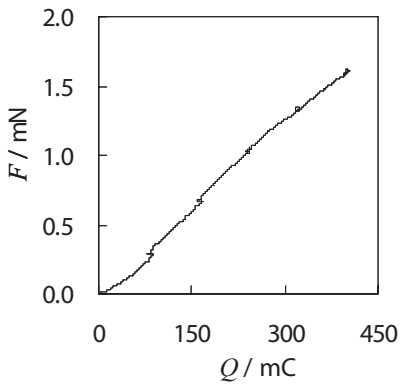


Figure 10: Q vs. F obtained by rearranging the data shown in Fig. 9.

4. CONCLUSIONS

It was confirmed that the force generated by the silver-plated Selemion in the largely dehydrated state was governed by the charge given to Selemion itself. The relationship between charge and force was represented by the almost straight line, as long as the redox reaction of silver layers on Selemion surfaces was induced. Such a correlation of force to the charge did not deteriorate even under the complicatedly controlled charge. Therefore we can speculate that the force generated by Selemion can be precisely controlled with the control of charge. This finding offers one reason that the fabrication of practical IEPM actuator must be possible in the future.

ACKNOWLEDGMENT

This research was conducted under the financial support by The Naito Research Grant.

REFERENCES

1. Tanaka, T., I. Nishio, S-T. Sun, and S. Ueno-Nishio, "Collapse of gels in an electric field," *Science*, Vol. 218, 467–469, 1982.
2. Oguro, K., Y. Kawami, and H. Takenaka, "Bending of an ion-conducting polymer film-electrode composite by an electric stimulus at low voltage," *Trans. J. Micromachine Soc.*, Vol. 5, 27–30, 1992.
3. Oguro, K., K. Asaka, and H. Takenaka, "Polymer film actuator driven by a low voltage," *Proc. of 4th International Symposium on Micro Machine and Human Science*, Nagoya, Japan, 39–40, 1993.
4. Asaka, K., K. Oguro, Y. Nishimura, M. Mizuhata, and H. Takenaka, "Bending of polyelectrolyte membrane-platinum composites by electric stimuli. I. response characteristics to various waveforms," *Polymer J.*, Vol. 27, 436–440, 1995.
5. Salehpoor, K., M. Shahinpoor, and M. Mojarrad, "Linear and platform type robotic actuators made from ion-exchange membrane-metal composites, smart materials technologies," *Proc. SPIE Smart Mater. Struct.*, Vol. 3040, 192–198, San Diego, 1997.
6. Shahinpoor, M., M. Mojarrad, and K. Salehpoor, "Electrically induced large amplitude vibration and resonance characteristics of ionic polymeric membrane-metal composites artificial muscles," *Proc. SPIE Conference*, Vol. 3041, 829–838, San Diego, 1997.
7. Iwamoto, K., G. Tuji, S. Yoshida, and M. Seno, "Bending behavior of bipolar membranes having a weakly acidic cation exchange membrane," *The Chem. Soc. Jpn.*, 425–429, 1997.

8. Bar-Cohen, Y., T. Xue, M. Shahinpoor, K. Salehpoor, J. Simpson, and J. Smith, "Low-mass muscle actuators using electroactive polymers (EAP)," *Proc. SPIE Smart Mater. Struct.*, Vol. 3324, 218–223, San Diego, 1998.
9. Oguro, K., N. Fujiwara, K. Asaka, K. Onishi, and S. Sewa, "Polymer electrolyte actuator with gold electrodes," *Proc. SPIE*, Vol. 3669, 64–71, 1999.
10. Popovic, S., "Design of electro-active polymer gels as actuator materials," PhD thesis, University of Washington, 2001.
11. Asaka, K., N. Mori, K. Hayashi, Y. Nakabo, T. Mukai, and Z. W. Luo, "Modeling of the electromechanical response of ionic polymer metal composites (IPMC)," *Proc. SPIE Annual International Symposium on Smart Structures/NDE 2004*, 172–181, 2004.
12. Tamagawa, H., F. Nogata, T. Watanabe, A. Abe, K. Yagasaki, and J.-Y. Jin, "Influence of metal plating treatment on the electric response of Nafion," *J. Mater. Sci.*, Vol. 38, 1039–1044, 2003.
13. Tamagawa, H., F. Nogata, and S. Popovic, "Roles of Ag redox reaction and water absorption inducing the Selemion bending," *J. Membrane Sci.*, Vol. 251, 145–150, 2005.
14. Tamagawa, H. and F. Nogata, "Dominant factor for the precise bending control of Selemion," *Sens. & Actuators: B: Chem.*, Vol. 120, 19–24, 2006.

Electro-optic Properties and Phase Behavior of Chiral-nematic Molecules

Kyongok Kang[1], **Samuel Sprunt**[2], **and Jan K. G. Dhont**[1]

[1]Forschungszentrum Juelich, IFF-Weiche Materie, D-52425, Juelich, Germany
[2]Department of Physics, Kent State University, Kent, OH, USA

Abstract— We will discuss electro-optic properties and phase-behavior of lyotropic chiral-nematic materials under electric fields.

(I) Fingerprint rolls of a chiral nematic between two antiparallel-rubbed ITO substrates are formed at low electric field strengths. This pattern can be photo-stabilized through UV polymerization of added monomers by means of an UV focal-conic Gaussian laser beam. The chiral nematic liquid crystalline molecules are still free to diffuse within the polymer-stabilized confining structure. The resulting stabilized pattern can be used as a highly diffractive electrically switchable polymer-stabilized cholesteric diffraction grating (PSCDG). The regular spatial variation of chiral-nematic material gives rise to strong light intensity in forward diffraction peaks, together with weaker intensity peaks of half the spacing from the polymer matrix. The electro-optic properties and dynamical modes of these gratings depend on the dynamics of the chiral-nematic liquid within the polymer matrix, which is found to depend on whether isotropic or mesogenic monomers are used. We found a strong effect of confinement on the dynamics of chiral-nematic order fluctuations in the mesogenic polymer network for scattering vectors along the helicoidal axis of the fingerprint-roll structures.

(II) Preliminary experiments on the phase-behavior of fd-virus suspensions under external electric fields will be shortly discussed. A phase diagram in the electric field-amplitude versus frequency plane will be presented. At low ionic strengths, electric field-induced polarization of the double layer induces an isotropic to chiral-nematic phase transition. At higher field strengths we observe melting of the chiral nematic state, which is probably due to either induced dipole-dipole interactions or the finite relaxation time of the double-layer charge density.

1. HIGHLY DIFFRACTIVE ELECTRICALLY SWITCH-ABLE POLYMER-STABILIZED CHOLESTERIC DIFFRACTION GRATINGS (PSCDGS)

Chiral macromolecules can be used to produce electro-optic active materials. Highly diffractive ($\sim 90\%$ diffraction efficiency) electro-responsive gratings can be made by UV-induced polymerization of mesogenic types of monomers in which chiral nematic molecules are embedded. Polymerization is most conveniently performed by means of an in-situ focal conic UV laser photo-polymerization under an electric field, where "controllable" sizes of fingerprint rolls are formed [1, 2]. The two images in Fig. 1 show an example of the typical optical morphology of the mesoscopic structure of such a Polymer-Stabilized Cholesteric Diffraction grating (PSCDG) as observed under crossed polarizers.

The left image is a higher magnification of a regular fingerprint roll structure shown in the morphology of the right side (mesogenic type) PSCDG [1]. A schematic of the distribution of polymer and liquid crystalline material within the polymer-stabilized chiral-nematic fingerprint is given below the images in Fig. 1. After photo-polymerization, the chiral-nematic molecules are still able to diffuse and reorient within the confined regular structure formed by the polymerized template. The structure of the chiral-nematic liquid crystals can thus be switched between a highly diffracting "On-state" at lower electric field and the minimum diffracting "Off-state" at higher electric field strength. The typical diffracting performance is demonstrated in Fig. 2. There are strong, primary diffraction peaks due to the regularly spaced chiral-nematic liquid crystal and weaker, secondary peaks due to the presence of polymer-network.

The switching properties of the grating are determined through the diffusive motion of the chiral-nematic molecules through the confining polymer network. The performance of switching between the diffracting "On-state" and "Off-state" is approximately a few tens of milli seconds [1]. A typical confining geometry as imposed by the reactive mesogenic polymer network is shown as a SEM image and a depolarized optical microscopy image in Fig. 3 (upper images). The SEM image shows the bare morphology of the polymer walls after removal of the chiral nematic liquid

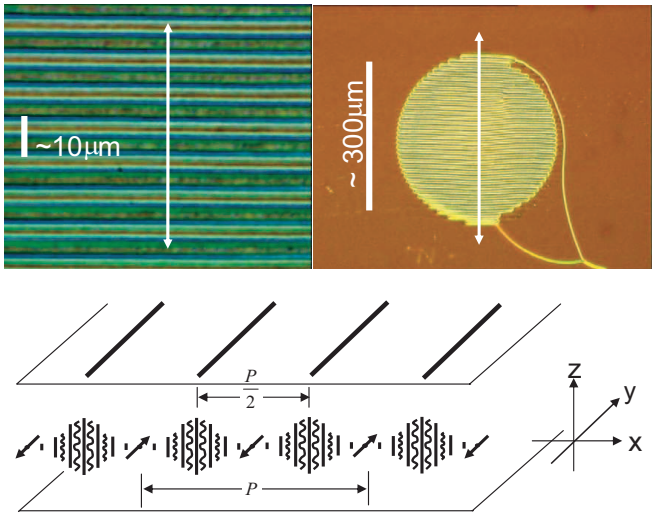


Figure 1: Typical depolarized optical morphology of PSCDGs made of mesogenic monomer RM257-based micrograting (see also Ref. [1]), and the brief sketch of the regular structure of polymer-stabilized chiral nematic fingerprint texture. The cell thickness is the same as the optical pitch of chiral nematics $d = 10\,\mu\text{m}$. The polarization direction of the UV light used to polymerize the momomer is along the $\hat{y}$-direction. Arrows in the upper images indicate the direction of the helicoidal $\hat{x}$ axis.

crystal material, which indicates the presence of dynamic curvature of the embedded chiral nematic material between the "polymerized walls". The dynamics of the chiral molecules within the polymer-stabilized fingerprint texture has been studied by means of dynamic light scattering [3]. Scattering studies have been performed for two different types of polymer-networks by isotropic (HDDA) and reactive mesogenic (RM257) monomers. When the scattering vector is perpendicular to the helicoidal axis, there is no significant difference in the relaxation process of chiral nematics

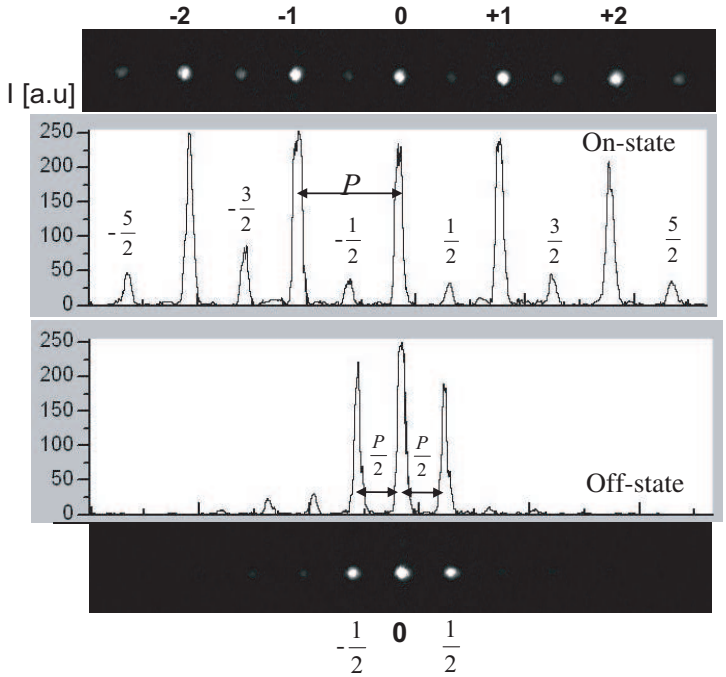


Figure 2: Typical forward transmitted diffraction spectrum of a PSCD state in the z-direction: Diffraction "On-state" occurs under a low electric field conditions ($|E| \sim 0.35\,[\text{V}/\mu\text{m}]$), while the minimum diffraction is shown as "Off-state" ($|E| \sim 1.60\,[\text{V}/\mu\text{m}]$) at high field strengths. The primary (integer) and secondary (half-integer) peaks are due to the optical pitch of chiral nematics (as $P = 10\,\mu\text{m}$). The (mesogenic RM-257 based) polymer network has shown a pitch equal to $P/2 = 5\,\mu\text{m}$. More details can be found in Ref. [1].

in both polymer networks. However when the scattering vector is along the helicoidal axis, there is a significant difference in the slow relaxation of thermal fluctuations of the chiral molecules, depending on the monomer type embedding the chiral nematic molecules, as can be seen from the scattered intensity auto-correlation functions given in Fig. 3 [3]. The chiral molecules relax faster in the isotropic (HDDA) polymer-network as compared to the reactive mesogenic (RM257) polymer-network. This is due to a more effective coupling of the chiral nematic molecules to a slow polymer mode for the reactive mesogenic type of polymer-network along the helicoidal axis [3].

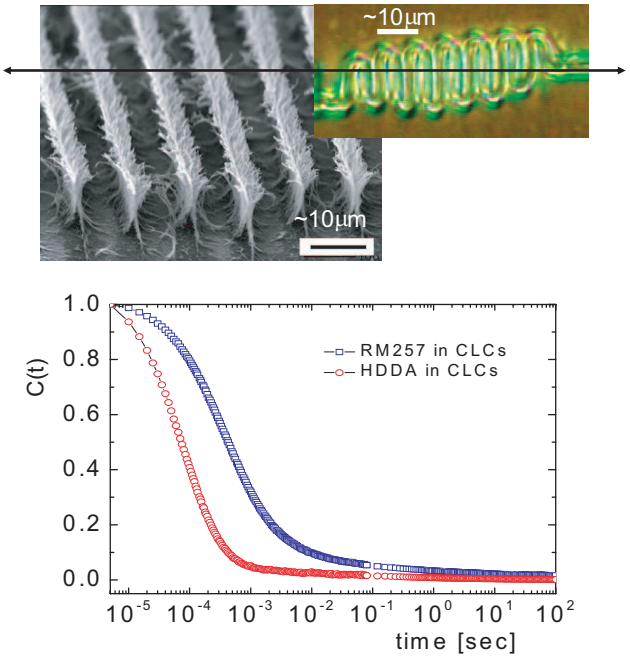


Figure 3: **Upper**: SEM image of bare mesogenic (RM257) polymer-network (courtesy of S. Kang et al.) and a depolarized optical morphology of polymer-stabilized chiral nematic fingerprint rolls (the arrow indicates the direction of the helicoidal axis). **Lower**: Normalized scattered intensity auto-correlation functions, which characterize the decay of thermal fluctuations of photostzbilized chiral nematic molecules along the helicoidal axis, where two different types of polymer-networks are shown for a given scattering vector ($q_x \sim 9.25\,\mu\text{m}^{-1}$).

2. PRELIMINARY RESULTS ON THE PHASE-BEHAVIOR OF FD-VIRUS SUSPENSIONS UNDER EXTERNAL ELECTRIC FIELDS

We recently started to study the electric-field induced phase behaviour of fd-virus suspensions in water at low ionic strength. Fd virus is a very long (880 nm) and thin (6.7 nm) rod-like particle with a relatively large persistence length (2500 nm). There is an extended double layer around these stiff colloidal rods at low ionic strengths, which can be polarized by means of an external electric field. The core of the virus is not electrically polarized by the moderate electric field strengths used in our study. It has been observed before [4], by means of birefringence studies, that at low fd-concentrations the rods tend to align along the electric field lines, as expected, but that at higher concentrations (a few time the overlap concentration) the direction of alignment changes with increasing field strength and/or frequency. The reason for this behavior is unknown, but is clearly related to interactions between the polarized double layers. The suspensions of fd-virus particles at low ionic strength exhibit an isotropic to chiral nematic (N*) phase transition at much higher concentrations (of about 50–100 times the overlap concentration). The interactions between the polarized double-layers are therefore expected to have a pronounced effect on the phase behavior just below the isotropic to chiral nematic (N*) transition concentration without an electric field. A preliminary phase diagram in the AC field-amplitude versus frequency plane is presented in Fig. 4 for a fixed fd-virus concentration and a low ionic strength. In the absence of the external electric field the dispersion is isotropic. With increasing AC electric field amplitude at a fixed frequency, we observe a transition from an isotropic to a chiral-nematic state. At higher field strengths, the chiral nematic (N*) phase melts. The reason for such a melting might be connected to a significant disruption of the double layer due to the large ion-currents past the core of the fd-virus particles.

Note that melting is also observed at fixed field amplitude on increase of the frequency. This might be due to the finite response time of the polarization of the double layer.

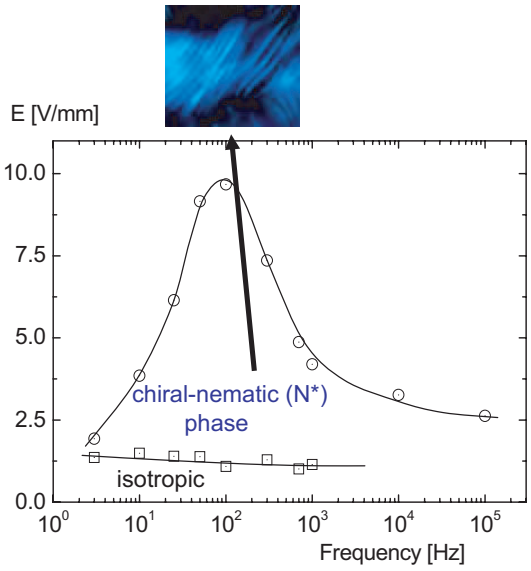


Figure 4: A preliminary phase diagram and a typical morphology of chiral nematic (N*) phase of fd-virus suspensions under AC external electric fields ([fd] ~ 2.0 mg/ml at 0.16 mM Tris/HCl buffer). The largest optical pitch of an electric field-induced N* phase is $\sim 10\,\mu$m as observed by depolarization microscopy.

3. CONCLUSIONS

Chiral nematic molecules are interesting both for their electro-optic properties and phase behavior under external electric fields. Chiral nematic fingerprint rolls that are formed under electric fields can be stabilized by means of photo-polymerization of a monomer matrix. The resulting regular, stable structures exhibit high diffraction efficiency. Due to the degree of "reorientation" freedom of the chiral nematic molecules, the electro-optic responsive diffraction gratings can be switched between highly diffracting "On-state" and the minimum diffraction "Off-state" by means of an external electric field.

We have also observed that a chiral-nematic can be induced in the isotropic dispersions of fd-virus suspensions at a low ionic strength by applying an external electric field. This may be due to interactions between polarized double layers. At high electric field strengths, the chiral-nematic (N*) phase melts. Further investigation of other fd-concentraions is useful for understanding the precise mechanisms of such phase behavior under an external electric field.

ACKNOWLEDGMENT

Part of this research was supported by the Office of Naval Research under Grant No. N00014-99-1-0899, and part by the Deutsch-Dutch collaborative SFB-TR6 and SoftComp.

REFERENCES

1. Kang, K., L. C. Chien, and S. Sprunt, "Polymer-stabilized cholesteric liquid crystal micrograt-ings: a comparison of polymer network formation and electro-optic properties for mesogenic and non-mesogeninc monomers," *Liq. Cryst.*, Vol. 29, No. 1, 9–18, 2002.
2. Kang, K. and S. Sprunt, "Light-controlled polymerization kinetics for the photopolymerization of cholesteric fingerprint rolls," *Mol. Cryst. Liq. Cryst.*, Vol. 466, 23–38, 2007.
3. Kang, K. and S. Sprunt, "Dynamics of distorted cholesterics in a mesogenic versus isotropic polymer network," *Phys. Rev. E.*, Vol. 72, 031702, 2005.
4. Kramer, H., C. Graf, M. Hagenbuche, C. Johner, C. Martin, P. Schwind, and R. Weber, "Electro-optic effects of aqueous fd-virus suspensions at very low ionic strength," *J. Phys. II*, Vol. 4, 1061–1074, France, 1994.

Design of Composite Electromagnetic Wave Absorber Made of Soft Magnetic Materials Dispersed and Isolated in Polystyrene Resin

K. Sakai, **Y. Wada**, **and S. Yoshikado**
Department of Electronics, Doshisha University, Japan

Abstract— Composite electromagnetic wave absorbers made of a soft magnetic material (permalloy or sendust) and polystyrene resin were investigated. The volume mixture ratio of magnetic material was varied in the range from 18 vol% to 75 vol%. The composites with the low volume mixture ratio of soft magnetic material absorbed more than 99% of electromagnetic wave power in the frequency rage from 1 GHz to 12 GHz. The values of the real part μ_r' of the relative complex permeability μ_r^* for both magnetic materials were less than unity at frequencies above approximately 6 GHz as the volume mixture ratio of magnetic material increased. This result suggests the possible realization of an electromagnetic wave absorber that can operate above 10 GHz.

1. INTRODUCTION

Electromagnetic waves with frequencies higher than 1 GHz are now more widely used for communication with the increasing use of telecommunication devices such as mobile phones. In the case of electric toll collection (ETC) systems, their operating frequency will increase from 5.8 GHz to more than 10 GHz in the future. Therefore, the development of an electromagnetic wave absorber suitable for these frequency bands is required.

The purpose of this study is to investigate the design of a practical composite absorber that operates in a wide frequency range above 1 GHz. The composite absorber consists of soft magnetic material particles, such as permalloy or sendust, and polystyrene resin. Both permalloy and sendust satisfy Snoek's limit at high frequencies [1] and have high permeability values in the frequency range above 1 GHz. These characteristics make it possible to fabricate an electromagnetic wave absorber suitable for this frequency band [2, 3]. In addition, it is expected that the composite materials have values of μ_r', the real part of the relative complex permeability μ_r^*, of less than unity, because in qualitative theoretical calculations it has been predicted that the μ_r' has to be less than unity for a constant relative dielectric constant ε_r'. This characteristic allows electromagnetic wave absorption above 10 GHz. These soft magnetic materials, however, have high electric conductivity. Thus, if the number of particles of such a magnetic material dispersed in the resin, such as polystyrene, exceeds the percolation threshold, magnetic particles will be in direct contact with each other and the average conductivity σ of the composite will increase markedly. Eventually, the reflection coefficient of the electromagnetic wave reflected by the composite will increase and the absorption characteristics will be degraded [4]. To prevent the increase in σ, we attempt to disperse and isolate the magnetic particles in the polystyrene resin so that they are not in contact with each other. To isolate the magnetic particles, the surface of each magnetic particle is coated with very fine polystyrene particles of less than approximately 1 μm diameter, which is less than that (approximately 20 μm diameter) of the magnetic particles.

2. EXPERIMENTS

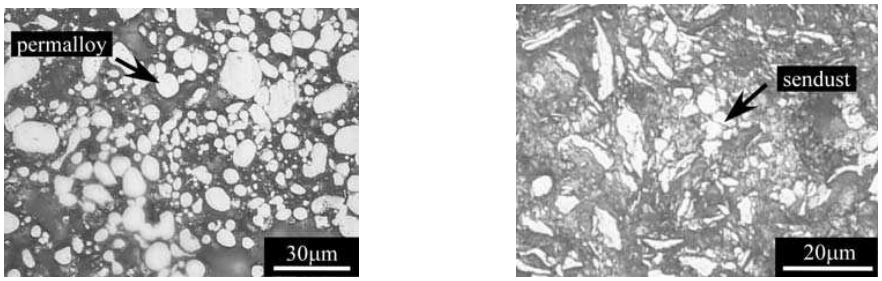
Chips of polystyrene resin of approximately 200 μm diameter were ground with ethanol by mechanical milling (MM) (Fritsch, P7) using a zirconia pot and zirconia balls of 1 mm diameter for 1.5 h. The rotation speed of the turntable was 600 rpm. The ratio of the rotation speeds of the milling pot to the turntable was 2:1. Permalloy (Ni 45%, Fe 55%) particles (grain type, average grain size of approximately 10 μm) or sendust (Al 5%, Si 10%, Fe 85%) particles (flake type, average grain size of approximately 20 μm) and ground polystyrene particles (average grain size of approximately 1 μm) were then mixed by MM for 30 min to coat the particles of the magnetic material with the polystyrene particles. After mixing, the powder mixture was heated to melt the polystyrene resin then hot-pressed at a pressure of 5 MPa into a pellet shape. Then, the pellet was cooled naturally to room temperature and processed to a toroidal-core shape (outer diameter of 7 mm and inner diameter of 3.04 mm) for use in a 7 mm coaxial line or to a rectangular shape (22.8 mm × 10.16 mm) for use in a waveguide at the X-band (8.2 GHz to 12.4 GHz). The sample was loaded into a coaxial line or rectangular waveguide while ensuring that there was no gap between the coaxial line

or rectangular waveguide and the toroidal-core sample. The complex scattering matrix elements, S_{11}^* (reflection coefficient) and S_{21}^* (transmission coefficient) were measured using a vector network analyzer (Agilent Technology 8722ES) by the full-two-port or one-port reflection method. The values of μ_r^* ($\mu_r^* = \mu_r' - j\mu_r''$, $j = \sqrt{-1}$) and the relative complex permittivity ($\varepsilon_r^* = \varepsilon_r' - j\varepsilon_r''$) were calculated from the data of both S_{11}^* and S_{21}^*. The return loss R for various sample thicknesses was calculated from the complex reflection coefficients Γ^* using the relation $R = 20\log_{10}|\Gamma^*|$.

3. RESULTS AND DISCUSSION

3.1. Dispersion State of Magnetic Material Particles and the Conductivity of Composite Materials

Surface optical microphotographs of the composites made of permalloy or sendust are shown in Fig. 1. Each particle of the magnetic material was surrounded by polystyrene resin even if the volume mixture ratio of the magnetic material exceeded the percolation threshold of 33 vol%, as shown in Fig. 1. This is because very fine polystyrene particles coated the surface of each particle of the magnetic material and melted polystyrene resin entered between the particles of the magnetic material. The values of ε_r'' increased gradually for both composite with the volume mixture of the magnetic material, even when it exceeded the percolation threshold. These results show that the particles of the magnetic material are isolated from each other and that the composites have a low value of σ.



(a) 75 vol% permalloy composite (b) 57 vol% sendust composite

Figure 1: Surface optical micro-photographs of composites: (a) 75 vol% permalloy and (b) 57 vol% sendust.

3.2. Frequency Dependence of μ_r' and Absorption Characteristics

Figures 2 and 3 show the frequency dependences of μ_r' and μ_r'' for the composites made of permalloy or sendust, respectively. In this study, a qualitative theoretical calculation was performed to investigate the absorption in the frequency range above 1 GHz. The measured values of μ_r^* for the composites made of permalloy or sendust and the calculated values of μ_r^* that satisfy the nonreflective condition Equation (1) are shown in Figs. 4 and 5, respectively [5].

$$1 = \sqrt{\mu_r^*/\varepsilon_r^*}\tanh\left(\gamma_0 d\sqrt{\mu_r^*\varepsilon_r^*}\right) \tag{1}$$

Here, γ_0 is the propagation constant in free space and d is the sample thickness. The value of ε_r' used for calculation was independent of frequency and the same as the measured value. ε_r'' was assumed to be zero. The measured values of μ_r' for the composite made of 33 vol% permalloy roughly agreed with the calculated values for $d = 3$ and 4 mm, as shown in Fig. 4, and those for the composite made of 25 vol% sendust almost agreed with the calculated values for $d = 4$ mm, as shown in Fig. 5, in the frequency range from 1 GHz to 10 GHz. On the other hand, the plots of the measured values of μ_r'' for the the composite made of 33 vol% permalloy intersected the calculated line near 6 GHz for $\varepsilon_r' = 10$ and $d = 3$ mm, and those for the composite made of 25 vol% sendust intersected the calculated line near 5 GHz for $\varepsilon_r' = 10$ and $d = 4$ mm, as shown in Figs. 4 and 5. Therefore, it is expected that the absorption of a large amount of electromagnetic wave power occurs at approximately 6 GHz and $d = 3$ mm for the composite made of 33 vol% permalloy and at approximately 5 GHz and $d = 4$ mm for the composite made of 25 vol% sendust, as shown in Fig. 6. Fig. 6 shows the absorption center frequency f_0 and the normalized -20 dB bandwidth (the bandwidth corresponding to the return loss of -20 dB is divided by f_0) of each composite. The value of -20 dB corresponds to the absorption of 99% of the electromagnetic wave power.

The composites made of permalloy or sendust showed a return loss of less than −20 dB for the optimum d value and the bandwidth was approximately 10%. Furthermore, the sample thickness for which the return loss becomes less than −20 dB was relatively thin compared with that of commercial absorbers. In particular, the composites made of sendust have some advantages over those made of permalloy because sendust contains no rare metals such as Ni and the mass density of the composites made of sendust is very small. For example, the mass density is approximately 1.5 for the composite made of 18 vol% sendust and approximately 1.8 that of 33 vol% sendust.

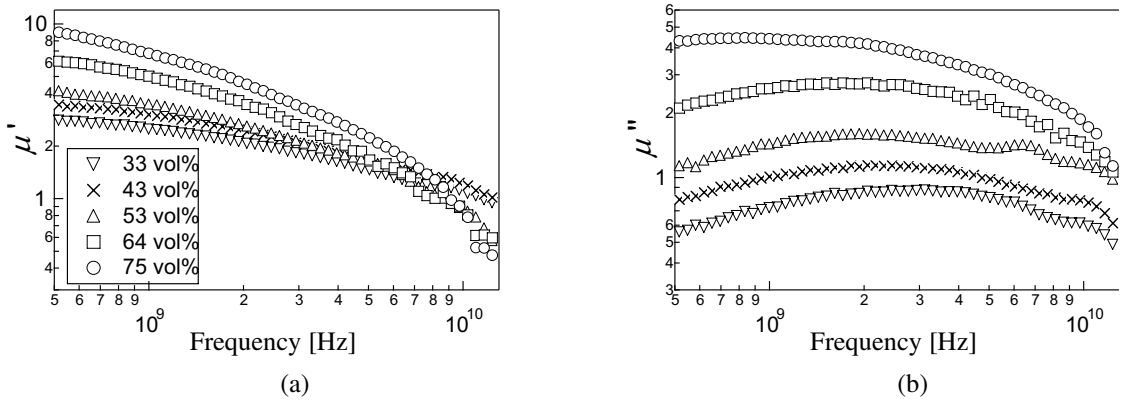


Figure 2: Frequency dependences of (a) μ_r' and (b) μ_r'' for composites made of permalloy and polystyrene resin.

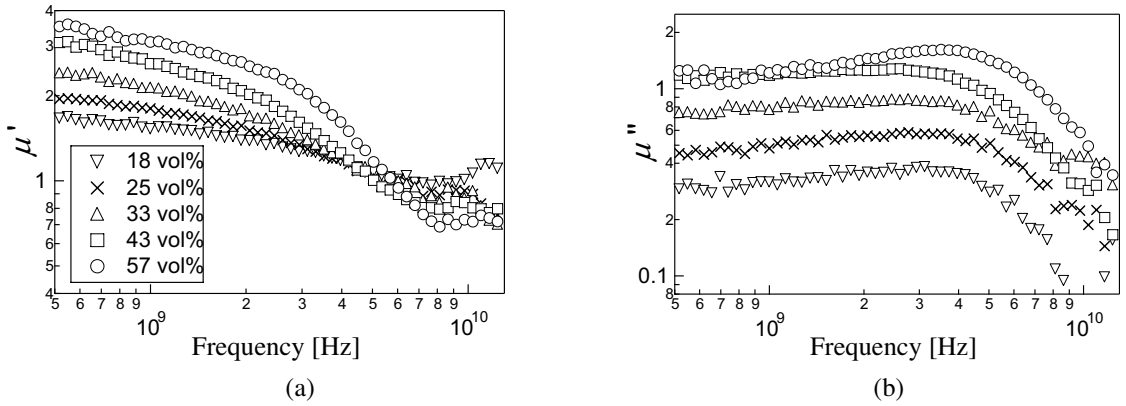


Figure 3: Frequency dependences of (a) μ_r' and (b) μ_r'' for composites made of sendust and polystyrene resin.

To investigate the absorption at frequencies near 10 GHz, the values of μ_r^* and ε_r^* for the composite made of 25 vol% sendust were measured by the reflection method using the rectangular waveguide at X-band frequencies because the accuracy of measurement was poor using the coaxial line due to the generation of higher-order modes. The frequency dependences of the return loss at X-band frequencies for the composite made of 25 vol% sendust are shown in Fig. 7. The absorbing center frequency f_0 increases from 8.5 GHz to 11.6 GHz as the sample thickness decreases from 2 mm to 1.5 mm. f_0 varies markedly over the narrow range of sample thickness as the frequency becomes higher than 8 GHz. The return loss was less than −20 dB near 11.6 GHz for the sample thickness of 1.5 mm, as shown in Fig. 7, and the bandwidth was approximately 6%. Therefore, the composite made of 25 vol% sendust can also be used as an electromagnetic wave absorber in the frequency range above 10 GHz.

The composites made of permalloy or sendust showed values of μ_r' of less than unity at frequencies above approximately 6 GHz as the volume mixture ratio of magnetic material increased, as shown in Figs. 2 and 3. It is speculated that this phenomenon is due to magnetic moments generated by an eddy current flowing on the surface of the particles of the magnetic material. Soft magnetic material is generally conductive and the skin depth δ of soft magnetic material, which is given by Equation (2), is less than the radius of the soft magnetic material particles at frequencies above approximately 5 GHz.

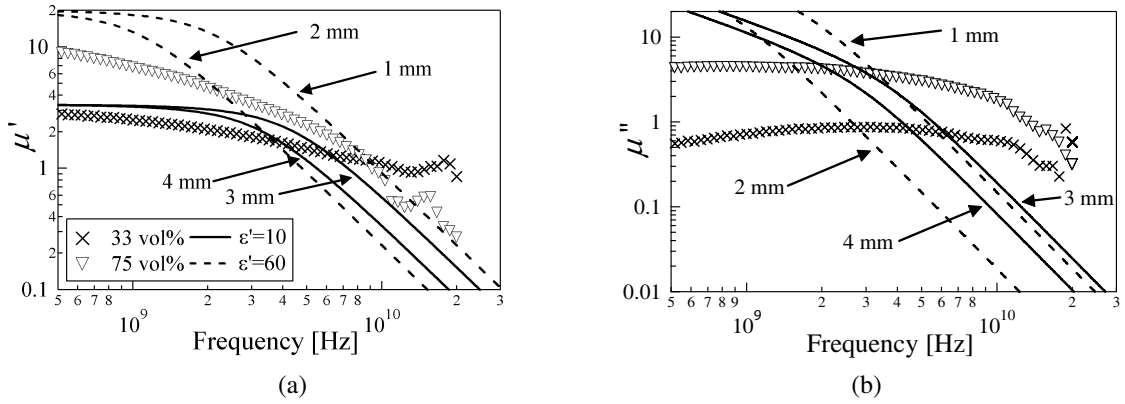


Figure 4: Measured and calculated values of (a) μ_r' and (b) μ_r'' using Equation (1) for composites made of permalloy.

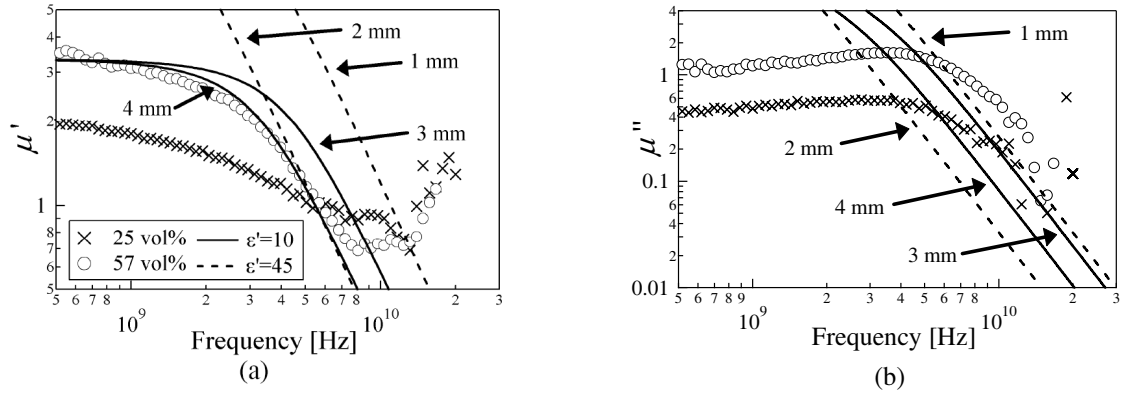


Figure 5: Measured and calculated values of (a) μ_r' and (b) μ_r'' using Equation (1) for composites made of sendust.

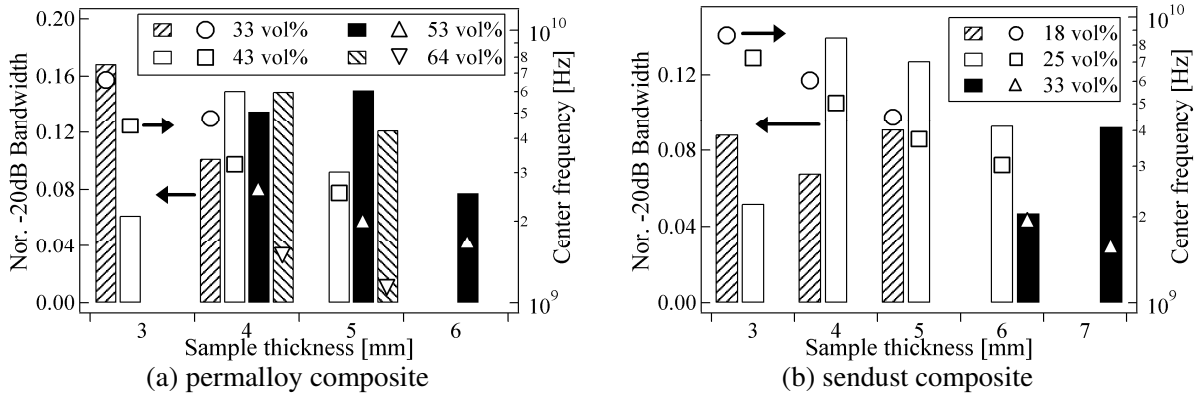


Figure 6: Values of normalized −20 dB bandwidth and center frequency f_0 of composites made of (a) permalloy and (b) sendust. Bars show the normalized −20 dB bandwidth and symbols show the absorption center frequency.

$$\delta = \sqrt{\frac{2\rho}{\omega\mu_0\mu_r'}} \tag{2}$$

Here, ω is the angular frequency and ρ is the resistivity. For example, δ for this permalloy is estimated to be 3.1 μm at 5 GHz because the resistivity of the permalloy is $2 \times 10^{-7}\,\Omega$m [6] and it decrease in proportion to $1/\sqrt{\omega}$. The eddy current flows in the layer of skin depth δ on a magnetic material particle and generates a magnetic moment antiparallel to incident magnetic field. Consequently, the value of μ_r' is reduced and sometimes becomes less than unity. This result suggests

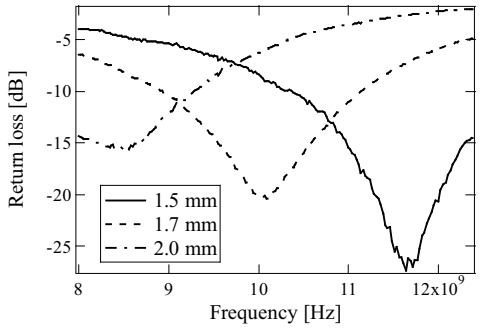


Figure 7: Values of return loss for the composite made of 25 vol%-sendust.

the possibility of electromagnetic wave absorption above 10 GHz, because Equation (1) indicates that the values of μ_r' should be less than unity in the high-frequency region. The values of $1-\mu_r'$ for various volume mixture ratios of permalloy and sendust are shown in Fig. 8. $1-\mu_r'$ increased with the increase in the volume mixture ratio of magnetic material. In particular, $1-\mu_r'$ for sendust was roughly proportional to the volume mixture ratio above 25 vol%. This result indicates that the effect of the magnetic moments generated by the eddy current becomes dominant as the amount of magnetic material increases. To examine this phenomenon, composites made of aluminum particles having a grain size of approximately 80 µm and polystyrene resin were prepared. Fig. 9 shows the values of $1-\mu_r'$ for various volume mixture ratios of aluminum. $1-\mu_r'$ increased proportional to the volume mixture ratio of aluminum particles. Although the grain size and skin depth of the aluminum particles are different from those of the magnetic material particles, a similar result to the composites made of permalloy or sendust was observed in the composites made of aluminum. Therefore, it is speculated that the reason why the vales of μ_r' become less than unity may be qualitatively explained by magnetic moment generated by the eddy current.

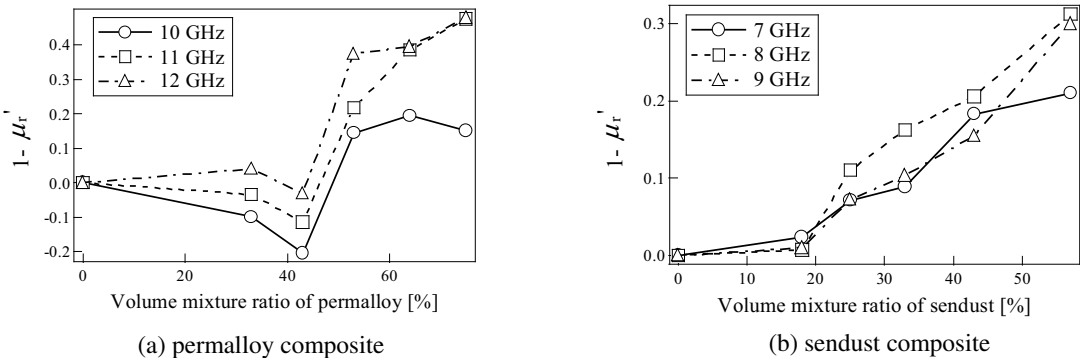


Figure 8: Values of $1-\mu_r'$ for various volume mixture ratios of (a) permalloy and (b) sendust.

For the composite with a volume mixture ratio of 75 vol% permalloy, the measured values of μ_r' almost agreed with the calculated values at frequencies from approximately 8 GHz to 20 GHz for $\varepsilon_r' = 60$ and $d = 1$ mm, as shown in Fig. 4(a). On the other hand, the measured values of μ_r'' agreed with the calculated values only at approximately 3.5 GHz. Therefore, the composites with a high volume mixture ratio of permalloy did not show a return loss of less than −20 dB in the measured frequency range. A similar result was obtained for the composites with a high volume mixture ratio of sendust. However, it is speculated from Figs. 4 and 5 that the lines showing the frequency dependence of the measured values of μ_r'' may intersect the calculated lines at a frequency above 10 GHz, which is outside the range of the measurement, because μ_r^* rapidly decrease in the high-frequency range. Therefore, it may be possible to show a return loss of less than −20 dB at frequencies above 10 GHz. Investigations of the absorption in this high-frequency range is now in progress.

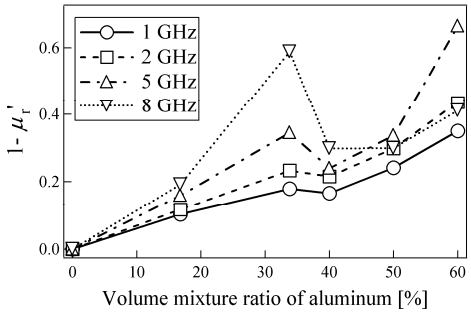


Figure 9: Values of $1 - \mu_r'$ for various volume mixture ratios of aluminum.

4. CONCLUSIONS

Composites with a low volume mixture ratio of soft magnetic material showed show a return loss of less than −20 dB in the frequency range from 1 GHz to 12 GHz. It is concluded that these composites are suitable for use in practical electromagnetic wave absorbers in the frequency range used for ETC systems or mobile phones.

The values of μ_r' for composites with a high volume mixture ratio of soft magnetic material were less than unity in the frequency range above approximately 6 GHz, and the absorption of electromagnetic wave power at frequency range above 10 GHz is expected.

REFERENCES

1. Snoek, J. L., "Dispersion and absorption in magnetic ferrites at frequencies above one megacycle," *Physca XIV*, Vol. 14, 207–217, 1948.
2. Matsumoto, M. and Y. Miyata, "Thin electromagnetic wave absorber for quasi-microwave band containing aligned thin magnetic metal particles," *IEEE Trans. Magn.*, Vol. 33, No. 6, 4459–4464, 1997.
3. Olmedo, L., G. Chateau, C. Deleuze, and J. L. Forveille, "Microwave characterization and modelization of magnetic granular materials," *J. Appl. Phys.*, Vol. 73, No. 10, 6992–6994, 1993.
4. Kusunoki, H., T. Kondo, and S. Yoshikado, "Fabrication and evaluation of composite electromagnetic wave absorbers," *Key Eng. Mater.*, Vol. 228–229, 315–318, 2002.
5. Musal, Jr., H. M. and H. T. Hahn, "Thin-layer electromagnetic absorber design," *IEEE Trans. Magn.*, Vol. 25, No. 5, 3851–3853, 1989.
6. Kasagi, T., T. Tsutaoka, and K. Hatakeyama, "Particle size effect on the complex permeability for permalloy composite materials," *IEEE Trans. Magn.*, Vol. 35, No. 5, 3424–3426, 1999.

An Unsupervised Classification Method for Polarimetric SAR Images Based on Inhomogeneous Markov Random Field and Graph Cuts

Xing Rong[1], **Jian Yang**[1], **Weijie Zhang**[1], **Wen Hong**[2], **and Fang Cao**[2]

[1]Dept. of Electronic Eng., Tsinghua University, Beijing 100084, China
[2]National Key Lab of Microwave Imaging Technology
Institute of Electronics, CAS, Beijing 100080, China

Abstract— A new unsupervised classification method is proposed for polarimetric SAR images to keep the spatial coherence of pixels and edges of different kinds of targets simultaneously. We consider the label scale variability of images by combining Inhomogeneous Markov Random Field (MRF) and Bayes' theorem. After minimizing an energy function using an expansion algorithm based on Graph Cuts, we can obtain classification results that are discontinuity preserving. Using a NASA/JPL AIRSAR image, we demonstrate the effectiveness of the proposed method.

1. INTRODUCTION

Due to the fact that the prior information about cluster center characteristics of SAR images cannot be obtained in many cases, it is necessary to develop unsupervised classification methods for POLSAR images. Cloude and Pottier proposed an unsupervised classification method based on their target decomposition theory [1]. Lee proposed a method based on the Wishart classifier [2]. Both the methods are based on the assumption that the pixels are independent, so the methods are sensitive to noise and lack of spatial coherence.

Markov Random Field (MRF) is a useful tool to incorporate the spatial interaction of pixels. MRF has been applied to classification in remote sensing [3, 4]. The standard MRF (homogeneous MRF) cannot represent the label scale variability because it is assumed that the coupling between pixel labels (the parameter β) is constant throughout the whole image. In some regions of an image, e.g., the ocean, the label (sort) of pixels is constant, so β is very large. On the other hand, in some other regions, e.g., many small bare grounds in a small forest, the labels of pixels are various, so β is very small. When we classify an image using the homogeneous MRF model, it is difficult to keep details of a region associated with a small β and prevent another region associated with a large β from being divided simultaneously. Therefore, it is necessary to incorporate the Inhomogeneous MRF [5] to represent the label scale variability of the image.

In this paper, we will develop a new iteration method consisting of two steps: (1) to fix the labeling (a classification result) and estimate the cluster parameters and the smoothness parameter β of the Inhomogeneous MRF in the energy function; (2) to fix the parameters and estimate a labeling that minimizes an energy function by using an expansion algorithm [6] based on Graph Cuts. Using the proposed iteration method, we can obtain good classification results.

2. INHOMOGENEOUS MRF MODEL

According to [5], we employ the Inhomogeneous MRF model for polarimetric SAR images in this section. The dimension of the image is $n \times m$, each site is denoted as s_{ij} $(1 \le i \le n, 1 \le j \le m)$, the set of labels (sorts) is $L = \{1, 2, \ldots, K\}$, the label field $\{X_{ij} | X_{ij} \in L\}$, Y_{ij} is the covariance matrix at s_{ij}, $Y = \{Y_{ij}\}$, Bayes' theorem yields $P(X|Y) \propto P(X)P(Y|X)$, so the unsupervised classification of polarimetric SAR images by using maximum a-posteriori probability (MAP) estimator is equivalent to estimate $X = \arg\max P(X)P(Y|X)$ for a given Y.

From [2], we have

$$P(Y|X) = \prod_{(i,j)\in s} P(Y_{ij}|X_{ij}) = \prod_{(i,j)\in s} \frac{n^{qn} \left|[Y_{ij}]\right|^{n-q} \exp\left(-tr\left(n\left[\Sigma_{ij}\right]^{-1}\left[Y_{ij}\right]\right)\right)}{K(n,\, q)\left|[\Sigma_{ij}]\right|^{n}}, \tag{1}$$

where $K(n,\, q) = \pi^{q(q-1)/2} \prod_{i=1}^{q} \Gamma(n-i+1)$, $\Gamma()$ represents the Gamma function, n is the number of looks, $q = 3$ for the reciprocal case. Σ_{ij} is the covariance matrix of the class which s_{ij} belongs to. $||$ and $tr()$ represent the determinant and the trace of a matrix, respectively.

From the Hammersley-Clifford theorem [7], we know that $P(x)$ can be described by a Gibbs distribution:

$$P(X) = \frac{1}{z}\exp(\beta V(X))\frac{1}{z}\exp\left(\beta\sum_{i,j}V_{ij}\left(N_{ij}\cup x_{ij}\right)\right), \tag{2}$$

where x_{ij} is a realization of X_{ij} at s_{ij}, N_{ij} is the neighborhood of s_{ij}, V is the potential function, β is the smoothness parameter, z is a constant for normalization. We use the Potts model: $V_{ij}(N_{ij}\cup x_{ij}) = n_{ij}(x_{ij})$, where $n_{ij}(x_{ij})$ is the number of pixels labeled x_{ij} in the neighborhood of the site s_{ij}. Thus the potential function V penalizes the labeling which does not satisfy the spatial coherence. As V is non-convex, we avoid overpenalizing the sudden change between the labels of neighboring pixels. So this model is *discontinuity preserving* [6], i.e., it can preserve the sharp jump at borders of regions while keeping smoothness inside homogeneous regions. When we use the Inhomogeneous MRF model,

$$P(X) = \frac{1}{z}\exp\left(\sum_{i,j}\beta_{ij}V_{ij}\left(N_{ij}\cup x_{ij}\right)\right) = \frac{1}{z}\exp\left(\sum_{i,j}\beta_{ij}n_{ij}\left(x_{ij}\right)\right) \tag{3}$$

Thus, maximizing $P(X|Y)$ is equivalent to minimizing $-\ln(P(X)P(Y|X))$. Removing terms which are independent of X, the problem becomes: to find a labeling x (a realization of X) for minimizing an energy function:

$$E(x) = \sum_{i,j} n\times\left(|[\Sigma_{ij}]| + tr\left([\Sigma_{ij}]^{-1}[Y_{ij}]\right)\right) + \sum_{i,j}-\beta_{ij}n_{ij}(x_{ij}) \tag{4}$$

In order to rapidly obtain the accurate result of classification, we have to precisely estimate the parameters Σ_{ij} and β_{ij} and rapidly find the labeling which minimizes the energy function $E(x)$. We use an iteration method consisting of two steps: (1) to estimate the parameters Σ_{ij} and β_{ij} by using the labeling obtained from the last iteration; (2) to fix the parameters Σ_{ij} and β_{ij} estimated in (1) and to find the labeling that minimizes the energy function $E(x)$ by using the expansion algorithm based on Graph Cuts. Relative to the standard EM approach, step (1) is an approximation to the E-step and step (2) is similar to the M-step [8].

3. ESTIMATION OF PARAMETERS

If the covariance matrix of the m-th cluster center is denoted as Σ_m, then the estimation of Σ_m is the mean of covariance matrices of the pixels corresponding to the label m in the last iteration [2].

The estimation of β_{ij} is intractable because β_{ij} appears in both the exponent of the probability density model and in the normalizing constant z. We divide the whole image into 16×16 equal blocks, and in each block β is regarded as a constant which is estimated by using the Pseudo Likelihood Estimator (PLE) [5] in each block. In this way, β can approximately represent the local label scale and its estimation is tractable.

From the Maximum Likelihood Estimator (MLE),

$$\begin{aligned}&\frac{\partial}{\partial\beta}P(X|Y) = 0 \Leftrightarrow P(Y|X)\cdot\frac{\partial}{\partial\beta}P(X) = 0\\ &\frac{\partial}{\partial\beta}\frac{1}{z(\beta)}e^{\beta\cdot V(X,\beta)} = 0\end{aligned} \tag{5}$$

Using the PLE, we have $P(X) = \prod_{ij}\frac{1}{z_{ij}}\exp(\beta\cdot n_{ij}(x_{ij}))$, where $z_{ij}\sum_{x=1}^{K}\exp(\beta\cdot n_{ij}(x))$, K is the number of classes.

Define $\xi = e^{\beta}$,

$$\frac{\partial}{\partial\beta}\ln P(X) = \sum_{ij}n_{ij}(x_{ij}) - \sum_{ij}\frac{\sum_x n_{ij}(x)\xi^{n_{ij}(x)}}{\sum_x \xi^{n_{ij}(x)}} = c - \sum_{ij}\frac{P_{ij}(\xi)}{Q_{ij}(\xi)} \tag{6}$$

where c is a constant, $P_{ij}(\xi)$ and $Q_{ij}(\xi)$ are polynomials which degrees are not more than the number of pixels in the neighborhood of s_{ij}. If we choose the standard 4 connected grid (as we did in later experiment), $P_{ij}(\xi)$ nd $Q_{ij}(\xi)$ have only 5 possible forms, respectively. Their degrees are not more than 4. Thus the estimation of β is converted to solve $c - \sum_{i=1}^{5} r_i \frac{P_i(\xi)}{Q_i(\xi)} = 0$, i.e.,

$$c - \left(r_1 \cdot \frac{4\xi^4}{\xi^4 + K - 1} + r_2 \cdot \frac{3\xi^3 + \xi}{\xi^3 + \xi + K - 2} + r_3 \cdot \frac{4\xi^2}{2\xi^2 + K - 2} + r_4 \cdot \frac{2\xi^2 + 2\xi}{\xi^2 + 2\xi + K - 3} + r_5 \cdot \frac{4\xi}{4\xi + K - 4} \right) = 0 \quad (7)$$

where r_i ($i = 1, 2, \ldots, 5$) is the number of pixels which neighborhood system is the i-th possible forms in the partition of the image ($1/(16\times16)$ of the whole image), and K is the number of classes. Using the labeling obtained in the last iteration, we can obtain c and r_i ($i = 1, 2, \ldots, 5$). It is easy to demonstrate that the left side of the Equation (7) is a monotonically decreasing function, so β can easily be obtained by using Newton-Raphson scheme to solve (7). In practice, from the characteristics of POLSAR images we constrain that $\beta \in (0, 3)$.

4. OPTIMIZATION OF LABELING

In each iteration, we estimate the labeling for minimizing the energy function $E(x)$ in Equation (4). The first term of $E(x)$ is called Data term, which measures the disagreement between x and the observed data. The second term is called Smoothness term, which measures the extent to which x is not piecewise smooth. As mentioned above, to make the labeling to be discontinuity preserving, we chose the Potts model because the potential function of the Smoothness term has to be non-convex. It was proved that the global minimum of $E(x)$ is NP-hard due to non-convexity of the potential function [6]. Therefore we have to find a suitable algorithm to compute the approximation of the global minimum. For both Iterated Conditional Modes (ICM) and Simulated Annealing (SA) the standard moves (i.e., only one pixel can change its label at a time) were adopted to find an approximation. Once a local minimum has been reached, the value of the energy function cannot be reduced by standard moves any more. But a local minimum may be far from the global minimum, so the result may be very poor. Although we can obtain the global minimum theoretically by SA in certain case, this algorithm is very slow and it was demonstrated in [9] that the result of SA may be very far from the optimum in practice.

In [6], an expansion algorithm was proposed to minimize energy functions based on Graph Cuts. In contrast to standard moves, the expansion algorithm allows many pixels to change their labels at a time, which is called one α-expansion move, i.e., labels of any pixels in the image can be converted into α at this step. Therefore, when a local minimum has been reached, we can still reduce the value of the energy function by using this algorithm. It was proved that the expansion algorithm gives a solution that is within a factor of two of the global minimum for Potts model and its runtime is short [6]. Thus we can obtain more precise results by using this algorithm than ICM and SA. In addition, the algorithm is more effective than the others.

There is a one to one correspondence between labelings x and partitions of image pixels. To find a labeling for minimizing the energy function, we adopt the expansion algorithm based on the following iterations [6]:

Step1. Start with an arbitrary labeling x.

Step2. Set success $:= 0$.

Step3. For each label $\alpha \in \{1, 2, \ldots, K\}$:
3.1 Find $x'' = \arg\min E(x')$ among x' within one α-expansion of x.
3.2 if $E(x'') < E(x)$, set $x := x''$ and success $:= 1$.

Step4. if success $= 1$, goto Step2.

Step5. return x.

An execution of Steps 2, 3, 4 is called a cycle, and an execution of Steps 3.1 and 3.2 is called an iteration. The key point is in 3.1: given a labeling x and a label α, how to find a labeling x'' which energy function is minimum among x' within one α-expansion move of x. It was shown in [6] that by setting a suitable graph we can convert this problem into finding the minimum cut in the graph we set. The minimum cut can be obtained by using a max-flow algorithm. For classification of a POLSAR image with 700×900 pixels, we obtained the minimum within 5 cycles, mostly within 3 cycles. It only took a few seconds.

5. CLASSIFICATION SCHEME

It should be pointed out that to estimate β we need to use the labeling obtained from the last iteration. If the labeling we used is far from the true result, the value of the estimated β may be far from the true β. For unsupervised classification, if the labels we assigned initially or obtained from the last iteration are the same throughout a certain region which labels are actually various, the value of the estimated β may be too large. To estimate a new labeling in the next iteration by using this β, details of the region may be lost, causing the next β still too large, i.e., once the weight of the smoothness term of a region is too large, the data term of this region may lose its influence forever. To overcome this problem, when the iteration number is odd, we reduce the value of β (e.g., let $\beta = \beta/5$) so that the data term could still exert influence even if the initial labeling was bad; when the iteration number is even, we use the value of the estimated β so that the spatial information can be used to obtain reliable results. The procedure of the proposed method is given as follows:

Step1: Initialize the labeling.

Step2: compute the corresponding covariance matrices of cluster centers Σ_m.

Step3: Divide the image into 16×16 equal blocks. For each partition, compute the coefficients of the Equation (7) by using the labeling obtained from the last iteration. Then compute β by using the Newton-Raphson scheme. Let $\beta = \beta/5$ if the iteration number is odd.

Step4: Set the energy function as (4) for the whole image by using the estimated Σ_m of the whole image and β of all partitions. Then find the new labeling by using the expansion algorithm based on Graph Cuts.

Step5: Go to Step2 unless termination condition is satisfied.

The iteration is terminated when the classification results are almost the same as those by the last iteration or the loop time reaches a pre-defined bound.

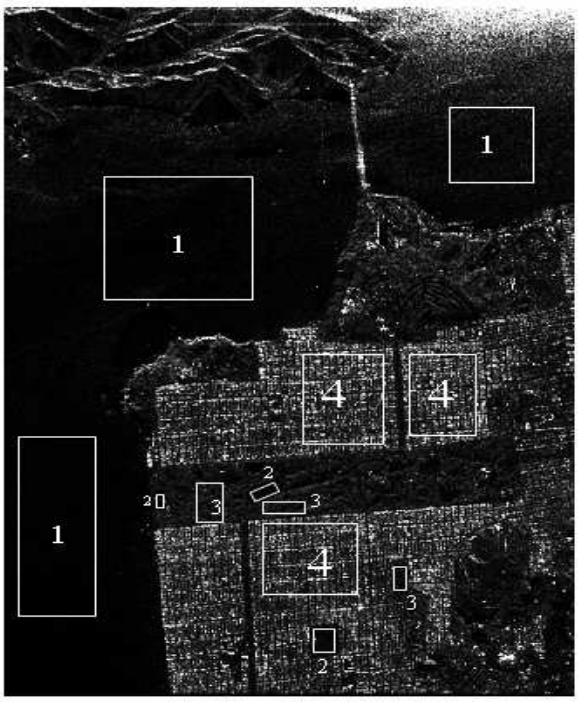


Figure 1: Span image.

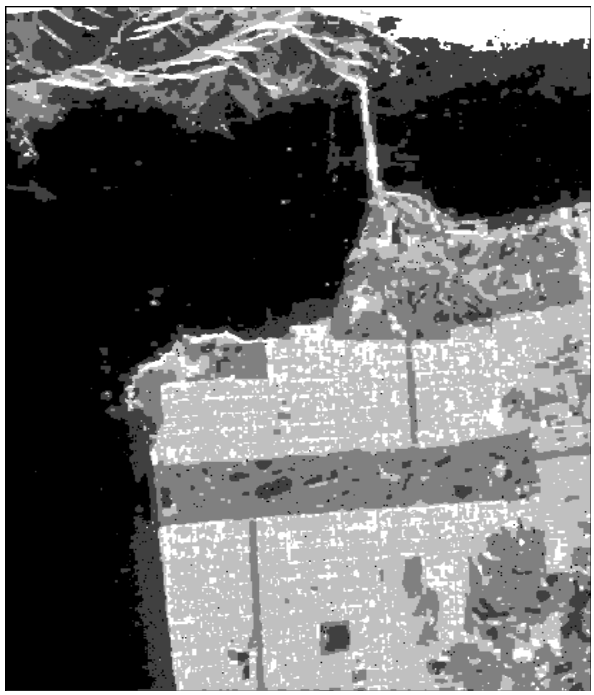


Figure 2: The result of our scheme.

6. EXPERIMENT RESULT

A NASA/JPL AIRSAR L-band image of San Francisco was used for illustrating the classification scheme proposed above. The 4-look calibrated SAR image consists of 700×900 pixels. The span image is shown in Fig. 1. Speckle filtering was applied before classification [10]. We chose $K = 5$ and classified the image into 5 classes by using the target decomposition theory [1] as the initialization

of the labeling. We obtain the classification results of the Wishart classifier [2] and our scheme after 4 iterations. The result of our scheme is shown as Fig. 2. From the figure we can see that in contrast to much noise of the result of the Wishart classifier, the result of our scheme shows the spatial coherence while keeping details of the image.

7. CONCLUSION

A new unsupervised classification scheme for polarimetric SAR images has been proposed. In this method, we considered the spatial information by using the Inhomogeneous MRF model that can represent the label scale variability. The classification results can be rapidly obtained by using an expansion algorithm based on Graph Cuts. The excellent performance of the proposed method has been illustrated by a NASA/JPL AIRSAR L-band image.

ACKNOWLEDGMENT

This work was supported in part by the Research Fund of National Key Lab of Microwave Imaging Technology, Institute of Electronics, Chinese Academy of Sciences.

REFERENCES

1. Cloude, S. R. and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sensing*, Vol. 35, No. 1, 68–78, Jan. 1997.
2. Lee, J. S., M. R. Grunes, et al., "Unsupervised classification using polarimetric decomposition and the complex Wishart classifier," *IEEE Trans. Geosci. Remote Sensing*, Vol. 37, No. 5, 2249–2257, Sept. 1999.
3. Solberg, A. H. S., T. Taxt, and A. K. Jain, "A Markov random field model for classification of multisource satellite imagery," *IEEE Trans. Geosci. Remote Sensing*, Vol. 34, 100–113, Jan. 1996.
4. Chen, Y., et al., "Supervised Polarimetric SAR Classification method using graph-cut," *IEEE Geoscience and Remote Sensing Letters.*
5. Cadez, I. and P. Smyth, "Modeling of inhomogeneous Markov random fields with applications to cloud screening," Technical Report UCI-ICS 98-21, University of California, 1998.
6. Boykov, Y., O. Veksler, and R. Zabih, "Fast approximate energy minimization via Graph Cuts," *Proc. IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 23, No. 11, 1222–1239, Nov. 2001.
7. Besag, J., "On the statistical analysis of dirty pictures," *J. Royal Statistical Soc., Series B*, Vol. 48, No. 3, 259–302, 1986.
8. Zabih, R. and V. Kolmogorov, "Spatially coherent clustering using graph cuts," *Proc. of CVPR'04*, 437–444, 2004.
9. Greig, D., B. Porteous, and A. Seheult, "Exact maximum a posteriori estimation for binary images," *J. Royal Statistical Soc., Series B*, Vol. 51, No. 2, 271–279, 1989.
10. Lee, J. S., M. R. Grunes, and G. De Grandi, "Polarimetric SAR speckle filtering and its implication on classification," *IEEE Trans. Geosci. Remote Sensing*, this issue, 2363–2373.

Spatial Distribution Pattern of MODIS-NDVI and Correlation between NDVI and Meteorology Factors in Shandong Province in China

Dongmei Song, **Peng Guo**, **and Hui Sheng**
Earth Resource and Information Institute, China University of Petroleum, Qing Dao 266555, China

Abstract— The vegetation index research was necessary for monitoring plant growth. Based on the data of 250 m spatial resolution NDVI (Normalized Difference Vegetation Index) of MODIS (Moderate Resolution Imaging Spectroradiometer), this paper analyzed the spatial distribution pattern of NDVI in Summer in Shandong Province in the east of China. With the data of six meteorologic sites and the average NDVI in January, in April, in July and in October in 2006 including the average atmosphere temperature, relative humidity, precipitation, sunlight hours and using the SPSS statistics software, the correlation between NDVI and meteorology factors was researched. The results were showed as follows: there was obvious character of spatial distribution pattern of NDVI in Shandong Province. The value of NDVI in plain region was higher than the value of mountain and hilly region, and moreover the distribution of value of NDVI was even in plain, in which the most of NDVI was greater than 0.7–0.8. In the mountain region, NDVI was falling off with the decreasing of the altitude, where NDVI was about 0.6–0.7. Deeply research also showed that NDVI was affected by human activity distinctly, and so the NDVI in the city was lower than 0.4, as such lower than that of the suburb. The atmosphere temperature and quantity of the precipitation were the two main factors affecting the change of NDVI, at the same time the seasonal sunlight hours was second-class factor that caused the change of the NDVI. The research results also indicated that with the decreasing of the latitude, the correlation between NDVI and temperature was decreasing too, on contrast, the correlation between NDVI and the quantity of precipitation was increasing.

1. INTRODUCTION

The vegetation is acting an important role in the earth ecological system, and it affects the energy balance between the earth and atmosphere, being the nature "link".

The vegetation growth has obvious seasonal variation characteristic, and become the "display" in the global climatic change. Therefore research on the correlation between the vegetation and the climatic change has become very important in the global change research. The vegetation index is a kind of value which can instruct the vegetation growing trend and the biomass (through linearity or non-linear combination way including addition, subtraction, multiplication, division with selecting the multi-spectra remote sensing data) [1–5]. The vegetation index is a simple, effective measure to the surface vegetation condition, and is the "bridge" between the remote sensing image and the vegetation research. Analyzing the spatial distribution pattern of vegetation index of Moderate Resolution Imaging Spectroradiometer (MODIS) in Shandong Province and correlation with climatic element has the vital significance to the Shandong Province ecological environment monitoring and the vegetation response to the climatic change.

Normalized Difference Vegetation Index (NDVI) is one of the best instruction factors of the vegetation growth condition and the degree of vegetation cover [4–9]. If an area is covered by the vegetation, then the value of the NDVI of this area is just a positive number, and also increases along with the vegetation cover increasing [1–6]. Since the 1990s a few researches on the NDVI spatial pattern and it with climatic element correlation have been carried in China, but majority of researches was based on NOAA/AVHRR-NDVI [10–13]. The MODIS vegetation index was improved on the basis of the NOAA/AVHRR vegetation index [13–16], and at present, it has been an effective method of monitoring global and regional scale of vegetation condition since MODIS-NDVI is more sensitive to vegetation condition owing to reducing the disturbances which exterior factors bring (for example atmospheric, solar elevation angle, cloud and so on). First band (red wave) and second band (near-infrared wave) of MODIS was used to calculate NDVI, and their space resolution is 250 m [1–3]. Terra-MODIS scans the same area of the earth two times every day, this kind of suitable space and time resolution can reflect the change characteristic of the vegetation very well [9–12]. This paper taking Shandong Province as a research region, discovered the spatial pattern characteristic of MODIS-NDVI, and analyzed correlation between the mean value of NDVI

in January, April, July, and October in 2006 and the corresponding meteorological factor with using data of six meteorology observation stations.

2. RESEARCH DATA AND METHOD

2.1. Brief Introduction of Research Area Condition

Shandong Province is located in the downriver of the Yellow River, east near to Bohai, Yellow Sea, is situated between the east longitude 114° 36'–122° 43' and the north latitude 34° 25'–38° 23', and holds the important status in the national economical pattern. The territory including the peninsula and the interior two parts, eastern part is peninsula hilly area; western and northern parts are the plain; and southern and central parts are the mountainous region.

Shandong Province belongs to the East Asian warm temperate zone monsoon climate. The increase of the quantity of the precipitation and heat take place at the same season. Spring and autumn are short, while winter and summer are longer. Average temperature is 11–14 °C, and average precipitation is generally 600–900 mm.

2.2. Data Origin and Processing

This research used vegetation index products (16 days-integrated and 250 m space resolution) which were provided by the American aerospace bureau (National Aeronautics and Space Administration, NASA). Format of the NDVI data was HDF, and the time of NDVI images were in January, in April, in July, in August, and in October in 2006. These NDVI data were already adjusted through the radiation method, geography localization and atmospheric adjustment. The meteorological data came from Chinese meteorology scientific data sharing network.

First using MRT tool software (MODIS Reprojection Tool) to complete the projection and the format conversion of NDVI images, and as a result, the projection of transformation images was: Geographic (Lat/Lon) WGS84, and the format of the images was GEOTIFF, and the sampling way was Bilinear. Then carried on the mask operation to obtain MODIS-NDVI images of Shandong Province by Shandong Province's administrative boundary image.

Taking six meteorological stands (Huimin, Chengshantou, Jinan, Weifang, Dingtao, Yanzhou) of Shandong Province as the center, and 30 km as the buffer radius, got buffer images of NDVI data in January, in April, in July, and in October in 2006.

2.3. Research Method

In order to analyzing the spatial distribution pattern of NDVI in summer–autumn, this research used ENVI 4.1 remote sensing processing software, calculated average NDVI data of the last ten days of August and beginning ten days of September (data integrated of the 225th–240th days, and 241st –257th days), and obtained the NDVI graduation distribution map.

Since some research indicated that the vegetation has the lag response to the meteorological factor [13], therefore this research carried on the correlation analysis between buffer NDVI images and month and seasonal climate factors including average temperature, the average relative humidity, the precipitation, the sunshine hours with applying SPSS software, using statistics method. For example, in order to analyze the correlation between NDVI in July and corresponding seasonal temperature, we need to calculated the average value of the temperature of May, June and July. After we finished all four seasonal calculation with the similar method, we can obtain a group of corresponding values, then we can calculated the correlation coefficient between NDVI and seasonal mean value of the climatic factor.

3. RESULT AND ANALYSIS

3.1. MODIS-NDVI Spatial Distribution Pattern in Shandong Province

According to the vegetation distribution characteristic in Shandong province, the vegetation index was divided into seven levels. Because the area where NDVI was lower than 0.4 was smaller, moreover, it was generally population concentration city, therefore NDVI between 0 and 0.4 was divided into a level. From Figure 1, the spatial distribution of vegetation index had the obvious region characteristic, as follows:

First, we can see this kind of spatial distribution characteristic of southwest and northwest plain area of Shandong province. This area was forming from flood silting by Yellow River, and the elevation in this area was about 50 meters, with the fertile soil, and the developed agriculture. Through the Shandong Province vegetation type map, we can know that in this area, most of the vegetation type was the crops, with few forest. Crop types and season change were main factors

that affected the vegetation index change. Because crops were planted nearly at the same time as well as the crops area certain widespread, vegetation index at the same time were quite consistent, and the NDVI value were all above 0.7. Especially, the NDVI of the Lubei plain and Jinghang canal basin were all achieve above 0.8. But the NDVI value of the city obviously reduced, that number in the population crowded area was lower than 0.4, and eventhough, at suburb, NDVI value was also only between 0.5 and 0.6.

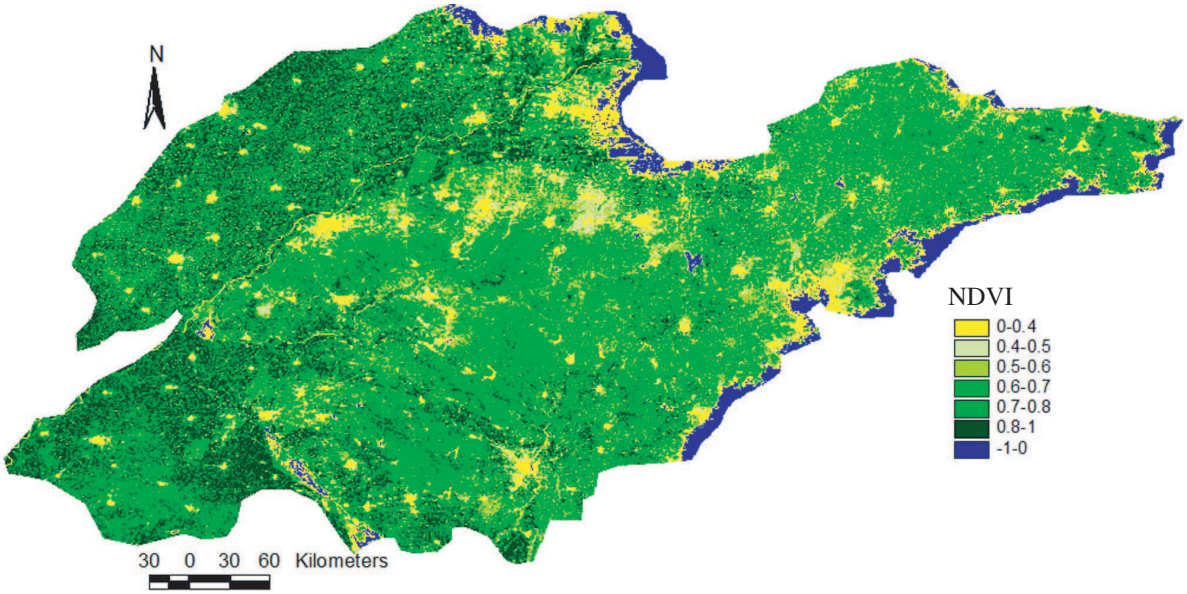


Figure 1: Spatial pattern of MODIS-NDVI in 2006 in Shandong Province.

Second, notice the characteristic of Yellow River Delta. The humanity activities intensely affected this area, and the soil was mild as well as with serious salinification. Most of vegetation types were the graminaceous grass and the weed meadow, lacking the natural forest vegetation, and there were a large area of primary or the secondary natural meadow and some willow tree, and other areas were all the farmland. Because this area was seriously salinification, with barren soil, the vegetation growing not good, therefore the vegetation index was lower and the value was about 0.4–0.6 Fig. 1 showed that in the large coastal shallow and tidelands, the NDVI value was below 0. In the meadow area of the inland, NDVI value was also below 0.4. Compared with former, the area where the NDVI value was higher than 0.7 was covered by the nature grass in the long area on both banks of the Yellow River.

Third, there was the obvious characteristic of southern, central and western Shandong Province and peninsula mountainous region. The topography is higher, and terrain was complex. In the area where the elevation is above 200 m, there was more quantity of moisture, and land was covered by forest, so the degree of forest cover was higher, therefore average NDVI was higher than the value of other place at the same time, and majority of NDVI in the area was between 0.7 and 0.8. However in the area where elevation is 50–200 m, the vegetation was mainly the patch of grass and the crops, so cover ratio was lower, there the NDVI value was majority between 0.6 and 0.7. In this area, another characteristic was that the NDVI value of the top of the mountainous area was highest and then with the deduce of the elevation NDVI decreased progressively to all around.

Fourth, should pay attention to the characteristic of low hilly area in eastern peninsula. The elevation of majority of this area is about 50–500 m, there is enough precipitation and quantity of the heat, the main vegetation type included the black pine, the Japanese red pine, the hemp oak tree, the locust tree, coniferous forest and the temperate zone fallen leaf foliage forest. However, in lower elevation area, majorities of farmland were the vegetation. As a result of marine climate adjustment, vegetation growth is luxuriant, the vegetation index was higher, majority of NDVI was all above 0.7-0.8.

In brief, in Shandong Province, NDVI value of all cities was all below 0.4. Not only the NDVI value of farmlands in plain was highest, but also this kind of spatial distribution of NDVI of farmland was even. In the mountainous area, the value of top of mountain was highest and then taking the mountainous top as the center, NDVI decreased progressively to all around with the

decrease of the elevation.

3.2. Correlation of NDVI and Climatic Factors

3.2.1. Correlation of NDVI and Climatic Factors at the Same Month

Comparing average NDVI of 30 km buffer area of weather stations with average temperatures, the average precipitation, the average relative humidity, average sunshine hours in the same month in 2006 year, and analyzing their relevance, correlation coefficients were calculated, the result was showed in Table 1.

Table 1: Correlation coefficient between NDVI and month mean value of the climatic factor.

Different region	average temperature	average relative humidity	average precipitation	average sunshine hours
Huimin	0.95	−0.024	0.605	0.773
Chengshantou	0.944	0.89	0.723	−0.428
Jinan	0.888	0.115	0.818	0.527
Weifang	0.836	−0.153	0.906	0.785
Dingtao	0.715	0.326	0.76	0.521
Yanzhou	0.853	0.312	0.856	0.208

There was remarkable correlation between average temperature in various regions and NDVI, in comparison, the correlation coefficient between NDVI and the average precipitation was generally lower than the former, even though correlation coefficient were also all above 0.6. Except correlation coefficient of Chengshantou regions was 0.89, other correlation coefficients were all smaller. Between NDVI and sunshine hours, except Chengshantou, other regions all presented the correlation, moreover, the area of correlation coefficient above 0.5 achieved 65%. There was not obvious relevance between NDVI and the relative humidity.

Table 2: Correlation coefficient between NDVI and seasonal mean value of the climatic factors.

Different region	average temperature	average relative humidity	average precipitation	average sunshine hours
Huimin	0.757	0.059	0.674	0.816
Chengshantou	0.717	0.933	0.7	0.448
Jinan	0.655	−0.081	0.662	0.933
Weifang	0.534	−0.112	0.279	0.927
Dingtao	0.384	−0.237	0.522	0.991
Yanzhou	0.587	−0.195	0.611	0.926

Table 1 also showed that the temperature and the precipitation were the main meteorological factors that affected NDVI changes in the Shandong province, which founding was consistent with that results in the publication literature by Li Bengang [7]. The further analysis indicated that the relativity of the temperature and the NDVI value reduced in the direction from the north to the south, however, the tendency of correlation of the precipitation and the NDVI increased. This indicated that, in Shandong Province, along with latitude reducing, dependence of the vegetation growth on the temperature reduced, however, dependence of the vegetation growth on the precipitation increased. There was no obvious rule to obey between NDVI and the temperature and the precipitation in the direction from east to west.

3.2.2. Correlation of NDVI and Seasonal Mean Value of the Climate Factors

Because the NDVI value has the accumulation characteristic [7], this research carried on the correlation analysis between NDVI and the season mean value of corresponding climate factors, and

the correlation coefficient can be seen in the Table 2.

Through comparing Table 1 and Table 2, the result showed that the relevance between NDVI and the same month mean temperature, and the average precipitation were higher than the value between NDVI and the seasonal mean value, but the relevance between NDVI and the same month sunshine hours was lower than the value between NDVI and the seasonal average sunshine hours.

4. CONCLUSION

In this paper, applying MODIS-NDVI in January, in April, in July, in October in 2006 in Shandong Province and the weather observation data, we analyzed the correlation between NDVI and climatic factors, draw the following main conclusions:

(1) Spatial pattern characteristic of MODIS-NDVI in Shandong Province in summer is obviously, NDVI value of the plain is the highest, and also distribution is even, NDVI of mountain region of the elevation above 500 is the second high. The MODIS-NDVI value is affected by the humanity activities remarkably, and in city area, the vegetation index is obviously lower than the other area covered by vegetation.

(2) In Shandong Province, MODIS-NDVI value is mainly affected by the temperature and the precipitation, therefore these two climatic factors are very important to the vegetation growing. Along with latitude reducing, the dependence of vegetation development upon the temperature decreases, on contrast, this dependence of vegetation growing upon the precipitation increases.

(3) Correlation between the NDVI and the same month mean temperature, and average precipitation are higher than the correlation between the NDVI and the seasonal mean value, but correlation between NDVI and the same month sunshine hours is lower than the value between the NDVI and the seasonal average sunshine hours, so we can draw another conclusion that the influence of sunlight hours to vegetation growing has accumulation characteristic and Just because of this, the sunshine hours length in whole plant growth season has a profounder influence to the plant growth.

ACKNOWLEDGMENT

This research was supported by National Natural Science Foundation of China (No. 40501073).

REFERENCES

1. Zhao, Y., *Principles and Methods of Remote Sensing Application Analysis*, Publishing house of Science, Beijing, 2001.
2. Cihlar, J., et al., "Relation between the normalized difference vegetation index and ecological variables," *Remote Sens. Environ.*, Vol. 35, No. 1, 279–298, 1991.
3. Di, L., D. C. Rundquist, and L. Han, "Modelling relationships between NDVI and precipitation during vegetative growth cycles," *Int. J. Remote Sens.*, Vol. 15, No. 10, 2121–2136, 1994.
4. Banair, A., et al., "A review of vegetation indices," *Remote Sens. Review*, No. 13, 95–120, 1995.
5. Yang, L. M., et al., "An analysis of relationships among climate forcing and time-integrated NDVI of grasslands over the U.S. northern and central Great Plains," *Remote Sens. Environ.*, No. 65, 25–37, 1998.
6. Han, A., "Method of abstracting the desertification sensitive areas through analysis of the time series of MODIS vegetation indices," *Rorest Resources Management*, No. 1, 57–60, 2004.
7. Zhao, B., C. Liu, J. Wang, et al., "Spatial and temporal change of MODIS-NDVI in Xilinguole Grassland," *Grassland of China*, Vol. 26, No. 1, 1–8, 2004.
8. Zhang, W. and Z. Gao, "Spatial variation of water/thermal elements and NDVI with altitudes in central and eastern Tibetan Plateau," *Geographical Research*, Vol. 25, No. 5, 877–886, 2006.
9. Gu, Z., J. Chen, and P. Shi, "Correlation analysis of NDVI difference series and climate variables in Xilingole Steppe from 1983 to 1999," *Acta Phytoecologica Sinica*, Vol. 29, No. 5, 753–765, 2005.
10. Liu, L. and F. Xiao, "Spatial temporal correlations of NDVI with precipitation and temperature in Yellow River Basin," *Chinese Journal of Ecology*, Vol. 25, No. 5, 477–481, 2006.
11. Zhang, J., W. Dong, and D. Ye, "The fresh evidence of vegetation cover to summer weather effect in China," *Chinese Science Bulletin*, Vol. 48, No. 1, 91–95, 2003.
12. Niu, J. and H. Hu, "The review of study on relationship between vegetation and environments in China," *Acta Scientiarum Naturalium Universitatis NeiMongol*, Vol. 31, No. 1, 76–80, 2000.

13. Li, B., "Correlation between AVHRR NDVI and climate factors," *Acta Ecologica Sinica*, Vol. 20, No. 5, 898–902, 2000.
14. Li, X., Y. Wang, and K. Li, "NDVI sensitivity to seasonal and interannual rainfall variations in northern China," *Acta Geographica Sinica*, No. 55, 82–89, 2000.
15. Chen, X., J. Li, and Z. Li, "Vegetation change in Yanchi of Ningxia and its relationship with climate change in recent years," *Acta Ecologica Sinica*, Vol. 26, No. 5, 1517–1522, 2006.
16. Wang, J., K. P. Price, and P. M. Rich, "Spatial patters of NDVI in response to precitation and temperature in central Great Plains," *International Journal of Remote Sensing*, No. 22, 3827–3844, 2001.

Investigation on the RCS Measurement Technique of Large Targets at Near Distance

Nan-Jing Li, **Wei-Jun Chen**, **Chu-Feng Hu**, **and Lin-Xi Zhang**
UAV Speciality Technique Key Labrotary of National Defense Technology
Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China

Abstract— In general, RCS measurement must meet the far-field condition, which need a vast measuring area or a costly compact range. To solve the problem, a new technique by extrapolation correction dealing with near-distance RCS measurement of aircraft targets is set up. The RCS of aircraft target under spherical wave illuminating can be adopted to extrapolate far-field RCS of them. The convolution arithmetic utilizing conductive board as a reference target deduces the phase-correction coefficient $g(x)$ between spherical and plane wave, and in fact the convolution calculation of $g(x)$ can be replaced by Fast Fourier Transform, which is related to the scattering field of the reference target measured in near-distance, thus the calculating procedure could be completed simply and quickly. By theoretical simulation for a simple target, the RCS extrapolated can be obtained correctly. Experimental results of complex aircraft target show that this new technique makes experiment agree to theory precisely, moreover, it permits the measurement distance reduced by 15 percent of the minimum distance of far-field. Also, there is no serious limitation of measurement for target in dimension of aperture and depth.

1. INTRODUCTION

Scattering testing is of great significance to find stealth target and its scattering characteristics research. Generally RCS testing needs huge testing field or complex expensive compact range system [2–5]. Based on the shape character of aircraft, a new method of near-field RCS testing is put up in this paper. By defining a phase correction coefficient $g(x)$ only, which amends the phase difference brought by the bending wave front, the far-field RCS of an aircraft is obtained. The correction coefficient $g(x)$ can be derived by a relation between near-field and far-field of a reference metal plate.

2. CONVOLUTION EXTRAPOLATION METHOD

Table 1 lists a number of targets (scaled models), the height of the fuselage is less than 1/5 of its length. So they can easily meet the far field RCS testing requirements in the vertical direction (Table 2 shows the allowed maximum height of a target at different frequencies to meet the far field RCS testing requirements). Therefore, for the near-field RCS tests, field distribution in the target range can be looked approximately as a cylindrical wave.

Table 1: The height and length of some scaled models.

Model	Global Hawk	B2	Yun 8	Dark Star
length/(m)	1.9	2.0	2.5	3.3
height/(m)	0.15	0.12	0.4	0.2
Height/length	0.079	0.061	0.18	0.064

Table 2: The maximum height allowed for testing targets at different frequencies.

Test Distance $R = 20$ m					
Frequency/GHz	1	2	5	10	18
maximum height/m	1.73	1.22	0.77	0.55	0.41

The cylindrical wave front in the one-dimensional direction (vertical direction, as shown in Fig. 1) is flat, but bending in the other dimension (azimuth direction x), therefore, a correction

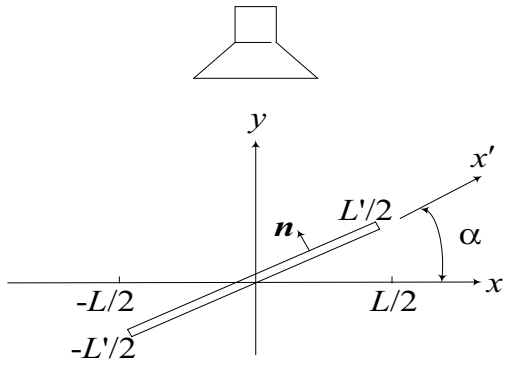


Figure 1: Metal plate rotates by α angle.

coefficient $g(x)$, which only contains the azimuth variable x, is defined to amend the bending wave front.

Consider a spherical wave irradiating on a metal plate, the length of which is L (width $\ll L$) , as shown in Fig. 1. The metal surface of the incident field is posed by the following two parts:

$$E_i\left(x'\right) = e_i\left[E_p\left(x'\right)e^{jkx'\sin\alpha}\right]\left[g\left(x'\right)\right] \tag{1}$$

The $E_p(x')e^{jkx'\sin\alpha}$ is the incident electric field of an ideal plane-wave irradiating on the metal surface. The $g(x')$ is the phase correction function of describing the actual deviation from plane-wave. The e_i is the united polarized vector of incident electric field E_i.

According to document [1], an expression exists as following:

$$E_p(x) \approx \frac{1}{g(x)} \cdot \mathrm{F}\left[E_s(u)\right] \tag{2}$$

where F[•] means Fourier transform, $u = 2k\tan\alpha$, k is the wave number. By taking inverse Fourier transform of expression (2),

$$E_{si}(u) \approx \mathrm{F}^{-1}\left[\frac{1}{g(x)}\right] \bigotimes E_s(u) \tag{3}$$

where $E_{si}(u)$ is the inverse Fourier transform of $E_p(x)$. With height and width being as same as the target under test, electing a metal plate as the reference plate, by the relationship between tested data of the plate at near-field and its far-field RCS theoretical value, $g(x)$ can be derived. From (2)

$$g(x) \approx \frac{\mathrm{F}^{-1}\left[E_s(u)\right]}{E_p(x)} \tag{4}$$

or

$$g(x) \approx \frac{\mathrm{F}^{-1}\left[E_{plane\_s}(u)\right]}{E_{plane\_\infty}(x)} \tag{5}$$

So (3) can be rewritten as

$$E_{si}(u) \approx \mathrm{F}\left\{\frac{E_{plane\_\infty}(x)}{\mathrm{F}^{-1}\left[E_{plane\_s}(u)\right]}\right\} \bigotimes E_s(u) \tag{6}$$

Certainly $E_{plane\_\infty}(x)$ is determined by

$$E_{plane\_\infty}(x) = \mathrm{F}^{-1}\left[E_{plane\_\infty}(u)\right] \tag{7}$$

Inversed Fourier transform of (6)is

$$\mathrm{F}^{-1}\left[E_{si}(u)\right] \approx \frac{\mathrm{F}^{-1}\left[E_{plane\_\infty}(u)\right]}{\mathrm{F}^{-1}\left[E_{plane\_s}(u)\right]}\mathrm{F}^{-1}\left[E_s(u)\right] \tag{8}$$

On the right side of (8), $E_{plane\_s}(u)$ is the measured value of reference plate at near-field, $E_{plane\_\infty}(u)$ is the far-field theoretical value of reference plate, $E_s(u)$ is the measured value of target at near-field, then $\mathrm{F}^{-1}[E_{si}(u)]$ can be gained by (8), and taking Fourier transform of $\mathrm{F}^{-1}[E_{si}(u)]$, $E_{si}(u)$ comes out. Although (6) include convolution operators, they can be replaced by some appropriate Fourier transforms and inversed Fourier transforms.

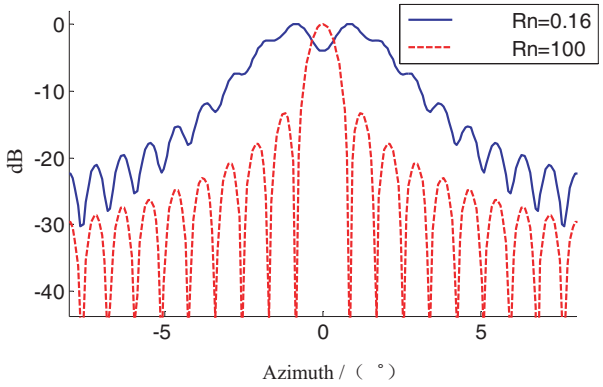


Figure 2: Normalized scattering pattern of Long plate ($L = 20\lambda$) in the near-field and far-field.

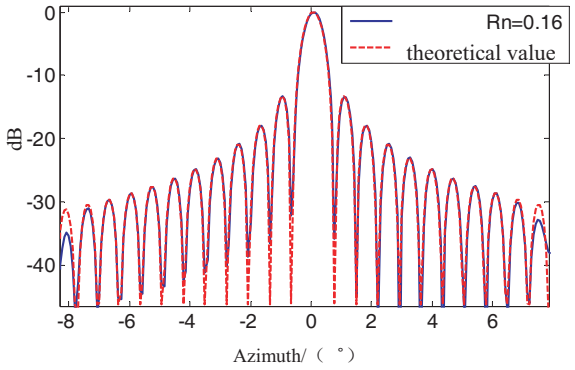


Figure 3: Normalized RCS pattern after convolution extrapolation compared with the theoretical value.

Figure 4: Global Hawk(1:10 Scaled Model). Wingspan: 1.8 m; Captain: 0.85 m; Height: 0.30 m

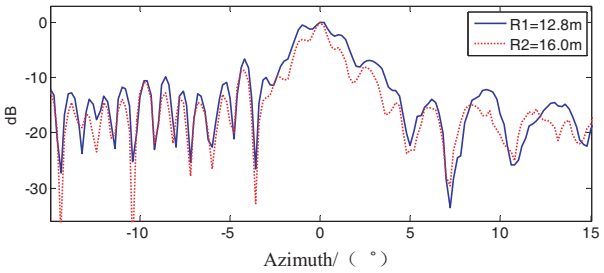


Figure 5: Normalized RCS pattern measured under different near distances.

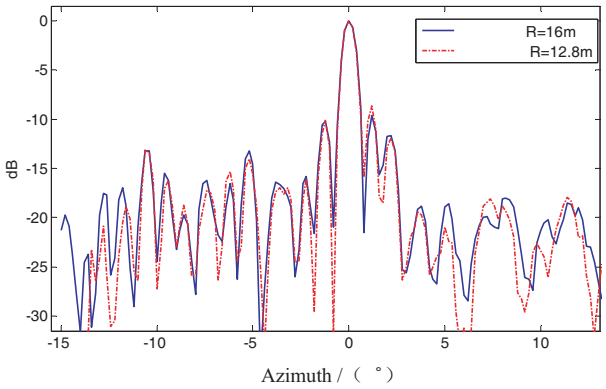


Figure 6: Normalized RCS pattern after convolution extrapolation.

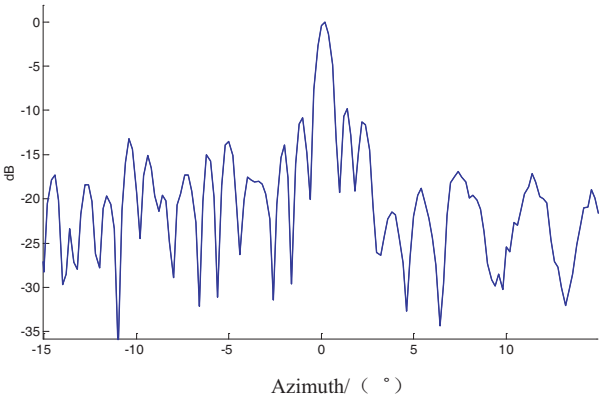


Figure 7: Normalized theoretical RCS pattern computed from physics optic method.

3. SIMULATION AND EXPERIMENT

Assume a metal plate with length $L = 25\lambda$, height $B = 4\lambda$ as reference target, and the other plate with length $L_1 = 20\lambda$, height $B = 4\lambda$ as target under test, the testing distance $R = 0.10R_0$ where $R_0 = 2L^2/\lambda$.

Figure 2 shows the normalized scattering pattern at near-field and far-filedbut after convolution extrapolation, as Fig. 3, the extrapolation result and far-filed result are very consistent.

Figure 4 is the indoor testing photo about Global Hawk (1 : 10 Scaled Model).

Testing distances are 12.8 m and 16 m respectively, and the testing frequency is 15 GHz. Fig. 5 gives two near-field measuring results. The normalized RCS pattern is a function of distance. Extrapolation results are shown as Fig. 6, within azimuth angle of $\pm 15^\circ$, the two extrapolation results are very consistent.

The application of convolution extrapolation technique provides a good agreement of both extrapolation results (see Fig. 6). Fig. 7 gives the theoretical results of the model, we can see that, at a certain angle scope ($\pm 15^\circ$) , the convolution extrapolation pattern is basically the same as the theoretical calculation result.

4. CONCLUSION

Convolution extrapolation technique uses a reference metal plate to gain the correction coefficient $g(x)$, and need not calculate the complex convolution by appropriate Fourier transform. From above results, the extrapolation technique used to investigate aircraft scattering has already shown high accuracy and great reduction of the test distance during RCS measurement, so it is valuable in the field of scattering measurement.

REFERENCES

1. Birtcher, C. R., C. A. Balanis, and V. J. Vokurka, "RCS measurements, transformations and comparisons under cylindrical and plane wave illumination," *IEEE Trans. Antennas Propagat.*, Vol. 42, No. 3, 329–333, 1994.
2. Hu, W., H. Sun, and X. Lv, "Research on RCS measurement under non-far field condition," *IEEE Trans. Antennas Propagat.*, Vol. 51, No. 3, 392–395, 2003.
3. Melin, J., "Measuring radar cross section at short distances," *IEEE Trans on Antennas Propagat.*, Vol. 35, No. 6, 991–996, 1987.
4. Falconer, D. G., "Extrapolation of near-field RCS measurements to the far zone," *IEEE Trans on Antennas Propagat.*, Vol. 36, No. 6, 822–829, 1988.
5. Chu, T.-H. and D.-B. Lin, "Microwave diversity imaging of perfectly conducting objects in the near-field region," *IEEE Trans on MTT*, Vol. 39, No. 3, 480–487, 1991.

Application of DSP in the Step-Frequency RCS Measurement System

Chufeng Hu, **Jiadong Xu**, **Nanjing Li**, **and Jin Cao**
National Defense Technology Institute, Northwestern Polytechnical University
Xi'an, Shaanxi 710072, China

Abstract— Step-Frequency RCS measurement system is a new and remarkable RCS measurement system, which can obtain the response of target at each frequency point in an anecho chamber. Comparing it with traditional way of CW RCS measurement, more information gains in the new system. IFFT, zero insertion and the use of additional windows are some of familiar DSP techniques, which can improve the results if they are applied to the measurement. Using IFFT, the frequency domain data of whole chamber obtained by the measurement can be interpreted into time domain value, and the unwanted signals can be removed by adding a gate of range. Inserting zero in the frequency domain for increasing the points of IFFT will be good for finding response of target in the time domain exactly. The measurement of finite frequency spectrum makes the cut of data, the use of windows can reduce the effect of side lobe. Two metal spheres are measured in the anecho chamber. Two sorts of diameter are 50 mm and 400 mm respectively. The former is used for a target, the latter for scaling. The span of frequency is 8 GHz $\sim$ 12 GHz, and the test frequency points are 801. Hamming windows are added to frequency domain data of the target and the scaler, then a number of zero as many as three times of frequency domain data are inserted for IFFT. A range gate is added to the target region where energy of target returns is much higher than the background. Returning to the frequency domain again, according to the results scaled, it shows that the accuracy of measurement is obviously improved. The value of RCS measurement is waved less than 1 dB around the center of frequency over 80%, in comparison with 4 dB waving over all test frequencies before, so it's significant that integrating DSP techniques mentioned above is specially useful for RCS measurement of low scattering targets.

1. INTRODUCTION

RCS measurement is necessary to research the electromagnetic scattering properties of radar target as well as to design stealth weapon system. The theory and algorithm of electromagnetic scattering computation can be verified by RCS measurement, moreover, it's difficult to compute electromagnetic scattering of complex objects, but the data can be obtained visually by RCS measurement. Step-Frequency RCS measurement system is a new and remarkable RCS measurement system, which can obtain the response of target at each frequency point in an anecho chamber. Some DSP technologies are applied to the Step-Frequency RCS measurement in this paper. The results of simulation and experiment show that the accuracy of measurement is improved.

2. THEORY

The Step-Frequency signal is a super width band radar signal which is changed by steps (shown in Figure 1). It consists of some pulses. The width of pulse can be adjusted according to the situation we need. The frequency of each pulse is different, and the interval between each frequency are constant. IFFT is used to process the return signal, then high resolution range imaging of target can be obtained as shown in Figure 2. Therefore, Step-Frequency signal is applied extensively to microwave imaging and target recognition.

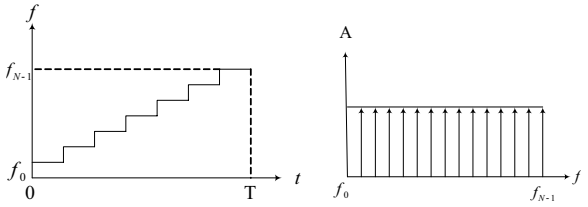


Figure 1: Step-Frequency signal.

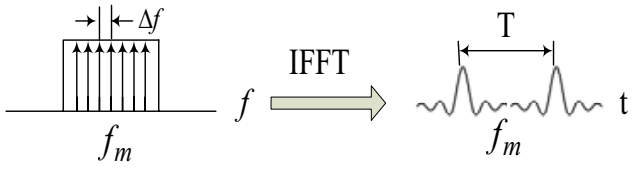


Figure 2: Relationship between frequency-domain and time-domain of Step-Frequency signal.

The high performance vector network analyzer can transmit this kind of signal using synthesizer of frequency, which can provide the combination of many discrete frequencies. The amplitude and phase at each sampling frequency keep consistent by using phase-locked loop and self-adaptation logical circuit, so it can be expressed as follow:

$$S(f) = \sum_{i=0}^{N-1} \delta\left(f_0 - i \cdot \Delta f\right) \tag{1}$$

where, f_0 is start frequency, Δf is interval of frequency. The time-domain form of the signal can be derived by IFFT:

$$\begin{aligned} s(t) &= \int_{f_0}^{f_{N-1}} \sum_{i=0}^{N-1} \delta\left(f_0 - i \cdot \Delta f\right) \exp(\mathrm{j}2\pi f t) df \\ &= e^{j2\pi[f_0+(N-1)\Delta f/2]t} \cdot \frac{\sin(\pi N \Delta f t)}{\sin(\pi \Delta f t)} \end{aligned} \tag{2}$$

From the above mentioned equation, the time-domain form of Step-Frequency signal is a series of Sa function whose carrier frequency is the center frequency. The repeatition period is $T = 1/\Delta f$. The accuracy of analysis measurement simulated from changing the points of IFFT and windows is as follows.

2.1. Number of Points for IFFT

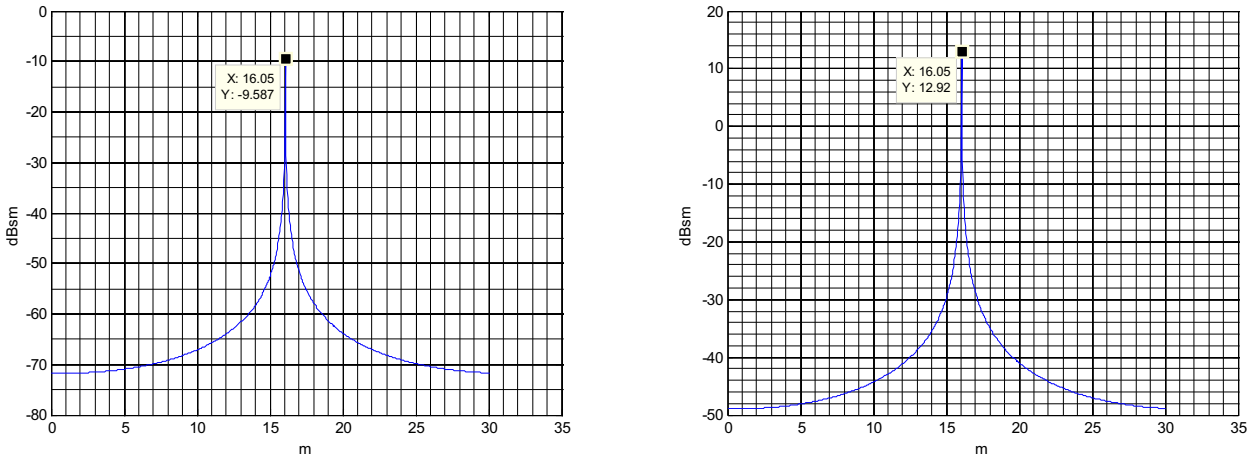
The targets of simulation are a metal square flat plate with width of 200 mm and a scaling metal sphere with diameter of 400 mm. The span of frequency is 4 GHz with 801 points lined. The distance of measurement is 16 m. Then the frequency-domain data are transformed into time-domain by IFFT with 801 points, the result is shown in Figure 3.

The theoretical RCS values of scaling sphere and metal flat plate at center frequency are −9.01 dBsm and 13.49 dBsm respectively, so the theoretical difference of RCS value of them is 22.50 dB, and according to Figure 3, the difference of value is 22.57 dB with 801 points. Then the points are increased to 3201, the result is shown in Figure 4.

The difference between scaling sphere and metal flat plate is 22.52 dB as shown in Figure 4, which is much closer to the theoretical value. The reason for that is that zero insertion in frequency-domain makes the sampling points more accurate in the time-domain, so it is beneficial to recognize the peak of target.

2.2. Windows Function

The process of IFFT above is actually using rectangular windows in frequency-domain, because there are a number of interference coming from background, then we can remove them away by



(a) (b)

Figure 3: The time-domain data transformed by IFFT with 801 points, (a) scaling sphere, (b) metal flat plate.

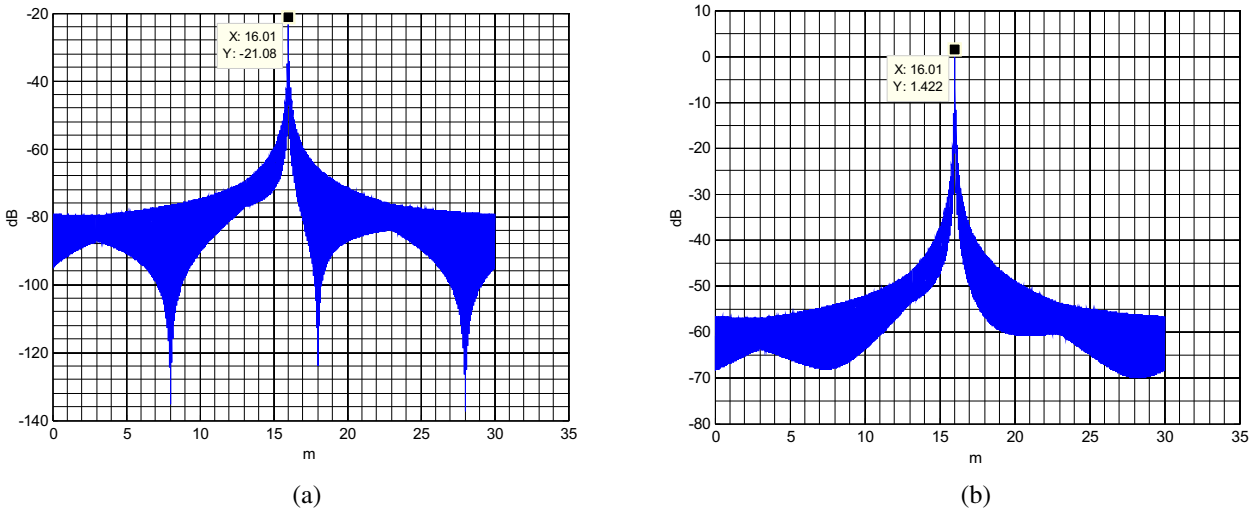


Figure 4: The time-domain data transformed by IFFT with 3201 points, (a) scaling sphere, (b) metal flat plate.

using a gate of range. After gating, the RCS value of a sphere in frequency-domain is different from the theoretical one, and the further away from the center frequency, the greater difference of RCS value comes. If we narrow the width of gate, the situation will be deteriorated shown as in Figure 5 .

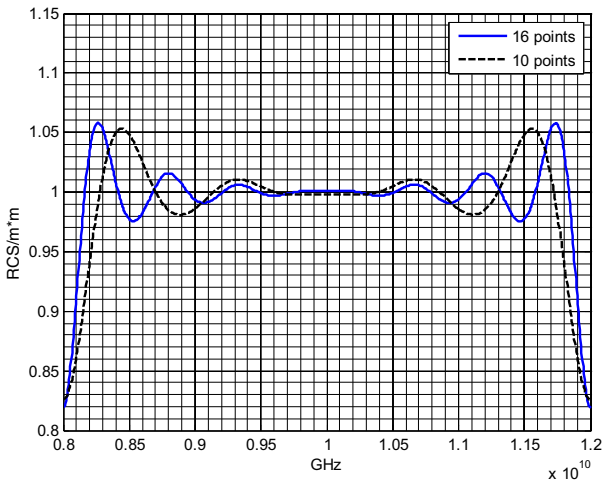


Figure 5: Effect of gate's width.

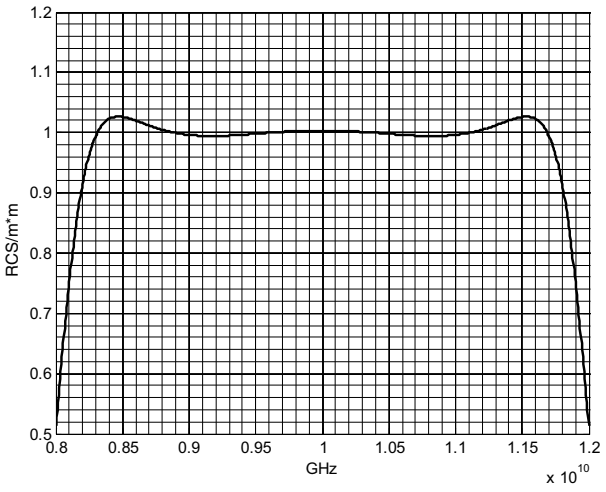


Figure 6: Effect of using Hamming windows.

In order to improve the accuracy, Hamming windows are used because they have low side- lobes. The line shown in Figure 6 is the result which we use Hamming windows in frequency-domain, then after gating and returning, it is natural to see that the value over 80% are identical with theoretical value, so the use of Hamming windows expands the accurate results in frequency-domain.

3. EXPERIMENTS AND ANALYSIS

The Step-Frequency RCS measurement system in the anecho chamber is shown in Figure 7. Two metal spheres whose diameter are 400 mm and 50 mm respectively, the former is used for scaling and the later is used for target.

The span of frequency is 4 GHz, the number of frequency points is 801, after using IFFT with 3201 points, a accurate time-domain response is come out, then we gate them appropriately, and return it to the frequency-domain, the value of target frequency response after scaling can be seen in Figure 8 comparing with theoretical data. However, if we use Hamming windows, the situation will be changed greatly as shown in Figure 9. Most of the results are close to theoretical value, so the accuracy of measurement is greatly improved by using these DSP techniques.

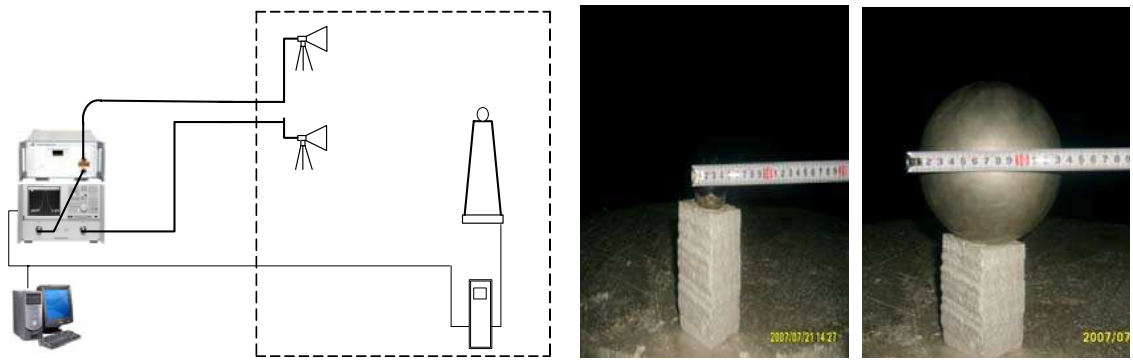


Figure 7: Step-Frequency RCS measurement system.

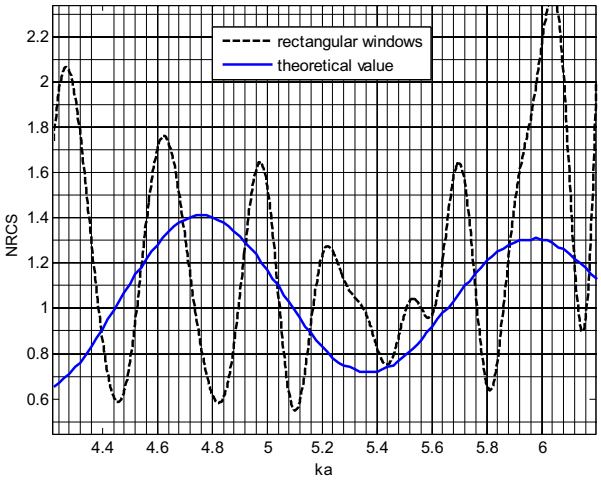


Figure 8: Frequency response of target with rectangular windows.

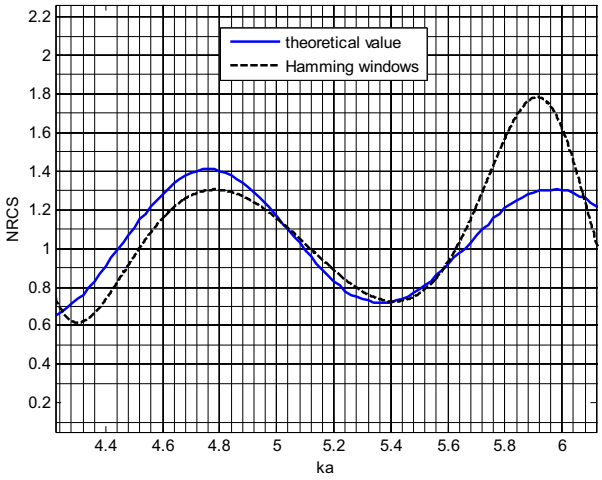


Figure 9: Frequency response of target with Hamming windows.

4. CONCLUSION

In the Step-Frequency RCS measurement system, if Hamming windows are applied to the original data in frequency-domain, and IFFT is used with sufficient points, then after adding a gate of range and returning back to frequency-domain, we can obtain the more precise results. It is great important to use DSP techniques in the low scatter RCS measurement.

REFERENCES

1. Dybdal, R. B., "Radar cross section measurements," *IEEE Transactions on Antennas and Propagation*, Vol. 75, No. 4, 498–516, 1987.
2. Currie, N. C., *Radar Reflectivity Measurements: Techniques and Applications*, Artech House, Norwood, MA, 1989.
3. Harris, F., "On the use of windows for harmonic analysis with the discrete Fourier transform," *IEEE Transactions on Antennas and Propagation*, Vol. 26, No. 1, 172–204, 1978.
4. Knott, E. F., J. F. Shaeffer, and M. T. Tuley, *Radar Cross Section*, Vol. 68, Artech House, 1985.

Electromagnetic Absorption by Conducting Fiber Filled Composite in the Centimeter- and Millimeter-wave Regions

L. Y. Liu[1], **L. Z. Wu**[1], **S. B. Pan**[2], **X. Wang**[3], **R. Z. Gong**[3], **and H. H. He**[3]

[1] School of Electrical and Electronic Engineering, Hubei University of Technology
Wuhan 430068, China
[2] Nonmetal Institute of Jinan, Jinan 250031, China
[3] Department of Electronic Science and Technology, Huazhong University of Science and Technology
Wuhan 430074, China

Abstract— We observed strong electromagnetic absorption by the conducing fiber filled foam composite with a very low fiber concentration in a wide frequency range. The strong absorption in the millimeter-wave region can be attributed to the magnetic property which originates from the interaction of fibers in high frequency region. The experimental data are in good agreement with the theoretical computations. With the aspect ratio of the fiber increasing and the conductivity decreasing, strong absorption would emerge throughout the centimeter- and millimeter-wave regions. The conducting fiber filled composite can be employed to develop very light microwave absorbing materials with strong absorbing ability in wide waveband.

1. INTRODUCTION

Microwave absorbing materials (MAMs) or electromagnetic interference (EMI) shielding materials are important for modern electronic technologies due to the increasingly severe electromagnetic pollution. Traditional MAMs are fabricated with granular fillers such as the ferrite and the magnetic metal particles. The loading ratios are usually the order of 10 vol%, therefore the MAMs are weighty. The frequency range with enough strong absorption is narrow, especially in the millimeter-wave region in which the intrinsic magnetic property of fillers is absent. For nonmagnetic granular fillers, the classical effective medium theories such as the Maxwell-Garnet Theory [1, 2] and the Bruggeman theory [2–5], regard the composite as nonmagnetic. When the magnetic property of fillers is absent, the wave impedance of the composite will mismatch much that of the free space, so the microwave reflectance is inevitably high. The conducting fiber filled composite can obtain high values of dielectric constant at a very low loading ratio [6–8]. With the frequency increasing, the fiber filled composite will acquire magnetic property due to the coupling among conducting fibers [6]. Therefore the wave impedance of the composite will match well that of the free space better in the millimeter region, and the microwave reflectance shall decrease.

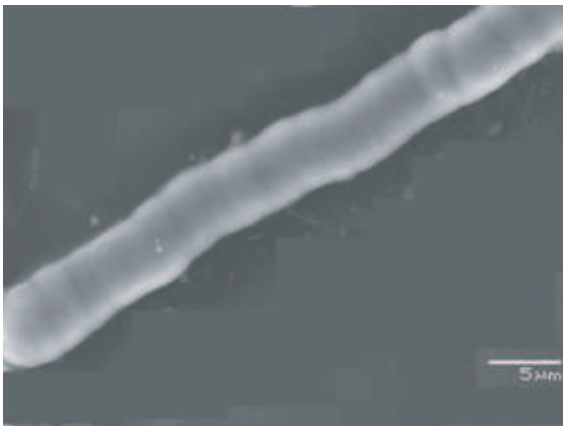


Figure 1: SEM micrograph of carbonyl iron fiber.

2. EXPERIMENT

The carbonyl iron fibers were fabricated by use of the magnetic-field-induced chemical decomposition method [9, 10]. The iron pentacarbonyl vaporized, and decomposed into spherical particles,

and the particles grew into carbonyl iron fibers under an adjustable magnetic field. The SEM micrograph of the carbonyl iron fiber is shown in Figure 1. The fiber filled composite was fabricated as follows. Polyol-polymethylene-isocyanate (PAPI) was filled into a tank (labeled as A). Polyol polyether, foam stabilizer, catalyst, blowing agent were mixed in another tank (labeled as B). Then the carbonyl iron fibers were dispersed into the mixture in B, and were mixed thoroughly. The PAPI was poured into B, and the components in B were mixed round $5 \sim 10$ s at 1800 rpm. When the color of mixture in B got white, the mixture was poured into a preheated mould. The mould was maintainted at about 80°C for 2 hours. Therefore the carbonyl iron fiber filled composite with a square plate shape (180 cm × 180 cm × 2 mm) was prepared. The microwave reflectance of the carbonyl iron fiber filled foam composite backed by a metal substrate was measured using the arch method. The reflectance was measured in $8 \sim 18$ GHz, $26.5 \sim 40$ GHz, and $75 \sim 110$ GHz. The diameter of the carbonyl iron fiber used is about 5 μm, and the length is about 1.5 mm, therefore the aspect ratio is around 300. The packing fraction of fibers is 0.046 vol%. The measured reflectance of the foam composite in the three regions is displayed in Figure 2. The absorption is weak in $8 \sim 18$ GHz and $26.5 \sim 40$ GHz. However, the absorbing ability of the foam composite is strong in $75 \sim 110$ GHz, with the reflectance being all less than -10 dB.

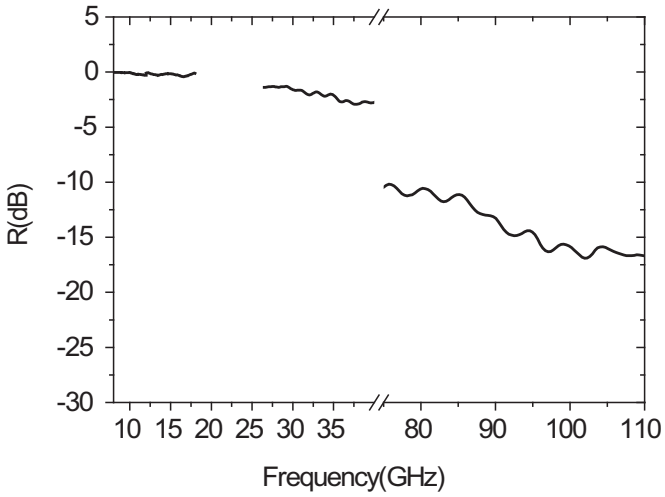


Figure 2: The measured microwave reflectance of the carbonyl iron fiber filled foam composite.

3. THEORY AND DISCUSSION

3.1. Theory

As well known, the microwave reflectance of the foam composite backed by a metal substrate is determined by [11],

$$R = 20 \lg \left| \frac{Z_{in} - 1}{Z_{in} + 1} \right|, \tag{1}$$

where Z_{in} is the normalized input impedance, being given by

$$Z_{in} = \sqrt{\frac{\mu_e}{\varepsilon_e}} \tanh \left(i \frac{2\pi d}{\lambda} \sqrt{\mu_e \varepsilon_e} \right), \tag{2}$$

where ε_e, μ_e and d are the effective permittivity, effective permeability, and thickness of the carbonyl iron fiber filled foam composite, and λ is the incident wavelength in free space. The effective permittivity of the composite can be calculated in terms of the effective medium theory (EMT). The classical EMT can be applied when the particles are spherical and their diameters are much less than the incident wavelength. If the particles are not spherical, spheroid can be used to simulate various nonspherical particles. In order to consider the skin effect in the case of high frequency, Lagarkov and Sarychev introduced the renormalized conductivity [6]. On the other hand, they revised the depolarization factor to take the interaction of fibers into account. Their theory, which is called scale dependent effective medium theory (SDEMT), anticipates a percolation threshold

that is linearly dependent upon the reciprocal of the aspect ratio of the conducting fiber. The anticipation is in agreement with the numerical simulations [12–14] and experiments [7]. The SDEMT equation is given by [6]

$$\frac{p}{3}\frac{f(\Delta)\varepsilon_m/\varepsilon_e}{1+f(\Delta)(\varepsilon_m/\varepsilon_d)(b^2/a^2)\ln(1+a\varepsilon_d/b\varepsilon_e)\cos\Omega}+3(1-p)\frac{\varepsilon_d-\varepsilon_e}{2\varepsilon_e+\varepsilon_d}=0, \tag{3}$$

where ε denotes the permittivity, subscript m indicates the conducting fiber, subscript d indicates the dielectric host, subscript e indicates the conducting fiber filled composite. $2a$ is the length of the fiber, b the radius of the fiber, p the packing fraction, and $f(\Delta)$ and Ω are given by

$$f(\Delta) = \frac{(1-i)}{\Delta}\frac{J_1[(1+i)\Delta]}{J_0[(1+i)\Delta]}, \tag{4}$$

$$\Omega^2 = \varepsilon_d(ak)^2\frac{\ln(a/b)+i\sqrt{\varepsilon_d}ka}{\ln(1+a\varepsilon_d/b\varepsilon_e)}, \tag{5}$$

where $\Delta = b\sqrt{2\pi\sigma_m\omega}/c$, is the ratio of the skin depth to the radius of the fiber, σ_m is the conductivity of the conducting fiber, J_0 and J_1 are Bessel functions of the zeroth and first order, ω is the radian frequency, c is the speed of light in the free space, and $k = \omega/c$.

According to the classical EMT, the effective permittivity and permeability of the composite are calculated respectively in terms of the dielectric permittivity and the magnetic permeability of the components. If the components are nonmagnetic materials, the composite are regarded as nonmagnetic. Bohren [15], and Mahan [16] ever pointed out that the composite consisting of relatively large nonmagnetic inclusions would show magnetic property. Bohren drew a conclusion that the magnetic permeability of the composite $\mu_{eff} = 1 + \dfrac{3ip}{2(kr)^3}[S(0) - S(\pi)]$, where $S(0)$ and $S(\pi)$ are the scattering amplitude of the embedded particle in the forward and the back direction respectively, k is the wave number in the host, and r is the radius of the particle. With the geometric size of the particle increasing, the scattered energy concentrate in the forward direction, therefore the composite consisting of relatively large nonmagnetic inclusions acquires magnetic property. The diameter of the carbonyl iron fiber is in the order of micrometer, which is much less than the incident wavelength. So the induced magnetic response is negligible [6, 17]. However, if we consider the coupling effect of the conducting fibers, the magnetic field in the incident electromagnetic wave will excite a circular current in high frequency, which comprises the displacement currents flowing between two nearest-neighbor fibers and the induced currents flowing in the fibers in the opposite direction. Every couple of the nearest-neighbor fibers obtains a magnetic moment, therefore the composite displays magnetic property. The magnetic moment of the couple of nearest-neighbor fibers $m \propto ad^2$, where d is the distance between two nearest-neighbor fibers. The magnetic moment of single fiber $m_1 \propto ab^2$, therefore the magnetic response of conducting fiber filled composite due to the coupling of the conducting fibers is much more evident than that induced by the magnetic moment of every single conducting fiber. When the conducting fibers are randomly distributed in the host, the iterative formula for the effective permeability of conducting fiber filled composite is given by [6]

$$\mu_e = 1 + \frac{2}{3}p\frac{a^2}{b^2}\frac{\varepsilon_d(ka)^2}{\ln(a/b)}\frac{\tan(ga)-ga}{(ga)^3}, \tag{6}$$

where

$$g = k\sqrt{\varepsilon_d\mu_e + i\frac{\varepsilon_d}{2\Delta^2 f(\Delta)\ln(a/b)}}. \tag{7}$$

3.2. Numerical Calculation and Discussion

Considering that the carbonyl iron fiber comprises some impurities such as Carbon [3, 4], their conductivity certainly decreases compared with pure iron. We assume that the conductivity of carbonyl iron fibers $\sigma_m = 2 \times 10^6\,\mathrm{S/m}$. The dielectric permittivity of the foam host approaches that of free space [18], and we put it to $\varepsilon_d = 1.05$. Therefore we calculate the effective dielectric permittivity and magnetic permeability of the carbonyl iron fibers filled foam composite in terms of the SDEMT equation (3) and the equation (6), as displayed in Figure 3(a). The calculated microwave reflectance (labeled as 1) in $8 \sim 110\,\mathrm{GHz}$ is displayed in Figure 3(b), being in good

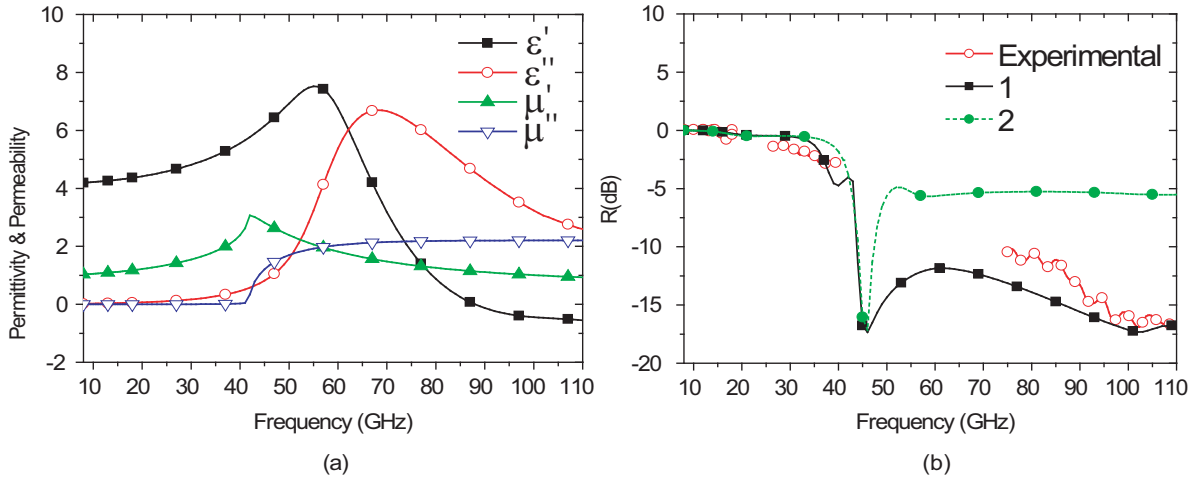


Figure 3: (a) The effective permittivity of carbonyl iron fiber filled foam composite calculated in terms of the SDEMT, and the effective permeability of the fiber filled foam composite calculated using Equation (6). (b) Curve 1: The reflectance calculated in terms of the effective permittivity and the effective permeability in (a); Curve 2: The reflectance calculated using the effective permittivity in (a) and the effective permeability equal to 1.

agreement with the experimental measurements. The intrinsic permeability of carbonyl iron [19] are less than 10 in $8 \sim 18$ GHz, and approaches unity in the millimeter-wave region. If we make use of the classical EMT to calculate the effective magnetic permeability of the carbonyl iron fiber filled foam composite, the effective magnetic permeability of the composite is almost equal to unity in $8 \sim 110$ GHz due to the very low packing fraction. The thus obtained microwave reflectance curve (labeled as 2) is displayed in Figure 3(b) as well. It is obvious that this curve is in disagreement with the experimental measurements, especially in $75 \sim 110$ GHz. In this region, the incident wavelength is comparable with the length of the conducting fiber, and the composite obtains an evident magnetic property as shown in Figure 3(a).

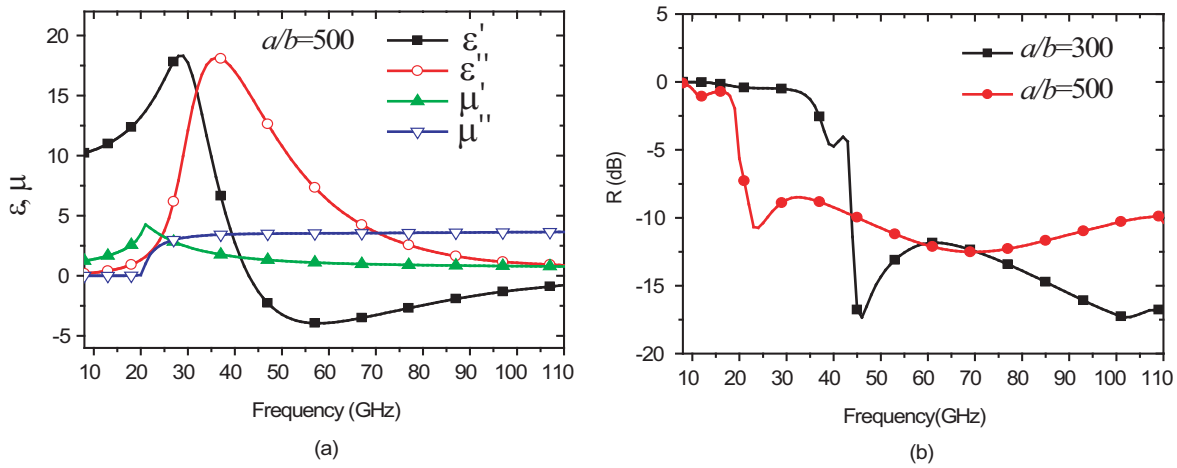


Figure 4: (a) The calculated effective permittivity and effective permeability of carbonyl iron fiber filled foam composite. The fiber aspect ratio is 500, but other parameters are all same as those in Figure 3. (b) Comparision of the calculated reflectance of the two composite consisting of conducting fibers with different aspect ratios.

As shown in Figure 2, the electromagnetic absorbing ability of the carbonyl iron filled foam composite is weak in the centimeter-wave region. This phenomenon can be attributed to the low aspect ratio and the high conductivity of the carbonyl iron fiber. The high aspect ratio is difficult to be achieved due to the brittleness of carbonyl iron fiber, so the aspect ratio of the fiber is 300. However, we set the aspect ratio to be 500, and calculate the effective permittivity, effective

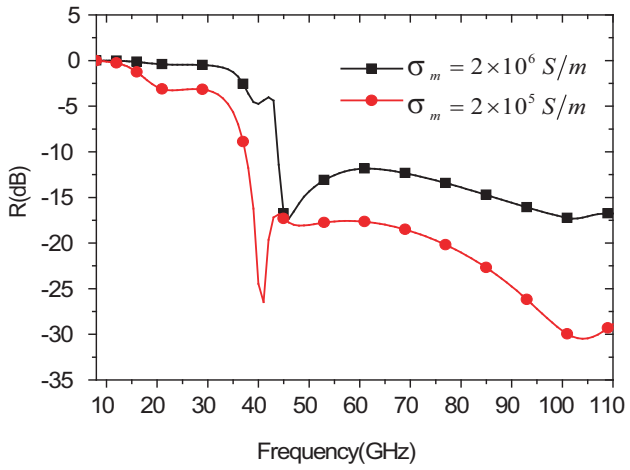


Figure 5: The reflectance of two piece of composite consisting of conducting fibers with different conductivities. The other parameters are same as those in Figure 2.

permeability, and reflectance as shown in Figure 4. When the conducting fiber is elongated, the frequency range where evident magnetic property emerges move to the low frequency region. The wave impedance of the composite match well that of the free space in the low frequency region, therefore the reflectance decrease also in this frequency region.

On the other hand, we can decrease the effective permittivity to make the wave impedance of the composite match well that of the free space. According to the SDEMT, we can decrease the effective permittivity of the composite by reducing the conductivity of the fiber. For example, we put the conductivity to 2×10^5 S/m, and compare the calculated reflectance with the reflectance calculated using $\sigma_m = 2 \times 10^6$ S/m. The results are displayed in Figure 5. With the conductivity of the fiber decreasing, the electromagnetic absorption property is improved, especially in the low frequency region.

Furthermore, we reduce the conductivity of the conducting fiber, and increase the aspect ratio simultaneously, and calculate the reflectance as shown in Figure 6. This composite shows a strong electromagnetic absorption property throughout the centimeter- and millimeter-wave regions. Therefore we can obtain excellent MAMs with very light weight and strong electromagnetic absorbing ability by adjusting the aspect ratio and conductivity of the conducting fiber.

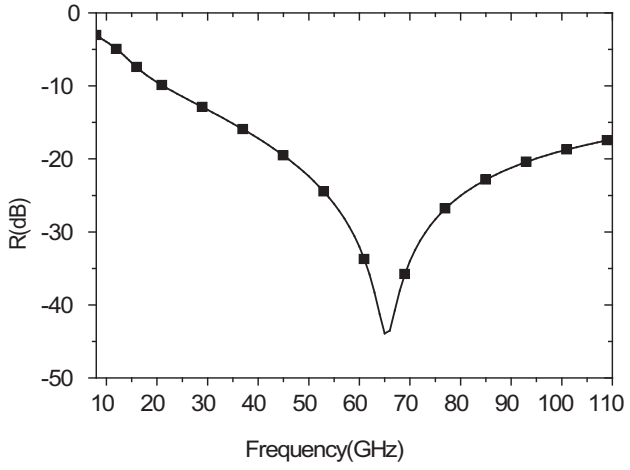


Figure 6: The calculated microwave reflectance of conducting fiber filled foam composite. We set parameters follows, fiber diameter $2b = 1\,\mu$m, length $2a = 2\,$mm (aspect ratio is 2000), conductivity $\sigma_m = 4 \times 10^5$ S/m, and packing fraction $p = 0.046\%$.

4. CONCLUSIONS

In summary, we studied experimentally and theoretically the microwave absorption property of carbonyl iron fiber filled foam composite in the centimeter- and millimeter-wave regions. We found that even at a very low packing fraction, the carbonyl iron fiber filled foam composite displayed a strong absorbing ability in a wide waveband, especially in the millimeter region. The theoretical approach, which gives the effective permittivity of the composite in terms of the SDEMT, and the effective permeability by considering the coupling of fibers, can predict the microwave reflectance which is in good agreement with the experiments. The numerical computation shows that we can obtain very light MAMs with strong absorbing ability in the centimeter- and millimeter-wave regions, by adjusting the parameters of conducting fiber such as aspect ratio and conductivity.

REFERENCES

1. Maxwell-Garnett, J. C., "Colours in metal glasses and in metallic films," *Philos. Trans. R. Soc.*, Vol. A203, 385–420, 1904.
2. Niklasson, G. A., C. G. Granqvist, and O. Hunderi, "Effective medium models for the optical properties of inhomogeneous materials," *Appl. Opt.*, Vol. 20, No. 21, 26–30, 1981.
3. Bruggeman, D. A. G., "Berechnung vershiedener physikalischer Konstanten von heterogenen Substanzen. I. Dielektrizitätskonstanten und Leitfähigkeiten der Mischkörper aus isotropen Substanzen," *Ann Phys (Leipzig)*, Vol. 24, 636–664, 1935.
4. Landauer, R., "Electrical conductivity in inhomogeneous media," *Electrical Transport and Optical Properties of Inhomogeneous Media*, edited by J. C. Garland and D. B. Tanner, *AIP. Conf. Proc.*, No. 40 (AIP, New York), 2–45, 1978.
5. Bergman, D. J. and D. Stroud, "Properties of macroscopically inhomogeneous media," *Solid State Phys.*, Vol. 46, 147–269, 1992.
6. Lagarkov, A. N. and A. K. Sarychev, "Electromagnetic properties of composites containing elongated conducting inclusions," *Phys. Rev. B*, Vol. 53, No. 10, 6318–6336, 1996.
7. Lagarkov, A. N., S. M. Matytsin, K. N. Rozanov, and A. K. Sarychev, "Dielectric properties of fiber-filled composites," *J. Appl. Phys.*, Vol. 84, No. 7, 3806–3184, 1998.
8. Lagarkov, A. N., A. K. Sarychev, Y. R. Smychkovich, and A. P. Vinogradov, "Effective medium theory for microwave dielectric constant and magnetic permeability of conducting stick composites," *Journal of Electromagnetic Waves and Applications*, Vol. 6, No. 9, 1159–1176, 1992.
9. Li, X. C., R. Z. Gong, Y. Nie, Z. S. Zhao, and H. H. He, "Electromagnetic properties of fe55ni45 fiber fabricated by magnetic-field-induced thermal decomposition," *Mater. Chem. Phys.*, Vol. 94, No. 2–3, 408–411, 2005.
10. Gong, R. Z. and X. C. Li, "Microwave electromagnetic characterization of carbonyl iron fibre prepared by magnetic-field-induced chemical decomposition," *Functional Materials*, Vol. 36, 1696–1697, 1700, 2005 (in Chinese).
11. Liu, L. Y., R. Z. Gong, Y. S. Cheng, F. G. Zhang, and H. H. He, "Emittance of a radar absorber coated with an infrared layer in the $3 \sim 5$ m window," *Opt. Express*, Vol. 13, No. 25, 10382–10391, 2005.
12. Balberg, I., N. Binenbaum, and C. H. Anderson, "Critical behavior of the two-dimensional sticks system," *Phys. Rev. Lett.*, Vol. 51, No. 18, 1605–1608, 1983.
13. Balberg, I., N. Binenbaum, and N. Wagner, "Percolation thresholds in the three-dimensional sticks system," *Phys. Rev. Lett.*, Vol. 52, No. 17, 1465–1458, 1984.
14. Balberg, I., "'Universal' percolation-threshold limits in the continuum," *Phys. Rev. B*, Vol. 31, No. 6, 4053–4055, 1985.
15. Bohren, C. F., "Applicability of effective-medium theories to problems of scattering and absorption by nonhomogeneous atmospheric particles," *J. Atmos. Sci.*, Vol. 43, No. 3, 468–475, 1986.
16. Mahan, G. D., "Long-wavelength absorption of cermets," *Phys.Rev. B*, Vol. 38, 9500–9502, 1988.
17. Podolskiy, V. A., "Plasmon modes and negative refraction in metal nanowire composites," *Opt. Express*, Vol. 11, No. 7, 735–745, 2003.
18. Liu, L., S. M. Matitsine, Y. B. Gan, and K. N. Rozanov, "Effective permittivity of planar composites with randomly or periodically distributed conducting fibers," *J. Appl. Phys.*, Vol. 98, 063512, 2005.
19. Olmedo, L., G. Chateau, C. Deleuze, and J. L. Forveille, "Microwave characterization and modelization of magnetic granular materials," *J. Appl. Phys.*, Vol. 73, No. 10, 6992, 1993.

Analysis of the Optimal Gap Width and Gap-to-gap Distance in π-mode Double-gap Cavities for Broadband Klystrons

Fu-Min Lin
Department of Physics, Shantou University, Guangdong 515063, China

Abstract— Single electron approximate approach is used to analyze the interaction process of ideal clustering electron beam with the resonant E-field in the gaps of π-mode double-gap cavities. The research shows that equal two gaps and the transit angle π of the gap-to-gap distance in π-mode double-gap cavities is not a suitable choice to get the maximal conversion efficiency. The formulae of the optimal gap width and gap-to-gap distance are put forward to design asymmetric π-mode double-gap cavities for higher conversion efficiency.

1. INTRODUCTION

Double-gap cavities are usually used as output cavities of broadband klystrons to widen working band or enhance output power [1–6]. An accurate design of the gap width and the gap-to-gap distance is very important to the efficiency of klystrons. Unsuitable gap width or distance between two gaps will cause parasitic oscillation or instability.

Equal two gaps with same width are generally used in π-mode double-gap cavities [7–10], but it may not be the optimal choice to get the maximal efficiency of energy exchange of well clustering electron bunch with resonant electro-magnetic field, because of the different speeds of electrons in two gaps. So we must calculate the different transit angle relative to the gap width of two gaps respectively. Also, it is not the best that the transit angle of the gap-to-gap distance in π-mode double-gap cavities is π, which has been chosen in conventional design of broadband klystrons. We will analyze the interaction process of well clustering electron beam with the resonant E-field in two gaps of a general asymmetric π-mode double-gap cavity and calculate the efficiency of energy exchange of motive electron with resonant electro-magnetic field with the single electron approximate approach. Then, it is possible to determine the optimal gap width and the gap-to-gap distance for higher conversion efficiency of kinetic energy of electron into energy of resonant electro-magnetic field of certain frequency.

2. THE ENERGY EXCHANGE OF IDEAL CLUSTERING BEAM WITH THE RESONANT FIELD IN GAPS

A general π-mode double-gap cavity works in TM-mode, and the resonant E-field in two gaps is always in opposite directions. Fig. 1 is the main section plane of the cavity, in which the directions of the E-field in two gaps are showed. When a clustering electron beam passes through a gap, it will exchange energy with the resonant electro-magnetic field. The conversion efficiency achieves the maximum in the case of that repeated electron clusters always pass through the gap in the direction of the resonant E-field and the repeated frequency of electron cluster equals to resonant frequency of the cavity. The kind of ideal clustering electron beam interacting with the resonant E-field in the gap for the maximum conversion efficiency is showed in Fig. 2. So it is reasonable to regard an ideal clustering electron beam as a series of periodically repeated single electron in calculation of the energy exchange of the clustering electron beam with the resonant electro-magnetic field of certainfrequency in the cavity.

Assuming that a electron with velocity V_0 enters into the first gap at the time of $t_0 = 0$, and the resonant E-fields of certain frequency in two gaps are expressed as

$$E_1 = E_0 \cos(\omega t + \varphi) \quad \text{(Gap 1)} \tag{1}$$

$$E_2 = -E_0 \cos(\omega t + \varphi) \quad \text{(Gap 2)} \tag{2}$$

where φ is the leading phase of E-field in first gap relative to the clustering electron beam. It is determined by the load or the gap impedance of the output cavity.

After the motive electron transits the Gap 1 at the time of t_1, the velocity of the electron becomes V_1. Then the electron will pass through the drift tube with constant velocity and enters

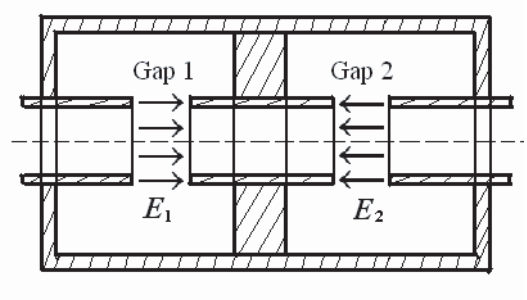


Figure 1: Main section plane of a double-gap cavity.

into the Gap 2 at the time of t_1+t_L. When the electron leaves the Gap 2 at the time of $t_1+t_L+t_2$, the velocity of the electron will change from V_1 to V_2. It is easy to get

$$V_1 = V_0 + \frac{eE_0}{m_e\omega}[\sin\varphi - \sin(\omega t_1+\varphi)] \tag{3}$$

$$V_2 = V_0 + \frac{eE_0}{m_e\omega}[\sin\varphi - \sin(\omega t_1+\varphi) - \sin(\omega t_1+\omega t_L+\varphi) + \sin(\omega t_1+\omega t_L+\omega t_2+\varphi)] \tag{4}$$

$$t_L = \frac{L}{V_0 + (eE_0/m_e\omega)[\sin\varphi - \sin(\omega t_1+\varphi)]} \tag{5}$$

$$d_1 = V_0 t_1 + \frac{eE_0}{m_e\omega^2}[\omega t_1 \sin\varphi + \cos(\omega t_1+\varphi) - \cos\varphi] \tag{6}$$

$$d_2 = V_0 t_2 + \frac{eE_0 t_2}{m_e\omega}[\sin\varphi - \sin(\omega t_1+\varphi) - \sin(\omega t_1+\omega t_L+\varphi)] + \frac{eE_0}{m_e\omega^2}[\cos(\omega t_1+\omega t_L+\varphi) - \cos(\omega t_1+\omega t_L+\omega t_2+\varphi)] \tag{7}$$

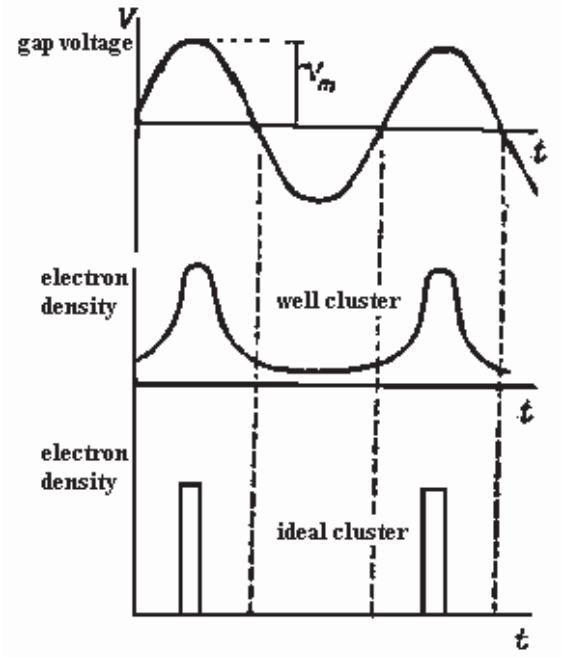


Figure 2: The ideal clustering electron beam.

When a π-mode double-gap cavity is used as an output cavity of a broadband klystron and has a suitable load and gap impedances, the V_2 expressed by Equation (4) will be smaller than V_0. Then the kinetic energy of electrons is converted into the energy of the resonant electro-magnetic

field of certain frequency. The conversion efficiency is

$$\eta = 1 - \left(\frac{V_2}{V_0}\right)^2 \tag{8}$$

3. THE OPTIMAL GAP WIDTH AND GAP-TO-GAP DISTANCE FOR HIGHER EFFICIENCY

It is seen from Equation (4) that the conversion efficiency would not achieve the maximum except that $\varphi = -\pi/2$, $\omega t_1 = \omega t_2 = \pi$, $\omega t_L = 2\pi$. But at the resonant frequency of the first gap, the leading phase φ of E-field relative to the clustering electron beam is zero. The general value of φ for all the suitable loads or gap impedances is $-\pi/2 < \varphi < \pi/2$. So in order to get higher conversion efficiency in a relative wide working band, we must choose the optimal gap width and gap-to-gap distance under the condition of $\varphi = 0$, which corresponds to the center of the working band of an output circuit. Then the optimal transit angles or the lengths of two gaps and the drift tube between two gaps can be calculated by Equations (4)–(7).

$$\theta_1 = \omega t_1 = \frac{\pi}{2}, \quad \theta_2 = \omega t_2 = \pi, \quad \theta_L = \omega t_L = 2\pi \tag{9}$$

$$d_1 = \frac{\pi V_0}{2\omega}\left(1 - \frac{2}{\pi}\frac{eE_0}{m_e\omega V_0}\right), \quad d_2 = \frac{\pi V_0}{\omega}\left(1 - \frac{2eE_0}{m_e\omega V_0}\right), \quad L = \frac{2\pi V_0}{\omega}\left(1 - \frac{eE_0}{m_e\omega V_0}\right) \tag{10}$$

The conversion efficiency in the center of the working band of an output circuit in the case of the optimal gap width and gap-to-gap distance is

$$\eta_{\text{opt}} = 1 - \left(1 - \frac{3eE_0}{m_e\omega V_0}\right)^2 \tag{11}$$

In the practical design of broadband klystrons, it is always that the conversion efficiency and the beam voltage or the original velocity V_0 of repeated electron clusters are predetermined. If so, the optimal gap width and gap-to-gap distance of π-mode double-gap cavities for higher conversion efficiency can be calculated by Equation (10) and Equation (11).

$$\frac{eE_0}{m_e\omega V_0} = \frac{1 - \sqrt{1 - \eta_{\text{opt}}}}{3} \tag{12}$$

$$d_1 = \frac{V_0}{3\omega}\left(\frac{3\pi}{2} - 1 + \sqrt{1 - \eta_{\text{opt}}}\right) \quad \text{or} \quad \frac{\omega d_1}{V_0} = \frac{\pi}{2} - \frac{1 - \sqrt{1 - \eta_{\text{opt}}}}{3} \tag{13}$$

$$d_2 = \frac{2\pi V_0}{3\omega}\left(\frac{1}{2} + \sqrt{1 - \eta_{\text{opt}}}\right) \quad \text{or} \quad \frac{\omega d_2}{V_0} = \frac{\pi}{3} + \frac{2\pi}{3}\sqrt{1 - \eta_{\text{opt}}} \tag{14}$$

$$L = \frac{2\pi V_0}{3\omega}\left(2 + \sqrt{1 - \eta_{\text{opt}}}\right) \quad \text{or} \quad \frac{\omega L}{V_0} = \frac{4\pi}{3} + \frac{2\pi}{3}\sqrt{1 - \eta_{\text{opt}}} \tag{15}$$

4. CONCLUSION

The paper analyzes the interaction process of ideal clustering electron beam with resonant electromagnetic field of certain frequency in an asymmetric π-mode double-gap cavity with single electron approximate approach, and puts forward the formulae of the optimal gap width and gap-to-gap distance to design asymmetric π-mode double-gap cavities for higher conversion efficiency. The results of Equations (9)–(15) may be beneficial to manufacture higher quality of broadband klystrons.

ACKNOWLEDGMENT

The work is supported by National Natural Science Foundation of China, No. 60671060.

REFERENCES

1. Vieks, A. E., R. S. Callin, G. Caryotakis, et al., "100 MW klystron development at SLAC," *Proceedings of the IEEE Particle Accelerator Conference*, Vol. 2, 798–800, San Francisco, CA, USA, 1991.
2. Lemke, R. W., M. C. Clark, and B. M. Marder, "Theoretical and experimental investigation of a method for increasing the output power of a microwave tube based on the split-cavity oscillator," *J. Appl. Phys.*, Vol. 75, No. 10, 5423–5432, 1994.

3. Kisanov, S. A., A. L. Klimov, S. D. Korovin, et al., "A vircator with electron beam premodulation bassed on high-current repetitively pulsed accelerator," *IEEE Trans. Plasma Sci.*, Vol. 30, No. 1, 274–285, 2002.
4. Lin, F.-M., Y. G. Ding, and B. Shen, "Analysis of characteristic of coupling two cavity," *Journal of Electronics & Information Technology*, Vol. 25, No. 7, 982–989, Jul. 2003.
5. Lin, F.-M. and Y. G. Ding, "Research and design of a-mode strongly coupling two-gap output cavity with four coupling slots suitable to low frequency broadband MBK," *Vacuum Electronics*, No. 5, 39–41, Oct. 2003.
6. Ding, Y., B. Shen, S. Shi, and J. Cao, "S-band multibeam klystron with bandwidth of 10%," *IEEE Trans. Electron Devices*, Vol. 52, No. 5, 889–894, 2005.
7. Ding, Y., Y. Zhu, X. Yin, et al., "100 kW L-band CW broadband multi-beam klystron," *Proceedings of the Fourth IEEE International Vacuum Electronics Conference*, 368–369, Seoul, South Korea, 2003.
8. Lin, F. M. and Y. G. Ding, "The research and design of a coaxial filter broadband output circuit for L-band multiple beam klystron," *Journal of Electronics & Information Technology*, Vol. 26, No. 5, 813–819, May 2004.
9. Lin, F.-M. and Y. Ding, "Research on an MBK output circuit of π-mode double gap cavity loaded by cutoff waveguide filter," *High Power Laser and Particle Beams*, Vol. 18, No. 12, 201–2018, 2006.
10. Yuan, W.-J., F.-M. Lin, and H.-Y. Wang, "Research on π-mode double-gap rectangular TM_{310} cavity," *5th Int. Conf. Micromave Millimeter Wave Technol. Proceedings*, 635–638, Guilin, China, 2007.

Analysis and Design of Power Generator on Passive RFID Transponders

Fan Jiang[1, 3], **Donghui Guo**[1, 2, 3], **and L. L. Cheng**[4]

[1]EDA Lab, Department of Physics, Xiamen University, Xiamen 361005, China
[2]Department of Electronic Engineering, Xiamen University, Xiamen 361005, China
[3]Xiamen RichIT Microelectronic Technologies LTD, Xiamen 361005, China
[4]Electronic Engineering Department of City University of Hong Kong, Hong Kong, China

Abstract— Rectifier is the essential part of passive RFID transponders. This paper based on two types of conventional rectifier, analysed several factors which influence the output voltage and Power Conversion Efficiency (PCE) of the rectifier. From the analysis results of these two types of rectifier, we proposed a modified version of the rectifier which has high output voltage and high PCE. Simulation results are also studied and presented in this paper.

1. INTRODUCTION

The radio frequency identification (RFID) system is a system makes use of the bidirectional wireless communication technology in order to identify the target and then process the relevant data. The system is mainly formed by two parts, it includes a RFID tag and a reader [1]. According to the power supply of RFID tag could be either one of the passive, semi-active or active power supplies. The passive RFID tag, due to its low cost and longer lasting life, hence it is more popular [2].

Passive tag does not consist of any power source in general and the tag obtains power energy from the electromagnetic waves sending by the RFID reader through the coupling antenna. The electromagnetic waves after rectification, voltage stabilization then the tag obtain the DC voltage which maintain the need of the normal operations of the RFID tag and its circuitry. Hence, DC voltage generation technique is the key technology for passive RFID tag.

The rectifier is the basic part of the DC voltage generation circuit of the RFID tag, the performance of tag relies on it. The main parameters of the rectifier affecting the performance of the RFID tags are the output voltage V_{out} and the Power Conversion Efficiency (PCE). The PCE is defined as the ratio of the output power P_{out} and the input power P_{in} of the rectifying circuit. (i.e., $PCE = P_{out}/P_{in}$).

Recently researches have paid attentions to this topic [3–5]. The RFID rectifier discussed in [4] is using the structure of a half wave rectifier and the PCE is low. The RFID rectifier discussed in [5] is a self-boosting circuit, with 3.16 V induced voltage on the tag antenna, the output voltage and PCE could reach 2.38 V and 70.1% respectively. However the structure is more complex and it uses more chip area. This paper is to discuss the factors that affect the PCE of rectifier, then proposed a new type of rectifier with the addition of two more MOSFET switches to the conventional circuit which help to reduce the power loss, hence increase the PCE and output voltage.

In order to explain the circuit design in detail, we first analysis of the functionalities and performances of two conventional rectification circuits, then a modified circuit will be proposed. Simulations and comparisons of the two types of conventional rectifier and the modified rectifier will be done.

2. CONVENTIONAL RECTIFICATION CIRCUIT

Figure 1 shows two types of conventional rectification circuits [5], they are mainly suitable for power source generator of the low and high frequency band passive RFID tag. In Figure 1(a) is the NMOS gate cross connection bridge rectification circuit and this circuit has two diodes connecting to the MOSFET, for this reason, the circuit will have a threshold voltage drop Vth from the antenna to the load capacitor. In Figure 1(b), it is using PMOS and NMOS device and has a cross gate connection structure. Although it solved the threshold voltage drop of the diodes connected to MOSFET, this circuit has to face another problem, such as the antenna voltage is smaller than the storage capacitor voltage, the storage charge will feedback to the antenna through the PMOS, thus it causes low PCE and a bigger voltage swing on the load.

Below is using the structure of the rectification circuit of Figure 1(b) as an example to analysis the factors affect the PCE. Figure 2 shows the transient equivalent circuit of the (N) MOS device

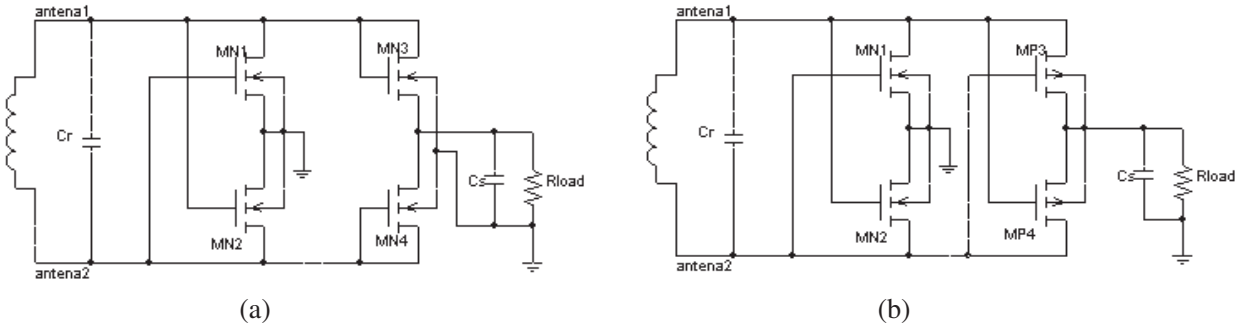


Figure 1: Two types of conventional rectifier circuit. (a) NMOS gate cross-connection bridge rectifier circuit. (b) PMOS, NMOS gate cross-connection bridge rectifier circuit.

(i.e., MOS1 model [6]), Dbd, Dbs are the substrate-drain diode and the substrate-source diode respectively. The rs, rd and rds are the source connection resistor, drain connection resistor and channel resistor respectively. The cgb, cgs, cgd, cbs and cbd represent the parasitic capacitances, comparing with the parameters above, rs and rd can be ignored [7]. Such that you can provide the transient equivalent circuit of the PMOS, NMOS gate cross-connection bridge rectifier circuit as shown in Figure 3.

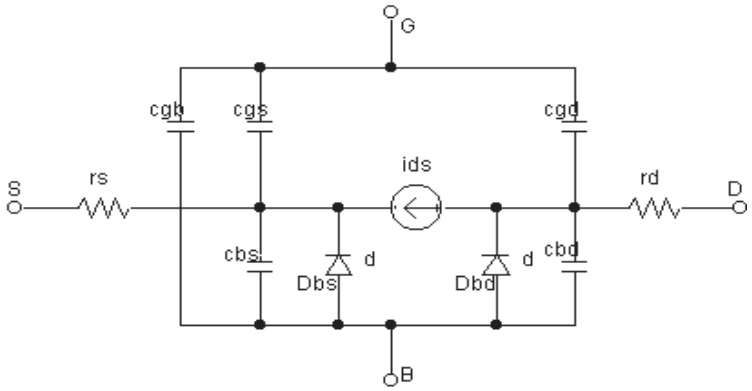


Figure 2: The transient equivalent circuit of a MOS device (i.e., MOS1 model).

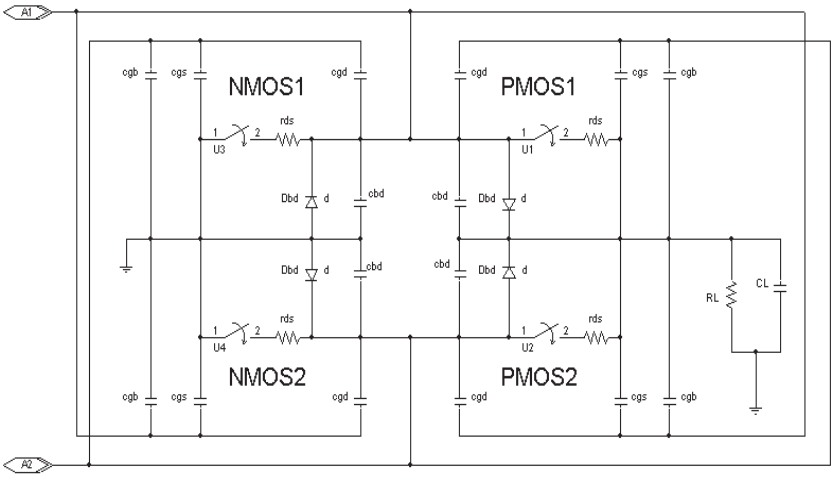


Figure 3: The transient equivalent circuit of PMOS, NMOS grid cross-connection bridge rectifier circuit.

Based on the working condition of an equivalent circuit as shown in Figure 3, we could find out the source that creates the power loss of this circuit:

1. The A1(A2) is a high voltage and A2(A1) is a low voltage, the PMOS1(PMOS2) and NMOS2 (NMOS1) are conductive, the circuit charge up the load capacitor through the conductive resistor rds. The main power loss is on the conductive resistor.
2. When A1(A2) is a high voltage and A2(A1) is a low voltage, if the voltage of the load capacitor CL is higher than the voltage on A1(A2), the charge on the load capacitor will go through the conductive resistor rds and parasitic capacitor then feedback to the antenna and it causes the power loss.
3. When the threshold of NMOS1(NMOS2) is greater than the conductive voltage of the parasitic diode Dbd, and the voltage of A1(A2) is smaller than the ground with the conductive voltage of the parasitic diode Dbd, the current will from the ground and pass through the diode to the antenna. This causes the current leakage of the substrate and the power conversion efficiency will drop.
4. PMOS and NMOS transistors are both treated as switches and the current pass through them will make the parasitic capacitor of the transistor being charged and then discharged, this will introduce power loss of the transistor.

3. MODIFICATION AND IMPROVEMENT OF THE CONVENTIONAL RECTIFIER CIRCUIT

According to the above analysis, we may improve the circuit in certain aspects,

1. In order to reduce or eliminate the threshold voltage drop from the antenna to the load capacitor, it can make use of the PMOS, NMOS gate cross connection bridge rectifier circuit.
2. In order to reduce the power loss on rds it can use a bigger size MOS transistor, but the bigger size MOS transistor will raise a larger parasitic capacitor. When the MOS transistor is conductive, the power loss will increase when the capacitor is charged and discharged. For this reason, the optimization of the size of the MOS transistor is needed.
3. The power loss due the current leakage could be reduced by using the substrate bias [8] technology.
4. Try hard to suppress the feedback current from the loading capacitor to the antenna.

This paper has shown the modified rectifier circuit as shown in Figure 4 and two more NMOSFET switches are added in front of the storage capacitor based on the circuit on Figure 1(b). The main purpose of these switches is to suppress the current feedback when the voltage on the antenna is lower than the voltage on CL.

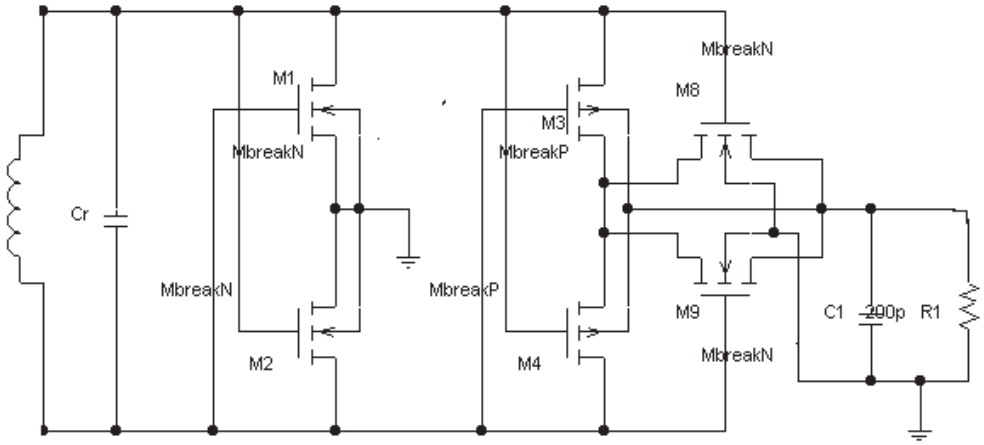


Figure 4: The rectifier circuit after improvement.

4. COMPARISONS OF THE SIMULATION RESULTS

The comparisons of the simulation results of two types of conventional rectifier circuits and the modified rectifier circuit will be shown below. The rectifier circuit input is complied with the standard of ISO/IEC 15693-2 and the sine-wave carrier frequency is decided to be 13.56 MHz, the amplitude is 3.16 V [5]. It is using the same size of the MOS transistors. The load capacitor is

200 p and the typical load resistor value is 45 KΩ [3] and CSMC 0.6 µm CMOS layout model is adopted [9]. Pspice is the tool for modeling the circuits and the simulations was done by the tool also. The simulation results are shown in Figure 5.

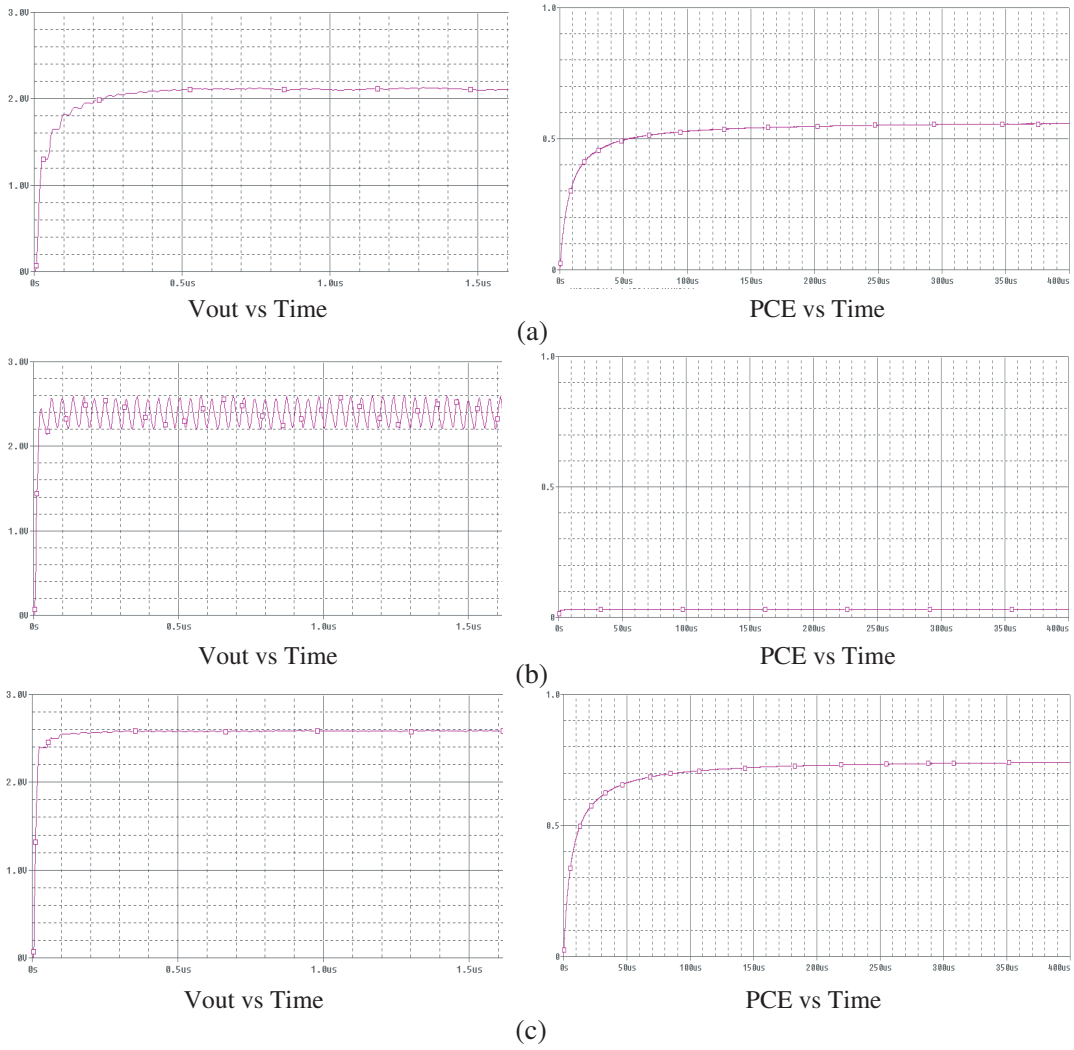


Figure 5: Simulation results for three types of rectifier circuits being discussed. (a) The NMOS gate cross-connection bridge rectifier circuit $V_{out} \approx 2.1$V and PCE $\approx 56\%$. (b) The PMOS, NMOS gate cross-connection bridge rectifier circuit of $V_{out} \approx 2.4$ V and PCE $\approx 3.1\%$. (c) Modified rectifier circuit of $V_{out} \approx 2.58$ V and PCE $\approx 74\%$.

Comparing of the simulation results, the modified rectifier circuit overcomes the threshold voltage drop of NMOS gate cross-connection bridge rectifier circuit. Hence, the output voltage can reach 2.58 V; simultaneously it overcomes the feedback current of the load capacitance of the PMOS, NMOS gate cross-connection bridge rectifier which causes the instability of the output voltage and the low efficiency of the power conversion. The PCE of the rectifier circuit hence can reach 74%.

In Figure 5(b), the simulation shows the PCE is very low because of the power loss on the conductive resistors and the parasitic capacitors due to the feedback current from the load capacitor.

The output voltage V_{out} and the PCE is limited by the size of the MOS transistor, due to its channel resistors and the parasitic capacitors.

From Figure 6, follows the change of the size of the MOS transistor, as the size increase, the channel resistance will become small. The voltage drop on the channel resistor will less, hence this will increase the output voltage V_{out}. These three types of rectifier have quite the same results. For PCE, as the size of the MOS transistor has increased, the power loss on the channel resistance will become less, however, it will introduce the parasitic capacitance simultaneously. Then, it will cause the power loss becomes larger. In this case, the channel resistance and the parasitic capacitance will be the major factors influence the value of PCE. In order to get the highest value

of PCE, an optimized size of the MOS transistor has to be found. In Figure 6(a) shows that when the channel width of the MOS transistor is $w = 15\,\mu m$, PCE = 56.3%. In Figure 6(c) shows that when the channel width of the MOS transistor is $w = 10\,\mu m$, PCE = 75.7%. In Figure 6(b), for PMOS, NMOS gate cross-connection bridge rectifier circuit, due to feedback current from the load capacitance to the antenna, the power loss due to the parasitic capacitance is higher than the channel resistance, therefore as the size of the MOS transistor increases, the PCE on the whole will decrease and the value of PCE is relatively low.

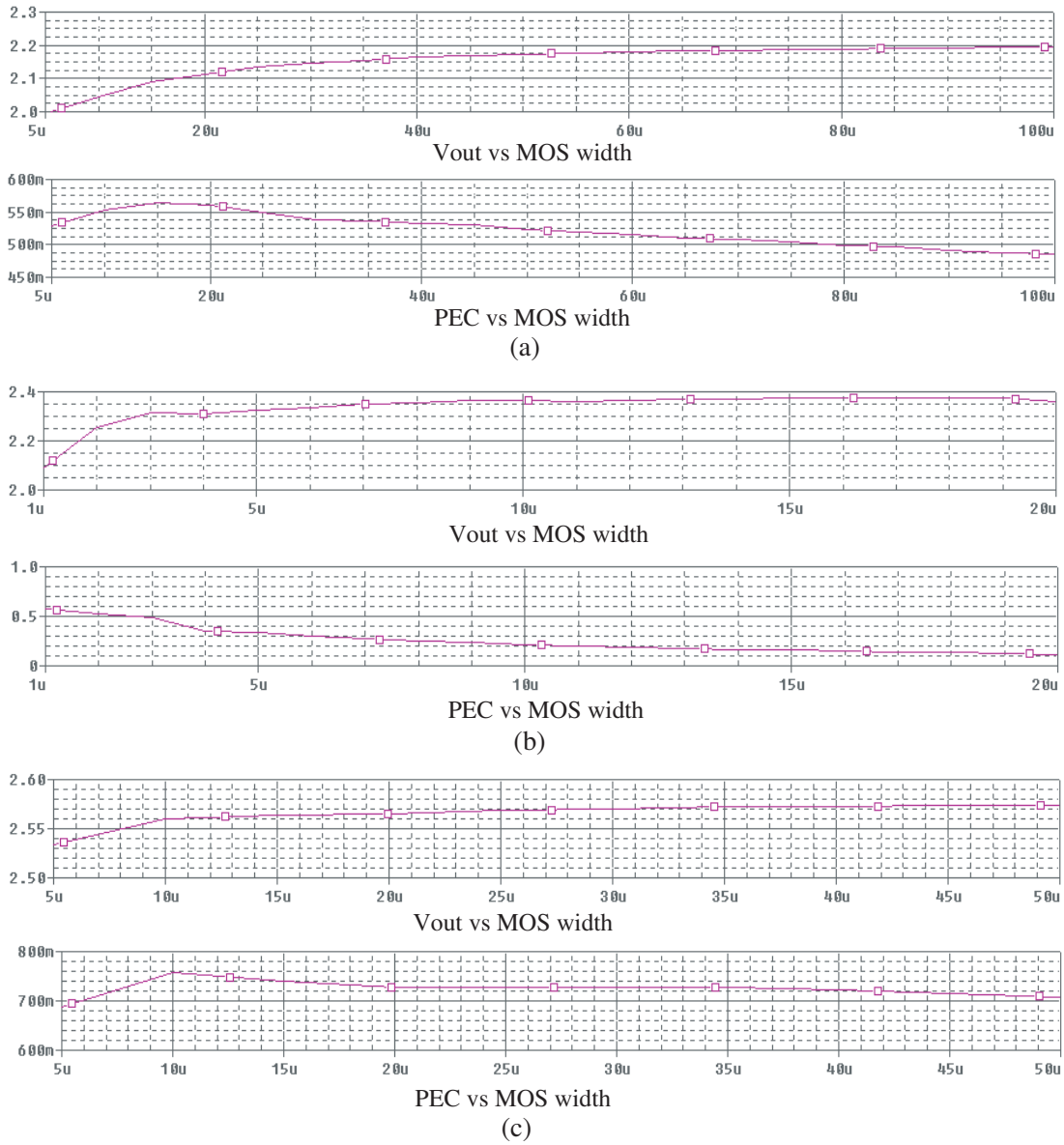


Figure 6: The relation between Output voltage V_{out}, the Energy Conversion Efficiency and MOS Transistor size. (a) NMOS gate cross-connection bridge rectifier circuit. (b) PMOS, NMOS gate cross-connection bridge rectifier circuit. (c) Modified rectifier circuit.

From the simulation results shown above, the performance of the modified rectifier circuit is higher than the two typical rectifier circuits. It has been significantly improved for V_{out} and PCE.

5. CONCLUSION

This paper has proposed a new type of rectifier circuit for RFID tag and it has done modification to the conventional rectifier circuit. It eliminates the threshold voltage drop from the antenna to the RFID chip such that it increases the output voltage. Moreover, two more MOS switches are added in order to suppress the feedback current from the RFID to the antenna. This improves

the PCE and the maximum PCE can reach 75.7%. The proposed circuit is compliance to the industrial standard of CMOS technology and is suitable for passive RFID tag chip application with low frequency and short distance.

REFERENCES

1. Finkenzeller, K., *RFID Handbook [Electronic Resource]: Fundamentals and Applications in Contactless Smart Cards and Identification*, translated by Rachel Waddington, Chichester, England; Wiley, Hoboken, N. J., c2003.
2. Huang, H.-P. and Y.-T. R. Chang, "Design for RFID system testing and applications," *2007 IEEE International Conference on Robotics and Automation*, 4594–4599, 10–14 April, 2007.
3. Zheng, Z., B. Jamal, and P. H. Cole, "Brief comparison of different rectifier structures for RFID transponders," http://www.m-lab.ch/auto-id/SwissReWorkshop/papers/BriefComparisonOfRectifierStructuresForRFIDtransponders.pdf.
4. Lu, C. and Y.-M. Li, "The RF interface circuits design of contactless IC cards," *4th International Conference on ASIC*, 278–280, Shanghai, China, 2001.
5. Hu, J.-Y., Y. He, and H. Min, "High efficient rectifier circuit eliminating threshold voltage drop for RFID transponders," *6th Internationl Conference on ASIC ASICON*, 607–610, Shanghai, China, 2005.
6. Abrahams, J. R., *Semiconductor Circuits: Theory, Design and Experiment*, Pergamon, Oxford, 1966.
7. Van, der T., J. Michael, and S. G. Chamberlain, "Buried-channel MOSFET model for SPICE," *IEEE Transactions on Computer-aided Design of Integrated Circuits and Systems*, Vol. 10, No. 8, 1015–1035, 1991.
8. Ghovanloo, M. and K. Najafi, "Fully integrated wideband high-currentr rectifiers for inductively powered devices," *IEEE Journal of Solid-state Circuits*, Vol. 39, No. 11, 1976–1984, 2004.
9. "6S06DPDM-CT02 process electrical design rule, Version♯: 3B02," CSMC Technologies Corporation.

An Injection-locked Millimeter Wave Oscillator Based on Field-emission Cathodes

Ming-Chieh Lin[1] **and Pu-Shih Lu**[1,2]

[1]NanoScience Simulation Laboratory (NSSL), Department of Physics
Fu Jen Catholic University, Taiwan, R.O.C.
[2]Institute of Optoelectronic Sciences, National Taiwan Ocean University, Taiwan, R.O.C.

Abstract— We propose an injection-locked millimeter wave generator based on field emission mechanism according to the previous successful designs of the THz and millimeter wave devices. The design of this injection-locked device has been carried out and verified by the numerical simulations. A bandwidth of about 60 MHz has been obtained and it shows the consistency with the prediction of Adler's equation. The Q factor of the resonant structure is also calculated. An electronic efficiency up to 7% can be achieved without any external magnetic field applied.

1. INTRODUCTION

The millimeter wave frequency region is part of the microwave range and is generally considered to extend from about 30 GHz to 300 GHz. Tunable high-power coherent radiation sources, have been extensively studied for many applications including high resolution radars, communications, materials processing, plasma heating, high energy particle accelerators and power transmission [1]. The most powerful sources available at low frequencies are klystrons, traveling wave tubes, magnetrons and crossed field amplifiers. For high power generation and amplification in the millimeter range, the well-known electron cyclotron masers (ECM) provide the solution [2]. Generally speaking, the weight and the volume of these devices are heavy and huge, and the external magnetic circuit is always needed.

In the previous work, we proposed a THz wave generator based on field emission cathodes [3–4]. However, the current density required to achieve oscillations of the THz device is too high to be available via today's technology. The average current density required is about 150 kA/cm^2. To the state of the art, the highest current density achievable is around 1600 A/cm^2. Such a high current density Spindt type field emission tips array has been demonstrated by MIT Lincoln Laboratory [5]. Therefore, we first adjust all corresponding parameters of the THz device, i.e., scale up the structure of the generator and scale down the frequency to Ka-band to reduce the start-oscillating current density [6]. The simulation results show that not only an electronic efficiency up to 9.16% can be obtained, but the average current density is successfully reduced down to an achievable range about 264 A/cm^2. This electromagnetic structure can be easily fabricated via the MEMS technology. These results motivate us to further study more physical characteristics, such as the spectral purity, the phase and frequency controllability, and the stability of tuning.

Controlling the phase and frequency in oscillators by means of injection locking has known for decades [7]. The idea is that this approach may be also applied to our device. In the present work, we have investigated an injection-locked millimeter wave oscillator based on field-emission cathodes by following the successful design of the millimeter wave devices. The verification of the injection-locked oscillator has been done using the 2D finite-difference time-domain (FDTD) particle in cell (PIC) code MAGIC (developed by ATK Mission Research, VA, US) [8]. A preliminary design has been carried out. The Q factors of the interaction resonant structure and the corresponding locking bandwidth have been determined.

2. PHYSICAL MODEL AND SIMULATION MODELS

2.1. Physical Model

Figure 1 shows the schematic of the injection-locked millimeter wave oscillator based on field-emission cathode we proposed. The anode consists of six coupled cavities and the cathode is some kind of field emission array (FEA). In our simulation, the effective work function of the cathode is set to be 0.2 eV. Two shaded rectangular areas indicated by "Ins." stand for insulators. A potential V_{in} is applied between the cathode and the anode. V_{in} contains a dc input voltage V_{dc} and an injected signal with a modulated frequency f. This electromagnetic structure can be easily fabricated via today's technologies such as MEMS. The material of the cavity is assumed to be the perfect conductor for facilitating the numerical simulations.

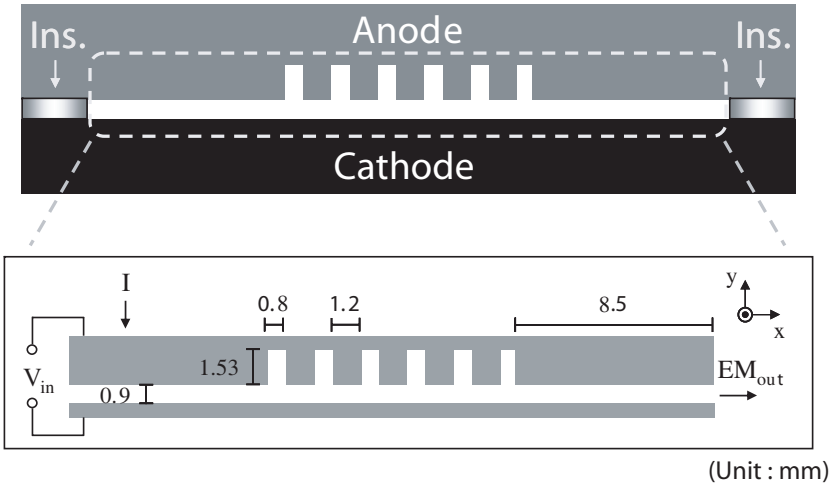


Figure 1: Schematic of the physical model. Two shaded rectangular areas indicated by "Ins." stand for insulators. The corresponding MAGIC simulation model is also shown.

2.2. Simulation Models

MAGIC is an electromagnetic particle-in-cell (EM PIC) code, based on the FDTD method for simulating plasma physics processes such as the processes that involve interactions between space charge and electromagnetic fields. Beginning from a specified initial state, the code simulates a physical process as it evolves in time. In the hot test, i.e., simulations including charged particles or electrons, the full set of time-dependent Maxwell's equations is solved to obtain electromagnetic fields. Then, the complete Lorentz force equation is solved to obtain relativistic particle trajectories, and the continuity equation is solved to provide current and charge densities for Maxwell's equations. This self-consistent approach is commonly referred to as EM PIC method, and is suitable for dealing with the interaction between charged particles and electromagnetic fields. The interaction between the charged particles is also included. In addition, the code has been provided with powerful algorithms to represent structural geometries, material properties, incoming and outgoing waves, particle emission process, real time observations of all physical quantities, and so forth. The real time observed data facilitate the post-processing for further understanding the physics inside.

The MAGIC code is employed for verifying the designs of the generation of millimeter waves with the electron-wave interactions and to do the hot tests of the field emission based millimeter wave generator. Therefore, the 2D MAGIC model for the design also shown in Figure 1 has been constructed for the hot tests.

2.3. Field Emission

The field emission is described by the Fowler-Nordheim Equation [9–20],

$$J = \frac{AE_s^2}{\phi_{eff}t(y)^2}\exp(\frac{-Bv(y)\phi_{eff}^{3/2}}{E_s}), \tag{1}$$

where A and B are the Fowler-Nordheim constants, and ϕ_{eff} is the effective work function assumed to be a constant allowed dependence on material and surface roughness [20–22]. One should note that the "effective work function" here is different from the work function as usually defined. The effective work function can be affected by the local electric fields [21–22]. The normal electric field at the cathode surface, E_s, is computed from the application of Gauss's law to the half-cell immediately above the emitting surface, or

$$\int \vec{E}\cdot d\vec{a} = q/\varepsilon_0, \tag{2}$$

$$E_cA_c + (-E_sA_s) = q/\varepsilon_0, \tag{3}$$

and

$$E_s = (E_cA_c - q/\varepsilon_0)/A_s, \tag{4}$$

where E_c is the electric field at the half-grid, A_c and A_s are the cell areas at half-grid and surface, respectively, and q is the existing charge in the half-cell. The functions $t(y)$ and $v(y)$ were Nordheim elliptical functions [18–19] and approximated by Spindt et al. [23] as below:

$$t(y)^2 = 1.1, \tag{5}$$

$$v(y) = 0.95 - y^2, \tag{6}$$

where $y = 3.79 \times 10^{-5} E_s^{1/2}/\phi_{eff}$ is the Nordheim parameter.

2.4. Adler's Equation

The relationship between the strength of the locking signal and the achievable bandwidth can be described and estimated by Adler's equation [7]:

$$\frac{\Delta f}{f_0} \approx \frac{1}{Q_e}\sqrt{\frac{P_{inj}}{P_{out}}} \tag{7}$$

where:

Δf = bandwidth, total frequency range over which locking will occur,

f_0 = resonant frequency of the free-running case,

Q_e = external quality factor (Q) of the millimeter wave oscillator,

P_{inj} = the power of the locking signal, and

P_{out} = output power.

3. SIMULATION RESULTS AND DISCUSSION

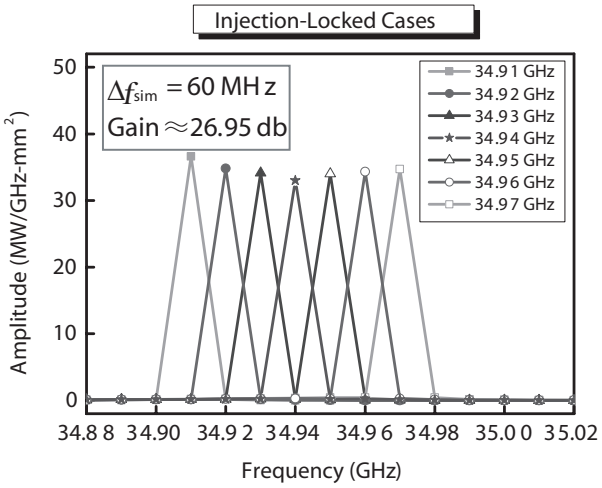


Figure 2: Power spectrum of successfully locked cases with injected frequency from 34.91 to 34.97 GHz.

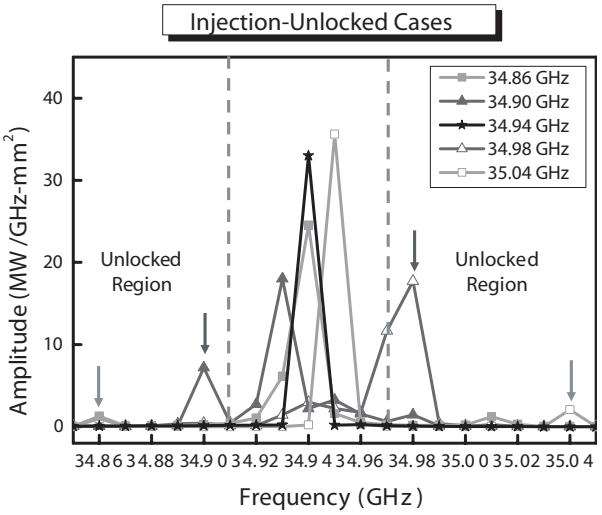


Figure 3: Power spectrum of failing to lock cases with injected frequency away from the central dashed region.

To investigate the injection locking characteristics of the device, the input voltage was set as $V_{in} = V_{dc} + V_{inj}{}^* \sin(2\pi ft)$. For the free oscillating operation, $V_{in} = V_{dc} = 16.75$ kv and $V_{inj} = 0$. In order to distinguish locked and unlocked states of operation, the fast Fourier transform of the output signal calculated up to $t_{max} = 100$ ns is analyzed. The frequency resolution is $df = 1/t_{max} = 10$ MHz. After doing Fourier transform of $V_d(t)$, we obtain the free oscillating frequency 34.94 GHz for the oscillator.

For the cases of injection locking, we increase the power of the injection signal by increasing the voltage V_{inj} up to 500 V at variant locking frequency from 34.86 GHz to 35.04 GHz with a step of 10 MHz. Figures 2 and 3 show the results of the successfully locked cases and fail to lock cases,

respectively. The bandwidth estimated is 60 MHz. To calculate the Q factor, we input a delta-function-like signal and measure the transmitted power. After doing the Fourier transform of the power, the Q factor can be determined to be about 35.48. According to the prediction of Adler's equation, the theoretical value of the bandwidth is about 80 MHz and our simulation is close to and consistent with the conventional prediction. The efficiencies of the injection-locked oscillator for each injection locked cases with different frequencies are around 7%.

4. CONCLUSIONS

In this paper, we propose an injection-locked millimeter wave oscillator based on field-emission cathodes and demonstrate the well phase/frequency controllability. The simulation results are consistent with the prediction of Adler's equation. An electronic efficiency up to 7% can be achieved with no external magnetic field applied. The preliminary results show that our design of the injection-locked oscillator seems to promise useful and compact Ka-band sources.

ACKNOWLEDGMENT

The work was supported by the National Science Council, Taiwan, R. O. C., under Grants NSC 96-2918-I-030-005 and NSC 96-2112-M-030-004-MY3, the National Center for Theoretical Sciences, and the National Center for High-performance Computing, Taiwan, R. O. C. which provides the MAGIC code and the computing resource.

REFERENCES

1. Gilmour, S., Jr., *Microwave Tubes*, Artech House, Norwood, MA, 1986.
2. Chu, K. R., "The electron cyclotron maser," *Rev. Mod. Phys.*, Vol. 76, No. 2, 489–541, 2004.
3. Lin, M. C., K. H. Huang, P. S. Lu, P. Y. Lin, and R. F. Jao, "Field-emission based vacuum device for the generation of THz waves," *J. Vac. Sci. Technol. B*, Vol. 23, No. 2, 849–852, 2005.
4. Lin, M. C. and P. S. Lu, "Interaction mechanism of a THz wave generator using a field emission cathode," *J. Vac. Sci. Technol. B*, Vol. 25, No. 2, 631–635, 2007.
5. Bozler, C. O., et al., "Arrays of gated field-emitter cones having 0.32 μm tip-to-tip spacing," *J. Vac. Sci. Technol. B*, Vol. 12, No. 2, 629–632, 1994.
6. Lin, M. C. and P. S. Lu, "Millimeter wave generator based on field emission cathode," *J. Vac. Sci. Technol. B*, Vol. 25, No. 2, 636–639, 2007.
7. Adler, R., "A study of locking phenomena in oscillators," *Proc. IRE*, Vol. 34, 351–357, 1946.
8. Goplen, B., L. Ludeking, D. Smithe, and G. Warren, "User-configurable MAGIC code for electromagnetic PIC calculations," *Computer Phys. Comm.*, Vol. 87, No. 1, 54–86, 1995.
9. Goplen, B., L. Ludeking, D. Smithe, and G. Warren, *MAGIC User's Manual*, Mission Research Corp., Newington, VA, MRC/WDC-R-409, 1997.
10. Fowler, R. H. and L. W. Nordheim, *Proc. R. Soc., Ser. A*, Vol. 119, 173, London, 1929.
11. Nordheim, L. W., *Proc. R. Soc., Ser. A*, Vol. 121, 626, London, 1928.
12. Nordheim, L. W., *Z. Phys.*, Vol. 30, 177, 1929.
13. Schottky, W., *Z. Phys.*, Vol. 14, 63, 1923.
14. Frank, N. H. and L. A. Young, "Transmission of electrons through potential barriers," *Phys. Rev.*, Vol. 38, 80–86, 1931.
15. Guth, E. and C. J. Mullin, "Electron emission of metals in electric fields III. The transition from thermionic to cold emission," *Phys. Rev.*, Vol. 61, 339–348, 1942.
16. Gogate, D. V. and D. S. Kothari, "Flow of energy in thermal transpiration for a bose-einstein and a fermi-dirac gas," *Phys. Rev.*, Vol. 61, 349–358, 1942.
17. Gasiorowicz, S., *Quantum Physics*, 2nd ed., Wiley, New York, 1996.
18. Murphy, E. L. and R. H. Good, "Thermionic emission, field emission, and the transition region," *Phys. Rev.*, Vol. 102, 1464–1473,1956.
19. Forbes, R. G., "Simple good approximations for the special elliptic functions in standard fowler-nordheim tunneling theory for a schottky-nordheim barrier," *Appl. Phys. Lett.*, Vol. 89, 113122, 2006.
20. Lin, M. C. and D. S. Chuu, " Quasistationary states of a relativistic field-emission-limited diode employing a high-transparency mesh anode," *Appl. Phys. Lett.*, Vol. 80, No. 22, 4262–4264, 2002.

21. Kim, C., B. Kim, S. M. Lee, C. Jo, and Y. H. Lee, "Electronic structures of capped carbon nanotubes under electric fields," *Phys. Rev. B*, Vol. 65, 165418, 2002.
22. Nation, J. A., et al., *Proceedings of the IEEE*, Vol. 87, No. 5, 865, 1999.
23. Spindt, C. A., I. Brodie, L. Humphrey, and E. R. Westerberg, "Physical properties of thin-film emission cathodes with molybdenum cones," *J. Appl. Phys.*, Vol. 47, 5248–5263, 1976.

A New Tunable Wideband Ring Filter with Merged Stubs and Miniaturized Geometry for Bluetooth Technology

M. S. Kheir[1] **and A. M. Abdin**[2]

[1]Department of Information Engineering, German University in Cairo, Cairo, Egypt
[2]Department of Communications Engineering, Shorouk Academy, Cairo, Egypt

Abstract— A new miniaturized bandpass filter geometry is hereby presented. The structure of this filter is based on microstrip annular ring resonators with a wide bandwidth and precise tuning capability. The center frequency of the filter is 2.45 GHz and the 3-dB Fractional Bandwidth (FBW) is 105% with a sharp out-of-band rejection. The filter geometry is based on a new technique that is to merge the stubs which allows considerable size reduction maintaining the same efficiency. The size of the designed filter, excluding the I/O wave ports, is $7 \times 7\,\mathrm{mm}^2$. Such features make it best suitable for Bluetooth communication applications. The validity of the design has been verified using two dissimilar electromagnetic full-wave numerical solvers. They are ADS-Momentum as a Method of Moments (MoM) and Ansoft-HFSS as a Finite Elements Method (FEM). There results have shown a very good agreement.

1. INTRODUCTION

Bluetooth is one of the promising technologies as a short-range connectivity solution for personal, portable, and hand-held devices [1]. In the same time, ring resonators are highly desired for integration purposes due to their miniature size and simple design methodology. However, microstrip structures usually suffer from the narrow bandwidth and poor out-of-band rejection. Few attempts to achieve wideband response using microstrip ring resonators have been reported in [2–4]. This article discusses a new tunable ring filter topology as well as investigating the possibility of miniaturizing its overall size and offering a wide bandwidth. The size of the new filter, excluding the I/O wave ports, is only $7 \times 7\,\mathrm{mm}^2$ instead of $10 \times 10\,\mathrm{mm}^2$ which is implemented beforehand in [5]. The initial and the modified filter geometries are shown in Figures 1 and 2 respectively. The new filter is designed on RT/Duroid 6010 substrate material with a dielectric constant ($\varepsilon_r = 10.2$) and a thickness ($h = 0.635\,\mathrm{mm}$). The filter geometry was originally based on the design discussed in [2] with some modifications on the resonator topology to reduce the size. The idea is simply to introduce via-holes at the edges of the short-circuited stubs as indicated in Section 2. Such technique can successfully increase the fractional bandwidth (FBW) up to 105% and provide a reasonably good insertion loss at the edges of the band. However, this design could be furtherly miniaturized by utilizing a new technique as will be declared in Section 3. Several investigations on the tuning parameters of the filter have been gone through and discussed in Section 4. The investigations show that the most significant parameters that affect the tuning of the filter are the stub-length and the via-hole radius.

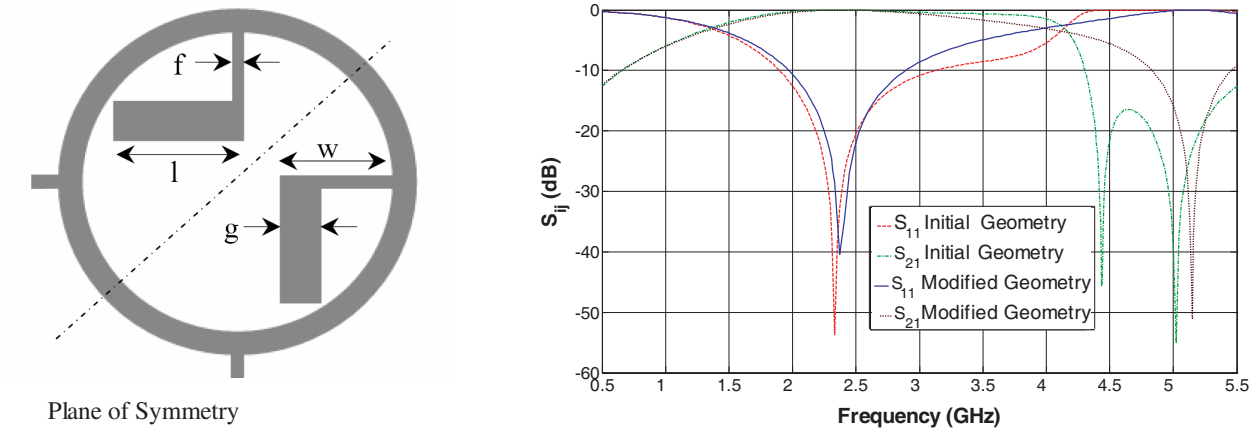
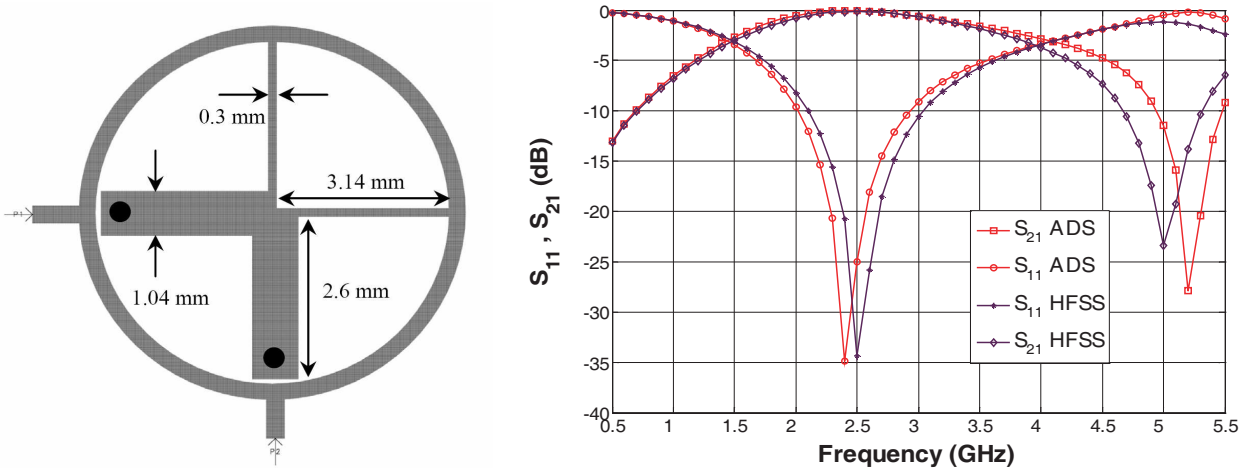


Figure 1: (a) The geometry of the filter discussed in [5], (b) S-parameters of both filters.

2. FILTER GEOMETRY

The design parameters and dimensions of the initial filter are discussed in [5]. The dimensions of filter in Figure 1 are as follows: the input port which can be of any length, $f = 0.3\,\text{mm}$, $w = 2.8\,\text{mm}$, $l = 3.25\,\text{mm}$ and $g = 1\,\text{mm}$. On the other hand, the dimensions of the modified filter are shown in Figure 2.



(a) (b)

Figure 2: (a) The geometry of the modified filter, (b) The simulation results.

A comparison between both of them is shown in Table 1. The size of the modified one is reduced by more than 50% and on the other hand the FBW is also decreased by 15%. The S-parameters of the two filters are shown in Figure 1(b). The return loss of the modified filter and the second one are $-40\,\text{dB}$ and $-54\,\text{dB}$ respectively. On the other hand the band edge frequency of the modified is 4.4 GHz while for the other one is 5.2 GHz.

Table 1: A comparison between the proposed filter and that one presented in [5].

Item	The proposed filter	The filter presented in [5]
Size (mm^2)	7×7	10×10
Return loss (dB)	-40	-54
FBW (%)	105	120
Tuning parameters	Stub-length and via-hole radius	Stub-length and via-hole radius

3. THE MODIFIED FILTER

The geometry of the modified structure is shown in Figure 2(a). The reduction technique is simple and based on merging the tuning stubs which consequently results in a reduced size with wide bandwidth. The modified geometry is a 49% of the original design. The proposed filter is simulated using ADS-Momentum as a Method of Moments (MoM) based solver and Ansoft-HFSS as a Finite Elements Method (FEM) based solver. The results are shown in Figure 2(b), where a very good agreement between both of them is observed.

4. TUNING PARAMETERS

Several studies on tuning parameters of the filter have been investigated. According to the achieved results it was clear that the most significant parameters that affect the capability of tuning are the stub-length and via-hole radius. The effect of tuning parameters are shown in Figure 3. Both the center frequency and the FBW increase as the length of the stub decreases, as shown in Figure 3(a). On the other hand, the center frequency as well as the bandwidth of the filter can be controlled by

the radius of both via-holes. Both the center frequency and the FBW increase as the radius of both via-holes increases, as depicted in Figure 3(b). It could be benefited for selecting the dimensions according the required frequency and the suitable FBW.

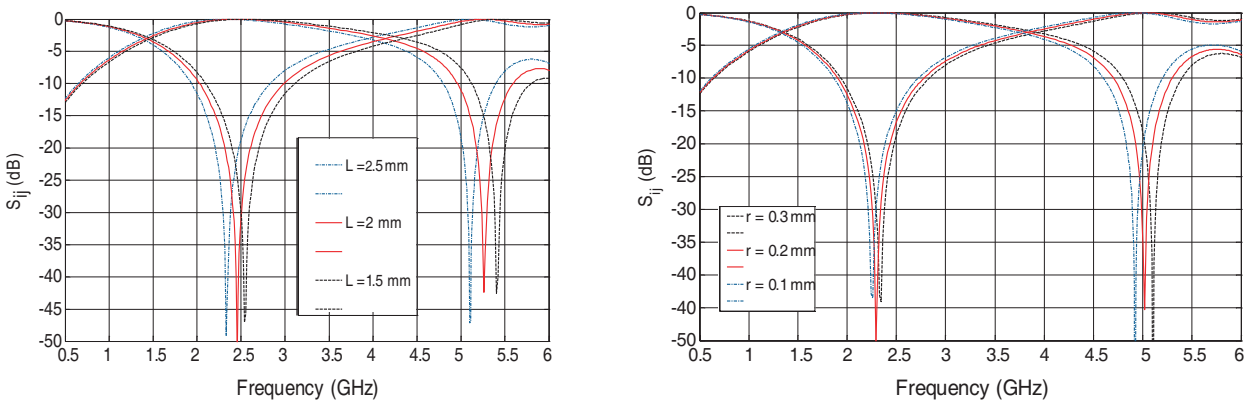


Figure 3: The effect of tuning parameters of the filter: (a) the effect of the via-hole radius. (b) the effect of the stub length.

5. CONCLUSIONS

A miniature bandpass filter design based on annular ring resonator structure has been introduced. The overall size and the wideband characteristics of the proposed filter made it suitable for wireless communication systems which require high bit-rate connectivity such as Bluetooth. According to the achieved results, it was concluded that the most significant parameters that affect the tuning of the filter are the stub-length and the via-hole radius. The results obtained from the two dissimilar numerical solving techniques have shown a very good agreement and verified the validity of the design.

REFERENCES

1. Chatschik, B., "Overview of the Bluetooth wireless technology," *IEEE Communications Magazine*, Vol. 39, No. 12, 86–94, 2001.
2. Hsu, C-Y., C-Y. Chen, and C-H. Huang, "UWB filter using a dual-mode ring resonator with spurious passband suppression," *Microwave Journal*, Vol. 48, No. 11, 2005.
3. Ishida, H. and K. Araki, "A design of tunable UWB filters," *Proc. Int. Workshop Ultra Wideband Systems*, 424–428, 2004.
4. L. H. Hsieh and K. Chang, "Compact, low insertion-loss, sharp-rejection, and wide-band microstrip bandpass filters," *IEEE Transactions on Microwave Theory Tech.*, Vol. 51, No. 4, 1241–1246, 2003.
5. Kheir, M. S., A. M. Abdin, and E. A. Hashish, "An extremely compact UWB microstrip annular ring filter for Bluetooth applications," *Proceedings of Innovations in Information Technology (IIT 2006)*, Dubai-UAE, 1–3, 2006.

A Circular Polarization Microstrip Stacked Structure Broadband Antenna

Huan-Cheng Lien, **Huei-Chiou Tsai**, **Yung-Cheng Lee**, **and Wen-Fei Lee**
Department of Electrical Engineering, Wufeng Institute of Technology
No. 117, Chian-Ku Rd., Sec. 2, Ming-Hsiung (621), Chiayi, Taiwan, R.O.C.

Abstract— This paper describes a single-feed S-Ring type stacked microstrip structure antenna for broadband applications. The antenna is designed for RHCP/LHCP at a center frequency of 2.1 GHz. The design of the antenna is aimed at obtaining both wider bandwidth on the impedance and better circular polarization AR (Axial-Ratio) for GPS (global positioning system) applications. The feeding technique and structures of the present antenna has been analyzed. The results of simulations and measurements results show that the antenna has an impedance bandwidth (VSWR.≦2) of 30% and a 3-dB AR bandwidth more than 23%.

1. INTRODUCTION

In many areas of wireless communications, both reducing the number and the miniaturization of antennas have been strongly required. Microstrip patch antennas are used in a variety of applications due to their some salient features [1].

The designs of conventional single-feed circular polarized microstrip antenna are available either of a square or circular patch on single planar. This paper describes a method by using a stacked structure for designing a circularly polarized antenna with single-feed, and a different impedance transformer method that further reduces the size of the patch.

The majority of current and future commercial and military applications typically use a circularly polarized antenna. A single-feed circularly polarized microstrip antenna allows a reduction in the complexity, weight, and have the additional advantage of small size and produce a completely planar antenna. Single-feed circularly polarized antennas are currently receiving much attention.

In general, the microstrip antenna has a narrow 3-dB AR bandwidth. Increased bandwidth can be achieved with a single patch [2, 3] or by using two (or more) stacked patches [4, 5], according to above-mentioned methods impedance bandwidths ranging from 20 to 35% have been obtained; however, it is unable to obviously increase the AR bandwidth.

The tradition the impedance matches feed networks requires with quarter-wave transformer or hybrid circuit. However, the introduction of the hybrid circuit requires more places and complicates the feeding structure. Further, this kind of antenna feed networks structure, if frequency deviation from the center frequency, the 90° phase difference can no longer be maintained to result in elliptical polarization even in the bore-sight direction.

In the practical design of a circularly polarization (CP) antennas, techniques for achieving wideband polarized characteristics, as well as impedance wideband and 3-dB AR bandwidths characteristics are important. For this reason, we design a different from traditional impedance transformer and employing stacked structure in this paper to overcome above perplexed problem at the same time.

2. ANTENNA STRUCTURE

By making use of stacked patch concept [6] and with two different (high/low) permittivity substrates [7], the AR bandwidth and the impedance bandwidth can be enhanced.

The proposed antenna was made up of two different permittivity substrates, microstrip feed line, impedance matching transformer and an annular ring radiate element. It is designed for a center frequency of 2.1 GHz for the TM_{10} mode [8]. The final constructed prototype is shown in Figure 1.

The antenna structure and the substrate material mainly affect antenna properties such as radiation efficiency, pattern and directivity etc. In the present configuration of antenna, we employ a lower permittivity substrate in the top layer for the radiate element and circularly signal couple element; besides, using a lower permittivity dielectric substrate situated in the second layer for the rectangular feed line. Finally, using a high permittivity dielectric substrate situated in the first layer for the ground plane. All of we mention antecedently, the high permittivity is FR4 substrate and

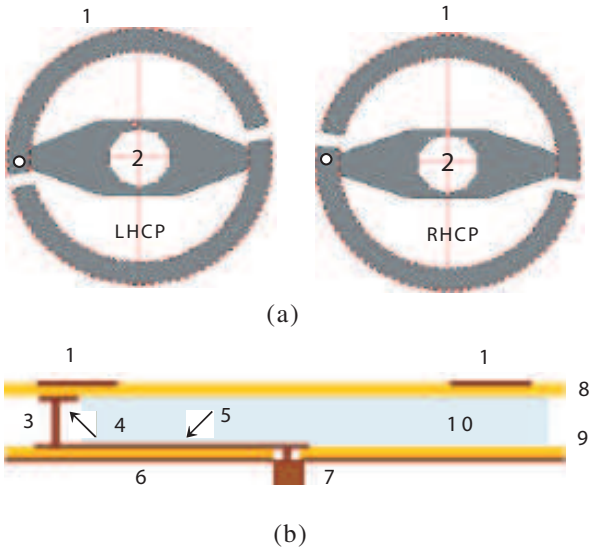


Figure 1: Antenna configuration of proposed, (a) Top view, (b) Side view.

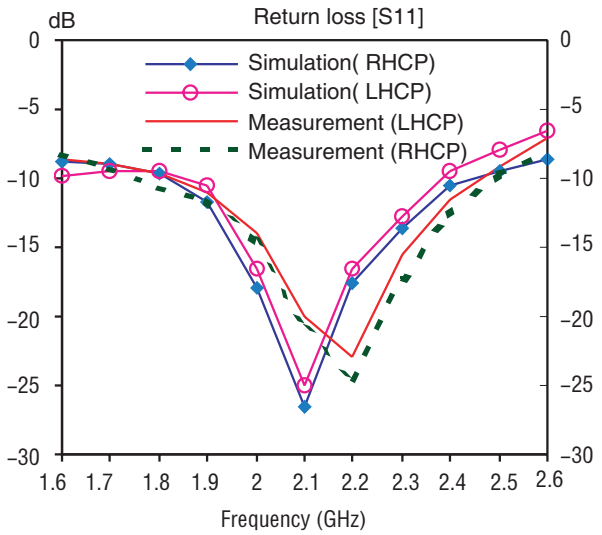


Figure 2: Return loss vs. frequence.

Table 1: The relative parameters of the proposed antennas (unit: mm).

1	S-Ring radiate element	$r_i = 28$, $r_o = 34$
2	Slot	$r = 6$
3	Cylindrical conductor	$h = 14.4$, $r = 0.8$
4	Circularparasitically element	$r = 2$
5	Rectangular microstrip feed line	$L = 16.36$, $W = 3$
6	Ground plane	60×60
7	SMA connector	
8	RO substrate	$\varepsilon_{r2} = 3.38$, tand $= 0.0025$, $d = 0.508$
9	FR4 substrate	$\varepsilon_{r1} = 4.4$, tand $= 0.0022$, $d = 1.6$
10	Polymer	$\varepsilon_r = \varepsilon_o = 1$, $h = 2$

the low permittivity is RO substrate, while the relevant parameters are list in Table 1, respectively.

The CP is obtained by choosing the suitable size of the radiating element to excit with equal amplitudes and a 90° phase shift. The right or left hand circularly polarization depends on choosing the feeding point positions. In the propose structure by cutting S-Ring component, Right Hand or Left-Hand Circular Polarized (RHCP/LHCP) is achieved, respectively. By adjusting the length of the cylindrical conductor from the radiating element to the feed line, the proposed antenna can achieve an important bandwidth broadening.

3. SIMULATION AND EXPERIMENTAL RESULTS

The simulated and measured return losses (−S11 in dB) of the proposed antenna are shown in Figure 2. The measurements show that the antennas are very well matched to the 50 Ω feed lines and the optimum impedance bandwidth of 30% with the −10.0 dB return loss has been obtained when the S-Ring gap is 8.5 mm ($\sim 0.07\lambda_o$). Good agreement between measurement and simulation could be noticed, confirming the wide frequency behavior of the new SRA feeding topology.

The AR result obtained from simulation and experiment are shown together in Figure 3 as a comparison. The experimental results AR value is less than 3 dB in excess of 23% bandwidth, which

is approximate simulation purpose, while the CPG (circular polarization gain) showed in Figure 4 is measured over 5 dBi across a frequency band between 1.8 to 2.5 GHz. Figure 5 depicts the far field patterns of the antenna.

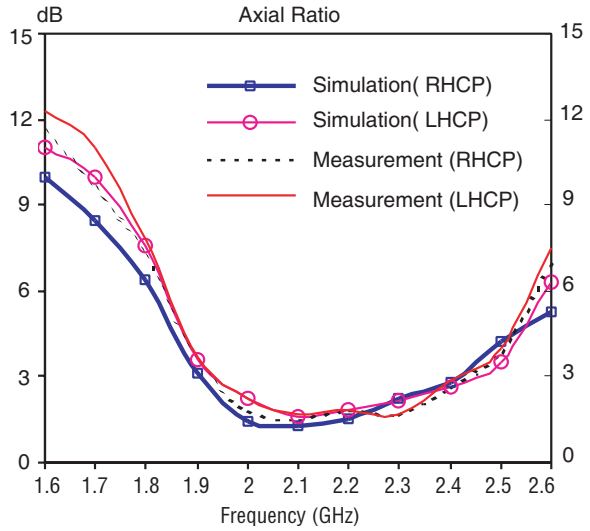


Figure 3: Axial vs. frequence.

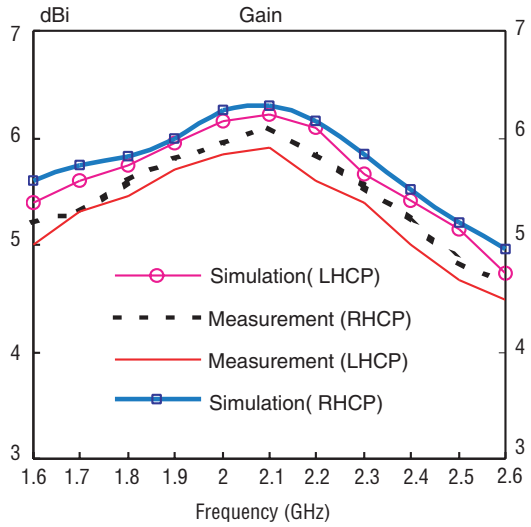


Figure 4: Gain vs. frequence.

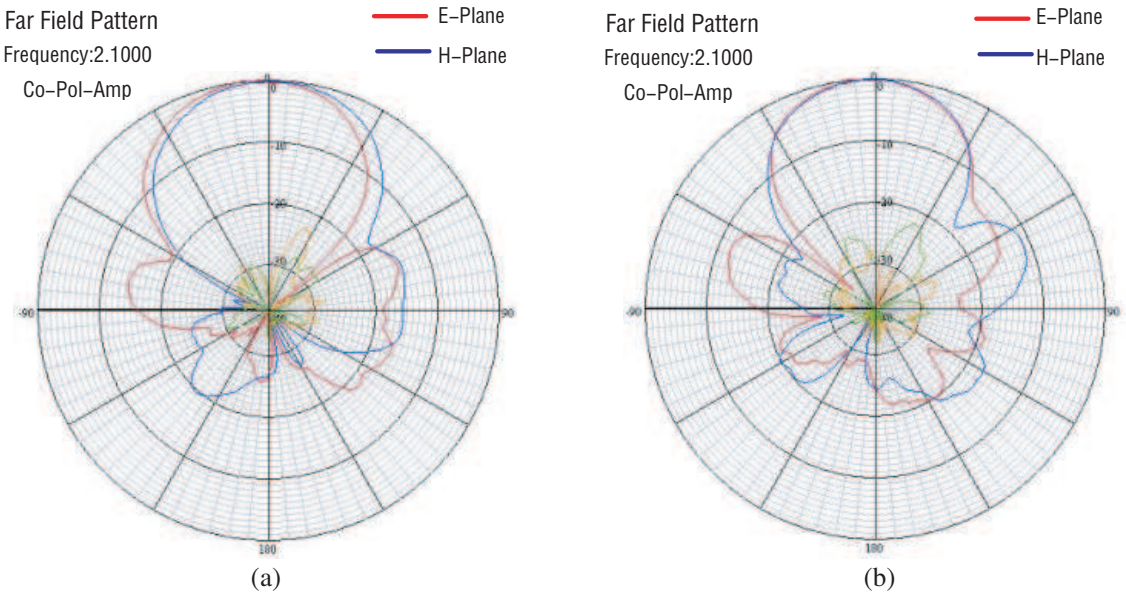


Figure 5: Depicts the far field patterns of the antenna, (a) RHCP, (b) LHCP.

4. CONCLUSIONS

We describe a method for designing circularly polarized stacked structure antennas with a new feeding topology. The excitation was done at a point from the back of the antenna substrate.

The propose impedance transformer permits a better purity of the CP and a wider impedance bandwidth with respect to conventional antennas and the proposed antenna is suitable for implementing low cost, high stable and well circular polarized GPS antenna.

The proposed antennas design is not only limited to the improvement on the impedance and AR bandwidths of the conventional but also the radiation characteristics such as CP gain.

REFERENCES

1. Pozar, D. M., "Microstrip antennas," *Proc. IEEE*, Vol. 80, 79–91, January 1992.

2. Targonski, S. D. and D. M. Pozar, "Design of wideband circularly polarized aperture coupled microstrip antennas," *IEEE Trans. Antennas Propagation*, Vol. 41, 214–220, February 1993.
3. Croq, F. and A. Papiernik, "Wideband aperture coupled microstrip subarray," *Electronics Letter*, Vol. 26, 1293–1294, August 1990.
4. Wang, J., R. Fralich, C. Wu, and J. Litva, "Multifunctional aperture coupled stack antenna," *Electronics Letter*, Vol. 26, 2067–2068, December 1990.
5. Croq, F. and D. M. Pozar, "Millimeter wave design of wide-band aperture coupled stacked microstrip antennas," *IEEE Trans. Antennas Propagation*, Vol. 39, 1770–1776, December 1991.
6. Chen, C. H., A. Tulintseff, and R. M. Sorbello, "Broadband two-layer microstrip antenna," *IEEE AP-S International Symposium Digest*, 251–254, 1984.
7. Waterhouse, R. B., "Stacked patches using high and low dielectric constant material combinations," *IEEE Trans. Antennas and Propagation*, Vol. 47, 1767–1771, December 1999.
8. Tsai, M. J. and N. G. Alexopoulos, "Electromagnetically coupled microstrip ring-type antennas of arbitrary shape," *IEEE AP-S Int. Symposium*, Vol. 1, 684–687, June 1995.

A Wide-band Circular Polarization Stacked Patch Antenna for the Wireless Communication Applications

Huan-Cheng Lien and Huei-Chiou Tsai
Department of Electrical Engineering, Wufeng Institute of Technology
No. 117, Chian-Ku Rd., Sec. 2, Ming-Hsiung 621, Chiayi, Taiwan, R.O.C.

Abstract— A wide-band 'corners-truncated rectangular' stacked patch antenna for use in the circular polarization applications was proposed. The antenna proposed in this paper an axial ratio of less than 3 dB and a VSWR of less than 2 : 1 were shown to be achievable over a 25% bandwidth for use in the wireless communication applications, and this antenna can achieves higher gain, lower side lobes and wider bandwidth compared to the traditional microstrip patch antenna.

1. INTRODUCTION

The most serious limitation of the circularly polarized (CP) microstrip antennas is its narrow bandwidth [1], to overcome its inherent limitation of narrow impedance and axial ratio (AR) bandwidth, many techniques have been research and development for the enhancement of microstrip antenna bandwidth e.g., for probe fed stacked antenna, slotted patch antenna, microstrip patch antennas on electrically thick substrate [2–7].

Increased bandwidth can be achieved with [8, 9] or [10, 11]. The AR bandwidth can be enhanced by [12, 13] concept. Hence, the present design of propose antenna is choice of a high permittivity substrate to situate the bottom layer for the feed line element and a lower permittivity substrate on the top layer for the radiate while an adoption of a higher air-spaced. The objective of the proposed design is to generate a CP and to improve the bandwidth of the patch antenna.

2. ANTENNA DESIGN

The proposed antenna configuration is shown in Figure 1, that was designed consists of four layers where first layer is ground plane, second layer is feed line, and the third and the fourth layer are radiate and a parasitic patch, respectively.

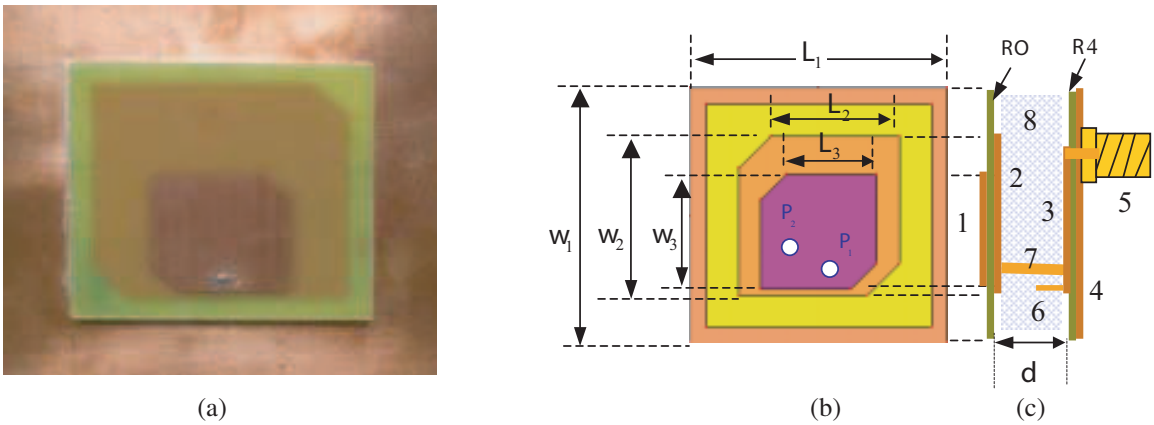


Figure 1: Geometry of the proposed antenna: (a) snapshot of proposed antenna, (b) top view, (c) side view.

By making use of stacked patch concept one can obtain enhanced impedance and using high permittivity can enhance the AR bandwidth. All of we mention antecedently, the stacked layers are FR4 and RO dielectric substrate.

In the present configuration of antenna we employ a lower permittivity of RO dielectric substrate in the top layer for the radiate element and a coupled parasitically patch, using a R4 dielectric substrate situated in the bottom layer for the feed line and the ground plane.

A less than quarter-wavelength impedance transformer, is placed at the back of the radiate patch, that is consists of two structures of the rectangular feed line and a cylindrical conductor, a

tuning of stub wire is place at the rectangular feed line end, one can improve matching capabilities by properly adjusting the stub. The signal is fed at back of bottom substrate to the impedance transformer, while the bottom side is being a ground plane for the antenna.

By properly adjusting above the mentions sizes (including the coupled parasitically patch, corners-truncated rectangular radiate element, rectangular feed line, cylindrical conductor, stub wire), then we can obtain a better and wider bandwidth of CP antenna. Figure 1 depicts the detailed antenna structure of the propose antenna, the values of various parameters involved in Figure 1 are given in Table 1.

Table 1: The relative parameters of the proposed antenna (UNIT: mm).

1	Coupled parasitically patch	$W_3 = L_3 = 13.28$
2	Rectangular radiate element	$W_2 = L_2 = 17.7$
3	Rectangular microstrip feed line	$W = 3,\ L = 4.12$
4	Ground plane	$W_1 = L_1 = 30$
5	SMA connector	
6	Stub wire	Radius = 1.6
7	Cylindrical conductor	Length = 10, Radius = 1.6
8	Polymer	$60 \times 60 \times 10$
FR4	Dielectric substrate	$\varepsilon_{r1} = 4.4,\ \tan\delta = 0.022$
RO	Dielectric substrate	$\varepsilon_{r2} = 3.38,\ \tan\delta = 0.0025$
$P_1,\ P_2$	Feed position	$(x_1, y_1, z_1) = (13, 8, 11.6)(x_2, y_2, z_2) = (17, 11, 11.6)$

3. SIMULATION RESULTS

Referring to the configuration shown in Figure 1, the right or left hand circularly polarization depends on choosing the feeding point positions, by changing the signal fed positions P_1, P_2, then the right or left hand CP can easily achieved. The proposed antenna is designed for Right-Hand CP (RHCP) radiation. The proposed antenna was designed and simulated at a center frequency of 2.3 GHz with the aid of IE3D software, which is based on the method of moments [14].

Figure 2 shows that the simulated and measured return losses (−S11 in dB) of the proposed

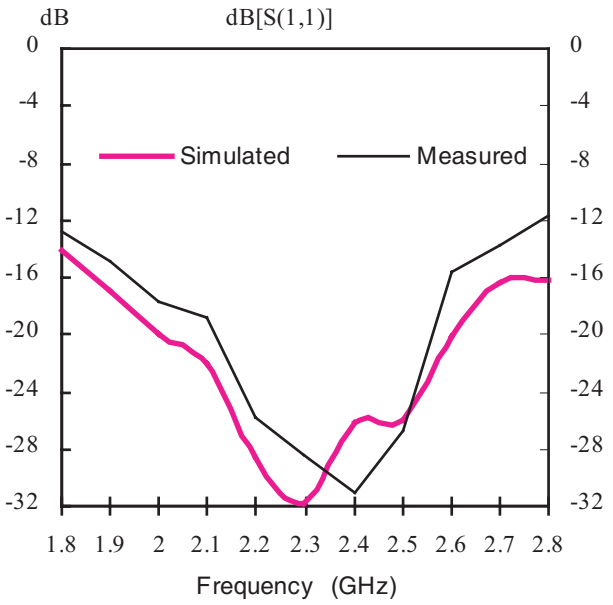


Figure 2: Measured and simulated return losses of the proposed antenna.

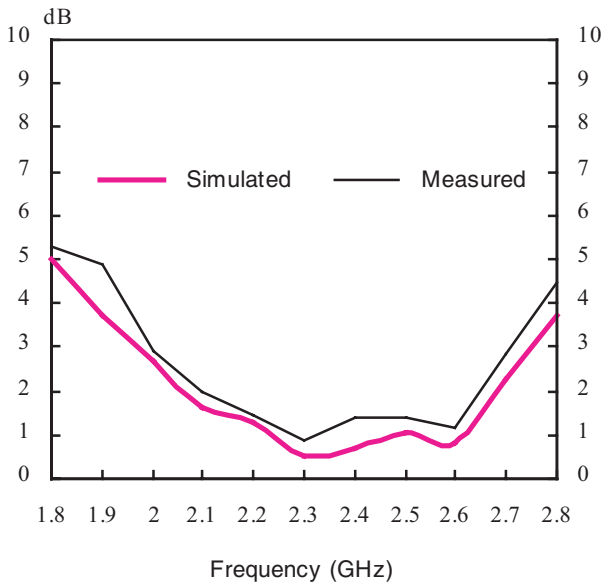


Figure 3: Measured and simulated axial ratios of the proposed antenna.

antenna are less than −20 dB over 2.0 GHz to 2.6 GHz. The measurements show that the antennas are very well matched to the impedance transformer and the optimum impedance bandwidth more than 30% with the −10.0 dB return loss has been obtained, when cutting corners-truncated rectangular is 8.5 mm ($\sim 0.07\lambda_o$). There is a frequency shift of about 100 MHz for measured return loss with respect to simulated results, which may be mainly caused by the fabrication tolerance as well as the possible uncertainty of in-house antenna assembly.

The measured and simulated 3 dB AR shown in Figure 3 covers the range of 2.0–2.7 GHz and the AR is lower than 2.5 dB across 2.0–2.65 GHz is in excess of 28% bandwidth. Figure 4 depicts that the measured CP gain is more than 5 dBi over 1.8 GHz to 2.6 GHz. The measured far field radiation patterns in two orthogonal planes (the x-z and y-z planes) at 2.3 GHz for the proposed antenna are shown in Figure 5, the symmetry and wide angular radiation patterns are observed.

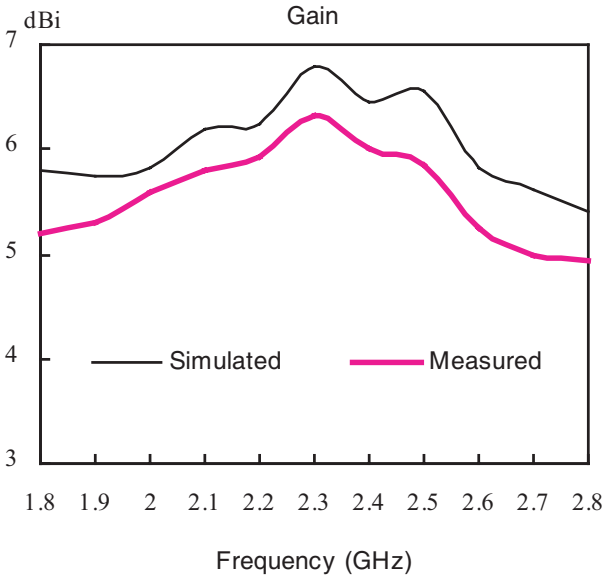


Figure 4: Measured and simulated gains of the proposed antenna.

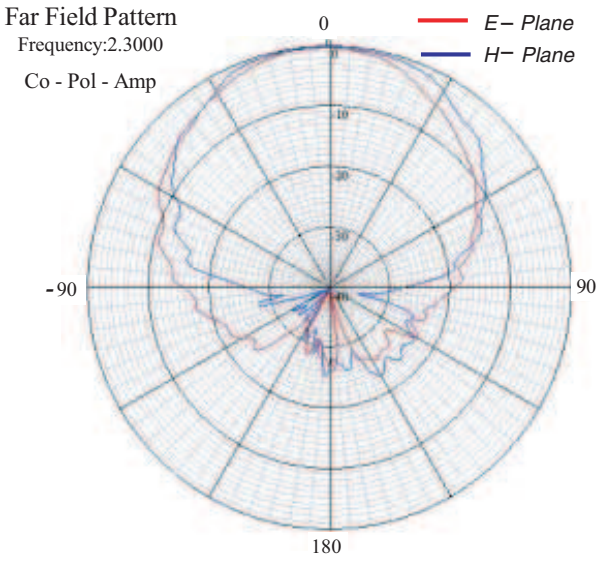


Figure 5: Measured radiation patterns of the proposed antenna.

4. CONCLUSIONS

In this paper, a broadband CP stacked probed-fed patch antenna has been made and measured. The measurement result has showed that good impedance matching and AR can be obtained with this method at the same time. The proposed antenna permits a better purity of the CP and a significant improvement impedance bandwidth with respect to conventional antennas. The structure is easily fabricated, and the effects of the locations and dimensions of the proposed antenna are very limited. The information presented in this paper will be suitable design and optimize the antenna for engineering applications.

REFERENCES

1. Pozar, D. M., "Microstrip antenna," *IEEE Proceedings*, Vol. 80, 79–81, Jan. 1992.
2. Herscovici, N., "A wide band single layer patch antenna," *IEEE Trans. on Antenna and Propagations*, Vol. 46, No. 4, 471–474, April 1998.
3. Griffin, J. M. and J. R. Forrest, "Broadband circular disc microstrip antenna," *Electronics Letter*, Vol. 18, 266–269, 1982.
4. Chen, Z. N. and Y. W. M. Chia, "Broadband probe fed L-shaped plate antenna," *Microwave and Optical Technology Letters*, Vol. 26, No. 3, 204–206, Aug. 2002.
5. Guo, Y. X., K. M. Luk, K. F. Lee, and Y. L. Chow, "Double U-slot rectangular patch antenna," *Electronics Letter*, Vol. 34, No. 19, 1805–1806, 1998.
6. Sze, J. Y. and K. L. Wong, "Broadband rectangular microstrip antenna with of toothbrush shaped slots," *Electronics Letter*, Vol. 34, No. 9230, 2186–2187, 1998.

7. Chen, Z. N., "Broadband probe-fed L-shaped plate antenna," *Microwave and Optical Letters*, Aug. 2000.
8. Croq, F. and A. Papiernik, "Wideband aperture coupled microstrip subarray," *Electronics Letter*, Vol. 26, 1293–1294, Aug. 1990.
9. Targonski, S. D. and D. M. Pozar, "Design of wideband CP aperture coupled microstrip antennas," *IEEE Trans. Antennas Propagation*, Vol. 41, 214–220, Feb. 1993.
10. Wang, J., R. Fralich, C. Wu, and J. Litva, "Multifunctional aperture coupled stack antenna," *Electronics Letter*, Vol. 26, 2067–2068, Dec. 1990.
11. Croq, F. and D. M. Pozar, "Millimeter wave design of wide-band aperture coupled stacked microstrip antennas," *IEEE Trans. Antennas Propagation*, Vol. 39, 1770–1776, Dec. 1991.
12. Waterhouse, R. B., "Stacked patches using high and low dielectric constant material combinations," *IEEE Trans. Antennas and Propagation*, Vol. 47, 1767–1771, Dec. 1999.
13. Chen, C. H., A. Tulintseff, and R. M. Sorbello, "Broadband two-layer microstrip antenna," *IEEE AP-S International Symposium Digest*, 251–254, 1984.
14. IE3D Software Release 8 developed by M/S Zeland Software, Inc.

Reflector Antenna with Artificial Magnetic Conductor Structure

Jwo-Shiun Sun[1], **Guan-Yu Chen**[1], **Cheng-Hung Lin**[2], **Kwong-Kau Tiong**[2], **and Y. D. Chen**[3]

[1]Department of Electronic Engineering, National Taipei University of Technology, Taiwan
[2]Department of Electrical Engineering, National Taiwan Ocean University, Taiwan
[3]Antenna and EMC Laboratory, High Tech Computer Corp. (HTC), Taiwan

Abstract— The design of a meta-material realization of artificial magnetic conductor (AMC) surfaces for a high-gain reflector antenna application is presented. Artificial materials of periodic dielectrics exhibiting an electromagnetic band-gap (EBG) performance have been proposed and applied planar inverted-F antenna co-design and measurement. The artificial dielectric material (ADM) can enhance antenna radiation performance, spread antenna bandwidth and improve antenna radiation gain and efficiency. The artificial defected dielectric material has useful characteristics of harmonic rejection, band suppression and surface wave suppression.

1. INTRODUCTION

Planar or corner periodic metallic array behave as AMC placed on a grounded dielectric substrate. In this paper the AMC operation of single-layer arrays with via hole is studied using a resonant cavity model and a new application to high-gain reflector antenna is presented. Planar AMC surfaces are used the ground plane in a 10 GHz dipole antenna with a partially reflective surface.

2. DESIGN & RESULTS

The AMC structure with periodic metallic array is shown in Fig. 1 which may reduce the surface wave and enhance the radiation performance. A magnetically loaded AMC material providing enhanced bandwidth has been developed. The characteristics such as electromagnetic band-gap, frequency selective surface and high impedance ground plane (HIGP) surface [1–3]. The AMC structure can reduce the surface wave and enhance radiation performance. The traditional reflector antenna [4] with PEC ground plane is shown in Fig. 2 which introduces comparable worse

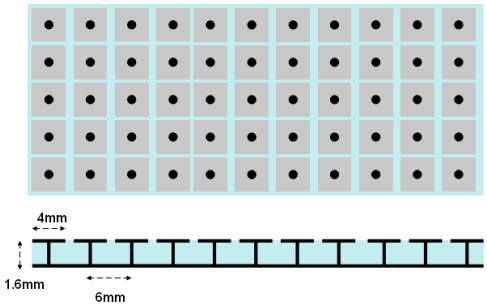


Figure 1: The artificial magnetic conductor (AMC) structure.

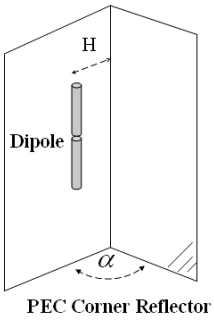


Figure 2: Traditional reflector antenna ($\alpha = 180°$, $H = 0.25\lambda$).

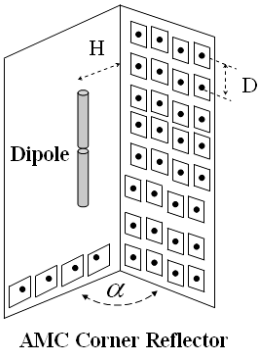


Figure 3: Proposed reflector antenna with AMC ($\alpha = 180°$, $H = 0.25\lambda$, $D = 6$ mm) structure.

cross-polarization. The proposed reflector antenna with AMC structure in Fig. 3 can reduce the surface wave interference, reduce cross-polarization and also improve the antenna gain (Fig. 4). The bandwidth and center frequency of AMC surfaces are investigated with full-wave analysis [5] and the qualitative predictions are validated.

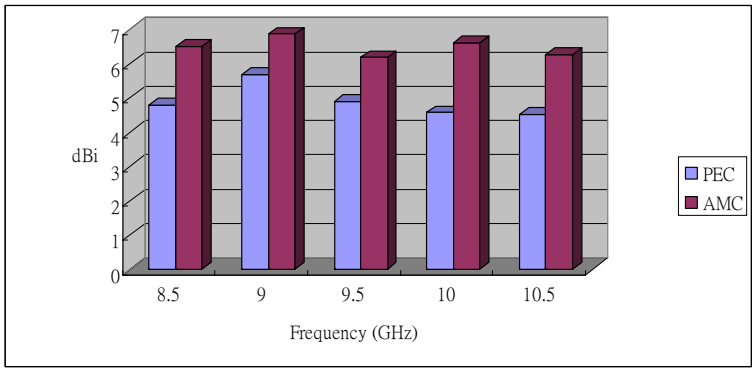


Figure 4: Compared data with PEC and AMC reflector.

3. CONCLUSIONS

The AMC structure of a high-impedance ground plane (HIGP) applied for a reflector antenna with dipole radiator has been developed. The designed antenna exhibits fairly high gain with the frequencies from 8.5 GHz to 10.5 GHz. It is believed to be useful for microwave antenna applications.

ACKNOWLEDGMENT

The authors would like to thank the Antenna and Wireless System Integration Department of High Tech Computer, Corp. (HTC), Taiwan for sporting the environment of measurement.

REFERENCES

1. Collin, R. E., *Field Theory of Guided Waves*, 2nd, IEEE Press, New York, 1991.
2. Soukoulis, C. M., *Photonic Band Gaps and Localization*, Plenum Press, New York, 1993.
3. Sievenpiper, D. F., "High-impedance electromagnetic surfaces," Ph.D. dissertation, Depart. Elect. Eng., Univ. California at Los Angeles, Los Angeles, CA, 1999.
4. Kraus, J. D. and R. J. Marhefka, *Antennas for All Applications*, McGraw-Hill, 2002.
5. AWR Microwave Office.

Design of Controlled RF Switch for Beam Steering Antenna Array

M. M. Abusitta, D. Zhou, R. A. Abd-Alhameed, and P. S. Excell
Mobile and Satellite Communications Research Centre, University of Bradford
Richmond Road, Bradford, West Yorkshire, BD7 1DP, UK

Abstract— A printed dipole antenna integrated with a duplex RF switch used for mobile base station antenna beam steering is presented. A coplanar waveguide to coplanar strip transition was adopted to feed the printed dipole. A novel RF switch circuit, used to control the RF signal fed to the dipole antenna and placed directly before the dipole, was proposed. Simulated and measured data for the CWP-to-CPS balun as well as the measured performance of the RF switch are shown. It has demonstrated the switch capability to control the beam in the design of beam steering antenna array for mobile base station applications.

1. INTRODUCTION

The use of adaptive array antennas for cellular base station application has recently become an active area of research and development [1–3]. Base station antennas normally radiate omni-directionally or in broad sectors, in which the most of the power is radiated in other directions than toward the user. This causes waste of power and interference for other users. Therefore new versions of base station antennas are now being made to overcome the problem by using antennas that have narrow steerable beams. These can give large increment in capacity, and the possibility of tracking mobile phones or vehicles.

In authors' previous work [4], a set of simple design procedures for beam steering single circular and concentric circular ring antenna arrays was proposed and analyzed theoretically (see Fig. 1). In the paper, the design theory was formulated and the results of the proposed analytical model, validated by a numerical model, were presented. Beam steering was achieved by implementing an ON/OFF system concept to excite only specific elements of the array dipole antenna. In this study, a following-up study was carried out on designing and implementing the RF switch for practically realising the beam steering using the proposed ON/OFF antenna array system. A novel design principle of RF switch, used to control the RF signal fed to each of the antenna elements in an array, was proposed. Moreover, a coplanar waveguide (CPW) to coplanar strip (CPS) transition was employed to feed the antenna element (i.e., printed dipole). Subsequently, performance of the CPW-to-CPS fed dipole antenna controlled by a RF switching circuit for duplex operation was evaluated and verified through hardware realisation. The measured results for the CPW-to-CPS balun and practical performance of the RF switch are shown in this paper.

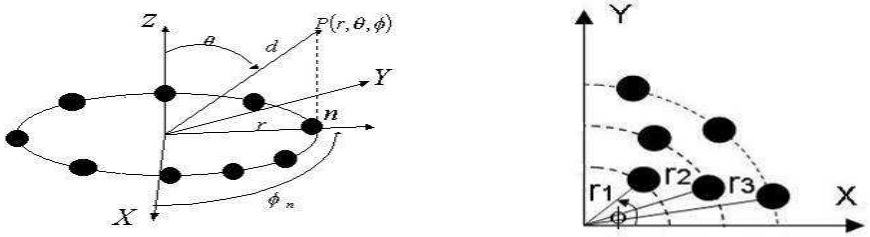


Figure 1: Circular array antenna, single circular ring arrays (left) and concentric circular ring arrays (right).

2. RF SWITCH AND PRINTED DIPOLE ANTENNA FOR BEAM STEERING ANTENNA ARRAY

2.1. CPW-to-CPS Baluns

Two back-to-back CPW-to-CPS balun (see Fig. 2) were examined using ADS simulator, which is based on the Method of Moment [5], in order to evaluate the balun performance such as, insertion and return loss at design frequency (GSM 1800 band). A CPW-to-CPS balun was chosen and designed for our application due to their several features such as low-loss, ease of fabrication and no need for via holes [6, 7]. The balun structure was mounted on Duriod material (ε_r = 2.5,

thickness $h = 1.524\,\text{mm}$, and $\tan\delta = 0.0019$). The measured insertion loss of the fabricated balun, achieved over the operating bandwidth from 1.47 GHz to 2.04 GHz, was found to be less than 1 dB as shown in Fig. 2. It is also noticeable that a reasonable return loss of 10 dB over the same frequency bandwidth. An excellent performance of the magnitude and phase imbalance between the two outputs of a single balun was observed within the intended operating band (the plot is not presented here).

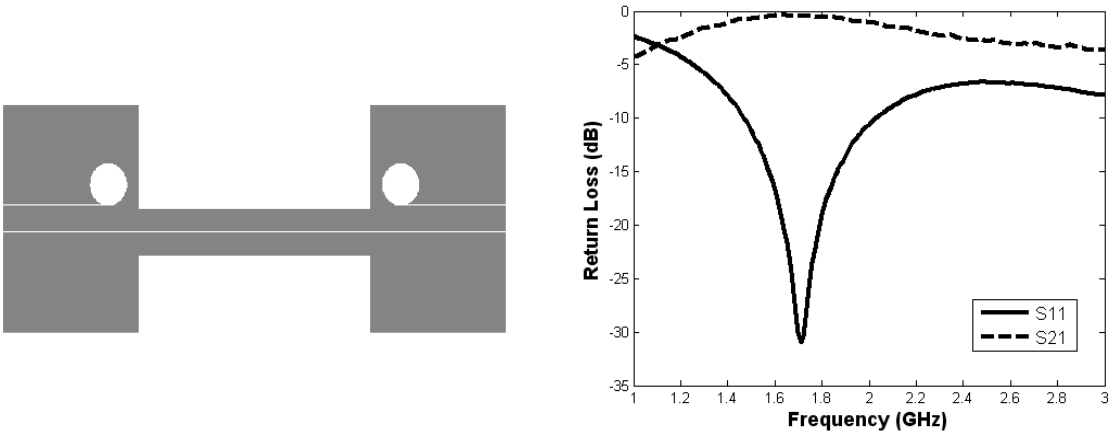


Figure 2: Layout of back-to-back balun (left) and the measured insertion loss and return loss of the balun studied.

2.2. Design of CPW-fed CPS Printed Dipole Antenna

For analysis, performance of the CPW-to-CPS fed dipole antenna was investigated with the aid of ADS. The layout of this dipole antenna is illustrated in Fig. 3, in which the width of the centre conductor is 4 mm and the gap is 0.2 mm. The diameter of the circular slot is 6.4 mm and the antenna length is 78.95 mm which corresponds to slightly less than half the wavelength (i.e., antenna resonates at around 1.84 GHz). For validation, a prototype of such a design was fabricated and tested. Return loss of the fabricated dipole antenna was measured and the result was compared to the data in prediction, as shown in Fig. 3. A bondwire was used to prevent unnecessary higher order modes generated at the discontinuities [8].

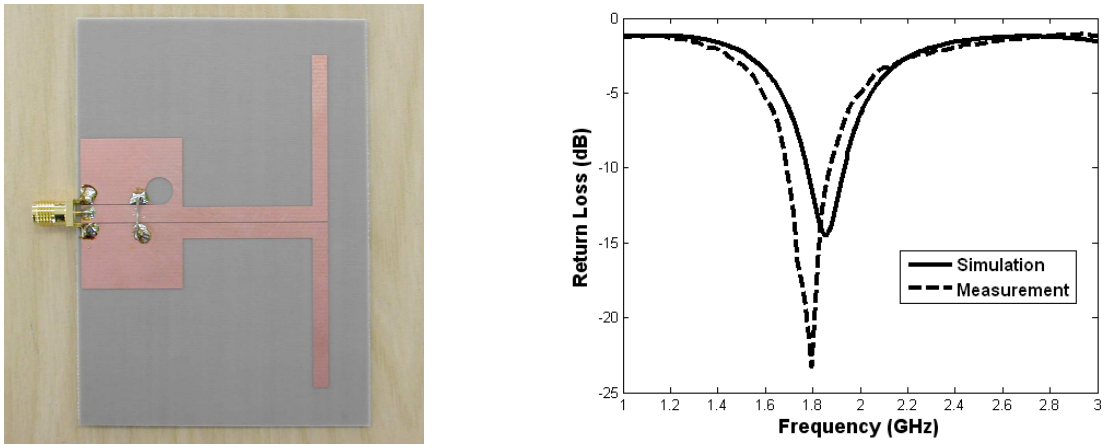


Figure 3: Layout of CPW-to-CPS balun integrated with printed dipole (left) and the measured return loss for this configuration studied (right).

2.3. Design Principle of RF Switch Circuit and Validation

RF switch is the integral part of modern communications system. Their application include well established areas such as radar and emerging areas such as smart (switched beam, phase and adaptive) antennas for terrestrial and satellite communication systems. The fundamental component in this switching is the operation of the RF p-i-n diode. The switches can be accommodated in

the beam forming network or adaptive control beam antenna array systems. An ON/OFF system concept for achieving antenna beam steering was practically implemented by a novel and simple RF switching circuit.

The proposed switch can be used for duplex operation and the circuit diagram is illustrated in Fig. 4. As can be seen, capacitors C1, C2, C3 and C4 are for dc blocking, and three diodes in the circuit with appropriate biasing voltage can be used to provide a function as RF switching. When both V1 and V2 are supplied with positive voltage, RF signal passes through forward biased diode D1 and transmit power to the antenna. There is no signal returned to the path through diode D3 since it is reverse biased. Therefore, the switch is ON and in RF transmission mode. On the contrary, when the power supply is given negative voltage both V1 and V2, diode D1 is in reverse biasing which can be effectively used for blocking the RF signal transmits to the antenna and the reflected RF signal can be eliminated through the diode D2 due to the fact that RF signal is shorted via a 50 ohm resistor (R1) to the ground. In this way, the switch is apparently OFF and in the RF reception mode since only RF signal path through diode D3 is turned on. Thus, a dual mode operation is realized.

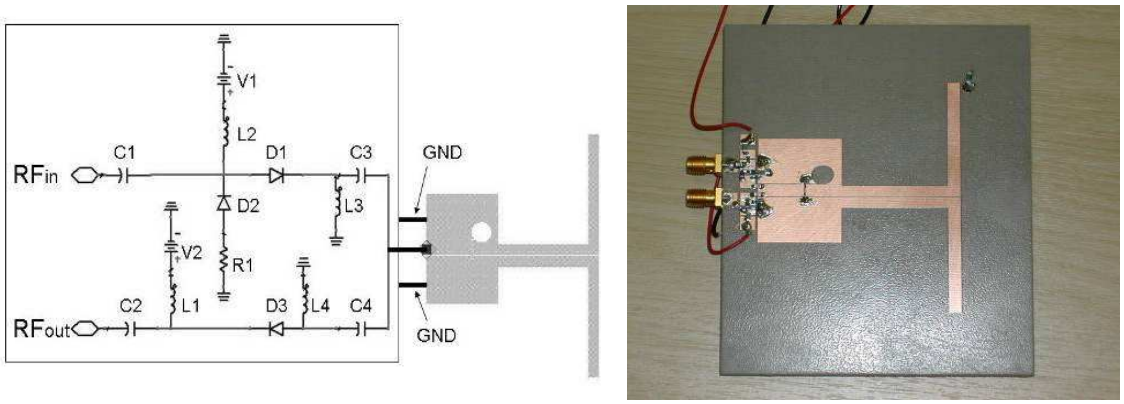


Figure 4: Overall circuit diagram of the RF switch integrated with balun and dipole (left) and prototype (right).

Prototypes of the RF switch integrated with back-to-back balun (see Fig. 5) and RF switch with balun and dipole (see Fig. 4) were fabricated and tested in order to validate the design theory of the proposed RF switching circuit.

The procedure of validation to the RF switch was carried out in two aspects (i.e., 'ON mode' and 'OFF mode'). For the purpose of simplifying the analysis, the fabricated prototype circuit in Fig. 4 was replaced by the back-to-back balun (see Fig. 5) for the evaluation purposes because it is well matched to the 50 ohm at the design frequencies and can be directly connected to the switch as a 50 ohm load. A practical measurement setup for evaluating this RF switch is illustrated in Fig. 5. To begin testing the 'ON mode' for the proposed switch, a positive voltage of 0.93 V was provided

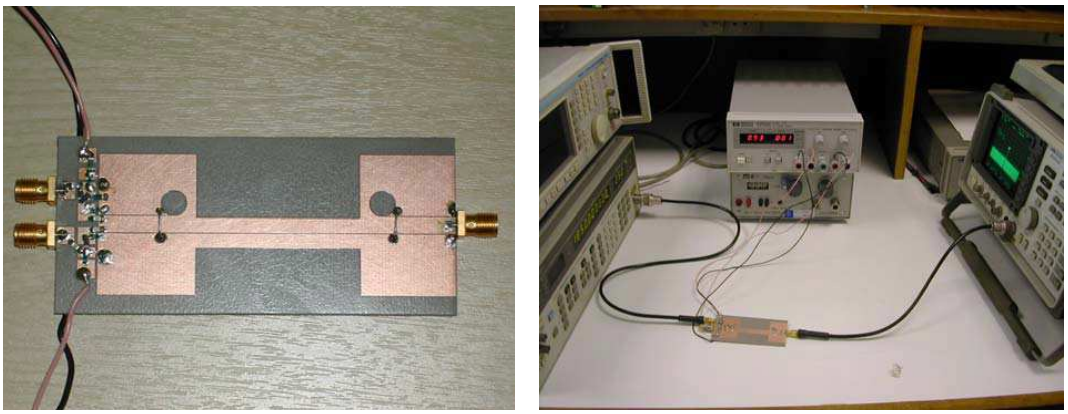


Figure 5: Photograph of the fabricated prototypes of the RF switch integrated with back-to-back balun (left) and measurement setup for RF switch evaluation (right).

to the V1 and V2 (see Fig. 4) and RF signal with power level of −20 dBm at single frequency of 1850 MHz was injected to the RF switch (RF_{in} port) from the signal generator. Subsequently, a RF output power with level of −23.67 dBm (see Fig. 6) from the back-to-back balun was observed on the spectrum analyser. Taking into account of the losses involved from the cable (1.33 dB) and the balun ($\leq$ 1 dB), the total insertion loss on the proposed RF swith was found to be approximately 1.5 dB when the switch is turned on.

On the contrary, in order to test the RF switch performance at 'OFF mode', the proposed switch was supplied with a negative voltage of −0.93 V to both V1 and V2 and RF signal with power level of −25 dBm at the same frequency was injected to the back-to-back balun. It is notable that a relatively less power was generated to test the RF switch in the receiving mode as the power level of the received signal is always small in the reception. As a result, a power level of −42.17 dBm was measured at the RF_{in} port. It implies an isolation performance with at least 15 dB was achieved for the proposed RF switch as shown in Fig. 7. It has to be noted that the other port (RF_{out}) in the switch was connected with a 50 ohm load in both cases.

It was found from the forgoing practical investigation that the proposed RF switch exhibits a relatively good performance at 'ON' and 'OFF' mode. As a consequence, it has demonstrated the capability used as the switch to control beam in the design of beam steering antenna array for mobile base station antennas applications. Therefore, a follow-up study on practical realisation to the mobile base station antenna with enhanced performance using the beam steering antenna array design principle in cooperation with the novel RF switch proposed in this paper will be carried out in the future work.

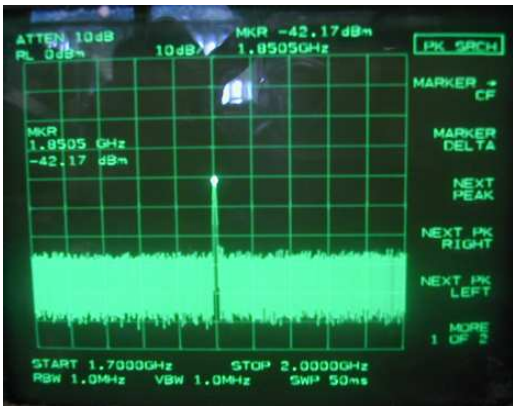


Figure 6: The output of the back to back baluns when the switch mode is off.

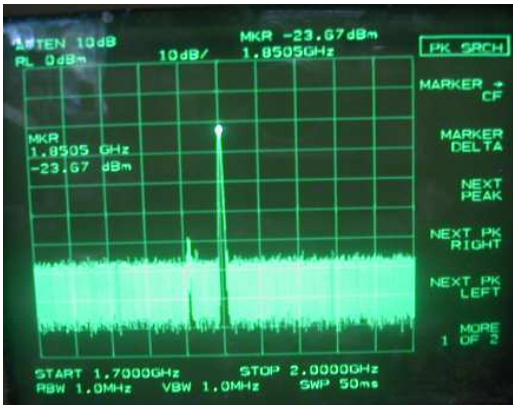


Figure 7: The output of the back to back baluns when the switch mode is on.

3. CONCLUSIONS

In this work, a complete analysis and design of the CPW-fed CPS balanced dipole antenna integrated with a RF switching circuit for antenna beam steering used in mobile base stations, was presented. The CPW-to-CPS balun and the dipole antenna, were investigated and the overall performance of the dipole antenna in collaboration with the RF switching circuit were analysed and evaluated. The predicted results indicating the design goal was well met. This is encouraging for practical implementation of these switchable dipoles in the design of beam steering antennas for the future work.

REFERENCES

1. Morishita, H., Y. Irie, S. Hara, Y. Nakaya, T. Toda, and Y. Oishi, "A beamforming method for a reactively steered adaptive array antenna with RF-MEMS device," *IEEE Topical Conference on Wireless Communication Technology*, 396–397, Oct. 2003.
2. Imamura, K. and H. Morishita, "Analysis of the mobile terminal adaptive array antenna in consideration of a case and elements," *IEEE Antennas and Propagation Society Symposium*, Vol. 3, 3195–3198, June 2004.
3. Nakaya, Y., T. Toda, S. Hara, and Y. Oishi, "MIMO receiver using an RF-adaptive array antenna with a novel control method," *IEEE International Conference on Communication*, Vol. 5, 2568–2572, June 2004.

4. Abd-Alhameed, R. A., N. T. Ali, P. S. Excell, M. K. Atiya, and C. H. See, "Beam steering antenna array for mobile base stations," *3rd IEEE International Conference on Systems, Signals & Devices SSD'05*, Sousse, Tunisia, Paper Ref. SSD05-CSP-118, March 2005.
5. Advanced Design System, Agilent Technologies, version ADS2005A.
6. Tilley, K., X.-D. Wu, and K. Chang, "Coplanar waveguide fed coplanar strip dipole antenna," *Electronics Letters*, Vol. 30, No. 3, 176–177, 1994.
7. Kolsrud, A. T., M.-Y. Li, and K. Chang, "Dual-frequency electronically tunable CPW-fed CPS dipole antenna," *Electronics Letters*, Vol. 34, No. 7, 609–610, 1998.
8. Thaysen, K. B. Jakobsen, and J. Appel-Hansen, "A wideband balun — How does it work?" *More Practical Filters and Couplers: A Collection from Applied Microwave & Wireless*, ISBN 1-884932-31-2, 77–82, Noble Publishing Corporation, 2002.

Ultra Low Side Lobe Level Synthesis with Particle Swarm Optimization for Symmetrical Non-uniform Linear Array Antennas

Xiaomiao Zhang[1], **Kwai Man Luk**[2], **Xue Bai**[1], **Yinhang Wang**[1], **and Jinyang Li**[1]

[1]National Key Laboratory of Antennas and Microwave Techniques, Xidian University
Xi'an 710071, Shaanxi Province, China
[2]Department of Electronic Engineering, City University of Hong Kong
83 Tat Chee Avenue, Kowloon, Hong Kong, China

Abstract— Ultra low side lobe level (ULSLL) array antennas have important applications in radar and communication systems, ULSLLs are usually realized with uniform linear array (ULA) or planar array antennas by excitation amplitude weighting, such as the current distributions resulted from Dolph-Chebyshev or Taylor synthesis. Lower SLL, however, implies lower aperture efficiency and larger excitation amplitude ratio (EAR), which makes the feed network more complicated and more difficult to fabricate; lower SLL is also very sensitive to the excitation amplitude/phase errors. In order to mitigate this difficulty, both the excitation amplitudes and the element positions are used for ULSLL synthesis, where the element positions act as extra freedoms for further SLL suppression. A symmetrical non-uniform linear array (SNULA) is proposed, the full information particle swarm optimization (FIPSO) based on the Von Neumann local neighborhood topology is adopted for ULSLL pattern synthesis, because it is effective for multi-dimensional, multi-local-minimum, non-linear complex problems and robust convergence. In the algorithm, the inertia weight is improved via the sigmoid limiting function; the boundary conditions are dealt with a randomized damp reflection. Two numerical examples are given, the first one is an 8-element -40 dB SLL array antenna, for comparison, a ULA with the same aperture size, element number and peak SLL is synthesized by Dolph-Chebyshev method, the results obtained show that they have nearly identical main beam and equal side lobes, but the SNULA has lower EAR and higher aperture efficiency than those of the ULA. The second example is a 12-element -45 dB SLL array, similar rules are obtained.

1. INTRODUCTION

Ultra-Low side lobe level (ULSLL) is usually realized with uniform linear array (ULA) by amplitude-only technique, where the excitation amplitude tapering (AT) is obtained by the aid of Dolph-Chebyshev or Taylor synthesize method. Lower SLL, however, implies lower aperture efficiency (AE) and larger excitation amplitude ratio (EAR), and it is very sensitive to the excitation amplitude/phase errors.

Density tapering (DT) technique is common used for limited side-lobe reduction of uniformly excited arrays [1, 2], several methods such as differential evolution algorithm (DEA) [3], genetic algorithm (GA) [4], and particle swarm optimization (PSO) [5–7] have been researched to optimize the element locations.

In this paper, the AT technique is added to the DT technique as another freedom to mitigate the difficult, which results in a symmetrical non-uniform linear array (SNULA), the full information PSO (FIPSO) is presented ULSLL pattern synthesis.

2. METHOD

An SNULA antenna with even number ($2N$) isotropic elements is shown in Fig. 1, the excitation amplitude are symmetric and in phase.

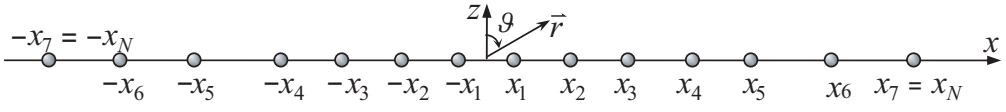


Figure 1: Symmetrical nonuniform linear array antennas ($N = 7$ for illustration).

The array factor (AF) is in the form,

$$F_a(\vartheta) = \frac{1}{N}\left\{\cos(kx_1\sin\vartheta) + \sum_{n=2}^{N} A_n\cos(kx_n\sin\vartheta)\right\} \tag{1}$$

where, ϑ is measured from the array broadside, $k = 2\pi/\lambda$ is the wavenumber and λ is the wavelength, the excitation amplitude of the 1st element has taken to be equal to 1.

The objective function is defined by,

$$f(A_2 \sim A_N;\, x_1 \sim x_N) = \max\left[20\log|F(\vartheta_m)|,\ \left.\frac{dF_a(\vartheta)}{d\vartheta}\right|_{\vartheta=\vartheta_m},\ m = 1 \sim M\right] \tag{2}$$

where M is the number of sidelobe, and ϑ_m is the location of mth sidelobe, there are $2N-1$ variables that span the solution space, obviously, it is a multi-dimension, non-linear complex optimization problem.

The synthesis is accomplished by the full information PSO (FIPSO), which is based upon the following two Equations [8],

$$\vec{v}_i \Leftarrow \chi\left[\vec{v}_i + \varphi\sum_{m=1}^{N_i} rand()\cdot\left(\vec{p}_{neighbor(m)} - \vec{x}_i\right)/N_i\right] \tag{3}$$

$$\vec{x}_i \Leftarrow \vec{x}_i + w\vec{v}_i \tag{4}$$

where $\vec{x}_i$ is position in the solution space of particle i, $\vec{v}_i$ is its velocity; χ is the constriction coefficient, and is typically equal to 0.729; φ is the acceleration constant and is set to 4.1; $rand()$ is a uniform random generator distributed within the interval $[0, 1)$; N_i is the number of influences among the particle i's neighbors, $\vec{p}_{neighbor(m)}$ is the position of its mth influence; w is the inertia weight.

The FIPSO offers that the subpopulation could search diverse region of the whole space, as individual particles is affected by its influences in the neighborhood, the local searching ability is improved, which avoids traping into local optimum; the information flows from one neighborhood to the next neighborhood, so the whole swarm shares and renews the information after an iteration.

A sigmoid damped limiting transformation is used to simulate the inertia weight in the iteration procedure,

$$w(t) \approx (w_s - w_f)/\{1 + \exp[\alpha(t - t_1)]\} + w_f \tag{5}$$

where t is the iterative step, w_s and w_f is the inertia weight at the beginning and end of iteration, respectively; α is the damp constant, which determines the decreasing slope; $t_1 \gg 1$ is the iterative step where the steepest decreasing of the inertia weight occurs, and $w(t_1) = (w_s + w_f)/2$.

A random mirror reflection relation is applied to the particle flying across the boundary in the d-dimension,

$$x_i^d \Leftarrow (1+\Gamma)b^d - \Gamma x_i^d \tag{6}$$

$$\Gamma = \Gamma_0\, rand(),\ \Gamma_0 \in [0,\ 1] \tag{7}$$

where x_i^d is the component of particle $\vec{x}_i$ in the d-dimension, b^d is the lower or upper boundary, Γ is the reflection coefficient, and Γ_0 is the maximum value.

3. SIMULATED RESULTS

The 1st example is an 8-element -40 dB SLL linear array, the resulted element positions, excitation amplitudes, EAR and aperture efficiency are listed in Table 1. Simulated AFs are shown in Fig. 2, where Dolph-Chebyshev pattern of the ULA with the same number of element and array length is plotted for comparison.

The 2nd example is a 12-element -45 dB SLL linear array, also, the results are listed in Table 1, and simulated AFs are shown in Fig. 3.

Table 1: Optimized parameters of the examples.

Examples	Methods	Element positions and excitation amplitudes							EAR	AE
		n	1	2	3	4	5	6		
1	PSO	x_n	0.259	0.871	1.569	2.294			4.86	0.82
		A_n	1.000	0.997	0.606	0.206				
	Chebyshev	x_n	0.328	0.983	1.639	2.294			6.97	0.76
		A_n	1.000	0.758	0.415	0.144				
2	PSO	x_n	0.261	0.782	1.409	2.133	2.893	3.668	7.74	0.81
		A_n	1.000	1.000	1.000	0.719	0.374	0.129		
	Chebyshev	x_n	0.333	1.001	1.667	2.334	3.001	3.668	12.5	0.72
		A_n	1.000	0.872	0.656	0.417	0.212	0.080		

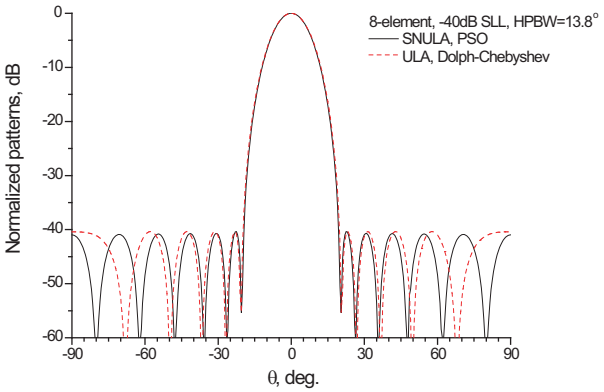


Figure 2: Array factors ($N = 8$, SLL $= -40$ dB).

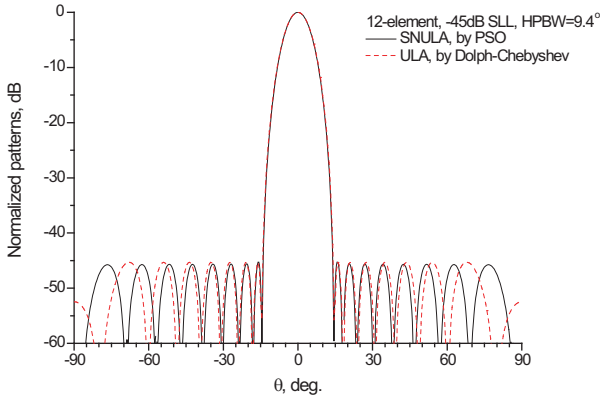


Figure 3: Array factors ($N = 12$, SLL $= -45$ dB).

4. CONCLUSION

By combination the AT technique and the DT technique, an SNULA antenna is proposed, the FIPSO is used for ULSLL synthesis. Compared with the ULSLL of conventional ULA under the conditions of the same HPBW and SLL, the simulated results of SNULA show that EAR is lower; the aperture efficiency is higher.

ACKNOWLEDGMENT

The authors would like express their thanks to the support of the NSFC of P. R. China under Grant (No. 60671056).

REFERENCES

1. Harrington, R. F., "Sidelobe reduction by nonuniform element spacing," *IRE Transactions on Antennas and Propagation*, 187–192, March 1961.
2. Hui, K. Y. and K. M. Luk, "Design of wideband base station antenna arrays for CDMA 800 and GSM 900 systems," *Microwave and Optical Technology Letters*, Vol. 39, No. 5, 406–409, December 2003.
3. Kurup, D. G., M. Himdi, and A. Rydberg, "Synthesis of uniform amplitude unequally spaced antenna arrays using the differential evolution algorithm," *IEEE Transactions on Antennas and Propagation*, Vol. 51, 2210–2217, September 2003.
4. Liu, A.-S., R.-B. Wu, Y.-C. Lin, and H.-J. Li, "Synthesis of nonuniformly spaced linear array for GSM/DCS/WCDMA base station application using genetic algorithm," *IEEE Antennas and Propagation Society International Symposium*, 137–141, 2004.
5. Khodier, M. M. and C. G. Christodoulou, "Linear array geometric synthesis with minimum sidelobe level and null control using particle swarm optimization," *IEEE Transactions on Antennas and Propagation*, Vol. 53, 2674–2679, August 2005.
6. Lee, K.-C. and J.-Y. Jhang, "Application of particle swarm algorithm to the optimization of unequally spaced antenna arrays," *Journal of Electromagnetic Waves and Applications*, Vol. 20, No. 14, 2001–2012, 2006.

7. Jin, N. and R.-S. Yahya, "Advances in particle swarm optimization for antenna design: real-number, binary, single-objective and multiobjective implementations," *IEEE Transactions on Antennas and Propagation*, Vol. 55, 556–567, March 2007.
8. Kennedy, J. and R. Mendes, "Neighborhood topologies in full informed and best-of-neighborhood particle swarm," *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, Vol. 36, No. 4, 515–519, 2006.

An Efficient Density Weighting Approach for Side-lobe Level Suppression of Linear Array Antennas

Xiaomiao Zhang[1], **Kwai Man Luk**[2], **Weiwei Song**[1], **Wei Zhao**[1], **and Yang Liu**[1]

[1]National Key Laboratory of Antennas and Microwave Techniques, Xidian University
Xi'an 710071, Shaanxi Province, China
[2]Department of Electronic Engineering, City University of Hong Kong
83 Tat Chee Avenue, Kowloon, Hong Kong, China

Abstract— Most likely, there are two techniques existed to reduce the side lobe level (SLL) of array antennas: The conventional excitation amplitude tapering (AT) and the unusual element space density tapering (DT). The AT technique, such as triangular, cosine, cosine-square and raised-cosine amplitude distributions, as well as the Dolph-Chebyshev or Taylor current coefficients, provides efficient means for low SLL (LSLL) and ultra-low SLL (ULSLL) pattern synthesis, and also affords significant dynamic ranges control of the aperture distributions. The feed network, however, becomes complicated and narrow bandwidth, the aperture efficiency decreased rapidly as the amplitude getting much tapered. The DT technique offers another way for limited SLL reduction while the aperture efficiency is maintained, the antenna is a uniformly excited non-uniformly spaced array with element positions thinned from the center element (odd number array) or center two elements (even number array), the main advantage is that it simplifies the bulk and lossy feed network. In this paper, A uniformly excited symmetrical array antenna, which is compose of a uniform central sub-array (central part) and two density tapered side sub-arrays (side parts), is presented for SLL suppression, a power function characterized by two-parameter is employed to express the increments between adjacent elements of the two side parts, the optimum values of the two parameters are searched by the Min-max method according to the peak SLL. Numerical examples are given for the cosine-element, one half wavelength least spacing, and 4-element central part arrays, 8-, 10-, and 12-element arrays are simulated; Dolph-Chebysheve liked patterns are observed, $-19.5\,\mathrm{dB}$, $-20.8\,\mathrm{dB}$ and $-21.4\,\mathrm{dB}$ SLLs are achieved, respectively.

1. INTRODUCTION

As an alternative sidelobe (SL) level (SLL) reduction technique, the nonuniformly (or unequally) spaced linear array technique has been investigated extensively in the last decades from theory to practice [1–5], a geometrically and electrical symmetrical array antennas is of particular interests due to its real pattern function and unique phase center [6, 7].

A uniformly excited symmetrical nonuniformly spaced linear array (SNULA) antenna with density weighting (DW) is presented, a power function is introduced to describe the spacing increment, and it provides an efficient approach for SLL reduction.

2. ANALYSIS

Figure 1 shows the array geometry for the cases: a) odd number of $2N+1$ elements, and b) even number of $2N$ elements, the elements are symmetrical about the array center. The array can be divided into 2 parts: 1) the central uniform spaced sub-array, and 2) the side non-uniformly spaced sub-array.

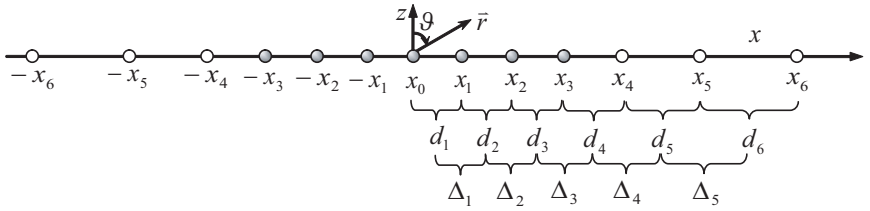
At first, the odd number array are discussed, x_n, $(n = 1 \sim N)$ is the position of the nth element along the x-axis, d_n is the distances between adjacent elements, and Δ_n is the increment between adjacent spacings,

$$d_n = x_n - x_{n-1}, \quad n \sim N \tag{1}$$

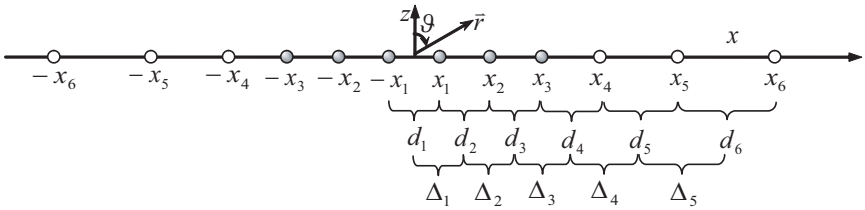
$$\Delta_n = d_{n+1} - d_n, \quad n \sim N-1 \tag{2}$$

Assume the space increments are increased monotonously according to a power function,

$$\Delta_n = \begin{cases} 0, \ n = 1 \sim N_1 - 1; \\ \delta(n - N_1 + 1)^p; \ n = N_1 \sim N - 1. \end{cases} \tag{3}$$



(a) Odd number array with $2N+1$ elements



(b) Even number array with $2N$ elements

Figure 1: Symmetrical linear array antennas ($N = 6$, $N_1 = 3$).

where $2N_1 + 1$ is the number of the central sub-array, δ is the increment parameter and p is the power. Application (3) to (2) and (1), one gets,

$$d_n = \begin{cases} d, \ n = 0 \sim N_1; \\ d + \delta \sum_{m=1}^{n-N_1} m^p; \ n = N_1 + 1 \sim N. \end{cases} \tag{4}$$

$$x_n = \begin{cases} nd, \ n = 0 \sim N_1; \\ nd + \delta \sum_{m=1}^{n-N_1} m^p; \ n = N_1 + 1 \sim N. \end{cases} \tag{5}$$

where d is the element space of the central sub-array. The array factor (AF) is the sum of the central sub-array factor $f_{a1}(\vartheta)$ and the side sub-array factor $f_{a2}(\vartheta)$,

$$F_a(\vartheta) = f_{a1}(\vartheta) + f_{a2} \tag{6}$$

$$f_{a1}(\vartheta) = \frac{1}{2N+1} \frac{\sin\left(k\frac{2N_1+1}{2} d \sin\vartheta\right)}{\sin\left(k\frac{1}{2} d \sin\vartheta\right)} \tag{7}$$

$$f_{a2}(\vartheta) = \frac{2}{2N+1} \sum_{n=N_1+1}^{N} \cos(k x_n \sin\vartheta) \tag{8}$$

where, ϑ is measured respect to the array broadside, $k = 2\pi/\lambda$ is the wavenumber and λ is the wavelength.

Secondly, the even number array with $2N$ elements is deduced in the same way,

$$d_n = \begin{cases} d, \ n = 0 \sim N_1 - 1; \\ d + \delta \sum_{m=1}^{n-N_1} m^p; \ n = N_1 \sim N. \end{cases} \tag{9}$$

$$x_n = \begin{cases} (n-0.5)d, \ n = 1 \sim N_1; \\ (n-0.5)d + \delta \sum_{m=1}^{n-N_1} m^p; \ n = N_1 + 1 \sim N. \end{cases} \tag{10}$$

$$f_{a1}(\vartheta) = \frac{1}{2N} \frac{\sin(kN_1 d \sin\vartheta)}{\sin\left(k\frac{1}{2} d \sin\vartheta\right)} \tag{11}$$

$$f_{a2}(\vartheta) = \frac{1}{N} \sum_{n=N_1+1}^{N} \cos\left(k x_n \sin\vartheta\right) \tag{12}$$

For small scaled central sub-array, $f_{a1}(\vartheta)$ has a broad main-beam and higher close-in SLs; while $f_{a2}(\vartheta)$ has a narrow main-beam and higher SLs, if the parameters (δ, p) are chosen appropriately, the sodelobes may be counteracted each other to some extent, meanwhile the outer SLs can be further suppressed by the element pattern, accordingly, the total array's SLLs are decreased.

For searching the optimum values of (δ, p), the objective function is defined by,

$$f = \max[20\log|\mathrm{F}(\vartheta)|, \vartheta \in \text{sidelobe region}] \tag{13}$$

The minimum value of f corresponds the optimum (δ_{opt}, p_{opt}).

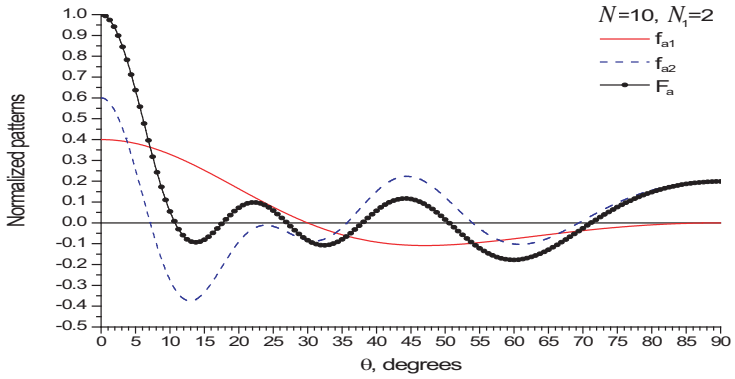
3. NUMERIC RESULTS

As an example, an even number $\cos\vartheta$-element array with $N = 10$, $N_1 = 2$ and $d = 0.5\lambda$ is simulated, the optimum parameters are (δ_{opt}, p_{opt}) = (0.0448, 1.853), the pattern is shown in Figure 2, Chebyshev-liked SLs with SLL= -20.77 dB is observed.

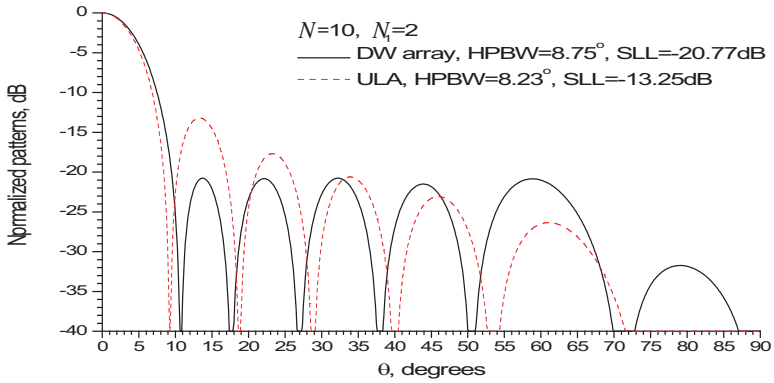
Several other arrays with $d = 0.5\lambda$ are computed and the SLLs are listed in Table 1.

Table 1: SLLs of several arrays ($\cos\vartheta$-element, $d = 0.5\lambda$,).

(N, N_1)	(5, 1)	(6, 2)	(7, 1)	(8, 2)	(9, 1)	(11, 2)
SLL (dB)	-17.30	-18.55	-18.90	-19.48	-20.24	-21.03
(N, N_1)	(12, 2)	(13, 2)	(14, 3)	(15, 3)	(16, 3)	
SLL (dB)	-21.42	-21.92	-22.00	-22.03	-22.57	



(a) Sub-array and array factors



(b) Total patterns

Figure 2: Pattern ($N = 10$, $N_1 = 2$).

4. CONCLUSION

The simulated results demonstrate that the spacing increments expressed as a power function may give an appropriate mode for density weighted linear array antennas for SLL suppression; approximated optimum SLL patterns can be obtained, for arrays no less than 9 elements, −20 dB peak SLL are achievable.

ACKNOWLEDGMENT

The authors would like to express their thanks to the support of the NSFC of P. R. China under Grant (No. 60671056).

REFERENCES

1. Harrington, R. F., "Sidelobe reduction by nonuniform element spacing," *IRE Transactions on Antennas and Propagation*, 187–192, March 1961.
2. Ishimaru, A., "Theory of unequally-spaced arrays," *IRE Transactions on Antennas and Propagation*, 691–702, November 1962.
3. Kumar, B. P. and G. R. Branner, "Design of unequally spaced arrays for performance improvement," *IEEE Transactions on Antennas and Propagation*, Vol. 47, 511–523, March 1999.
4. Chair, R., K. M. Luk, and K. F. Lee, "Wideband L-probe-fed 4-element array with unequal spacing," *Proceedings of APMC2001*, 938–941, 2001.
5. Yamada, Y., S. Takubo, and Y. Ebine, "An unequally spaced array antenna for mobile base stations," *IEEE AP-S*, 432–435, 2001.
6. Tomiyasu, K., "Combined equal and unequal element spacings for low sidelobe pattern of a symmetrical array with equal-amplitude elements," *IEEE Transactions on Antennas and Propagation*, Vol. 39, 265–266, February 1991.
7. Haupt, R. L., "Partial nonuniform spacing of array elements," *IEEE AP-S*, 1708–1711, 1991.