# Microstrip Antennas for Direct Human Skin Placement for Biomedical Applications

**Sudhir Shrestha**, **Mangilal Agarwal**, **Joshua Reid**, **and Kody Varahramyan**
Integrated Nanosystems Development Institute (INDI)
Department of Electrical and Computer Engineering
Indiana University — Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA

**Abstract**— Microwave breast imaging (MBI) uses low power and longer wavelength signals (compared to X-ray mammography) to obtain information about breast tissues, and promises a safer and more accurate modality for regular breast scanning. One of the major impediments of this technology has been a large signal reflection from the breast skin. In this paper, we present a method to reducing the effect of signal reflection from the breast skin by placing the antenna in-contact with the breast skin. In the reported method, the skin is considered a layer of the antenna substrate, and the effect of having the antenna in contact with the skin is included in the antenna design itself. Thus, the design allows placing the antenna on the breast skin. This reduces the signal scattering from the skin and more transmitted signal is irradiated on the tumor, thus, increasing the tumor detection sensitivity. Design and simulation in Ansoft High Frequency Simulation Software (HFSS) is presented. The simulation results show that the current density in the breast skin is much less while the current density in the tumor is much higher when the antenna is placed directly on the breast skin, compared to when the antenna is placed away from the skin. While, in this paper, the antenna is discussed in the context of microwave breast imaging, the concept of placing an antenna directly on human skin has a wide range of biomedical applications.

## 1. INTRODUCTION

In 2009, more than 192 thousand new cases of breast cancer were expected in the United States and more than 40 thousand women were estimated to lose their lives [1]. In 2007, 1.3 million new cases of breast cancer and 465 thousand deaths were estimated worldwide [2]. Early detection of breast cancer by regular breast screening has been shown to sharply reduce the breast cancer related mortality and increase the survival rate [3]. Despite the benefits of the regular breast cancer screening, only 66.5% women of forty years age and older conducted mammogram screening in 2005. This has been particularly due to the health risk associated with the X-ray used in mammography and its accuracy in detecting the early breast tumors [4]. Recently, the U.S. Preventive Services Task Force (USPSTF) has updated its guidelines, recommending a regular mammogram screening at the age of fifty and above, once every two years. This is a change from its previous recommendation of yearly screening at the age of forty [5]. This change has fueled additional debates on the health risk of X-ray mammography, its accuracy, and the merits of using it as the primary screening modality. However, the benefits of early breast tumor detection are well accepted. Thus, the development of a safer and more accurate breast scanning and imaging modality has been pursued for a relatively long period, with microwave breast imaging being one of the most viable technologies [6]. In MBI technique, an antenna transmits a microwave signal to the breast and the scattered signal is received and analyzed to extract dielectric properties of the tissues.

Another key challenge in breast cancer detection and treatment has been to design and develop a standard method that heightens efficiency coincided with nondestructive cancer psychoanalysis. The overall ideal for this innovative concept is to detect with potential benefits of adequate depth penetration via microwave imaging, while minimizing cumulative side-effects to healthy tissue due to ionized radiation. Due to the complexity that coincides with a successful methodology, these biological systems must endure heterogeneous characteristics [7]. Key functions for microwave-based breast cancer diagnosis would be (i) low health risk, (ii) the ability to detect breast cancer at a curable stage, (iii) is sensitive to tumors and specific to malignancies, (iv) involves minimal discomfort for lesion tolerability to women, and (v) provides easy to interpret, objective, and consistent results [8]. Such medical devices that posses adequate depth penetration while avoiding the use of ionized radiation and breast compression would be of great interest for malignant purposes. Further, the infrastructure of microwave breast imaging for cancer detection contains a bevy of algorithms alongside geometric configurations, which can overcome the problems associated with current methodologies.

Microwave imaging of biological compositions have developed an enormous amount of attention because of its ability to access the breast for imaging (e.g., image reconstruction algorithms, computational methods, increased computing supremacy.) MBI infrastructure consists of several deficiencies that conclude ex-vivo circumstances, electromagnetic (EM) signals coupled with gigahertz signal frequencies [9]. Further, the electromagnetic phenomenon is an approach that mirrors the electromagnetic properties of breast tissue allowing for a more efficient exam. Therefore, as mentioned before, the coupling with the signals in gigahertz range are considerable because of the significant absorption and scattering that occurs during EM exposure to the breast [9]. In particular, the latter two features of self-propagating waves could lead to enormous amplification of one particular attribute, or complicated instrumentation from the need to place transmitting and receiving RF transmitters to avoid spatial resolution. For example, significant electromagnetic loss in tissue, and it's extremely high contrast with surrounding air [9]. Furthermore, the development of microwave patch transceivers (antennas) operating at 2.45 GHz is known to reduce cluttered data, producing well-localized images or real and imaginary parts of the wave numbers [10]. Concluding all of which may be etched to comply with current breast cancer research, resulting with a more promising approach.

Despite the popularity, there are several proposed ideals that could potentially become historical due to promising results via microwave imaging. Such methodologies consist of ultrasound and digital mammography, magnetic resonance imaging (MRI), position emission tomography, and electrical impedance scanning. According to previous studies, the above cutting edge portrays some success in reference to cancer detection, but in accordance to early detection and unnecessary mastectomies, significant approaches are in high regard. For example, MRIs, a computer aided method that uses a magnet to create detailed images of internal structures without the use of radiation. In compliance with the rest (e.g., mammograms), the above treats with radioactive techniques which could pose significant side-effects, and worsen a patient's current condition. Therefore, an understanding of the breast tissue and skin diplomacy is documented to help one better understand how current methods detect as well as acknowledge between healthy and malignant tissue. In the field of enhancing breast detection applications, there are substantial interests in the development of new and effective approaches to demonstrate the use of microwave imaging for tumor detection. The interactions of EM and gigahertz signals are mainly being investigated for a significant and consistent contrast between malignant and other breast tissues [8].

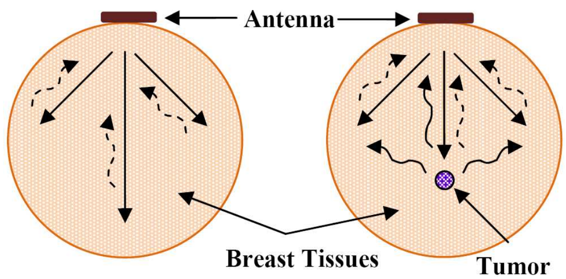
One of the major impediments of MBI technology has been a large signal reflection from the breast skin. Many techniques to reducing the signal reflection at the breast skin have been employed [11–13]. Most of the approaches use an intermediate solution between the antenna and the breast, but no significant success has been achieved thus far. In this paper, we present the design and simulation of microwave breast imaging using 2.54 GHz signal. HFSS (ver. 12) has been used for the modeling and simulation. While, the experimental implementation of the concept is still in progress, the simulation results presented shows that the direct placement of the antenna on the breast skin can significantly increases the sensitivity of the MBI systems.

## 2. THEORY

The microwave breast imaging technique utilizes the signal scattering by an object when the object is illuminated by an electromagnetic signal. The signal scattering by an object depends on various factors, including the environment, signal strength, and the material properties of the object [14]. For a given signal source and the environment, the scattered signal depends on the electrical properties of the object, especially dielectric and conductivity. This principal is utilized to detect the tumor in the breast using microwave signals. The breast tumors have very distinct electrical properties (higher dielectric permittivity and higher conductivity), which allows them to detect by analyzing the scattered signals. As shown in Figure 1, the amount of signal scattered by a breast tumor is higher than that of normal breast tissues, which can be received by a separate antenna or the property change of the transmitting antenna due to the scattered signals, can be analyzed and utilized for the tumor detection.

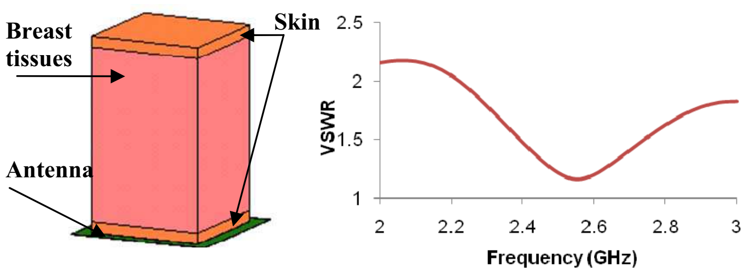
## 3. MODELING AND SIMULATION

The modeling and simulation was performed in Ansoft electromagnetic simulation softwares. Preliminary simulation and proof of concept were conducted in Ansoft Designer, while the complete simulation was conducted in Ansoft HFSS. Figure 2 shows a preliminary design in Ansoft Designer and antenna voltage standing wave ratio (VSWR) obtained for the design.




(a) (b)

Figure 1: Schematic representations of breast, antenna and tumor showing signal scattering in (a) a normal breast and (b) a breast with a tumor.




(a) (b)

Figure 2: (a) Antenna and breast layers modeled in Ansoft Designer. (b) VSWR vs. frequency plot of the antenna system (including breast layers) as shown in (a).

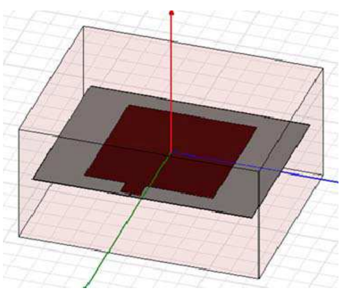
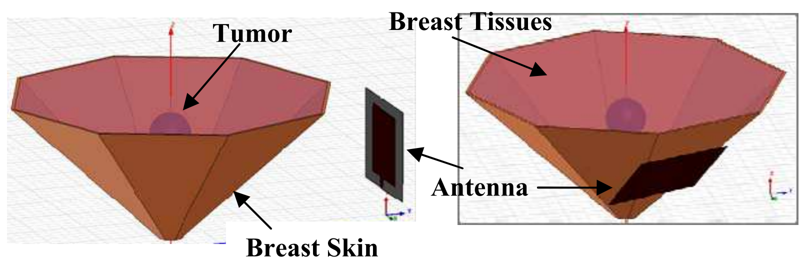


Figure 3: Rectangular microstrip antenna model in HFSS.




(a) (b)

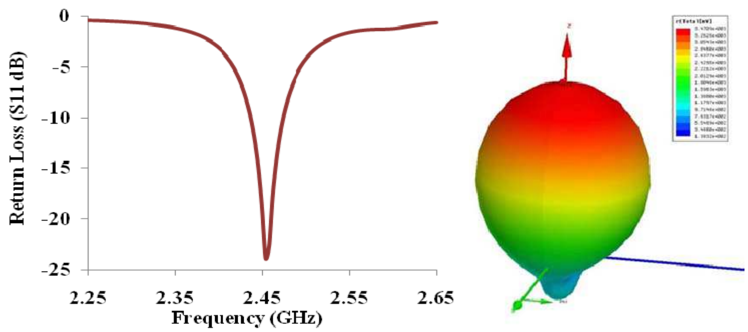
Figure 4: The developed breast models (with a tumor at the center) simulated in HFSS. (a) Antenna placed 5 cm away from the breast (10 cm from the center), case-I. (b) Antenna placed in contact with the breast (on the breast surface), case-II.

A rectangular microstrip antenna designed in HFSS was considered. A flexible copper substrate (Pyralux FR9151 CU CLAD) was selected as the antenna substrate. With the substrate parameters given, a rectangular antenna for 2.45 GHz with 50 Ω input impedance was designed (shown in Figure 3). With the help of the simulation software, the antenna was optimized to resonate at 2.45 GHz. The simulation and the experimental results were compared, and the simulation model was optimized to match with the experimental results. Breast model was developed in HFSS which is represented as a 10 cm wide and 5 cm high cone (with eight facets). The breast volume consist of breast skin of 5 mm width (dielectric coefficient, $\varepsilon_s = 39$, and conductivity, $\delta_s = 1.1\,\mathrm{S/m}$ [15]) and remaining volume consisting of breast fatty tissues (dielectric coefficient, $\varepsilon_f = 4.49$, and conductivity, $\delta_f = 0.59\,\mathrm{S/m}$ [16]). The developed breast model was simulated, (i) placing an antenna 5 cm away from the breast (case-I) and (ii) placing the antenna in contact with the breast (case-II), shown in Figures 4(a) and (b), respectively. The developed models were simulated for both the cases of antenna placement, for healthy breasts as well as for breasts with a tumor.

## 4. RESULTS AND DISCUSSION

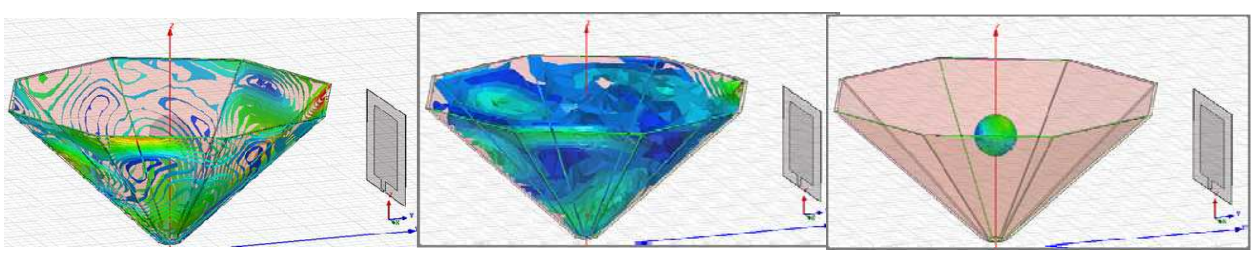
The rectangular microstrip antenna (shown in Figure 3) was simulated in Ansoft HFSS and various parameters of the antenna, including return loss, impedance, radiation pattern, and gain pattern were evaluated. After tuning the antenna for 2.45 GHz, the antenna was fabricated and tested. With the help of the experimental results the antenna simulation model was optimized to match with the experimental results. The return loss pattern and the radiation pattern of the optimized antenna are shown in Figures 5(a) and (b) respectively.

The optimized antenna model was used to simulate with the breast models developed in HFSS. The simulation was performed for the cases where the antenna was placed away (5 cm) from the breast (case-I, Figure 4(a)) and when it was placed on the breast surface (in contact with the breast skin, case-II, Figure 4(b)). Simulation models were analyzed for both the cases of healthy breast as well as for a breast with a tumor at the center. The tumor is represented as a 10 mm diameter sphere with dielectric permittivity $(\varepsilon_t) = 50$ and conductivity $(\delta_t) = 4\,\mathrm{S/m}$ [15]. When an antenna is placed on the surface of the breast, the performance characteristics of the antenna undergo changes. To minimize the effect of the breast skin contact with the antenna, the antenna was optimized to resonate again at 2.45 GHz. The optimization was performed with the help of the simulation model where the breast skin was considered as one of the substrate layer of the antenna. The optimized antenna was used for the simulation with the breast. As the signal scattered by an object depends on the induced current density in that object, the current density in the tumor as well as in the breast skin and fatty tissues are compared. The simulation results showing the



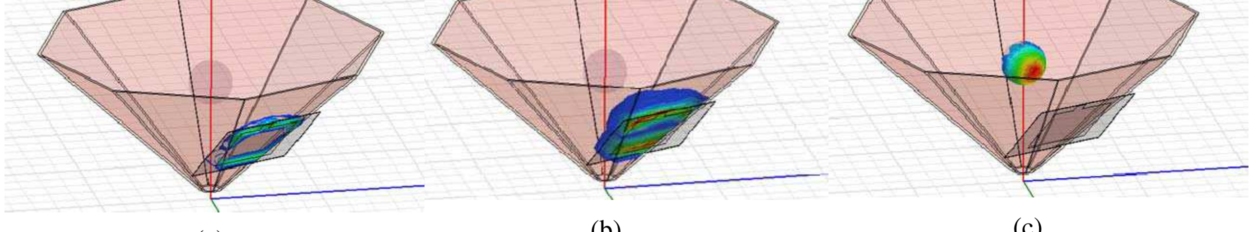

(a) (b)

Figure 5: Simulation results of the designed antenna. (a) Return loss pattern. (b) 3-D radiation pattern.



(a) (b) (c)

Figure 6: Simulated current densities in (a) skin, (b) fatty tissues, and (c) tumor, when the antenna is placed 5 cm away from the breast.

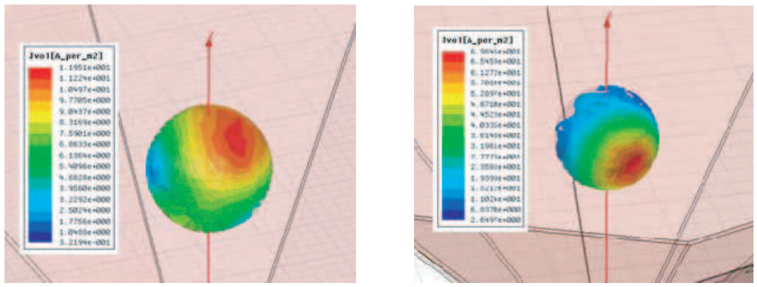


(a) (b) (c)

Figure 7: Simulated current densities in (a) skin, (b) fatty tissues, and (c) tumor, when the antenna is placed on the breast skin.

current densities in the breast skin, fatty tissues, and tumor, for both the cases (case-I and case-II) are presented in Figure 6 and Figure 7, respectively.

A comparison between the current densities in the breast tumor for the two cases (case-I and case-II) is shown in Figure 8. It is observed that the highest current density in the tumor in Case-II ($69.65\,\mathrm{A/m^2}$) is about six times higher than the current density in Case-I ($11.95\,\mathrm{A/m^2}$), which increases the visibility of the tumor to the antenna by the same factor.



(a) (b)

Figure 8: Current densities in breast tumors for, (a) case-I (maximum current density = $11.95\,\mathrm{A/m^2}$, (b) case-II (maximum current density = $69.65\,\mathrm{A/m^2}$).

The simulation results show that the visibility of the tumor or the sensitivity of the antenna to the tumor can be increased by placing an antenna in contact with the breast skin. The placement of the antenna on the breast skin affects the characteristics of the antenna. Thus, the antenna needs to be designed such that the effects of breast skin contacts are already considered in the design.

## 5. CONCLUSION

The simulation models developed and the results presented show that the sensitivity of the tumor detection increases when the antenna makes contact with surface of the breast as opposed to when it is placed away from the breast. It is due to the fact that, the breast skin has electrical properties comparable to the breast tumors and when the antenna is placed away, most of the received scattered signal is from the breast skin. This results in masking of the tumor scattered signal by strong skin scattered signals. For the case when the antenna is placed on the breast surface, the effect of the skin can be included in the antenna design, improving the signal strength received from the tumor. The simulation results show that the current densities inside the tumor is about six times higher when the antenna is in contact compared to when it is placed 5 cm away from the breast, and thus this provides us a proof of the concept presented. In this paper, the prospect of an antenna that can be placed on the breast skin has been presented and a simple antenna design for the case has been discussed. However, the design of an antenna which will be affected minimally when the properties of the skin vary, the experimental implementation of such an antenna, and the complete modeling and the experimental implementation of the system are still in progress. The results obtained from the further wok will be reported in the conference presentation.

## REFERENCES


1. "Cancer statistics 2009, a presentation from the american cancer society," American Cancer Soc., 2009.
2. "Global cancer, facts and figures 2007," National Cancer Inst., 2009.
3. Smith, R. A., D. Saslow, K. A. Sawyer, W. Burke, M. E. Costanza, W. P. Evans, R. S. Foster, E. Hendrick, H. J. Eyre, and S. Sener, "American cancer society guidelines for breast cancer screening: Updated 2003," *CA: A Cancer J for Clinician*, 141–169, 2003.
4. Epstein, S. S., R. Bertell, and B. Seaman, "Dangers and unreliability of mammography: Breast examination is a safe, effective, and practical alternative," *International Journal of Health Services*, 605–615, 2001.
5. "Screening for breast cancer: U.S. preventive services task force recommendation statement," *Annals of Internal Medicine*, Vol. 151, No. 10, 716–727, Nov. 2009.
6. Nass, S. J., I. C. Henderson, and J. C. Lashof, "Mammography and beyond: Developing technologies for the early detection of breast cancer," *National Cancer Policy Board, Institute of Medicine*, National Research Council, 2001.

7. Sabouni, A., S. Noghanian, and S. Pistorius, "Microwave tomography for breast cancer detection: Study of dispersion and heterogeneity effects," *Symposium on Antenna Technology and Applied Electromagnetics (ANTEM)*, Montreal, Canada, Jul. 16–19, 2006.
8. Fear, E. C., S. C. Hagness, P. M. Meaney, M. Okoniewski, and M. A. Stuchly, "Enhancing breast tumor detection with near field imaging," *IEEE Microwave Magazine*, Vol. 3, 48–56, 2002.
9. Poplack, S. P., T. D. Tosteson, W. A. Wells, B. W. Pogue, P. M. Meaney, A. Hartov, C. A. Kogel, S. K. Soho, J. J. Gibson, and K. D. Paulsen, "Electromagnetic breast imaging: Results of a pilot study in women with abnormal mammograms," *Radiology*, 350–359, 2007.
10. Fear, E. C. and M. A. Stuchly, "Microwave detection of breast cancer," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 48, 1854–1863, 2000.
11. Suri, J., R. M. Rangayyan, and S. Laxminarayan, "Emerging technology in breast imaging and mammography," *American Scientfic Publishers*, 2007.
12. Fear, E. C., "Microwave imaging of the breast," *Technology in Cancer Research and Treatment*, Vol. 4, No. 1, 69–82, Feb. 2005.
13. Meaney, P. M., M. W. Fanning, D. Li, S. P. Poplack, and K. D. Paulsen, "A clinical prototype for active microwave imaging of the breast," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 48, No. 11, 1841–1853, Nov. 2000.
14. Shrestha, S., M. Balachandran, M. Agarwal, L.-H. Zou, and K. Varahramyan, "A method to measure radar cross section parameters of antennas," *IEEE Trans. Antenna Prop.*, Vol. 56, No. 11, 3494–3500, Nov. 2008.
15. Padhi, S., F. Liu, B. K. Li, N. Shuley, and S. Crozier, "On the accurate modeling of a complex antenna for breast tumor detection using a hybrid MOM/FDTD approach," *Proc. of the 29th An. Int. Conf. of the IEEE EMBS*, 6636–6639, France, Aug. 23–26, 2007.

# Design of Small-sized and Low-cost Front End to Medical Microwave Radiometer

**O. Klemetsen[1], Y. Birkelund[1], P. F. Maccarini[2], P. Stauffer[2], and S. K. Jacobsen[1]**

[1]University of Tromso, Norway
[2]Duke University Medical Center, USA

**Abstract**— We have investigated the possibility of building a Dicke radiometer that is inexpensive, small-sized, stable, high sensitivity and consists of readily available microwave components. The selected frequency band is at 3–4 GHz and can be used for breast cancer detection, with sufficient spatial resolution. We have found microwave components that are small ($< 5\,\text{mm} \times 5\,\text{mm}$) and provide sufficient sensitivity. We have built two different Dicke radiometers: One is of conventional design with Dicke switch at front end to select antenna or noise rererence and the other with a low noise amplifier before the Dicke Switch. We have tested this concept with simulations and built prototypes. The two designs provide a gain of approximately 50 dB, and bandwidth of about 500 MHz. One of the designs has a stability $\mu > 1$ and the other design provide instability $\mu < 1$ for a part of the pass band. The prototypes are tested for sensitivity after calibration in two different known temperature waterbaths. The results show that the design with the low noise amplifier before the Dicke switch has 36% higher sensitivity than the other design with Dicke switch in front.

## 1. INTRODUCTION

Over the past four decades, research in microwave radiometry has been conducted for use in medical applications [1]. Detection of breast cancer, hyperthermia, inflammation and lung oedema are some applications for medical use [2]. Microwave radiometers can be assembled in different ways. The two most common are total power radiometer and Dicke radiometer [3]. A total power radiometer consists of a high gain low noise amplifier (LNA) followed by a power meter or a square law detector and integrator. This radiometer is very sensitive to amplification drift. A Dicke radiometer uses a switch in front of the LNA to select between the sensing antenna and a known noise reference, as shown in Fig. 1. When switching faster than the drift in gain, the drift is mostly canceled out [4]. The goal of the present design is to determine whether it is possible to create an inexpensive and small sized front stage to a radiometer, using available commercial surface mount device (SMD) components, that is stable against oscillations and has a usable sensitivity for radiometric measurement of temperature at deepth in tissue. It is previously shown that a LNA before the Dicke switch gives a better signal-to-clutter ratio (SCR) [5]. We want to find out if it is possible to have a LNA before the Dicke switch for both antenna and the reference signal.

## 2. THEORETICAL BACKGROUND

A microwave radiometer is an instrument that can measure temperatures inside the body. The measuring principle is to quantify the thermal emitted power over a given frequency band. Blackbody spectral radiance at temperature $T$ at all frequency is given by Planck's law. Planck's law

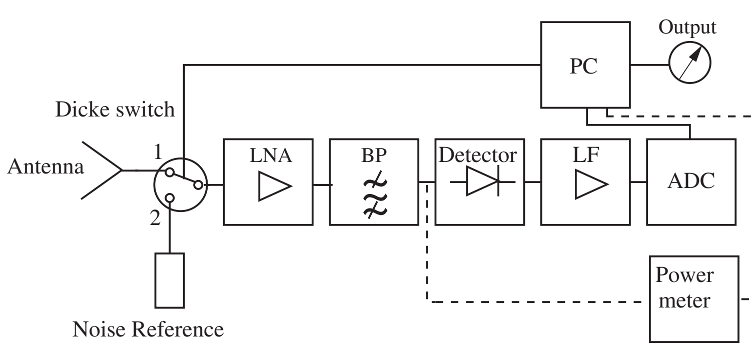



Figure 1: Block diagram of Dicke radiometer.

can be approximated by Rayleigh-Jean's law, and from which the noise power $P$ over a frequency band $\Delta f$ can be written as [4]:

$$P = kT\Delta f. \tag{1}$$

where $k$ is is Boltzmann's constant. The theoretical noise temperature $T_e$ of a cascade system is given by:

$$T_e = T_{e1} + \frac{T_{e2}}{G_1} + \frac{T_{e3}}{G_1 G_2} + \dots + \frac{T_{eN}}{G_1 G_2 \dots G_{N-1}} \tag{2}$$

where $T_{e,i}$, $G_i$ are the noise temperature and gain of the individual stages. From (2) the noise performance of the first stage is critical. The theoretical sensitivity of an ideal radiometer with no gain fluctuation is [4],

$$\Delta T = \frac{T_a + T_e}{\sqrt{\Delta f \tau}}, \tag{3}$$

where $\Delta T$ is the minimum detectable change in the radiometric antenna temperature $T_a$ and $\tau$ is the integration time in the low pass filter (integrator). For a Dicke radiometer (Fig. 1), the theoretical sensitivity is given by:

$$\Delta T_{\min} = \frac{2(T_a + T_e)}{\sqrt{\Delta f \tau}}. \tag{4}$$

A conventional Dicke radiometer front end has antenna, noise reference, Dicke switch, low noise amplifier and a bandpass filter (BP). The other part is the square law detector, low frequency (LF) amplifier, analog to digital converter (ADC) and a controller PC, or a power meter connected to the front end. With the Dicke switch attached directly to the antenna, the radiometer can be used as a total power radiometer.

A stable amplifier cascade configuration is important in order to avoid internal oscillation that interferes with the extremely low radiated power received by the antenna. The stability of a system can be described using the Edward-Sinsky stability factor defined as [6]:

$$\mu = \frac{1 - |S_{11}|^2}{|S_{22} - S_{11}^* \cdot \Delta| + |S_{12} \cdot S_{21}|} > 1 \tag{5}$$

where $S_{ij}$, $\Delta$ are elements and the determinant of the $S$ parameter matrix, respectively. The system is unconditionally stable if $\mu > 1$.

## 3. RADIOMETER DESIGN

Equation (1) shows that the noise power $P$ depends primarily on the temperature $T$ and bandwidth $\Delta f$. For human body temperatures, and for a practical bandwidth of 300–500 MHz, the noise power over that bandwidth is low. To detect this low power, it is crucial to have a sensitive instrument like a radiometer. From previous work [5], we have found that the frequency band around 3.5 GHz is disturbed less by EMI than other low GHz frequency bands. This frequency range provides lower penetration depth in the human body compared to lower GHz frequencies. Howerver of this band has been show suitable for detecting superficial breast cancer [7]. The challenge is to find a LNA balancing the trade offs between lowest possible noise, highest gain, low power consumption and low cost. Several LNAs were considered, but our choice was a LNA from Hittite. A combination of low

Table 1: Specification of used SMD components.

| Device | Name | Gain [dB] | NF [dB] | $T_e$ [K] | Frequency [GHz] |
|---|---|---|---|---|---|
| LNA | HMC593LP3E (Hittite) | 19 | 1.2 | 92.3 | 3.3–3.8 |
| BP (LP part) | LFCN-3400+ (Mini-Circuits) | −1.03 | 1.03 | 77.6 | 0–3.4 |
| BP (HP-part) | HFCN-3100+ (Mini-Circuits) | −1.01 | 1.01 | 75.9 | 3.4–9.9 |
| Dicke switch | CSWA2-63DR+ (Mini-Circuits) | −1.2 | 1.2 | 92.3 | 0.5–6 |
| DC block capacitor | GQM1885C1H470JB01 (Murata) | −0.014 | 0.014 | 0.94 | 3.5 |

pass (LP) and high pass (HP) filters gave appropiate 500 MHz band width. The steep slope of the frequency cutoff at the edges of the pass band was further enhanced by the use of two consecutive filters of the same type. The Dicke switch was another challenge of the design. We want a Dicke switch with lowest possible insertion loss, high isolation, 3.3–3.8 GHz bandwidth, small size, and single supply voltage. Our choice was a switch from Mini-Circuits. The switch requires a DC block capacitor 47 pF before and after. Our choise was a capacitor from Murata. A single LNA does not provide enough gain to get the power into the required range. Based on the component specifications listed in (Table 1), for the LNA, filter and switch we find that a configuration of three LNAs ($3 \times 19$ dB) and four filters ($4 \times (-1)$ dB) should give a reasonable amplification in this high frequency part. The bandpass filter gives a theoretical bandwidth $\Delta f = 500$ MHz.

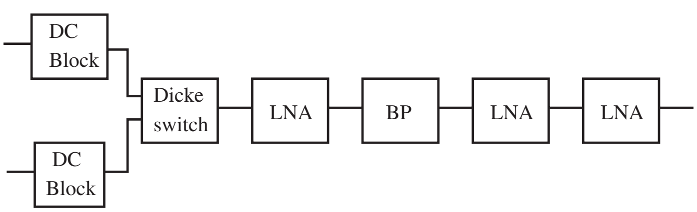
### 3.1. Different Design

The first stage of the radiometer has many issues that can cause problems; oscillation, stability, system noise, temperature and gain drift and electromagnetic interference (EMI). An optimum radiometer front end has high sensitivity, low noise temperature, stable, low gain drift and low surface component temperature. For flexibility and comparability a Dicke front end radiometer should also be able to run as a total power radiometer. A full radiometer can be designed in many ways. We want to design the Dicke concept with use of a powermeter and a PC with Labview as a post-detector and to drive the Dicke switch. The design was first simulated. Design #1: Conventional Dicke radiometer with the DC block capacitor and Dicke switch in front, one LNA, the bandpass filter and 2 LNA in cascade as a booster amplifier as shown in (Fig. 2(a)). Design #2: The proposed Dicke radiometer with the LNA in front, the switch, 2 LNA in cascade and the bandpass filter as shown in Fig. 2(b). The last design does not need DC block capacitors because there is an internal DC block in the LNA. The radiometers have been characterized using $S$-parameters, system noise (2), sensitivity (3), stability factors (5), mean and standard deviation of a known temperature source.

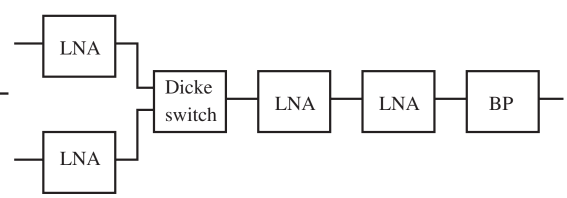
## 4. RESULTS

### 4.1. Simulation

In the simulation, $S$-parameters were given by touchstone files from the component manufactuer's web site. Simulations were performed by importing individual touchstone files for every single block in the design, and simulating the circuit's $S$-parameters over the frequency 1–6 GHz. The theoretical noise temperature of the cascade system given by (2) and parameters used from (Table 1)




(a) Design #1.




(b) Design #2.

Figure 2: Block diagram of design #1 and #2.

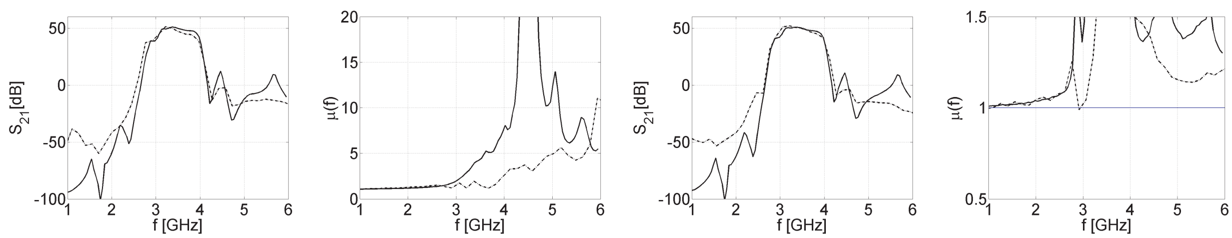
Table 2: Theoretical, simulation and measurement parameters for the radiometers.

| | Parameter | design #1 | design #2 |
|---|---|---|---|
| Theoretical system noise temperature | $T_e$ | 217.2 K | 95.1 K |
| Simulated noise power from $S$-parameters | $P_s$ | −34.25 dBm | −34.01 dBm |
| Simulated from measured $S$-parameters | $P_{sm}$ | −35.26 dBm | −33.40 dBm |
| Measured noise power | $P_m$ | −33.50 dBm | −32.80 dBm |
| Theoretical sensitivity, total power radiometer | $\Delta T$ | 0.016 K | 0.012 K |
| Theoretical sensitivity, Dicke radiometer | $\Delta T_{\min}$ | 0.032 K | 0.024 K |
| Measured sensitivity, total power radiometer | $\Delta T$ | 0.064 K | 0.046 K |
| Measured sensitivity, Dicke radiometer | $\Delta T$ | 0.098 K | 0.062 K |

gave $T_e$ for each design. The system noise temperature is clearly dependent on the first stage. The expected output power from the system when using $T_a = 290\,\mathrm{K}$ can be found by integrating over frequencies in the simulated system $|S_{21}|^2$ and multiplying by $k(T_a+T_e)$, given by (1) this will gives the simulated power in Watt, as in: $P_s = 10\log\left[\sum |S_{21}|^2 k(T_a + T_e)\right]$ in dBm. The theoretical noise temperature using (2), expected output power, and sensitivity using (3), (4) are listed in (Table 2).

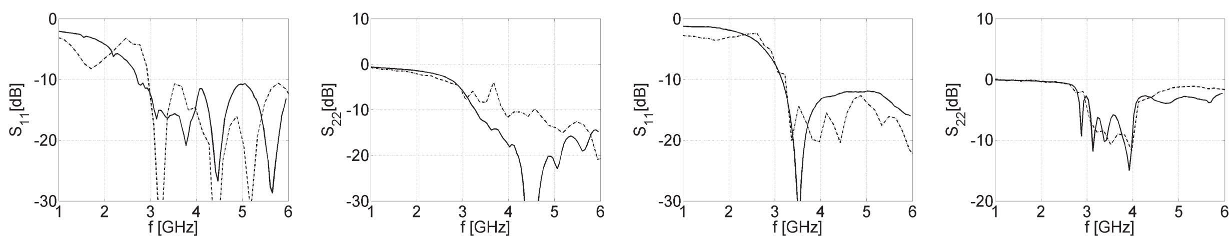
### 4.2. Prototype Testing

The prototype design was built with a printed circuit board (PCB) RO4350B, thickness of 0.254 mm and a copper layer of 35 microns. The system was tested in the Agilent- E5071C network analyzer and E4419B power meter with the E4412A sensor. The Dicke switch is controlled by an USB-3114 device from Measurement Computing. The measurements are done with the front end blocks from the (Fig. 1) and shown in (Fig. 2). The measurment were carried out by generating touchstone files from network analyzers and calculates the stability and the expected output power $P_{sm}$. A time-series measurements was conducted with a $50\,\Omega$ load as antenna at a temperature of $T_a = 290\,\mathrm{K}$, connected to the power meter to get the system's operation over time and the measured power $P_m$ and other system operation parameters listed in Table 2. The simulated and measured forward transmissin $S_{21}$ and stability factor $\mu$ are given in (Fig. 3) for the two designs. The input reflection coefficient $S_{11}$ and output reflection coefficient $S_{22}$ are given in (Fig. 4). Sensitivity of the radiometer design is found by calibrating the radiometers with two different and known temperatures in water bath, $T_h$, $T_c$. Calibration parameter $s = \frac{P_h - P_c}{T_h - T_c}$ is the slope of the output power $P_h$ and $P_c$ as




(a) #1. $S_{21}$ (b) #1. $\mu$ (c) #2. $S_{21}$ (d) #2. $\mu$

Figure 3: $S_{21}$ and $\mu$ for design #1 and #2. Solid line: Simulated values. Dotted line: Measured value.




(a) #1. $S_{11}$ (b) #1. $S_{22}$ (c) #2. $S_{11}$ (d) #2. $S_{22}$

Figure 4: $S_{11}$ and $S_{22}$ for design #1 and #2. Solid line: Simulated values. Dotted line: Measured value.

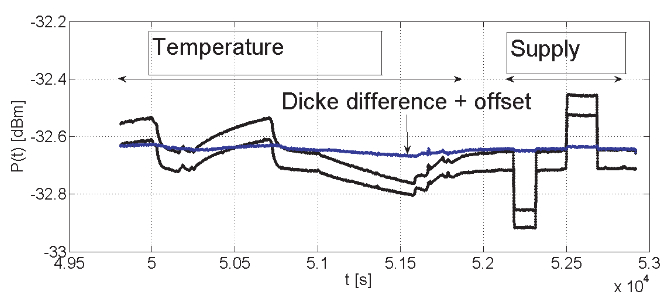



Figure 5: Measurement of the Dicke concept. The first part is the temperature change with a Peltier element of the radiometer temperature. The last part is a voltage change of $+/-2\%$.

a function of temperature $T_h$, $T_c$. We can then find the sensitivity $\Delta T = \frac{\sigma_{P_o}}{s}$ for the prototype radiometers and are given in (Table 2). $\sigma_{P_o}$ is the measured standard deviation to the measured output power $P_o$. The Dicke consept was also testet for gain variation caused by temperature changes on the radiometeret performed with a Peltier element and by changing the supply voltage (Fig. 5).

## 5. DISCUSSION

The proposed designs gave the expected gain over the selected frequency band. Design #1 was stable $\mu > 1$ for all frequencies. The other design #2 was instable $\mu = 0.99$ for frequency 2.91 GHz and stable for all the other frequencies. The instability can be caused by parasitic effects between the different LNA's. The problem can be solved by increasing the isolation between the different LNA's. This isolation can be performed with attenuators (pads) or by placing the various selected parts of the bandpass filter between the LNA. The input reflection coefficient is in the expexted range. The output reflection coefficient (#1) has a peak in the center frequency 3.5 GHz, and care must be taken for the following stage, usually a square law detector. The noise temperature $T_e$ was better for design #2 than for design #1 and was expected since the first LNA in design #2 has a gain of 19 dB. The chosen thichness of the PCB is a challenge, because the soldered connector are easily broken apart from the PCB, and a new design requires a multi-layer PCB for better mechanical stability. Although the design is chosen for minimum power consumption, it requires a heatsink for the least possible influence from ambient temperature that gives unwanted temperature variations and thus gain drift. And in the design where the LNA are in the front, it is important that they has the same temperature and are drifting in the same way since the Dicke consept do not work at those. Ongoing research is looking at the possibility of using small heatpipe to conduct heat away from the circuit as a possible solution to reduce temperature drift. Sensitivity for the prototype was worse than expected, and can come from manual soldering of SMD components with more solder flux than in a professional soldering process. The design where the low noise amplifier is before the Dicke switch for both antenna and reference input, has 36% better sensitivity than the design where the Dicke switch is in front. From (Fig. 5) we can conclude that the Dicke consept is working as expected. The output from the Dicke radiometer is stable when we are changing the temperature of the radiometer and supply voltage.

## REFERENCES


1. Mizushina, S., H. Ohba, K. Abe, S. Mizoshira, and T. Sugiura, "Recent trends in medical microwave radiometry," *IEICE Trans. Commun.*, Vol. E78-B, No. 6, 789–798, 1995.
2. Leroy, Y., A. Mamouni, J. V. D. Velde, B. Bocquet, and B. Dujardin, "Microwave radiometry for non-invasive thermometry," *Automedica*, Vol. 8, 181–202, 1987.
3. Aitken, G. J. M., "A new correlation radiometer," *IEEE Transactions on Antennas and Propagation*, Vol. 16, No. 2, Mar. 1968.
4. Ulaby, F., R. Moore, and A. Fung, *Microwave Remote Sensing, Microwave Remote Sensing Fundamentals and Radiometry*, 1st Edition, Vol. 1, Artec House, 685 Canton Street, Norwood, MA 02062, USA, 1981.
5. Jacobsen, S. and O. Klemetsen, "Improved detectability in medical microwave radio- thermometers as obtained by active antennas," *IEEE Transactions on Biomedical Engineering*, Vol. 55, No. 12, 2778–2785, Dec. 2008.
6. Edwards, M. and J. Sinsky, "A new criterion for linear 2-port stability using a single geometrically derived parameter," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 40, No. 12, 2303–2311, Dec. 1992.
7. Bocquet, B., J. C. van de Velde, A. Mamouni, Y. Leroy, G. Giaux, J. Delannoy, and D. Delvalee, "Microwave radiometric imaging at 3 GHz for the exploration of breast tumors," *IEEE Trans. Microw. Theory Tech.*, Vol. 38, No. 6, 791–793, Jun. 1990.

# The Application of the Hilbert-Huang Transform in Through-wall Life Detection with UWB Impulse Radar

**Zijian Liu**[1], **Lanbo Liu**[1, 2], **and Benjamin Barrowes**[2]

[1]School of Engineering, University of Connecticut, Storrs, CT 06269-2037, USA
[2]USACE Engineer Research and Development Center
Cold Regions Research and Engineering Laboratory, Hanover, NH 03755, USA

**Abstract**— Hilbert-Huang Transformation (HHT) is a powerful tool for nonlinear and non-stationary data analysis. In this paper, a dataset using an ultra-wide (UWB) impulse radar system with central frequency of 1 GHz was collected for life motion detection behind a cinder block wall. To extract the information of life motions such as breathing and heartbeats from the raw data, we first applied the empirical mode decomposition (EMD), the first step of HHT to decompose the signal (background signal included) into a family of the intrinsic mode functions (IMFs). We then apply Hilbert spectral analysis (HSA) to get the frequency spectra of different IMFs. After dividing by the spectrum of the background radar record (equivalent to de-convolving the background record in the time domain), we found that breathing appear as a spectral peak at 0.2–0.4 Hz and heart beating appears as 1.0–1.2 Hz. This is coinciding with real condition. Our preliminary results show that the HHT technique provides significant assistance in signal processing for the detection of human targets behind opaque obstacles.

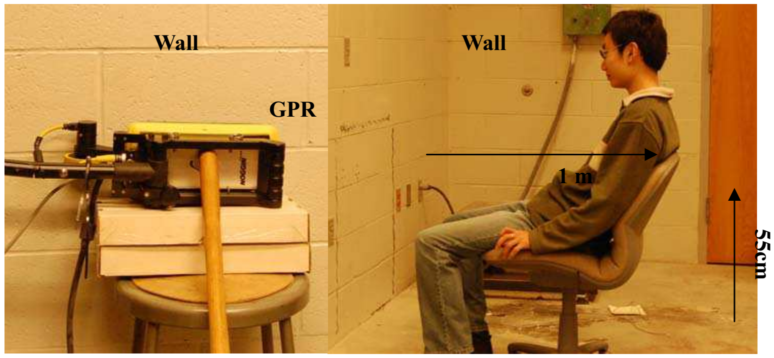
## 1. INTRODUCTION

Hilbert-Huang Transformation (HHT) is a novel digital signal processing technology based on the combination of the empirical mode decomposition (EMD) and the Hilbert spectral analysis (HSA). It is designed specifically for analyzing nonlinear and nonstationary data [1]. In this paper we first design a through-wall life detection experiment and then using HHT to analysis the dataset collected by ultra-wide band (UWB) impulse radar system. Our effort contributes to the research of live human beings detection, the one is becoming increasing important in modern society. Its purpose is to identify life-being located behind the obstacle, which has significant meaning on the application of military, rescue operations under conditions of extraordinary situations or other related field. In common condition, even a person stays quietly without doing anything can still appear to have tiny movement caused by breath and heart beat. These displacements can be captured by EM wave detection. Some research shows that UWB impulse radar system was efficient of capturing human breath and heart-beat movement [2, 3]. Based on these, E. Zaikov explored an experiment using UWB radar for trapped- people detection [4]. In which the tiny breath and heart-beat movement was easily submerged by noise when people was motionless. By applying HHT preliminary into analysis, Ram M. Narayanan [5] presented the results with several potential breath-signal peaks in frequency domain in 0.4–0.8 Hz. However real breath signal still can not be localized; at the same time the extraction of heart beat, only 1–2 mm chest movement related, is failed to achieve. Here we are striving to overcome these difficulties.

## 2. EXPERIMENT DESIGN

The through-wall human cardioasperation detection experiments were conducted in a laboratory with ciderbrick walls (no reinforced steel bars) as shown in Figure 1. The Ground Penetration Radar (GPR) (Sensors and Software 1-GHz Nuggins system) is used as the UWB impulse source and receiving system. The recording time window length for each recording trace is chosen to be 16 ns, long enough for the radar waves to be reflected from an object in the air within 2-meter radius. GPR is located with its antenna closely stick to one side (Figure 1(a)) of the wall. On the other side of the wall, a chair is put in front of the wall, the distance between which and wall is set to be 1 meter. The subject we choose has good health condition with average heart beat frequency 1.05–1.2 Hz (65–72 beats/min). The distance between subject's cardiac and ground floor is measured to be 55 cm, when he is sitting in the chair. Taking this as a reference, the location of GPR is set to validate that both antenna and the cardiac stay in same horizontal line (Figure 1(b)).

The GPR system collected 1 recording trace per 0.1 sec, with 8 stacking to minimize the random error. At the beginning of the experiment, subject is absent; data with the 1024 recording traces




(a) (b)

Figure 1: Photos of experiment setup. (a) and (b) are device set on two sides of the wall, the location of GPR is set to validate that both antenna and the cardiac stay in same horizontal line.

(nearly two minutes) is collected by GPR as a background reference signal. It will be used in data analysis later. After that, subject is asked to sit peacefully on the chair with his shoulders and back tightly against the back of the chair to stop body displacement. Now we collect data in three statuses: normal breathing, holding their breath, and repeatedly speaking the words "one, two, three". The breath frequency of normal breath and speaking are measured to be around 0.2–0.4 Hz (12–24 times/min) and 0.4–0.5 Hz (24–30 times/min) respectively. And for breath holding, breath frequency can be seen as 0. For the first and the third status the total number of recording traces is well above 1024, for about 2 minutes. For the status of holding breath the number of recording traces is less than 256, i.e., generally less than 30 seconds.

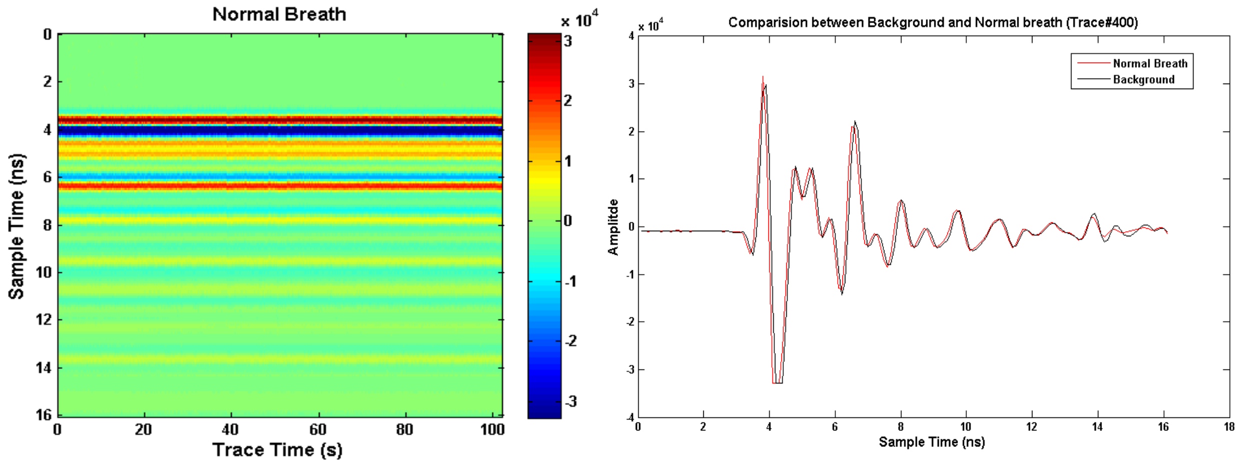
## 3. HHT DATA PROCESSING AND RESULT ANALYSIS

### 3.1. Empirical Mode Decomposition (EMD) and Back Ground De-convolution

In experiment mentioned before, four datasets are totally collected: background signal, normal breathing, breath holding and speaking. Here taking dataset of normal breathing (Figure 2(a)) as an example, which is made of 1024 time traces. Each trace is generated from one pulse. Because the time period between two traces is 0.1 s, in $X$-direction the total measured time is 102.4 s. The $Y$-direction shows the length of each trace. 161 points is contained in one trace with the time window of 16.1 ns, so the time period between two points is 0.1 ns in $Y$ direction. In Figure 2(a), with the existence of wall strong reflection appears from 0.3 ns to 0.7 ns. With the consideration of void velocity of EM wave and the distance of 1 m between wall and human, the reflection signal of human physiology movement should appear after 12 ns. This is also shown in Figure 2(b) precisely, taking the example of trace No. 400, when compare this trace between normal breath and background reference, obvious phase distortion appears after 14 ns. This is caused by reflection and velocity change when EM wave passing human body. Based on this, we only focus the dataset from 14 ns to 16 ns, which includes signal of breathing and heart beating. However, due to multiple refection from wall and test environment, data collected is too complex to recognize human physiology movement directly, which is a relative small portion submerged into complex background. If this complex signal can be divided into several simple patterns, the small target portion we concerned will be easier to be extracted.

Here we apply EMD, one step of HHT into data analysis. EMD sees a signal trace as a muster of many coexisting simple oscillatory modes with the same number of extreme and zero crossings, symmetric envelopes and significantly different frequencies. These modes are called intrinsic functions (IMFs). What EMD does is to separate those IMFs from the original signal one by one, until the residue is monotonic [1]. Now EMD is applied to the data with the sample time from 14ns to 16ns, and the same IMFs of different sample time are summated and averaged. By doing this, the signal noise ratio of signal is increased and the uncertainty of signal is reduced. At the same time, dataset of background reference is derived into IMFs as the same way as what mentioned before. After letting each IMF of normal breath de-convoluted by identical IMF from background, we derive the result of normal breath with 7 IMFs included (Figure 3).

From IMFs of normal breath (Figure 3), we can find out that EMD always extract IMF from the highest frequency to lowest. The tiny peaks with relative high frequency appearing on IMF7 and




(a) (b)

Figure 2: (a) is the dataset of normal breathing, in which $X$-axis represents different traces from No. 1 to No. 1024, and the time period between two traces next to each other is 0.1 s. $Y$-axis shows the length of each trace. 161 points is contained in one trace with the total time window of 16.1 ns. (b) is one trace plot for trace No. 400 from the datasets of both normal breath (red) and background (blue). Phase distortion appears after sample time of 14 ns.

IMF5 are totally coming from errors of de-convolution rather than EMD. Here the first component, IMF1 appears to have a similar shape of radio wave, and its frequency is much higher than any kinds of physiological movements. It may mainly relate to the carrier wave radiated from GPR, the same as IMF2. From IMF3 and IMF4, 9–13 peaks can be counted in every 10 seconds; this coincides with the frequency of heart beating (1.1–1.2 Hz). However, as the displacement caused by cardiac movement is tiny, distinct feature can not be identified in either IMF3 or IMF4. In IMF5 and IMF6, the movement with obvious amplitude and every two to three peaks in 10 seconds is clearly shown. We believe this is caused by breath movement (0.2–0.4 Hz), the most obvious movement in chest area. In IMF7, both amplitude and frequency is very low, this is not taking into consideration.

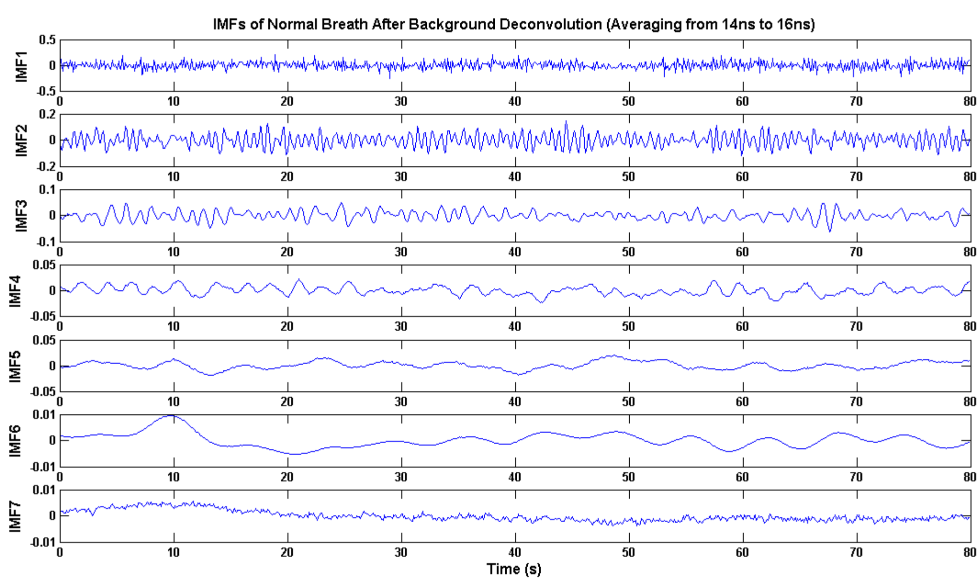



Figure 3: IMFs of normal breath after background de-convolution.

### 3.2. Hilbert Spectral Analysis (HSA)

EMD is the data processing approach used in time domain. It is not enough for further quantitative analysis for physiology features we concerned, especially for heart beating with tiny amplitude. So it is necessary to transform IMFs we derived into the spectrum of frequency domain. As to this point, traditional method used by most people before is fast Fourier transformation (FFT). But for non-stationary and instantaneous signal analysis, HSA, the other step of HHT, is a better solution than FFT. The principle of HSA is clearly shown in [Huang and Wu, 2008] [1]. Here we make a comparison between HSA and FFT by using original (without de-convolution) IMF 5 of normal breath and breath holding, the result is shown in Figure 4. It is clear that for both normal breathing and breath holding, the peaks derived by HAT are narrower and more identical than which derived from FFT. Another advantage coming from HSA is that the shape of the spectrum will not be significantly effected by the length of original trace in time domain. This can be seen from breath holding in Figure 4(b), in which the time window is only 256 points. After 0.3 Hz, compare to continuous fake peaks generated by FFT, spectrum of HSA is much more smooth and realistic.

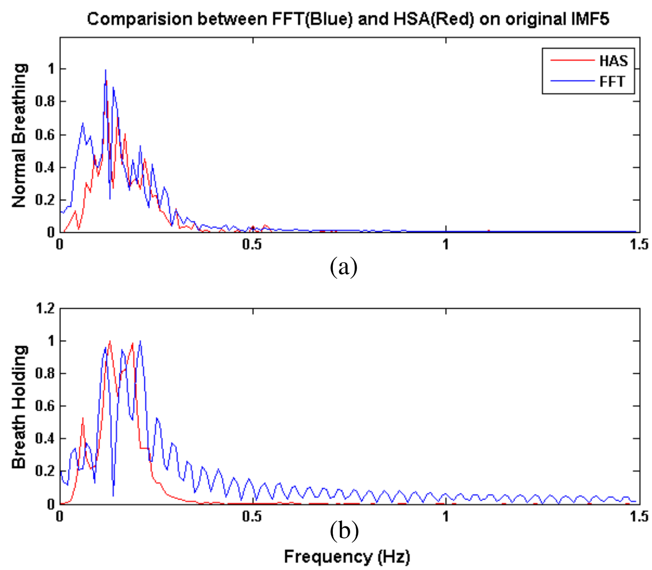
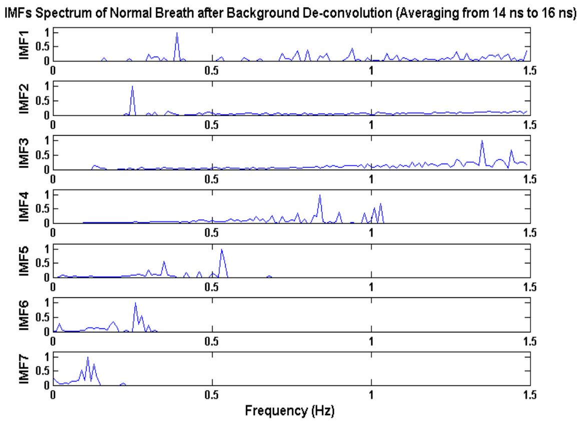



Figure 4: Comparison between FFT (blue) and HSA (red) spectrum of Original (without de-convolution) IMF5. (a) displays differences for dataset of normal breathing. (b) shows the condition of breath holding.

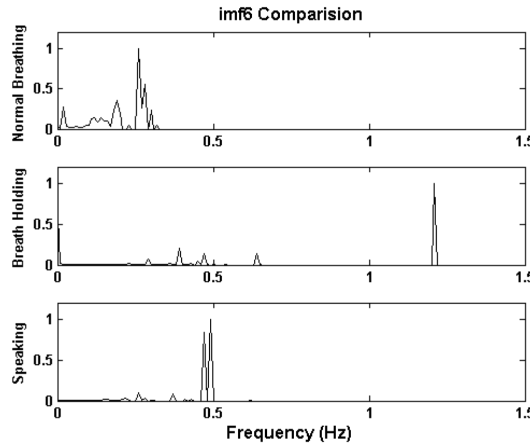
By using HSA on IMFs shown in Figure 3, we get frequency spectrum of 7 IMFs shown in Figure 5(a). Note that in IMF 4 obvious peak appears around the location of 1.05 Hz. We believe this can be seen as the feature of subject's heartbeat. Frequency component with large amplitude of 0.3 Hz is observed in IMF 2, 6 and IMF 1, 3, 4, 5, 7 with tiny peak. This may relates to human breathing. Besides these, there are still others peaks appear in different IMFs, the prediction of the origin can come about harmonic frequency, suppose the frequency of heart beating and breathing are 1.1 Hz and 0.3 Hz respectively. The production of these two movements will produce new frequency of 1.4 Hz and 0.8 Hz. This can explain the peaks appear in IMF3 and IMF4. If the harmonic wave is strong enough, it may effect on the original wave to produce secondary harmonic wave, such as the peak of IMF5 in 0.5 Hz. This prediction is waiting for further validation.

### 3.3. IMFs Compare among Normal Breathing, Breath Holding and Speaking

After processing datasets of breathing holding and speaking in the same way illustrated in part 3.2 and 3.3, we get frequency spectrum of 7 IMFs for two statuses respectively. After making comparison among statuses for each IMF, we observe obvious differences on chest movement appear in IMF6 (Figure 5(b)). When compare with normal breathing, the frequency of 0.3 Hz disappear in the spectrum breath holding. This is believed to cause by the seizing of chest movement. On the contrast, the higher peak of breathing appears at 0.5 Hz at the spectrum of speaking shows that when people speak, both the amplitude and frequency of chest movement increase.




(a)




(b)

Figure 5: (a) is frequency spectrum of 7 IMFs for normal breathing derived by HSA. Frequency component of breathing is obvious in IMF2 and IMF6, and heart beat can be seen in IMF4. (b) shows comparison on IMF6 among normal breathing, breath holding and speaking. With seizing chest movement, the frequency of 0.3 Hz, which coincides with breathing, disappears in the middle spectrum.

## 4. CONCLUSIONS

In this paper it proves that by using HHT, the combination approach of EMD and HSA, feature of human breathing and heart beating can be successfully extracted from datasets collected by UWB impulse radar system. For the frequency spectrum derived by HSA from different IMFs, distinct frequency relates to different physiology movement can be well recognized. After analysis, these components are highly coinciding with the real human movement. This shows the significant potential of HHT approach for through-wall life detection improvement.

### REFERENCES

1. Huang, N. E. and Z. Wu, "A review on Hilbert-Huang transform: Method and its application, to geophysics studies," *Rev. Geophys.*, Vol. 46, No. 2, RG2006, 2008.
2. Bugaev, A. S., V. V. Chapursky, S. I. Ivashov, V. V. Razevig, A. P. Sheyko, and I. A. Vasilyev, "Through wall sensing of human breathing and heart beating by monochromatic radar," *Proceedings of Tenth International Conference on Ground Penetrating Radar*, Vol. 1, 291–294, June 2004.
3. Yarovoy, A. G., L. P. Ligthart, J. Matuzas, and B. Levitas, "UWB radar for human being detection," *Aerospace and Electronic Systems Magazine, IEEE*, Vol. 21, 22–26, November 2006.
4. Zaikov, E., J. Sachs, M. Aftanas, and J. Rovnakova, "Detection of trapped people by UWB radar," *German Microwave Conference, GeMiC, 2008*, 240–243, Hamburg, Germany, Mar. 2008.
5. Narayanan, R. M., "Through-wall radar imaging using UWB noise waveforms," *Journal of The Franklin Institute*, Vol. 345, 659–678, 2008.
6. Liu, L., R. Mehl, W. Wang, and Q. Chen, "Applications of the hilbert-huang transform for microtremor data analysis enhancement," *Journal of Environmental and Engineering Geophysics*, in review.

# Compressive Through-focus Imaging

**Oren Mangoubi**[1] **and Edwin A. Marengo**[2]

[1]Yale University, USA
[2]Northeastern University, USA

**Abstract**— Optical sensing and imaging applications often suffer from a combination of low resolution object reconstructions and a large number of sensors (thousands), which depending on the frequency can be quite expensive or bulky. A key objective in optical design is to minimize the number of sensors (which reduces cost) for a given target resolution level (image quality) and permissible total sensor array size (compactness). Equivalently, for a given imaging hardware one seeks to maximize image quality, which in turn means fully exploiting the available sensors as well as all priors about the properties of the sought-after objects such as sparsity properties, and other, which can be incorporated into data processing schemes for object reconstructions. In this paper we propose a compressive-sensing-based method to process through-focus optical field data captured at a sensor array. This method applies to both two-dimensional (2D) and three-dimensional (3D) objects. The proposed approach treats in-focus and out-of-focus data as projective measurements for compressive sensing, and assumes that the objects are sparse under known linear transformations applied to them. This prior allows reconstruction via familiar compressive sensing methods based on 1-norm minimization. The proposed compressive through-focus imaging is illustrated in the reconstruction of canonical 2D and 3D objects, using either coherent or incoherent light. The obtained results illustrate the combined use of through-focus imaging and compressive sensing techniques, and also shed light onto the nature of the information that is present in in-focus and out-of-focus images.

## 1. INTRODUCTION

In traditional analog imaging, images are only acquired in focus, discarding additional information present in out-of-focus images. Recent research [1] suggests that one can significantly increase the amount of object information collected per detector by capturing images for not one but several focal planes (through-focus imaging). In conventional imaging one usually places the object in focus, and captures the respective image in the associated image plane. However, this may require using a large sensor array. If, in addition, one captures out-of-focus data, then the number of sensors can be reduced while maintaining the same image quality as the in-focus case. Similarly, in imaging three-dimensional objects one usually employs a particular "best focal plane" and captures the respective in-focus image. However, if one captures data for other focal planes then one can achieve comparable resolution as the "best focal plane" case with less sensors, and fully 3D imaging.

If, in addition, the object under investigation is known to be sparse when represented in a given basis or dictionary, or generally under a given linear transformation applied to it (such as the gradient operator, as pertinent to certain piecewise constant objects [2]), then one can implement compressive sensing inversion algorithms [2,3] to increase the resolution per sample ratio. We propose a method that treats the information in multiple through-focus images as projective measurements for compressive sensing, allowing a greater resolution per detector ratio than possible with either conventional through-focus imaging [1] or compressive sensing (of conventional in-focus data) alone. The proposed compressive through-focus imaging is illustrated in the reconstruction of canonical 2D and 3D objects, using either coherent or incoherent light. The obtained results illustrate the combined use of through-focus and compressive sensing techniques, and shed light onto the nature of the information that is present in in-focus and out-of-focus images. Information about sparse objects appears to be concentrated in completely out-of-focus planes for coherent light and in near-focus planes for incoherent light.

## 2. OPTICAL SYSTEMS

We consider a general imaging system that is characterized by unit-impulse response or Green's function $h(\mathbf{r}, \mathbf{r}'; p)$ where $\mathbf{r}$ and $\mathbf{r}'$ denote image and object coordinates, and $p$ denotes system parameters. For example, for the simple lens system in Figure 1, $p = (f, z_1, z_2)$ where $f$ denotes the lens focal length, $z_1$ is the distance from the object plane (for a 2D object) or a given plane in the object (for a 3D object) to the lens, and $z_2$ is the distance from the lens to the detector

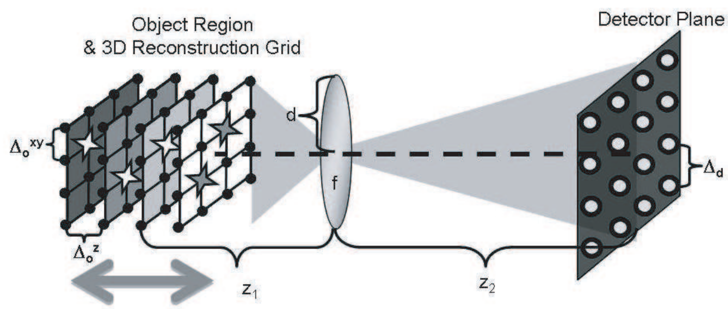



Figure 1: Lens-based through-focus imaging system for imaging of 2D or 3D objects.

plane. In the following we will explain the proposed compressive through-focus imaging assuming the particular lens-based system in Figure 1; however, clearly the key idea in this paper, that of using reconfigurable system states as a way to create compressive measurements, applies to more general systems as long as they exhibit degrees of (controllable) reconfigurability. By capturing data in different sensor positions and for different system configurations, and processing the data holistically, including priors, it is possible to maximally exploit the available sensing resources. Next the focal length is assumed to be constant while the lens position is varied to create different system states and thereby capture data corresponding to different states.

Two modalities are of interest: coherent and incoherent imaging. In coherent imaging, involving, e.g., a secondary source that is induced at a scatterer in its interaction with coherent light (due to a coherent source like a laser), the field at a detector at position $(u, v)$ in the detector plane that is due to an extended object characterized by object wavefield $U_{\text{obj}}(\mathbf{r}')$ is given by

$$U_{\text{det}}(u, v; p) = \int d\mathbf{r}' h\left[(u, v)\,,\, \mathbf{r}';\, p\right] U_{\text{obj}}\left(\mathbf{r}'\right). \tag{1}$$

Above the object coordinates $(\mathbf{r}')$ are in 2D space for thin (2D-approximable) objects such as transparencies and in 3D space for more general 3D objects. The detectors measure only the magnitude of this field, but by using reference beams one can also measure the phase. For incoherent imaging, involving primary or secondary incoherent sources, the corresponding relation is

$$I_{\text{det}}(u, v; p) = \int d\mathbf{r}' \left|h\left[(u, v),\, \mathbf{r}';\, p\right]\right|^2 I_{\text{obj}}\left(\mathbf{r}'\right) \tag{2}$$

where $I_{\text{det}}(u, v) = \langle |U_{\text{det}}(u, v)|^2\rangle$ and $I_{\text{obj}}(\mathbf{r}') = \langle |U_{\text{obj}}(\mathbf{r}')|^2\rangle$, where $\langle\cdot\rangle$ denotes average. In through-focus imaging, the data are captured for several "in" and "out-of" focus states, as defined by the distances $z_1$ and $z_2$ which correspond to different positions of the lens relative to the object and the detector plane. The next section outlines how the data are processed to create images.

## 3. COMPRESSIVE IMAGING

Importantly, in the coherent and incoherent optical systems described by (1) and (2) the mapping from the object function (i.e., the wavefield $U_{\text{obj}}$ in the coherent case and the intensity $I_{\text{obj}}$ in the incoherent case) to the data is linear. Then we can interpret the data as linear projections of the object to be imaged with known functions. In particular, defining the inner product

$$\langle g_1 | g_2\rangle = \int d\mathbf{r}' g_1^*\left(\mathbf{r}'\right) g_2\left(\mathbf{r}'\right) \tag{3}$$

then the data in (1) and (2), for the given set of detector positions $(u, v)$ (say $M$ such detectors) and system states $p$ (say $N$ such through-focus states), are the projective measurements

$$U_{\text{det}}(u,\, v;\, p) = \langle h^*(u,\, v,\, \cdot;\, p) | U_{\text{obj}}\rangle \tag{4}$$

and

$$I_{\text{det}}(u,\, v;\, p) = \left\langle |h^2|(u,\, v,\, \cdot;\, p) | I_{\text{obj}}\right\rangle \tag{5}$$

If the object is representable in a known basis, say (in the coherent case)

$$U_{\rm obj}\left(\mathbf{r}'\right) = \sum_{s=1,\,2,\,\ldots} \beta(s) B_s\left(\mathbf{r}'\right), \tag{6}$$

where $B_s$ are the basis functions and $\beta(s)$ are the basis coefficients of $U_{\rm obj}$, then the inverse problem corresponds to estimating $\beta(s)$ from the captured $MN$ data. The usual approach without sparsity priors is to find the solution of minimum 2-norm,

$$\hat{\beta} = \operatorname{argmin}||\beta||_2, \quad U_{\rm det}(u,\,v;\,p) = \sum_s \beta(s)\left\langle h^*(u,\,v,\,\cdot;\,p)|B_s\right\rangle, \tag{7}$$

but if it is known that the sought-after object is sparse then one can implement

$$\hat{\beta} = \operatorname{argmin}||\beta||_1, \quad U_{\rm det}(u,\,v;\,p) = \sum_s \beta(s)\left\langle h^*(u,\,v,\,\cdot;\,p)|B_s\right\rangle \tag{8}$$

which gives a exact or an approximate solution if the inner products $\langle h^*(u,\,v,\,\cdot;\,p)|B_s\rangle$ obey certain conditions [2, 3]. Generally, for a given sparsity, the number of projective measurements that are required to reconstruct the sparse signal is governed by the well-known restricted isometry property of compressive sensing. In the present case, the projective measurements are of a particular form that imposes constraints in the so-called coherence between the sparsity basis $\{B_s\}$ and the projectors $\{h^*(u,\,v,\,\cdot;\,p)\}$. In general, the lower the coherence as measured by the highest value of the inner product between the functions $B_s$ and $h^*$, the smaller the required amount of data. Also, to avoid redundance, the selected projections ($h^*$) should be linearly independent. Finally, perhaps a basis where the object is sparse is not known, but its gradient is known to be sparse (as for many practical extended objects, see [2]). Then one can apply the sparsity constraint to the so-called total variation (TV) of the object function [2], minimizing its 1-norm in the inversion.

## 4. COMPUTER ILLUSTRATIONS

To illustrate, we consider imaging of 2D and 3D objects from through-focus data captured for different lens and/or object positions at a fixed detector array. The forward and inverse results were obtained using the analytical results above and standard Fourier optics [4] along with suitable discretization of the equations (computational grids) as illustrated in Figure 1. For 2D objects we kept the object-detector distance fixed and changed only the lens position (the lens-detector distance $z_2$). Reconstructions of 2D objects with coherent light were performed with an in-focus magnification of one, simulating an everyday camera. Reconstructions of 2D objects with incoherent light were performed with the object plane in the far field, simulating a telescope. Reconstructions of 3D objects with both coherent and incoherent light were performed with an in-focus magnification of 100, simulating a high-powered microscope. To imitate a microscope stand, in the 3D case multiple through-focus pictures were acquired by moving the entire object back and forth in the $z$-direction

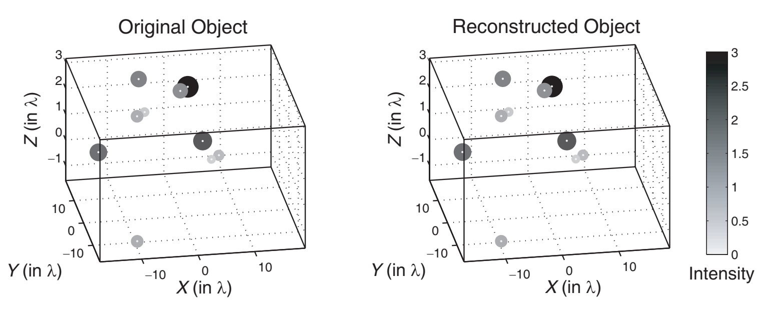



Figure 2: Incoherent imaging of ten point sources on a $24 \times 24 \times 4$ grid. Only sixteen detectors acquire pictures from eight evenly-spaced lens positions on both sides of the in-focus position $z_1 = 158\lambda$ (based on the object plane shown in Figure 1). Radius and shading of outer circle indicate intensity. Inner circle indicates the exact location of the point-source. (Detailed values used in the simulation, all in values of $\lambda$: $z_2 = 1.5798 \times 10^4$, $z_1 \in \{153.39,\, 154.66,\, 155.92,\, 157.18,\, 158.45,\, 159.71,\, 160.97,\, 162.24\}$, $f = 156.41$, $d = 1.5798 \times 10^4$, $\Delta_o^{xy} = 1.5798$, $\Delta_o^z = 1.5798$, $\Delta_d = 1.0734 \times 10^3$).

while keeping the lens and detector plane positions fixed. The results of a reconstruction of ten incoherent point sources are shown in Figure 2. The results of a reconstruction of ten coherent point sources are shown in Figure 3. The results are encouraging.

The TV-based inversion approach is illustrated in Figure 4. The object is a (2D) transparency formed by 4 shapes of uniform field value, taken to be unity inside the shapes and zero outside. Images were obtained using four methods: a) the conventional minimum 2-norm solution using all the through-focus data, b) the compressive sensing minimum 1-norm solution using through-focus data, c) the compressive sensing minimum TV 1-norm solution using only in-focus data, and d) the minimum TV 1-norm solution using through-focus data. Methods (c) and (d) visibly outperformed methods (a) and (b). Furthermore, when adopting the TV approach, the inversion based on through-focus data was also clearly superior to the one based on in-focus data only, confirming the additional information content in through-focus data. Although the through-focus data consisted of 164 samples while the strictly in-focus data consisted of 169 samples, the error in the TV-based reconstruction using through-focus data was noticeably smaller than the in-focus one.

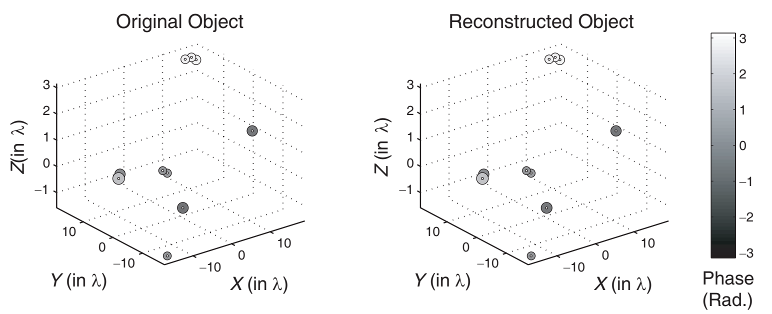



Figure 3: Coherent imaging of ten point sources lying on a discrete $24\times24\times4$ grid. Sixteen detectors acquired pictures from eight evenly-spaced lens positions on both sides of the in-focus position of $z_1 = 158\lambda$ (based on the object plane shown in Figure 1). Radius of outer circle is proportional to intensity and shading of outer circle indicates phase. White inner circle indicates the exact location of point-source on the grid. (Detailed values (in lambda): $z_2 = 1.5798 \times 10^4$, $z_1 \in \{153.39, 154.66, 155.92, 157.18, 158.45, 159.71, 160.97, 162.24\}$, $f = 156.41$, $d = 1.5798 \times 10^4$, $\Delta_o^{xy} = 1.5798$, $\Delta_o^z = 1.5798$, $\Delta_d = 1.0734 \times 10^3$).

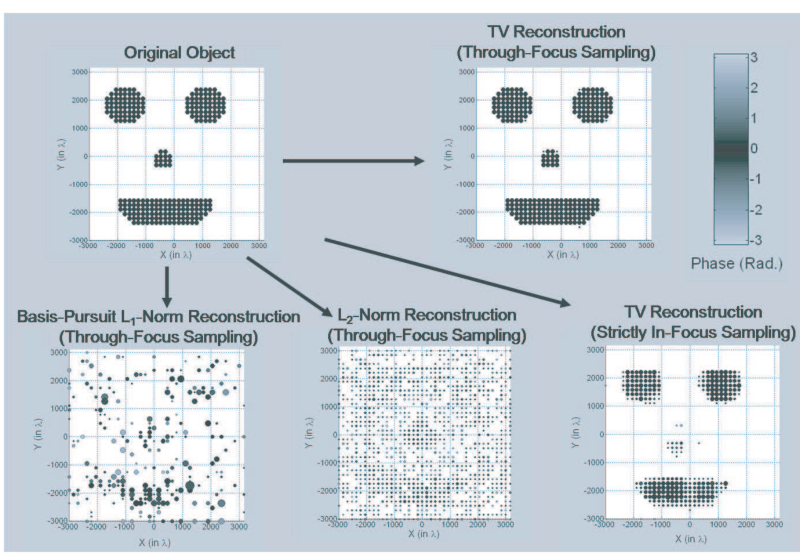



Figure 4: Through-focus imaging by minimization of the object's TV 1-norm. The through-focus data was acquired using 4 detectors at 41 evenly spaced lens positions centered at the in-focus position, while the strictly in-focus data was acquired using 169 detectors at a single in-focus object position, with both detector setups covering an area of $(6.32\times10^3\lambda)^2$. In the plots, circle radius is proportional to intensity while shading of the outer circle indicates phase. For clarity, reconstructed points with magnitudes smaller than 0.1 were not plotted. (Values (in lambda): $f = 7.90 \times 10^4$; $d = 3.16 \times 10^4$; $\Delta_o^{xy} = 157.98$. Through-focus: $z_2$ ranging from $0.0385 \times 10^5$ to $3.1210 \times 10^5$; $z_1 = 3.16 \times 10^5 - z_2$; $\Delta_d = 3.16 \times 10^3$. In-focus: $z_2 = 1.5798 \times 10^5$; $z_1 = 1.5798 \times 10^5$; $\Delta_d = 486.09$).

## 5. DISCUSSION AND CONCLUSION

The proposed compressive through-focus imaging approach was validated for both 2D and 3D objects and for different lens system configurations. After carrying out many examples, we concluded that the key factors governing the image quality are 1) the effective linear independence of the projective measurement vectors (mapping from the object, as given in the grid or Dirac delta basis, to the data at the different sensors and focal states), and 2) the coherence between the projective measurement basis and the grid or Dirac delta basis adopted for the object, which is known to play a key role in compressive sensing. The first aspect was investigated via the singular value decomposition. It was found that if the through-focus positions are all very close to a given in-focus position then the degree of linear independence of the projective measurement vectors is low. The linear independence is generally greater as the through-focus positions are farther apart. For coherent light the best strategy is to allow the through-focus positions to include out-of-focus positions over a broad separation. For incoherent imaging, it is also convenient to separate as much as possible the through-focus positions but they must remain relatively close to the in-focus position. Out of focus information is more limited in the incoherent case.

In summary, we showed that through-focus imaging and compressive sensing can be combined to reduce the number of samples, and specifically the number of photodetectors necessary to reconstruct sparse objects. While conventional in-focus imaging requires as many detectors as pixels in the acquired image, the number of samples required for compressive through-focus imaging can be much smaller since it is restricted only to the object's sparsity. By repeatedly reconfiguring the lens system setup to acquire multiple samples with each detector, compressive through-focus imaging allows a fuller exploitation of physical resources. The through-focus nature of compressive through-focus imaging holds additional advantages for microscopy. Although it is difficult to acquire an in-focus image of a 3D object in conventional microscopy, our results suggest that compressive through-focus imaging can reconstruct entire 3D objects by exploiting prior information like sparsity. We plan to continue developing further the ideas presented in this work, including the use of compressive sensing methods based on total variation (TV) that apply to certain extended objects.

## ACKNOWLEDGMENT

This research is supported by the National Science Foundation under grant 0746310.

## REFERENCES

1. Attota, R., T. Germer, and R. Silver, "Through-focus scanning-optical-microscope imaging method for nanoscale dimensional analysis," *Optics Letters*, Vol. 33, 1990–1992, 2008.
2. Candès, E. J., J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, Vol. 52, 489–509, 2006.
3. Donoho, D. L., "Compressed sensing," *IEEE Trans. Inform. Theory*, Vol. 52, 1289–1306, 2006.
4. Goodman, J. W., *Introduction to Fourier Optics*, Roberts & Co. Publishers, Greenwood Village, CO, USA, 2004.

# A Krylov Subspace Approach to Parametric Inversion of Electromagnetic Data Based on Residual Minimization

**E. Balidemaj and R. F. Remis**
Laboratory of Electromagnetic Research
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands

**Abstract**— In this paper we present a Krylov subspace technique and use residual minimization to efficiently solve parametric electromagnetic inversion problems. We exploit the shift-invariance property of Krylov subspaces to compute total fields inside a homogeneous object for a whole range of contrast values. As soon as these fields are found, we can determine the corresponding scattered fields in a straightforward manner. This approach allows us to solve the inverse problem by simply inspecting an objective function which measures the discrepancy between the measured and modeled scattered field data.

## 1. INTRODUCTION

Many iterative inversion methods use forward solutions in every inversion step. For homogeneous objects, forward problems can be solved efficiently by exploiting the shift-invariance property of Krylov subspaces. To be specific, solving a forward problem for a homogeneous object using an integral equation approach essentially amounts to evaluating a matrix resolvent in which the inverse of the contrast coefficient acts as the resolvent parameter. The action of this resolvent function on a given vector can be evaluated very efficiently by exploiting the shift-invariance property of a Krylov subspace. After only a single run of a Krylov subspace method, field approximations can be constructed for a whole range of contrast coefficients. In previous work [1], we constructed these field approximations using the Arnoldi algorithm in combination with a Full Orthogonalization (FO) approach. The drawback of this FO approach is that, for a given contrast, the norm of the residuals corresponding to successive field approximations is not necessarily a nonincreasing function. In this paper, we remedy this problem by constructing field approximations for which the 2-norm of the corresponding residual is minimum (similar to the well known Generalized Minimum Residual or GMRES method, [3]). For a whole range of contrast parameters, we only need to solve small-scale least-squares problems and for a given contrast coefficient, the 2-norm of the residual is guaranteed to be nonincreasing as the number of iterations increases. Since all Krylov basis vectors need to be stored in the Arnoldi algorithm, we may run into storage problems especially for electrically large objects. A restarted version of the above method as proposed in [2] may then be applied.

Given the Krylov field approximations, the scattered field at specified receiver locations can be computed in a straightforward manner and the inverse scattering problem can actually be solved by inspecting a (possibly nondifferentiable) objective function which measures the discrepancy between the true and modeled scattered field data.The effectiveness of the method is illustrated by a number of numerical examples for single and multi-frequency data.

## 2. INTEGRAL REPRESENTATIONS AND THEIR DISCRETIZED COUNTERPARTS

We consider $E$-polarized electromagnetic fields in a configuration that is invariant in the $z$-direction. An object with a finite support $\mathbb{D}$ is located in vacuum and is characterized by a conductivity $\sigma$, a permittivity $\varepsilon = \varepsilon_0\varepsilon_\mathrm{r}$, and a permeability $\mu_0$. Clearly, the object has no magnetic contrast and the conductivity $\sigma$ and the relative permittivity $\varepsilon_\mathrm{r}$ of the object are unknown. It is our objective to retrieve these medium parameters from scattered electric field data measured outside the support of the object. Specifically, for the scattered field at a receiver location with position vector $\boldsymbol{\rho} = \boldsymbol{\rho}^\mathrm{rec} \notin \mathbb{D}$, we have the so-called data equation

$$E_z^\mathrm{sc}(\boldsymbol{\rho}^\mathrm{rec},\omega) = \zeta \frac{\mathrm{i}k^2}{4} \int_{\boldsymbol{\rho}'\in\mathbb{D}} H_0^{(1)}(k|\boldsymbol{\rho}^\mathrm{rec} - \boldsymbol{\rho}'|) E_z(\boldsymbol{\rho}',\omega)\,\mathrm{d}A, \tag{1}$$

where i is the imaginary unit, $k = \omega/c_0$ is the wavenumber ($c_0$ is the electromagnetic wave speed in vacuum), $H_0^{(1)}$ is the Hankel function of the first-kind and order zero, $\zeta$ is the contrast coefficient given by

$$\zeta = \varepsilon_{\mathrm{r}} - 1 + \mathrm{i}\frac{\sigma}{\omega\varepsilon_0}, \tag{2}$$

and $E_z$ in the integrand is the total electric field strength inside the object. The latter field is unknown, but we do know that it has to satisfy the object equation

$$E_z(\boldsymbol{\rho},\omega) - \zeta\frac{\mathrm{i}k^2}{4}\int_{\boldsymbol{\rho}'\in\mathbb{D}} H_0^{(1)}(k|\boldsymbol{\rho}-\boldsymbol{\rho}'|)E_z(\boldsymbol{\rho}',\omega)\,\mathrm{d}A = E_z^{\mathrm{inc}}(\boldsymbol{\rho},\omega), \tag{3}$$

with $\boldsymbol{\rho} \in \mathbb{D}$ and where $E_z^{\mathrm{inc}}$ is the known incident field. Eq. (3) is a Fredholm integral equation of the second kind for the total field $E_z$ inside the object if the contrast coefficient is known. Also note that the total field in $\mathbb{D}$ is equal to the incident field and the scattered field at the receiver location vanishes if $\zeta = 0$ (no object is present in $\mathbb{D}$). From this moment on we exclude this trivial case and consider nonzero contrast coefficients only.

We discretize Eqs. (1) and (3) on a uniform grid and use pulse expansions for the electric field strength (see [4], for example). We obtain the discretized data equation

$$E_z^{\mathrm{sc}}(\boldsymbol{\rho}^{\mathrm{rec}},\omega) \approx \zeta\mathbf{r}^T\mathbf{u} \tag{4}$$

and the discretized object equation

$$(\mathbf{I} - \zeta\mathbf{G})\mathbf{u} = \mathbf{u}^{\mathrm{inc}}. \tag{5}$$

In these equations, $\mathbf{r}$ is the receiver vector, $\mathbf{u}$ contains the field approximations for the total field inside the object, and $\mathbf{u}^{\mathrm{inc}}$ contains incident electric field values sampled on the grid. Matrix $\mathbf{G}$ is a matrix representation of the Hankel function (Greens function) and its action on a vector can be computed using the FFT. Finally, we call the right-hand side of Eq. (4) modeled scattered field data and set $v^{\mathrm{sc}} = \zeta\mathbf{r}^T\mathbf{u}$.

The way we retrieve the permittivity and conductivity of the object is by inspecting a certain objective function. Specifically, we first assume that minimum and maximum relative permittivity and conductivity values can be found such that $\varepsilon_{\mathrm{r;min}} \le \varepsilon_{\mathrm{r}} \le \varepsilon_{\mathrm{r;max}}$ and $\sigma_{\mathrm{min}} \le \sigma \le \sigma_{\mathrm{max}}$, where $\varepsilon_{\mathrm{r;max}} \ge \varepsilon_{\mathrm{r;min}} \ge 1$ and $\sigma_{\mathrm{max}} \ge \sigma_{\mathrm{min}} \ge 0$. The above intervals define our region of interest in the complex $\zeta$-plane. This region is discretized using a uniform grid, and for each grid value we solve the object Eq. (5) for the total field $\mathbf{u}$. Having found this field, we substitute $\mathbf{u}$ in the expression for $v^{\mathrm{sc}}$ to obtain the corresponding modeled scattered field at the receiver location. We repeat this procedure for all values on the $\zeta$-grid. A collection of modeled scattered fields is obtained in this way and all these fields are compared with the true measured field using the objective function

$$F = \frac{|E_z^{\mathrm{sc}}(\boldsymbol{\rho}^{\mathrm{rec}},\omega) - v^{\mathrm{sc}}(\omega)|^2}{|E_z^{\mathrm{sc}}(\boldsymbol{\rho}^{\mathrm{rec}},\omega)|^2}. \tag{6}$$

Other, possibly nondifferentiable objective functions, can be used as well. If the modeled scattered field is such that $F < \epsilon$, where $\epsilon > 0$ is a user specified tolerance, then the contrast coefficient corresponding to this field is a solution of our inverse problem.

## 3. RESIDUAL MINIMIZATION

The approach outlined in the previous section is practical if we can efficiently solve the discretized object equation for a large number of different contrast coefficients. This can be achieved by exploiting the shift-invariance property of Krylov subspaces. Loosely speaking, what this amounts to is that if we solve the forward problem for a particular worst-case contrast coefficient, then very little extra work is needed to obtain the total field for all other contrast coefficients in our $\zeta$-domain of interest.

We start by introducing the contrast source $\mathbf{w} = -\zeta\mathbf{u}$ (the minus sign is included for convenience only) and we rewrite the object equation as

$$\mathbf{A}(\alpha)\mathbf{w} = \mathbf{u}^{\mathrm{inc}} \quad \text{where} \quad \mathbf{A}(\alpha) = \mathbf{G} + \alpha\mathbf{I} \tag{7}$$

and $\alpha = -1/\zeta$ (recall that $\zeta \neq 0$). Suppose now that we use the Generalized Minimum Residual method (GMRES method, see [3]) to solve this system for a particular value of $\alpha$. After $m$ iterations we have the field approximation $\mathbf{w}_m$ available and the corresponding residual is given by $\mathbf{r}_m =$

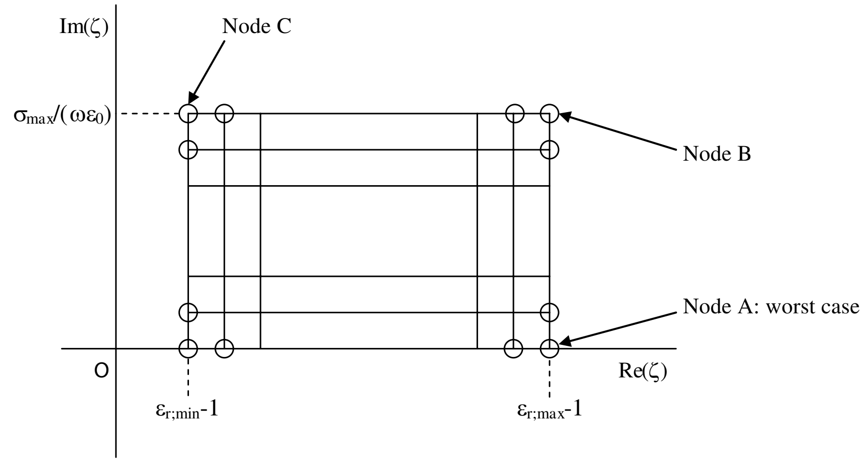



Figure 1: Grid of $\zeta$-values of interest ($\sigma_{\min} = 0$).

$\mathbf{u}^{\text{inc}} - \mathbf{A}(\alpha)\mathbf{w}_m$. The question now is: which value of $\alpha$ (or, equivalently, which value of $\zeta$) requires the largest number of iterations to reach a prescribed and sufficiently small tolerance level for the normalized 2-norm of the residual? It is reasonable to assume that the contrast coefficient for which $|\zeta|$ is maximum is the worst-case contrast coefficient (node B in Figure 1). This, however, turns out not to be true. Excessive numerical testing shows that the worst-case $\zeta$ is given by $\text{Im}(\zeta) = 0$ and $\text{Re}(\zeta) = \varepsilon_{\text{r;max}} - 1$. In other words, the worst-case configuration consists of a lossless object for which the relative permittivity is maximum (node A in Figure 1). This is illustrated in Figure 2 where we show the normalized 2-norm of the residual vector for contrast coefficients located at nodes A, B, and C in Figure 1 ($\mathbf{r}_0 = \mathbf{u}^{\text{inc}}$). The object is a square block with a side length equal to the wavelength of the electromagnetic field. Out of the three cases, the contrast coefficient corresponding to node A produces the worst convergence behavior.

Having identified the worst-case contrast coefficient, we write the corresponding $\alpha$ parameter as $\tilde{\alpha}$ and the system of Eq. (7) is written as

$$[\mathbf{A}(\tilde{\alpha}) + \beta\mathbf{I}]\,\mathbf{w} = \mathbf{u}^{\text{inc}} \quad \text{where} \quad \beta = \alpha - \tilde{\alpha}. \tag{8}$$

Following [2], we call the above system with $\beta = 0$ the seed system, and we call Eq. (8) with a $\beta \neq 0$ an add system. We run the GMRES algorithm only for the worst-case seed system to reach a prescribed tolerance for the normalized 2-norm of the residual vector. The field approximations for the add systems follow by solving small scale least-squares problems. Specifically, after $m$ iterations of the Arnoldi algorithm with the seed system matrix $\mathbf{A}(\tilde{\alpha})$, we have the summarizing equation

$$\mathbf{A}(\tilde{\alpha})\mathbf{V}_m = \mathbf{V}_{m+1}\tilde{\mathbf{H}}_m, \tag{9}$$

where $\tilde{\mathbf{H}}_m$ is an $m+1$-by-$m$ upper Hessenberg matrix containing the recurrence coefficients of the Arnoldi algorithm, and the columns of matrix $\mathbf{V}_m$ form an orthonormal basis of the Krylov subspace generated by $\mathbf{A}(\tilde{\alpha})$ and the incident field vector $\mathbf{r}_0 = \mathbf{u}^{\text{inc}}$. Now let $\mathbf{w}_m$ belong to this Krylov subspace and take this vector as an approximation to $\mathbf{w}$ in Eq. (8). We write $\mathbf{w}_m = \mathbf{V}_m\mathbf{c}_m$ and the corresponding residual is given by

$$\mathbf{r}_m = \mathbf{u}^{\text{inc}} - [\mathbf{A}(\tilde{\alpha}) + \beta\mathbf{I}]\,\mathbf{w}_m = \mathbf{V}_{m+1}\left\{\|\mathbf{u}^{\text{inc}}\|\mathbf{e}_1^{(m+1)} - \left[\tilde{\mathbf{H}}_m + \beta\begin{pmatrix}\mathbf{I}_m\\ -\mathbf{0}-\end{pmatrix}\right]\mathbf{c}_m\right\}, \tag{10}$$

where $\|\cdot\|$ denotes the 2-norm and $\mathbf{e}_1^{(m+1)}$ is the first column of the $m$-by-$m$ identity matrix. Since the basis vectors are orthonormal, the 2-norm of the residual is given by

$$\|\mathbf{r}_m\| = \left\|\,\|\mathbf{u}^{\text{inc}}\|\mathbf{e}_1^{(m+1)} - \left[\tilde{\mathbf{H}}_m + \beta\begin{pmatrix}\mathbf{I}_m\\ -\mathbf{0}-\end{pmatrix}\right]\mathbf{c}_m\right\| \tag{11}$$

and we observe that for each new value of $\beta$ we need to solve a small scale least-squares problem only. To summarize, GMRES solves the seed and add systems simultaneously, since the Krylov

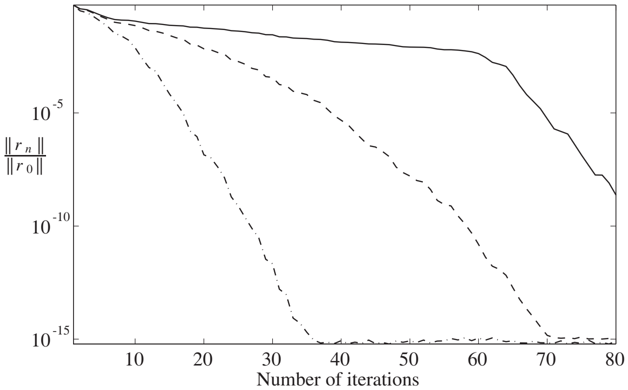



Figure 2: Normalized 2-norm of the residual for contrast coefficients corresponding to node A (solid line), node B (dashed line), and node C (dashed-dotted line).

subspaces generated with $\mathbf{r}_0 = \mathbf{u}^{\mathrm{inc}}$, $\mathbf{G}$, $\mathbf{A}(\tilde{\alpha})$ and $\mathbf{A}(\alpha) = \mathbf{A}(\tilde{\alpha}) + \beta\mathbf{I}$ are all the same. More precisely, we have

$$\mathbb{K}^m[\mathbf{A}(\tilde{\alpha}), \mathbf{r}_0] = \mathbb{K}^m[\mathbf{A}(\alpha), \mathbf{r}_0] = \mathbb{K}^m(\mathbf{G}, \mathbf{r}_0) \tag{12}$$

and the dimension $m$ is determined by the worst-case contrast coefficient. The object equation for this worst-case coefficient is the seed system and as soon as the 2-norm of the corresponding residual is sufficiently small, we have essentially solved all add systems as well.

The use of Arnoldi's method may become prohibitive for electrically large objects (objects large compared with the wavelength of the electromagnetic field) and problems requiring a large number of Arnoldi iterations, since all Krylov basis vectors need to be stored and the computational load also increases with $m$. Using a restart version of GMRES then comes to mind, but implementing such a procedure is not straightforward, since we may loose the shift-invariance property of the Krylov subspaces. Suppose, for example, that we restart the GMRES method after $p$ iterations. Denoting the residual of the seed and add system by $\mathbf{r}_p(\tilde{\alpha})$ and $\mathbf{r}_p(\alpha)$, respectively, we have

$$\mathbb{K}^m[\mathbf{A}(\tilde{\alpha}), \mathbf{r}_p(\tilde{\alpha})] \neq \mathbb{K}^m[\mathbf{A}(\alpha), \mathbf{r}_p(\alpha)], \tag{13}$$

in general. In [2], it is proposed to force the residual vectors of the seed and add systems to be collinear, that is, we require that $\mathbf{r}_p(\alpha) = \gamma_p \mathbf{r}_p(\tilde{\alpha})$ for some $\gamma_p$. With this condition it is possible to exploit the shift-invariance property of the Krylov subspaces and restarts are possible as well. We have implemented the restart version based on this idea to avoid possible storage problems. For details about the algorithm and a discussion on when the collinearity property exists, see [2].

## 4. NUMERICAL RESULTS

To retrieve the medium parameters of a square block located in vacuum, we illuminate this block by electromagnetic waves with a frequency of $f = 42\,\mathrm{MHz}$. The side length of the block is equal to $\lambda$, where $\lambda$ is the wavelength in vacuum that corresponds to this frequency. A single-source/single-receiver unit is symmetrically located above the block. The distance between the block and the source/receiver unit is $\lambda/2$. Furthermore, the true medium parameters are $\varepsilon_{\mathrm{r}} = 4.5$ and $\sigma = 5$ mS/m and for the minimum and maximum medium parameters we take $\varepsilon_{\mathrm{r;min}} = 1$, $\varepsilon_{\mathrm{r;max}} = 10$, $\sigma_{\min} = 0\,\mathrm{mS/m}$, and $\sigma_{\max} = 10\,\mathrm{mS/m}$. Finally, the domain of interest in the complex $\zeta$-plane is discretized using a $40 \times 40$ grid. We therefore need 1599 forward solutions (the scattered and total field for $\zeta = 0$ are known). The seed system has a contrast coefficient $\zeta = 9$ and there are 1598 add systems to be solved. The resulting objective function on the domain of interest is shown in Figure 3. The true medium parameters are easily recognized, but there are other minima as well. We therefore carry out two additional experiments at two different frequencies, namely, at $f = 30\,\mathrm{MHz}$ and $f = 36\,\mathrm{MHz}$. The objective function becomes

$$F = \sum_{k=1}^{3} m_k \frac{|E_z^{\mathrm{sc}}(\boldsymbol{\rho}^{\mathrm{rec}}, \omega_k) - v^{\mathrm{sc}}(\omega_k)|^2}{|E_z^{\mathrm{sc}}(\boldsymbol{\rho}^{\mathrm{rec}}, \omega_k)|^2} \quad \text{with} \quad \sum_{k=1}^{3} m_k = 1. \tag{14}$$

We take $m_k = 1/3$ for $k = 1, 2, 3$. We evaluated this objective function by computing forward solutions for each frequency separately and for each frequency we used the Arnoldi algorithm with

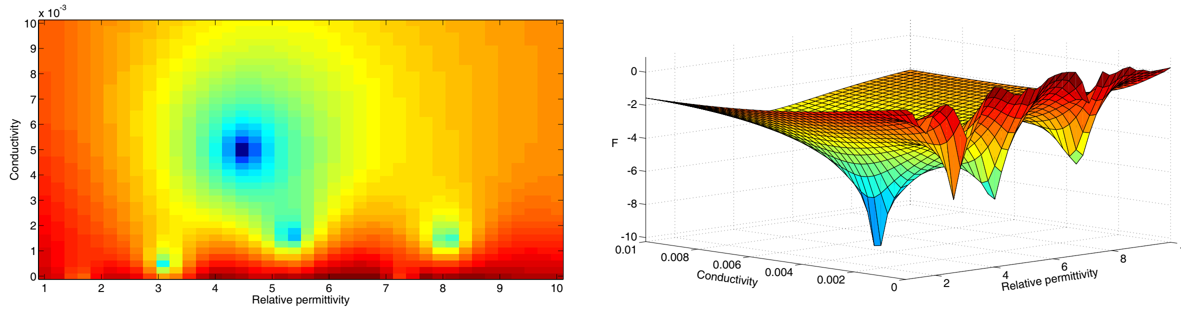



Figure 3: Top view (left) and surface plot (right) of the objective function $F$ on the domain of interest for a single frequency of $f = 42\,\text{MHz}$.

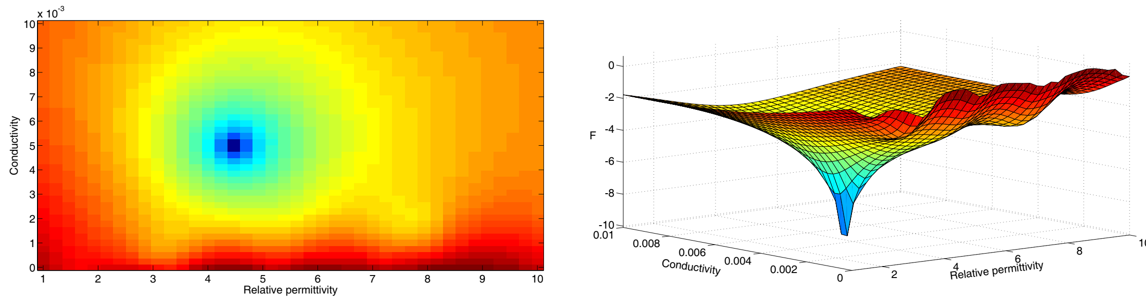



Figure 4: Top view (left) and surface plot (right) of the objective function $F$ on the domain of interest using three frequencies ($f = 30$, $f = 36$, and $f = 42\,\text{MHz}$).

and without restarts to obtain the total fields inside the object. Both approaches produce the same objective function, of course, and this function is shown in Figure 4. The global minimum corresponding to the true medium parameters remains where it is and a number of false minima have disappeared. We still have local minima around the $\sigma = 0$ axis, but the objective function at these minima is much larger than at the global minimum. Moreover, the objective function is smoother than in the single frequency case.

## 5. CONCLUSIONS

We have presented an efficient electromagnetic inversion method for homogeneous objects. The method exploits the shift-invariance property of Krylov subspaces and through residual minimization it very efficiently solves for the total field inside the object. Because of this property, we can solve the inverse problem by simply inspecting a user-defined objective function. Although we have presented our results for a differentiable objective function, other possibly nondifferentiable objective functions may be used as well.

In this paper we have considered homogeneous objects only. For inhomogeneous objects we can still try to match the scattered field due to a homogeneous object to the scattered field generated by an inhomogeneous object. This topic is very important in many different areas most notably in effective medium theory. Presently we are using the technique presented in this paper to investigate to what extent it is possible to find an effective homogeneous scatterer for homogeneous objects with randomly distributed inclusions.

## REFERENCES

1. Budko, N. V. and R. F. Remis, “Electromagnetic inversion using a reduced-order three-dimensional homogeneous model,” *Inverse Problems*, Vol. 20, No. 6, S17–S26, 2004.
2. Frommer, A. and U. Glässner, “Restarted GMRES for shifted linear systems,” *SIAM J. Sci. Comput.*, Vol. 19, No. 1, 15–26, 1998.
3. Saad, Y., *Iterative Methods for Sparse Linear Systems*, 2nd Edition, SIAM, Philadelphia, 2003.
4. Zwamborn, A. P. M. and P. M. van den Berg, “A weak form of the conjugate gradient FFT method for two-dimensional TE scattering problems,” *IEEE Trans. Antennas and Propagat.*, Vol. 39, No. 6, 953–960, 1991.

# Antenna Modeling Issues in Quantitative Image Reconstruction Using a Flexible Microwave Tomography System

**Nikola Petrovic**[1], **Tommy Henriksson**[1,2], **and Magnus Otterskog**[1]

[1]School of Innovation, Design and Engineering, Mälardalen University, Box 883, 721 23 Västerås, Sweden
[2]Laboratoire des Signaux et Systèmes, SUPÉLEC
3 rue Joliot-Curie, F-91192 Gif-sur-Yvette Cedex, France

**Abstract**— Antenna modeling is a significant issue with respect to the image reconstruction process in a microwave tomography system. Herein, the radiated field from the antenna is modeled as a vertical polarized cylindrical wave, in the numerical incident field model of the algorithm. Two different monopole antenna designs are compared with the incident field model in both amplitude and phase, using data conducted by a flexible microwave imaging system. Finally, the incident field's impact on the reconstructed images by the Newton Kantorovich algorithm, using measured data from both antennas, is presented.

## 1. INTRODUCTION

Quantitative microwave imaging has been extensively studied in the past years as an alternative technique in biomedical imaging, with a strong potential in early stage breast cancer detection [1]. Many encouraging results have been presented for the application and even prototype systems have been developed for obtaining clinical results by using either a radar-based technique [2] or microwave tomography [1]. In microwave tomography the inverse scattering problem is solved to retrieve the complex permittivity profile of the biological object from the measured scattered field and the applied incident field. This can be done in two different approaches, solving the nonlinear inverse problem or using different linear approximations, like the Born or Rytov approximations. The linear approach is computationally efficient and suitable for smaller objects with low contrast [3]. However, in biomedical applications the objects are often large with a high dielectric contrast, limiting the usability of the linear approximations. Therefore, more often different iterative methods are used to solve the nonlinear inverse scattering problem. Since, this is a highly computational-heavy process the three-dimensional problem is often simplified into a two-dimensional (2-D) problem [4, 5]. Still today there are many practical issues to be considered in the quantitative microwave imaging problem including: modeling errors, coupling effects, calibration, measurement errors, etc.

In this study, the nonlinear inverse scattering problem is solved, with a Newton-based iterative optimization scheme, the Newton-Kantorovich (NK) algorithm [6]. The measured scattered field is iteratively compared with the computed field from the direct problem with the complex permittivity profile estimation and an applied incident field model. Consequently, the solution is highly sensitive to model errors in the incident field. This paper focus on this model error's impact on the reconstructed quantitative image using data from a flexible robotic microwave imaging system, developed at Mälardalen University (MDH). This study is conducted during the development of the imaging system and the first quantitative images of a 2-D breast phantom are obtained.

## 2. TOMOGRAPHY SYSTEM OVERVIEW

The robotic microwave imaging system is developed as a flexible experimental platform, where one of the applications is breast imaging. Using a robot controlled system different system geometries can be investigated by measuring the scattered field with a single transmitting/receiving antenna-pair, thus avoiding the mutual coupling that can occur when an antenna-array is used. The receiving antenna can be positioned along planar, cylindrical and spherical surfaces with a high accuracy, with a relative position error less than 0.1 mm. In addition, the wide-band system gives the possibility to investigate different frequency bands for different imaging applications with a relatively fast data acquisition time, 0.5 s per measurement point when the robot is used in a very slow movement mode for highest precision of the positioning.

The main parts of the system, as shown in Fig. 1, are an ABB robot which controls the mechanical positioning of the receiving antenna, a water-tank (2 m in diameter) with an object-fixture and step motor for multi-view measurement, a vector network analyzer (VNA), and a developed Matlab$^{\mathrm{TM}}$ control interface collecting field data.

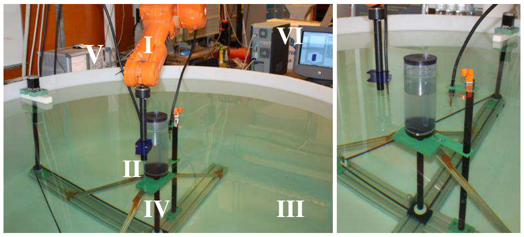



Figure 1: Experimental setup of the robot controlled data acquisition system. The robot (I) with the receiving antenna (II), water-filled tank (III) with fixture (transmitting antenna, rotational board, object and step motor) (IV), VNA (V), and the PC with a Matlab$^{\mathrm{TM}}$ control interface (VI).

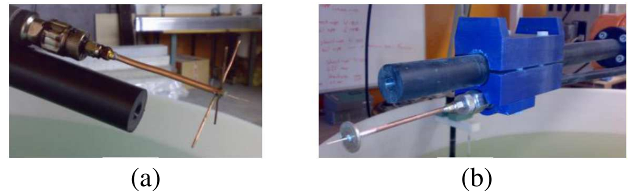



Figure 2: The receiving monopole antennas mounted on the robot arm. (a) Antenna 1: with a wire ground plane, (b) antenna 2: with a circular ground plane.

### 3. ANTENNAS AND EXPERIMENTALL SETUP

In this study, two different monopole antenna designs are compared with the numerical incident field model. Both antennas, depicted in Figs. 2(a) and 2(b), are simple monopoles consisting of a semi rigid coaxial cable which has a physical length of 11 mm of the protruding inner conductor and tuned for tap water with a complex permittivity of $\varepsilon^* = 76 + j7.3$ (at 1140 MHz and 21°C), estimated from a measurement based method [6] using the same antennas. It has been shown that the monopole antenna can be easily modeled as a line source in a 2-D imaging problem [7]; also the fact that the monopole antenna is an efficient radiating element, cheap and easy to manufacture gives good motivation for using them in the tomography system. When immersed into water the antennas are fairly broad banded from 950 MHz–1450 MHz with return loss less than 10 dB. However, in this study only a single frequency of 1140 MHz is used for both antenna setups.

By using the monopole antenna with a ground plane increases the antenna gain, thus improves the efficiency of the imaging system. The difference between the antennas is the design of the ground-plane. In the first prototype (Fig. 2(a)), it is simply designed with four wires forming a horizontal cross, and the second setup (Fig. 2(b)) uses a circular ground plane. Since, the first design do not obtain a rotational symmetric radiation pattern in the horizontal plane, the orientation of the receiving antenna must be kept in a specific angel towards the transmitter to obtain a incident field as close as possible to the model. This process was done by manually tuning the angular position of the receiver and transmitter, around their own axis, to obtain a cylindrical waveform fitting the incident field model. In Fig. 3, the amplitude of the measured incident field is depicted, during the tuning process to best fit the simulated field. As expected, the second design obtains a more rotational symmetric radiation pattern in the horizontal plane and is easier matched with the computed values, without having to manually tune the directions of the antennas.

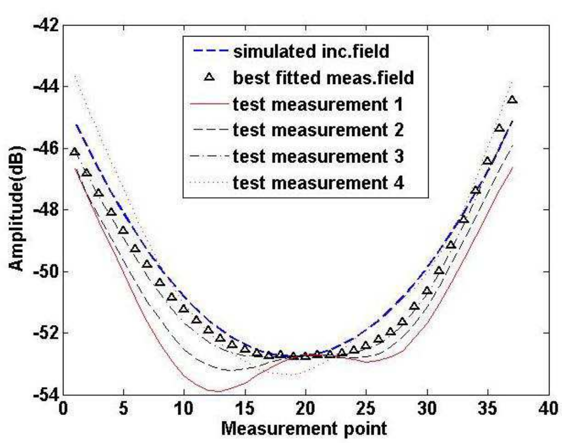



Figure 3: Amplitude of the incident measured field for antenna 1, during the tuning process.

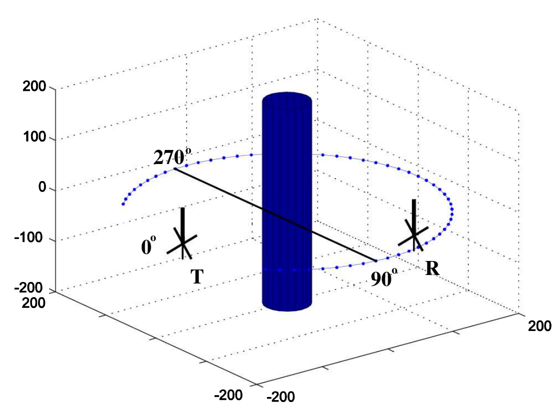



Figure 4: The experimental setup geometry.

The scattered field is, herein, measured along a circular arc with a radius of 120 mm, in the horizontal plane with vertically polarized monopole antennas, considering a two dimensional transverse magnetic case (2D-TM). The robot arm moving the attached receiving antenna in 37 points using a 5° angular step sweeping 180° from 90° to 270° (Fig. 4). The transmitter is positioned at 0° on the same circular arc. The object is rotated with an angular step of 9° using the step motor controller to obtain a multi-view examination in 40 views.

In this study, a simple breast phantom is placed in the centre of the system. It consists of two different PVC cylinders with a diameter of 110 mm to hold the normal breast tissue ($\varepsilon^* = 35 + j7.3$) and 20 mm to hold the tumor liquid (saline water $\varepsilon^* = 58 + j15$), with 2.5 mm and 2 mm tick PVC structure, respectively. The breast tissue is simulated with the mixture of the surfactant Triton X-100 and deionized water considering an average of the three categories of adipose breast tissues presented in paper [8] by Lazebnik et al..

## 4. FIELD VALIDATION AND IMAGE RECONSTRUTION

As mentioned earlier, the NK algorithm is iteratively minimizing the error between the measured and the computed scattered field from a numerical incident field model in the direct problem. As a result, the solution is highly sensitive to model errors of the incident field inside the object region. The numerical incident field must agree with the measured field from the experimental setup. The radiated field from the transmitting antenna is simply implemented as a vertical polarized cylindrical wave emitted by a line source, as

$$E_{inc}(r) = -\frac{\pi}{2} f \mu_0 H_0^{(1)}(k_1|r - r'|) \tag{1}$$

The $H_0^{(1)}(k_1|r - r'|)$ term is the zero-order Hankel function of the first kind, $f$ is the frequency, $k_1$ is the wavenumber of the background medium, and finally $r$ and $r'$ represents the observation and source point, respectively. A validation of the measured incident field (in one view) can be done by comparing it with the computed incident field at the receiving points, with the assumption that when a good agreement is obtained also the incident field inside the object region must fit well. In Figs. 5 and 6, the measured amplitude and phase is compared with the numerical model for antenna 1 and 2, correspondingly. One can see that the second antenna design has a much better fit in both amplitude and phase compared to first design.

However, in the image reconstruction process it is the incident field inside the object that is of major importance. Therefore, the image reconstruction from the multi-view data of the breast phantom using both antennas designs will be compared. In this way the impact of the incident field model error inside the object region due to the antenna choice can be investigated, in terms of artifacts in the reconstructed images. Fig. 7 shows the reconstructed real and imaginary permittivity profiles during the first three iterations, starting from the initial guess of a breast without tumor (Fig. 7(a)), for both antenna designs. One can see that results is better with the

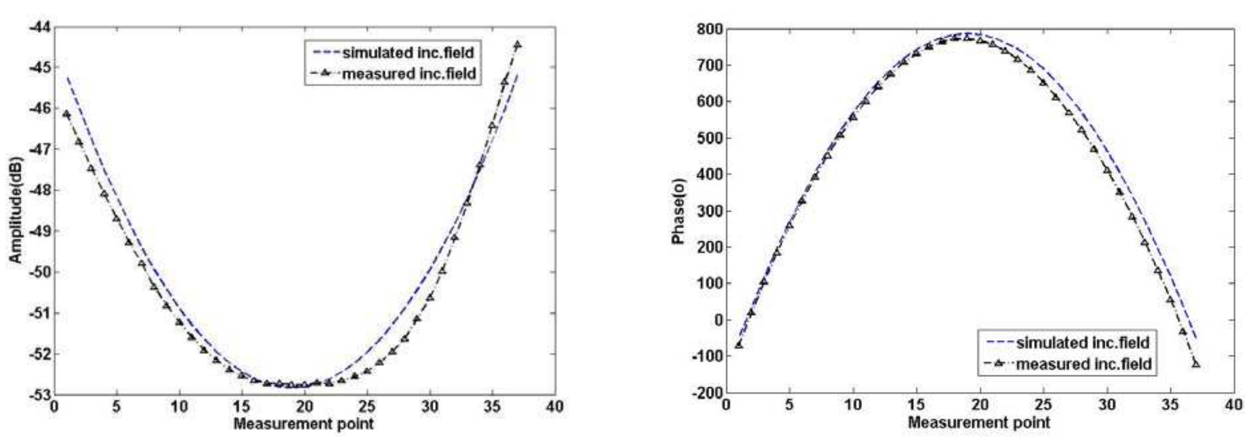
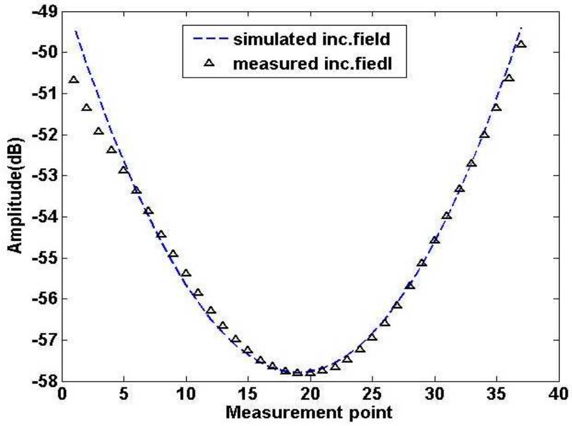



Figure 5: The incident field comparison between measured field for antenna 1 (Fig. 2(a)) and the simulated field, amplitude (left) and unwrapped phase (right). The measurement are obtained at a temperature of 21.5°C for the background-medium giving an estimated complex permittivity of $\varepsilon^* = 75.6 + j7.3$.



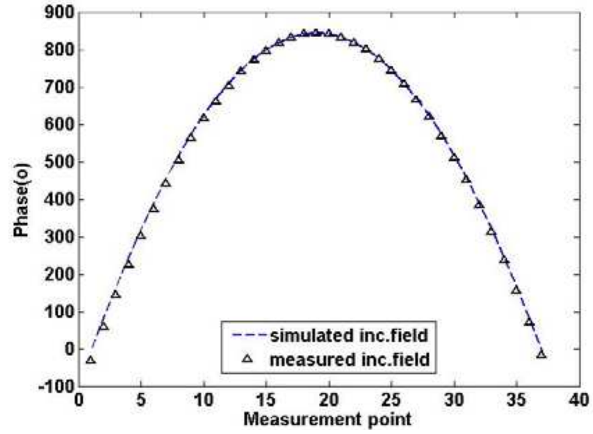




Figure 6: The incident field comparison between measured field for antenna 2 (Fig. 2(b)) and the simulated field, amplitude (left) and unwrapped phase (right). The measurement are obtained at a temperature of 18.8°C for the background-medium giving an estimated complex permittivity of $\varepsilon^* = 78 + j8.3$.

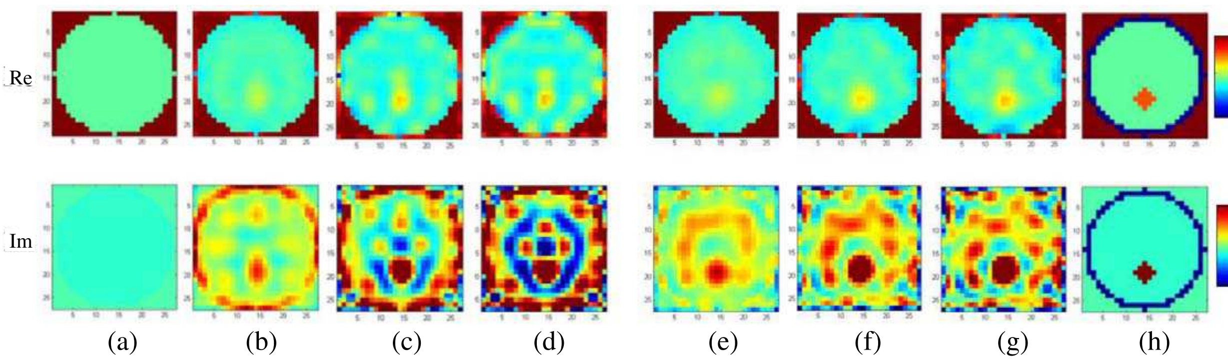



Figure 7: Reconstructed real and imaginary permittivity profile of the phantom using both antennas (Figs. 2(a) and 2(b)), starting with (a) the initial guess, (b), (c) and (d) iteration 1, 2, 3 for antenna 1, respectively, (e), (f) and (g) iteration 1, 2, 3 for antenna 2, respectively and (h) expected profile.

second antenna design, the reconstructed images has much less artifacts, especially in the imaginary part. The tumor phantom is clearly reconstructed even if some artifacts appear in the imaginary part. These results indicate the importance of choosing an antenna design that enables a minimal incident field model error in the algorithm.

## 5. CONCLUSIONS

In this paper, two different monopole antenna designs have been validated in a flexible microwave tomography system. The results show how the selection of antenna design impacts the error between the measured incident field and the numerical incident field model. By comparing the reconstructed images of a simple breast phantom it was shown how the quantitative image is affected by this model error. Using an antenna design that minimizes the error between the model and the measured values leads to a better reconstruction of the object. This confirms the importance of minimizing the model error inside the object region.


## ACKNOWLEDGMENT

The authors would like to thank Robotdalen for the support with the industrial robot in our experimental setup.


## REFERENCES


1. Paulsen, K. D. and P. M. Meaney, *Alternative Breast Imaging*, The Springer International Series in Engineering and Computer Science, 778, 2005.
2. Klemm, M., I. Craddock, J. Leendertz, A. Preece, and R. Benjamin, "Experimental and clinical

results for breast cancer detection using UWB microwave radar," *Proceedings of IEEE Antennas and Propagation Society International Symposium*, San Diego, USA, 7–12, July 2008.

3. Pichot, C., L. Jofre, G. Peronnet, and J.-C. Bolomey, "Active microwave imaging of inhomogeneous bodies," *IEEE Trans. Antennas and Propag.*, Vol. 33, 416–423, Apr. 1985.
4. Abubakar, A., P. M. van den Berg, and J. J. Mallorqui, "Imaging of biomedical data using a multiplicative regularized contrast source inversion method," *IEEE Trans. Microw. Theory Tech.*, Vol. 59, No. 7, 1761–1771, Jul. 2002.
5. Souvorov, A. E., A. E. Bulyshev, S. Y. Semenov, R. H. Svenson, A. G. Nazarov, Y. E. Sizov, and G. P. Tatsis, "Microwave tomography: A two dimensional newton iterative scheme," *IEEE Trans. Microw. Theory Tech.*, Vol. 46, 1654–1659, Nov. 1998.
6. Henriksson, T., N. Joachimowicz, C. Conessa, and J.-C. Bolomey, "Quantitative microwave imaging of inhomogeneous objects using a planar 2.45 GHz System," *IEEE Trans. Instr. Meas.*, to appear.
7. Meaney, P. M., M. W. Fanning, D. Li, S. P. Poplack, and K. D. Paulsen, "A clinical prototype for active microwave imaging of the breast," *IEEE Trans. Microw. Theory Tech.*, Vol. 48, No. 11, 1841–1853, Nov. 2000.
8. Lazebnik, M., E. L. Madsen, G. R. Frank, and S. C. Hagness, "Tissue mimicking phantom materials for narrowband and ultrawideband microwave applications," *Physics in Medicine and Biology*, Vol. 50, 4225–4258, 2005.

# A Two Layers Multi-scale Bi-dimensional SPM Model for the Study of Radar Backscatter Behavior on Semi-arid Soil Subsurfaces

**L. Bennaceur Farah**[1], **I. R. Farah**[2], **R. Bennaceur**[3], **I. Hosni**[1], **and M. R. Boussema**[1]

[1]LTSIRS, ENIT, Tunisia
[2]RIADI, ENSI, Tunisia
[3]LMPC, FST, Tunisia

**Abstract**— In this paper, we propose to simulate the SAR response of semi-arid subsurfaces. We characterize the soil surfaces and subsurfaces by a two layer geo-electrical model. The upper layer is described by its dielectrical constant, thickness, a multi-scale bi-dimensional surface roughness model by using the wavelet transform and the Mallat algorithm, and volume scattering parameters. The lower layer is described by its dielectric constant and multi-scale surface roughness. To compute surface, subsurface and volume scattering, we consider a two layers multi-scale bi-dimensional Small perturbations model. In this study, each surface of the two layers surface is considered as a band limited fractal random process corresponding to a superposition of a finite number of one dimensional Gaussian processes each one having a spatial scale. We investigated the dependence of backscattering coefficient on roughness multi-scale parameters and soil moisture parameters for different incident angles by a sensitivity analysis. This sensitivity analyses is the first step of an inversion procedure.

## 1. INTRODUCTION

The retrieval of information related to physical surface parameters is a major objective of many studies in remote sensing investigations. In that context, modeling radar backscattering through natural surfaces has become an important theme of research and active remote sensing and has shown its utility for many applications in hydrology, geology, astrophysics, etc.

Previous works [9] have studied the potential of low-frequency imaging SAR to map the shallow subsurface of Mars by comparing the results obtained for comparative test sites used to extrapolate their models to the Martian case. Considering the penetrations capabilities of L band SAR in dry sand, they developed a two layer integral equation model to simulate radar backscattering on arid subsurfaces. They used the classical statistical description of natural surfaces and characterized roughness by statistical parameters namely correlation length and standard deviation.

However, the weakness of the classical description of natural surfaces is the large spatial variability which affects the correlation function and make classical roughness parameters very variable [2, 5, 7]. In that context, many previous works have suggested that natural surfaces are better described as self affine random processes ($1/f$ processes) than as stationary processes. In previous works, we have analysed radar backscattering on multi-scale bi-dimensional surfaces [1, 4]. This novel multi-scale bi-dimensional description does not depend on classical roughness parameters standard deviation and correlation length but on new parameters related to multi-scale surfaces properties. In the first section, we present the multi-scale surface description.

In the next section, we present the study of the impact of this multi-scale roughness description and soil moisture on radar backscattering using our two layers multi-scale bi-dimensional small perturbation model (SPM) by investigating the sensitivity of backscattering to the new surface parameters and to the dielectric constant.

## 2. A TWO DIMENSIONAL MULTI-SCALE DESCRIPTION OF NATURAL ROUGHNESS

In this section, we describe and discuss the employed surface model. In this paper, we model natural roughness as a multi-scale process having a $1/f$ spectrum with a finite range of spatial scales going from a few millimeters $b$ ($b \leq \frac{\lambda}{10}$) to several meters ($B \leq$ resolution cell) [1, 2]. The surface is considered as a superposition of a finite number of one-dimensional gaussian processes each one having a spatial scale [1] characterized by:

$$z_p(x) = \sum_{m=-p_1}^{p_2} \sum_{n=-\infty}^{+\infty} z_n^m \psi_n^m(x/L)$$

where $z_n^m$ is a collection of gaussian random independent variables with variance $\gamma_0^2 2^{-m\nu}$, $x$ a normalized distance with respect to an arbitrary length $L = 2^b b$ and $\psi_n^m$ a collection of orthonormal wavelet (4th Daubechies).

As natural roughness changes from one direction to another, one-dimensional profile are insufficient. Thus, bi-dimensional profiles are required to describe more adequately natural surfaces. Wavelet theory can be extended from one-dimensional to two-dimensional case using the separable dyadic multi-resolution analysis introduced by Mallat [1, 4, 9]]. The bi-dimensional wavelet transform gives us respectively the vertical wavelet component (1), the horizontal wavelet (2) component and the diagonal wavelet component (3) of the height $z_p^i$, (where $i$ = Vertical, Horizontal or Diagonal) considered as a $1/f$ process over a finite range of spatial scales going from an inner spatial scale $b$ of a few millimeters to an outer spatial scale $B$ of several meters.

$$z_p^V(x,y) = \sum_{m_x=0}^{p}\sum_{m_y=0}^{p}\sum_{n_x=-\infty}^{+\infty}\sum_{n_y=-\infty}^{+\infty} z_{n_x}^{m_x}\, z_{n_y}^{m_y}\, \psi\left(\frac{2^{m_x}}{B}x - n_x\right)\phi\left(\frac{2^{m_y}}{B}y - n_y\right) \tag{1}$$

$$z_p^H(x,y) = \sum_{m_x=0}^{p}\sum_{m_y=0}^{p}\sum_{n_x=-\infty}^{+\infty}\sum_{n_y=-\infty}^{+\infty} z_{n_x}^{m_x}\, z_{n_y}^{m_y}\, \phi\left(\frac{2^{m_x}}{B}x - n_x\right)\psi\left(\frac{2^{m_y}}{B}y - n_y\right) \tag{2}$$

$$z_p^D(x,y) = \sum_{m_x=0}^{p}\sum_{m_y=0}^{p}\sum_{n_x=-\infty}^{+\infty}\sum_{n_y=-\infty}^{+\infty} z_{n_x}^{m_x}\, z_{n_y}^{m_y}\, \psi\left(\frac{2^{m_x}}{B}x - n_x\right)\psi\left(\frac{2^{m_y}}{B}y - n_y\right) \tag{3}$$

where $z_{n_x}^{m_x}$ and $z_{n_y}^{m_y}$ are a collection of uncorrelated zero mean Gaussian random variables [1, 4].

Their associated autocorrelation function (ACF) is given by the following Equations (4) and (5):

$$\rho^i(x,y,x+\xi,y+\eta) = \left\langle z_p^i(x,y)\;\; z_p^i(x+\xi,y+\eta\,)\right\rangle \tag{4}$$

and the standard deviation is given by

$$s^2 = r_c^H(0,0) = r_c^D(0,0) = r_c^V(0,0) \tag{5}$$

The roughness multi-scale parameter $\nu$ is related to the fractal dimension and $\gamma$ is related to the standard deviation and the number of spatial scales is equal to $P$.

## 3. A TWO LAYERS MULTI-SCALE BI-DIMENSIONAL SPM MODEL FOR THE STUDY OF RADAR BACKSCATTER BEHAVIOR ON SEMI-ARID SOIL SUBSURFACES

The main purpose of our study is to develop an inversion model for soil moisture and multi-scale roughness parameters retrieval over on semi-arid soil subsurfaces using remotely sensed data. The first step of this study presented in this paper, consist to model radar backscattering over a two layer geoelectrical model [9] by taking into account volume scattering. Each layer is described as a multi-scale bi dimensional surface using our multi-scale description. The radar backscattering coefficient can be expressed as the sum of a single scattering component and a multiple scattering component (6).

$$\sigma_{qp}^0 = \sigma_{qp}^S + \sigma_{qp}^M \tag{6}$$

$q$ and $p$ indicate the polarization state of the emitted and received wave respectively $H$ for horizontal polarization and $V$ for vertical polarization. In this present work, multiple scattering component is neglected.

We have considered a two layer backscattering model which can describe backscattering from Mars surfaces which are characterized by a superficial dry layer of thickness $d$ over a second dry or wet layer of basaltic bedrock. We also take into account volume diffusion.

The total backscattering coefficient is the contribution of the two layers and a volume component (7).

$$\sigma_{qp}^0(\theta) = \sigma_{S1qp}^0(\theta) + \sigma_{V1qp}^0(\theta) + \sigma_{SS2qp}^0(\theta) \tag{7}$$

$$\text{with}\quad \sigma_{SS2qp}^0(\theta) = \frac{\cos(\theta)}{\cos(\theta_t)}\, T_{12}\, T_{21}\, e^{-\frac{2\kappa_e d}{\cos(\theta_t)}}\, \sigma_{S2qp}^0(\theta_t) \tag{8}$$

$$\sigma_{qp}^V(\theta) = \frac{1}{2}\frac{\kappa_s}{\kappa_e}\, T_{12}\, T_{21}\left(1 - e^{-\frac{2\kappa_e d}{\cos(\theta_t)}}\right)\cos(\theta)\, P_{qp} \tag{9}$$

where $\theta$ is the incidence angle and $\theta_t$ is the transmission angle, $P_{qp}$ has a value of 1.5 for the copolar case $\kappa_s$ and $\kappa_e$, are the diffusion and extinction coefficients given by Fung in [5]

$$\sigma^0_{Siqp} = \frac{k}{4}\exp(-2k^2)(\cos^2\theta)s^2|I_{qp}|W(-2k\sin\theta, 0) \tag{10}$$

$i = 1$ for the first layer and 2 for the second layer and $|I_{qp}|$ given by [5] and

$$W^{(n)}\left(-2k_x, 0\right) = \frac{2}{\pi}\int_0^\infty\int_0^\infty \left(\frac{r_c^i(\xi,\eta)}{r_c^i(0,0)}\right)^n \cos(2k_x\,\xi)d\xi d\eta \tag{11}$$

where $W^{(n)}$ is the Fourier transform of the $n$th power of the multi-scale autocorrelation function given by Mattia in [7] with $n = 1$ for the SPM model [1, 4]. Surfaces are characterized by the dielectric constant related to soil moisture, the albedo, the optical depth and surface roughness. Previous works used classical statistical parameters namely correlation length and standard deviation in the expression of the autocorrelation function $W$. The principal aim of this study is to use the multi-scale surface description in the backscattering coefficient.

## 4. SENSITIVITY ANALYSIS

We have in a first step considered the $VV$ polarization and studied the sensitivity of radar backscattering and angular trends for different multi-scale roughness, for different dielectric constants of each layer. We present this sensitivity study in Figure 1.

### 4.1. Sensitivity to Multi-scale Roughness Parameters

We have simulated the angular trends from 20 to 60 degrees of the backscattering coefficient for different roughness parameters. We kept $\gamma_0$ at 0.2 cm (Figure 2) and varied the fractal parameter $\nu_1$ from 1.3 to 2.1 in $VV$ polarization for ten spatial scales.

Surfaces with $\nu$ between 1.7 and 2.1 can be considered as smooth where as surfaces with $\nu_1 = 1.1$ are quite rough. For all the simulations, the backscattering coefficient decreases with the incidence angle. We notice that the backscattering coefficient increases when $\nu$ increases. We also notice an increasing trend of the backscattering coefficient with the multi-scale parameter $\gamma$ because the surface is rougher (Figure 2(b)).

### 4.2. Sensitivity to Soil Moisture

Soil moisture is related to the complex dielectric constant $\varepsilon$. In the Figure 3, we have represented radar backscattering as a function of the module of the complex permittivity of the second layer is. We notice that the backscattering coefficient $\sigma_0$ decreases for the low values of $\varepsilon_1$ and $\varepsilon_2$ and after passing by a minimum and then increases. This can be explained by the fact that when the layers are dry corresponding to a lower humidity and as a consequence a lower dielectric constant the penetration of the signal is more important and the backscattered signal is lower. As the

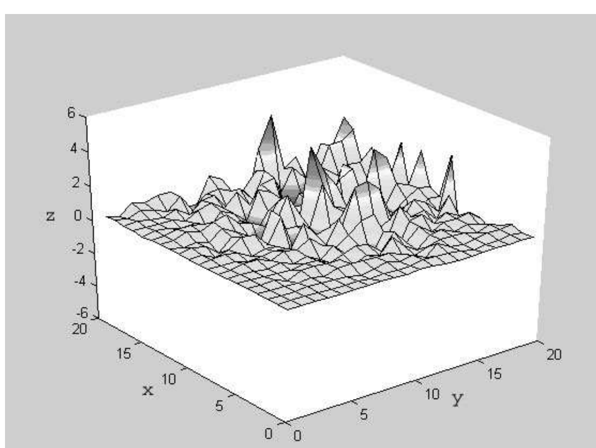
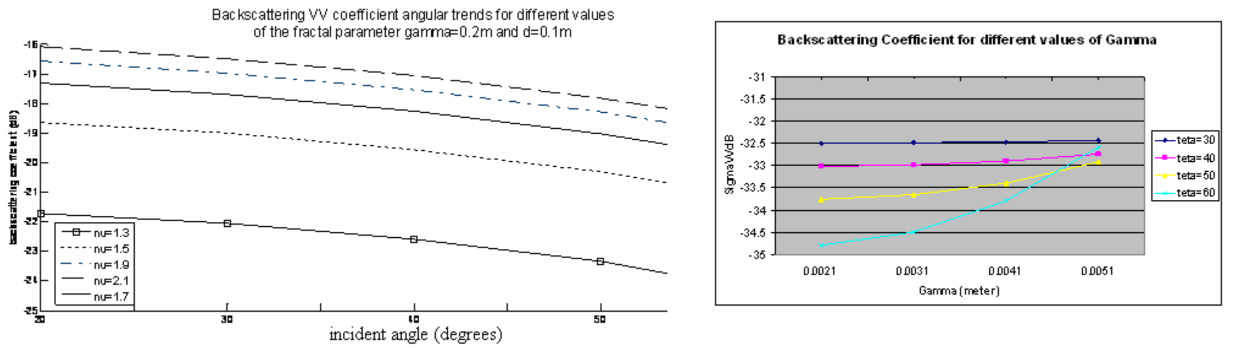



Figure 1: 3D representation of a multi-scale surface with multi-scale parameters $\nu_x = \nu_y = 1.1$; $\gamma_x = \gamma_1 = 0.2$ cm; $\gamma_y = \gamma_2 = 0.3$ cm. $Z_{\max} = 3.7$ cm.

dielectric constant increases, the surfaces and subsurface become wetter and the backscattered signal increases because the penetration is lower.




(a) (b)

Figure 2: Backscattering coefficient dependence on fractal parameter $\nu$.

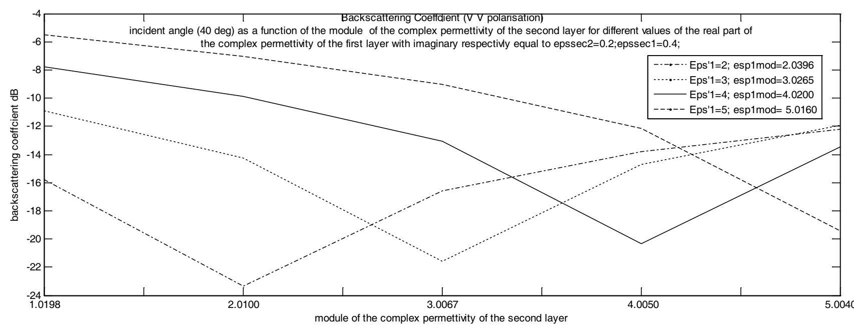



Figure 3: Backscattering coefficient as a function of the module of the complex permittivity of the second layer for different values of the complex permittivity of the first layer.

## 5. CONCLUSIONS

In this study, we have used a multi-scale roughness description using the wavelet transform and the Mallat algorithm to describe natural surface roughness and investigated the impact of this description on radar backscattering using a two layers multi-scale SPM model through a sensitivity analysis of backscattering coefficient to the multi-scale roughness parameters. We noticed an increasing trend of radar signal with the dielectric constant due to a less penetration of the signal. The overall objective of this work is to predict correctly surface and volume scattering in order to retrieve roughness and soil moisture parameters by inverting radar polarimetric signals in a future work.

## REFERENCES

1. Bennaceur, L., Z. Belhadj, and M. R. Boussema, "A study of radar backscattering multi-scale bi dimensional surface," *Proceedings IEEE IGARSS 2002*, Toronto, Canada, June 2002.
2. Davidson, M., T. Le Toan, F. Mattia, G. Satalino, T. Manninen, and M. Borgeaud, "On the characterization of agricultural soil roughness for radar sensing studies," *IEEE Trans. Geosc. Rem. Sens.*, Vol. 38, 630–640, 2000.
3. Daubechies, I., *Ten Lectures on Wavelets*, CBMS-NFS Lecture Notes, NR. 61, SIAM, 1992.
4. Farah, L. B., I. R. Farah, R. Bennaceur, Z. Belhadj, and R. Boussema, "Soil multi-scale roughness parameters and soil moisture retrieval from radar backscattering using a neural network technique," *PIERS Proceedings*, Tokyo, Japan, August 2–5, 2006.
5. Fung, A. K., *Microwave Scattering and Emission Models and Their Applications*, Artech House, 1994.
6. Mallat, S. G., "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, 674–693, July 1989.

7. Mattia, F. and T. Le Toan, "Backscattering properties of multi-scale rough surfaces," *Journal of Electromagnetic Waves and Applications*, Vol. 13, 493–527, 1999.
8. Paillou, P., Y. Lasne, E. Heggy, J. M. Malezieux, and G. Ruffie, "A study of P-band synthetic aperture radar applicability and performance for Mars exploration: Imaging subsurface geology and detecting shallow moisture," *Journal of Geophysical Research*, Vol. 111, 2006.

# Soil Moisture Retrieval Using Data Cube Representation of Radar Scattering

**Seung-Bum Kim and Eni G. Njoku**
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

**Abstract**— A time-series algorithm is proposed to retrieve surface (from surface down to 1 m depth) soil moisture using the simulated radar data. The time-series approach uses co-polarized ($VV$ and $HH$) backscattering coefficient ($\sigma^0$) values. Temporal averaging is applied to reduce the radar measurement noise. To the extent that the surface roughness does not change within the time-series window, the reduction of the noise enables the retrieval of the roughness. With the roughness estimate, subsequently soil moisture is retrieved. The proposed retrieval is performed using 'data cubes'. The data cubes relate soil moisture and $\sigma^0$, and are lookup tables with the dimensions of soil moisture, roughness, and vegetation water content (VWC). The cubes are generated by the first-order small perturbation model and the discrete scatterer model for the grass vegetation. A Monte-Carlo analysis demonstrates that the soil moisture may be retrieved within the error better than $0.06\,\mathrm{cm^3/cm^3}$ up to about $3\,\mathrm{kg/m^2}$ VWC using six time-series records, although presently assuming that the radar model correctly describes the surface scattering processes.

## 1. INTRODUCTION

Radar returns from the land surface consist of scattering by the ground, the vegetation, and the interaction between the two scattering media. Major parameters determining the radar backscattering coefficient ($\sigma^0$) are soil moisture, surface roughness properties, and vegetation biomass. For bare or short vegetation surfaces, semi-empirical or analytical models of scattering were constructed (e.g., [1, 2]). Inversion of these models is available in either closed-form or approximations. Although the bare surface study is an important basis of soil moisture retrieval, most of the global land surfaces are vegetated. Often the effect of the vegetation scattering is formulated empirically based on limited sets of field observations (e.g., [3]). Whether the empirical forms apply well to more general cases needs examination before retrieving over larger areas.

To extend the soil moisture retrieval beyond the bare surface and beyond the validity range of the empirical forms, one may utilize a radar scattering model that evaluates the scattering from the vegetation as well as from the ground. The radar scattering models may describe the scattering processes in more general forms than empirical forms do. $\sigma^0$ from the scattering model may be presented with three axes, soil moisture, roughness, and vegetation water content. We call such presentation as 'data cube' (see Section 2 for more details of the data cube). This paper studies a soil moisture retrieval strategy using the data cubes.

A previous study of the soil moisture retrieval using the data cubes indicated that the surface roughness needs to be known well, in order to retrieval soil moisture with reasonable accuracy ($0.06\,\mathrm{cm^3/cm^3}$, [4]). Instead, when the roughness becomes an unknown during the retrieval, the soil moisture estimate error quickly exceeds $0.1\,\mathrm{cm^3/cm^3}$ most likely because the roughness has a strong effect on $\sigma^0$. However, a good knowledge of roughness is hard to obtain. In the present study, the roughness is retrieved using time-series observations of $\sigma^0$. The retrieved roughness value is then used to improve the soil moisture retrieval.

## 2. METHODOLOGY

### 2.1. Data Cube Representation of Radar Scattering

The proposed retrieval is performed using 'data cube' representation of the radar scattering processes. The data cubes relate soil moisture and $\sigma^0$, and are lookup tables with the dimensions of the parameters most influential in characterizing $\sigma^0$. The number of dimension is set to 3, because too large dimension means too many degrees of freedom during retrieval. These parameters are bare surface roughness ($ks$, with $k$ and $s$ being the wave number and rms height of the bare soil surface respectively), volumetric soil moisture ($m_v$), and VWC. The correlation length and correlation function of the bare surface are also important in characterizing the scattering from the bare surface. Different sets of the data cubes may be generated for each of the choices of the correlation length and function, and the optimal set may be chosen such that the $m_v$ retrieval performs best. VWC is a lump-sum parameter representing various vegetation parameters such as dielectric constant, orientation and number of scatterers, and layers of canopy. It is practical to let VWC to represent all the vegetation information.

The data cubes are generated by a radar scattering model for a vegetation-covered soil layer, where the inputs to the model are $m_v$, $ks$, VWC, and other parameters describing bare surface and vegetation, and the outputs are multi-channel $\sigma^0$. The scattering model consists of three components:

$$\sigma^0(ks, m_v, VWC) = \sigma^s(ks, m_v)\exp(-2\tau(VWC)\cos\theta) + \sigma^{sv}(ks, m_v, VWC) + \sigma^v(VWC) \quad (1)$$

Here $\sigma^0$ on the left-hand side of the equation represents the total backscattering coefficient in $HH$, $VV$, and $HV$ polarizations as a function of $ks$, $m_v$, and VWC, $\sigma^s$ represents the scattering from the bare soil surface modified by the two-way vegetation attenuation, $\sigma^v$ is the scattering of the vegetation volume, and $\sigma^{sv}$ denotes the scattering interaction between the soil and vegetation. The vegetation attenuation is modeled with the opacity, $\tau$, as a function of VWC. The attenuation is further modified by the incidence angle, $\theta$. The scattering model used in this study incorporates the 1st order small perturbation model for the bare surface and the discrete scatterer model for the vegetation (SPM/DSM) as implemented [5] for the grass vegetation surface. Figure 1 shows the data cubes for the co-pol $\sigma_0$ obtained with the SPM/DSM model. The SPM is valid up to $ks$ of 0.3 and other surface scattering models are under development to extend the range of $ks$. In general, co-pol $\sigma^0$ increases with $ks$, $m_v$, and VWC. Also $\sigma^0_{HH}$ is typically smaller than $\sigma^0_{VV}$ over the low-vegetation surfaces. These are clearly shown by the data cubes in Figure 1.

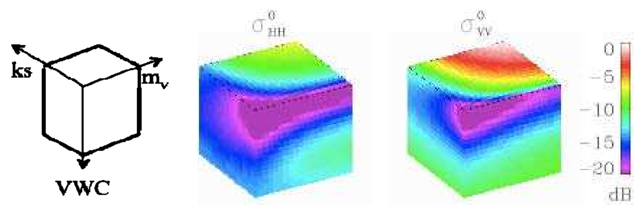



Figure 1: Data cubes of co-pol $\sigma^0$ generated by the SPM/DSM radar scattering model. The ranges of the axes are 0.1 to 2 (dimensionless), 0 to $0.6\,\mathrm{cm^3/cm^3}$, and 0 to $5\,\mathrm{kg/m^2}$ for $ks$ ($k$ and $s$ are the wave number and rms height, respectively), $m_v$ (soil moisture), and VWC (vegetation water content), respectively. $\sigma^0$ values saturate to white and purple color. The SPM is valid up to $ks$ of 0.3 and other surface scattering models are under development to extend the range of $ks$.

### 2.2. Soil Moisture Retrieval with the Time-series $\sigma_0$ and Data Cubes

Retrieval is in principle the inversion of the forward scattering processes. Given $\sigma^0$ observations, the goal is to estimate $ks$, $m_v$, and VWC. It is always advantageous to reduce the degree of freedom during retrieval: in this respect, VWC may be retrieved from the cross-pol $\sigma^0$ or be provided based on optical images. This study proposes to estimate $ks$ first, reduce the degree of freedom, and then retrieve $m_v$ using the estimate of $ks$. The retrievals of VWC and $ks$ let us choose the correct slices of VWC and $ks$ in the data cube. A slice refers to a 2-D plane perpendicular to an axis of the cube. Then the retrieval of $m_v$ becomes a search within a one-dimensional series of $\sigma^0$ that is provided by the intersection of the VWC slice and the $ks$ slice.

The choice of the $ks$ slice (equivalently, $ks$ retrieval) is based on the assumption that $ks$ does not change over a period of time-series $\sigma^0$ measurements. The assumption would hold in the absence of heavy rainfall or harvest. Despite the changes in VWC and $m_v$, $\sigma^0$ measurements will be located on the same $ks$ slice of the data cubes if there were no radar measurement noise. The radar measurement noise is denoted by 'Kp noise' with the value of Kp representing $1\sigma$ of the Gaussian noise. The presence of the Kp noise displaces the $\sigma^0$ measurements away from the same $ks$ slice. The retrieval of $ks$ is to determine the same single slice for $ks$ where all the time-series $\sigma^0$ are located. Noting that the displacement is caused by the Gaussian Kp noise with zero-mean, temporal averaging will reduce the effect of the noise for the purpose of $ks$ estimate. $ks$ is estimated such that the time-series $\sigma^0$ at the estimated $ks$ slice is closest to the $\sigma^0$ observations.

The retrieval of $ks$ and $m_v$ is formulated as follows using $N$ time-series sets of two co-pol $\sigma^0$ measurements (Eqs. (2a)–(2c)). There are $2N$ independent observations and $N + 1$ unknowns consisting of $Nm_v$ and one $ks$ value. The values of $ks$ and $m_v$ are discrete on their axes in the data cube. Therefore the retrieval of these parameters is to find the corresponding indexes ($iks_{rtr}$ and $imv_{rtr}$, where the subscript $rtr$ denotes retrieval). First, at each candidate of $iks_{rtr}$, $m_v$ is estimated using each of the time-series pair of $\sigma_0$ measurements such that the difference between 'retrieved' and 'measured' $\sigma^0$ is minimized in the least-square sense. The difference is denoted by $\Delta\sigma^0_{mv}(imv_{rtr}(i))|_{iks=iks\_rtr}$ in Eq. (2a), where $i$ is the index representing the number of time-series records and $i$ ranges from 1 to $N$. In these equations, $ich$ refers to the index for radar channels and $w$ denotes weights. This process provides an estimate of $m_v$, $imv_{rtr}(i)$, that is temporary and not final yet. Here the 'retrieved' $\sigma^0$ refers to $\sigma_0$ evaluated by the estimates of roughness and soil moisture ($iks_{rtr}$ and $imv_{rtr}$, respectively). Although the expression of 'measured' is used, actual measurements of $\sigma^0$ has not yet been used in this study, but instead the measured $\sigma^0$ is simulated using the known value of $ks$, $m_v$ and VWC. These known values are used as 'truth' when assessing the retrieval performance, and therefore are denoted by the subscript 'true' in Eq. (2b). The measured $\sigma^0$ includes the multiplicative Kp noise.

$$\Delta\sigma^0_{mv}(imv_{rtr}(i))\big|_{iks=iks\_rtr} = \sum_{ich}^{HH,VV} w_{ich}\Delta\sigma^0_{ich}(iks_{rtr}, imv_{rtr}(i)) \tag{2a}$$

$$\Delta\sigma^0_{ich}(iks_{rtr}, imv_{rtr}) = (\sigma^0_{ich,mea}(iks_{true}, imv_{true}, ivwc_{true}) - \sigma^0_{ich,rtr}(iks_{rtr}, imv_{rtr}, ivwc_{true}+\delta))^2 \tag{2b}$$

$$\Delta\sigma^0_{ks}(iks_{rtr}) = \frac{1}{N}\sum_{i=1}^{N} w_i\Delta\sigma^0_{mv}(imv_{rtr}(i))\big|_{iks=iks\_rtr} \tag{2c}$$

$\delta$ represents a random noise accounting for the 10% error of the VWC knowledge ($ivwc_{true}$). Figures 2(a)–(c) provide an example of how the $m_v$ retrieval works for one pair of $\sigma_0$ measurements. The plots illustrates the determination of $imv_{rtr}(i)$. $\sigma^0_{HH}$ increases with soil moisture and roughness while its contours span from lower-left to upper-right corners of the image (Figure 2(a)). Along the contour the difference between the retrieved and measured $\sigma^0$, $\Delta\sigma^0_{mv}(imv_{rtr}(i))|_{iks=iks\_rtr})$, stays the same and therefore the difference has the contours of similar shape (Figure 2(b)). Note that Figure 2(a) shows one channel but Figure 2(b) is constructed using two channels, which explains why the shapes of the contours are not identical between the two figures. The behavior of the distance for the $m_v$ retrieval ($\Delta\sigma^0_{mv}$) with respect to soil moisture is shown in Figure 2(c), where the

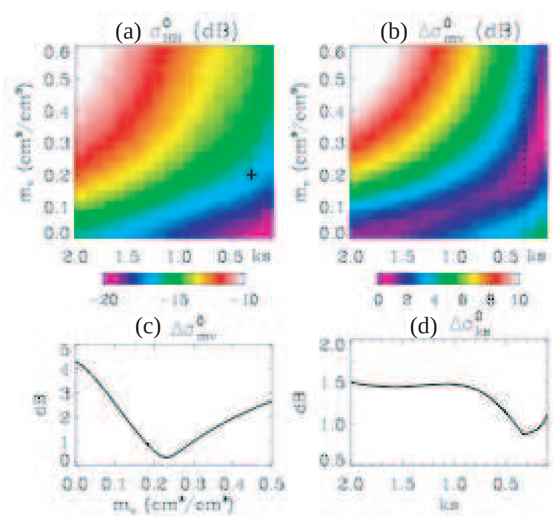



Figure 2: Example of soil moisture ($m_v$) retrieval at one time instance, and of roughness ($ks$) retrieval over the entire time-series. In (a), '+' sign indicates the location of the truth $m_v$ ($0.2\,\mathrm{cm^3/cm^3}$) and truth $ks$ (0.3) on $\sigma^0_{HH}$. $\sigma^0_{HH}$represents the data cube sliced at VWC of $2\,\mathrm{kg/m^2}$. (b), (c) illustrate the retrieval of $m_v$. (b) shows the difference between the retrieved and measured $\sigma_0$ after the difference is accumulated over the two channels, which is denoted by $\Delta\sigma^0_{mv}$ in Eq. (2a). The measured $\sigma_0$ is the truth ('+' sign in (a)) multiplied by the Kp noise. The dotted line marks the retrieved $ks$ value. (c) shows $\Delta\sigma^0_{mv}$ of (b) along the dotted line in (b), where the minimum $\Delta\sigma^0_{mv}$ gives the $m_v$ retrieval. (d) illustrates the $ks$ retrieval, which searches the minimum of $\Delta\sigma^0_{ks}$ defined in Eq. (2c). The Kp noise level is 17% ($1\sigma$ of the Gaussian noise) and 10% error was introduced to VWC knowledge.

minimum $\Delta\sigma^0_{mv}$ gives the $m_v$ retrieval of $0.23\,\mathrm{cm}^3/\mathrm{cm}^3$. The truth $m_v$ was $0.2\,\mathrm{cm}^3/\mathrm{cm}^3$, resulting in $0.03\,\mathrm{cm}^3/\mathrm{cm}^3$ retrieval error.

The roughness ($ks$) is now retrieved so as to minimize the average over the entire time span of the differences between retrieved and measured $\sigma^0$ (denoted as $\Delta\sigma^0_{ks}(iks_{rtr})$ in Eq. (2c)). Once the $ks$ is retrieved, which is denoted as $iks_{rtr}$, the corresponding indexes for $N$ number of $m_v$ values have already been estimated by Eq. (2a). An example of the $ks$ retrieval is shown in Figure 2(d). $\Delta\sigma^0_{ks}$ varies between 0.5 dB and 1.5 dB. The $ks$ retrieval is 0.32 where the minimum $\Delta\sigma^0_{ks}$ is found. The truth $ks$ was 0.3.

## 3. MONTE-CARLO SIMULATION OF RETRIEVAL

Monte-Carlo analyses are performed to assess the retrieval performance in various soil conditions. There are $6 \times 11$ nodes of the Monte-Carlo experiment: $m_v$ increases from 0 to $0.5\,\mathrm{cm}^3/\mathrm{cm}^3$ by 0.1, VWC increases from 0 to $5\,\mathrm{kg/m}^2$ by 0.5, and $ks$ is fixed. Two Gaussian random variables characterize the Monte-Carlo simulation: the ancillary VWC information with 10% error ($1\sigma$) and the Kp noise with 17% ($1\sigma$) multiplicative error. 300 realizations of the two Gaussian random variables are made per node (the retrieval performance was similar between 300 and 3000 realizations). 6 time-series records are used and the simulation is performed with $ks$ of 0.3 that corresponds to 1 cm rms height of the surface.

The Monte-Carlo analysis of the time-series $m_v$ retrieval is shown in Figure 3(a). The retrieval error increases with $m_v$, because the sensitivity, $\partial\sigma^0/\partial m_v$, decreases with $m_v$ (Figure 3(b)). The attenuation of the surface scattering by the vegetation becomes more significant with the higher VWC. As a result the sensitivity decreases and the retrieval error becomes larger.

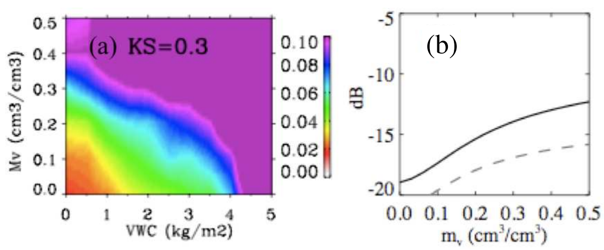



Figure 3: (a) rms errors from the Monte-Carlo simulation of the time-series $m_v$ retrieval. The error of greater than $0.1\,\mathrm{cm}^3/\mathrm{cm}^3$ saturates to the purple color. (b) The response of $VV$ (solid) and $HH$ (dash) $\sigma^0$ to $m_v$ with $0.3ks$ and $1.1\,\mathrm{kg/m}^2$ VWC.

## 4. CONCLUSION

The "data-cube time-series" algorithm is proposed to retrieve the soil moisture ($m_v$) using L-band radar observations with the Soil Moisture Active and Passive (SMAP) mission as an example. The algorithm relies on the data cube representation of radar scattering. Through the temporal averaging at the same location on the ground, the effect of radar measurement noise (Kp) on the estimate of the surface roughness is reduced and as a result the roughness retrieval improves significantly. With the $ks$ estimate, $m_v$ is retrieved such that the estimated $\sigma^0$ matches the observed $\sigma^0$ best. The proposed approach is tested with the Monte-Carlo simulation. $m_v$ may be retrieved with a rms error of $0.06\,\mathrm{cm}^3/\mathrm{cm}^3$ or better up to the vegetation water content (VWC) of about $3\,\mathrm{kg/m}^2$. The ancillary VWC information has 10% error ($1\sigma$), the Kp noise is a 17% ($1\sigma$) multiplicative error, and 6 time-series records are used as input. The retrieval error will improve since the SMAP radar offers the Kp noise down to 13% level. However, the errors in the data cube representation of the surface scattering have not been accounted for yet, which will increase the retrieval error.

## ACKNOWLEDGMENT

The authors would like to thank SMAP Science Definition Team members for valuable suggestions. This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

## REFERENCES

1. Dubois, P. C., J. J. van Zyl, and E. T. Engman, "Measuring soil moisture with imaging radars," *IEEE Trans. Geosci. Remote Sens.*, Vol. 33, 915–926, 1995.

2. Shi, J. C., J. R. Wang, A. Y. Hsu, P. E. O'Neill, and E. T. Engman, "Estimation of bare surface soil moisture and surface roughness parameter using L-band SAR image data," *IEEE Trans. Geosci. Remote Sens.*, Vol. 35, 1254–1266, 1997.
3. Joseph, A. T., R. van der Velde, P. E. O'Neill, R. H. Lang, and T. Gish, "Soil moisture retrieval during a corn growth cycle using L-band (1.6 GHz) radar observations," *IEEE Trans. Geosci. Remote Sens.*, Vol. 46, 2365–2374, 2008.
4. Kim, S. B., E. G. Njoku, and J. J. van Zyl, "Monitoring surface soil moisture and freeze-thaw state with the high-resolution radar of the Soil Moisture Active/Passive (SMAP) mission," *IEEE Radar Conference 2010*, Washington DC, 2010, CD-ROM & http://ieeexplore.ieee.org.
5. Arii, M., "Retrieval of soil moisture under vegetation using polarimetric radar," 201, Ph.D. thesis, California Institute of Technology, 2009.

# Azimuthal Signature of Coincidental Brightness Temperature and Normalized Radar Cross-section Obtained Using Airborne PALS Instrument

**Andreas Colliander[1], Seungbum Kim[1], Simon Yueh[1], Mike Cosh[2], Tom Jackson[2], and Eni Njoku[1]**

[1]Jet Propulsion Laboratory, California Institute of Technology, USA
[2]USDA Agriculture Research Service, USA

**Abstract**— Coincidental airborne brightness temperature (TB) and normalized radar-cross section (NRCS) measurements were carried out with the PALS (Passive and Active L- and S-band) instrument in the SMAPVEX08 (SMAP Validation Experiment 2008) field campaign. This paper describes results obtained from a set of flights which measured a field in 45° steps over the azimuth angle. The field contained mature soy beans with distinct row structure. The measurement shows that both TB and NRCS experience modulation effects over the azimuth as expected based on the theory. The result is useful in development and validation of land surface parameter forward models and retrieval algorithms, such as the soil moisture algorithm for NASA's SMAP (Soil Moisture Active and Passive) mission. Although the footprint of the SMAP will not be sensitive to the small resolution scale effects as the one presented in this paper, it is nevertheless important to understand the effects at smaller scale.

## 1. INTRODUCTION

The airborne PALS (Passive and Active L- and S-band) instrument was used in the SMAPVEX08 (SMAP Validation Experiment 2008) campaign [2] to measure the brightness temperature (TB) and normalized radar cross-section (NRCS) at L-band during a period of several days over an area spanning tens of kilometers in Maryland and Delaware. SMAP (Soil Moisture Active and Passive) is a NASA mission to measure global soil moisture and boreal land surface freeze/thaw state [1]. The satellite will carry radar (active) and radiometer (passive) L-band instruments that will perform simultaneous and coincident measurements of the Earth's surface. The airborne PALS instrument is an airborne simulator for SMAP.

Electromagnetic theory and experimental results indicate that over certain type of surface and landcover both radiometer and radar measurements may experience signal modulation over the azimuth angle with respect to the measurement location. This effect may be due to reflection symmetry of the surface (e.g., [3]) or the Bragg scattering, or the combination of the two effects. In both cases, the first order requirement for this azimuthal signature to manifest itself in the measured signal of natural targets is essentially a periodic structure on the surface, sub surface (within penetration depth), or in the layer above surface of the measured medium. Over ocean surfaces the reflection symmetry phenomenon has actually been utilized to measure the wind vector (e.g., [4, 5]). However, in the observation of the water content in the soils the azimuthal signature may cause degradation in the retrieval if the effect is not accounted for in the forward modeling. At the same time, a critical factor in the azimuthal signature affecting the measurement is the ratio between the size of the periodic surface and the instrument footprint on the ground and orientation of periodic area, or areas. A large footprint effectively averages the surface features over the entire footprint area and, when considering realistic surface conditions distributions, the net effect of possible periodic surface structures within the footprint may reduce to negligible.

It is not expected that the resolution scale of SMAP, which ranges from 3 km of the synthetic aperture radar to 40 km of the radiometer, would experience significant azimuth effects such as the ones observed in this or many other experiments. However, it is nevertheless important for the mission algorithm development to understand the fine resolution scattering and emission behavior. These results can be used, for example, for model development and validation. To this end, this paper presents results obtained on one of the SMAPVEX08 campaign days when the PALS instrument measured a crop field in 45 degree steps over the azimuth. In the analysis the azimuthal behavior of both radar and radiometer measurement signals are quantified and the geophysical explanation discussed.

## 2. THEORETICAL BACKGROUND

The symmetry properties of Maxwell's equations affect the scattering and emission coefficients of media with reflection symmetry. Natural objects with periodically structured surface are generally reflection symmetric. The co- and cross-polarized radar cross-section and horizontally and vertically polarized brightness temperature have even symmetry and the third and fourth Stokes parameters have odd symmetry [3].

On the other hand, constructive summation of scattered waves takes place when the orientation of the structures and the relationship between the wavelength and periodicity is aligned in certain way. This is called Bragg scattering. Whereas the reflection symmetry causes smooth signature over the azimuth angle the Bragg scattering causes sharp enhancement of backscattering at the particular azimuth angle for, for example, crop rows.

The equations used for modeling the effect of the reflection symmetry are usually formulated as second-order Fourier series. For brightness temperature the equations are written as [6]:

$$T_{B,p} = T_{B,p0} + T_{p1}\cos\phi + T_{p2}\cos 2\phi \quad (1)$$

$$T_{B,q} = T_{q1}\sin\phi + T_{q2}\sin 2\phi \quad (2)$$

where $T_B$ stands for brightness temperature; $T_{p/q1/2}$ stand for the harmonic coefficients, in which $p$ denotes either first or second Stokes parameter, or modified Stokes parameter [7], and $q$ either third or fourth Stokes parameter; and $\phi$ stands for the azimuth angle with respect to the periodic surface. And for radar cross-section the equation can be written as [3, 8]:

$$\sigma^0 = A_0 + A_1\cos\phi + A_2\cos 2\phi \quad (3)$$

where $A_{0/1/2}$ stand for the harmonic coefficients.

On the other hand, the condition for the row structure of the surface to cause Bragg scattering can be formulated as:

$$n\lambda = 2d\sin\theta \quad (4)$$

where $n$ stands for integer denoting the multiples of the wavelength, $\lambda$ stands for the wavelength, $d$ stands for the spacing of the periodicity, $\theta$ stands for the incidence angle.

It is expected that the response of PALS in SMAPVEX08 azimuth experiment is composed of these two components: reflection symmetry and Bragg scattering.

## 3. EXPERIMENT

The SMAPVEX08 soil moisture field experiment took place in Maryland and Delaware from September 29 to October 13, 2008. This study focuses on a set of flights carried out over one crop field growing soy beans.

In the campaign the PALS instrument was mounted on Twin Otter aircraft with incidence angle of about 40°. The flight altitude over the field was 1 km where the antenna with 20° beamwidth allowed footprint of about 450 × 600 m. The radar of PALS operates at about 1.26 GHz and the radiometer at 1.413 GHz. In the SMAPVEX08 campaign the PALS instrument was flown with an Agile Digital Detector (ADD) for mitigation of Radio Frequency Interference (RFI) [2]. The resolution of PALS radiometer and radar are specified in $< 0.2$ K and $< 0.2$ dB range, respectively. PALS flew over the field in varying azimuth angles in 45° steps over the full 360°. Figure 1 shows the ground tracks of the footprint centers over the field.

The soy beans on the field were on mature state at the time of the experiment. The spacing between the rows was about 20 cm. The water content in the soil was relative high being slightly less than $0.3\,\mathrm{cm^3/cm^3}$.

Figure 2 shows the row spacing which causes Bragg scattering under the conditions of the PALS measurements in SMAPVEX08 campaign as a function of the distance from the aircraft nadir point. The antenna beam and the beam incidence angles within the main lobe is collocated with the horizontal axis (the distance from nadir point) and marked with red color.

## 4. RESULTS

The radar and radiometer measurements over the footprint intersection were binned based on the azimuth angle. The measurements were collected over a circular area with radius 250 m (see the magenta circle in Figure 1) and binned in 8 azimuth bins. Brightness temperatures and radar

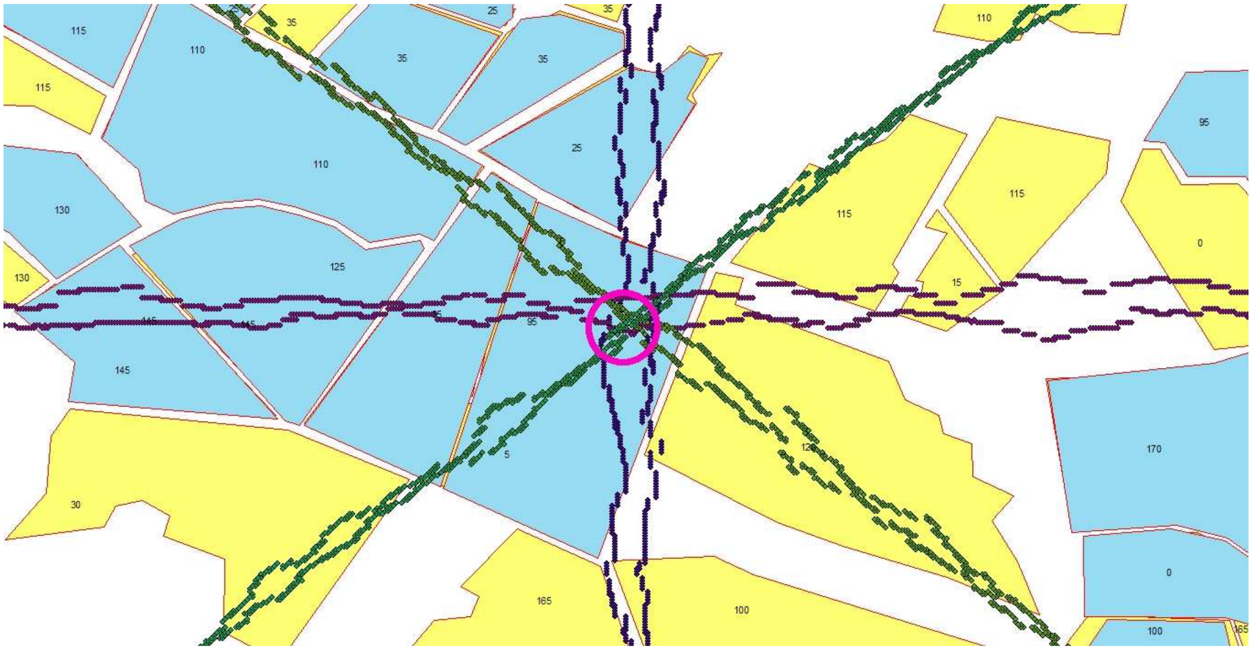


Figure 1: The ground tracks of the footprints in the 1-km altitude azimuth experiment. The magenta circle shows the area over which the data is analyzed.

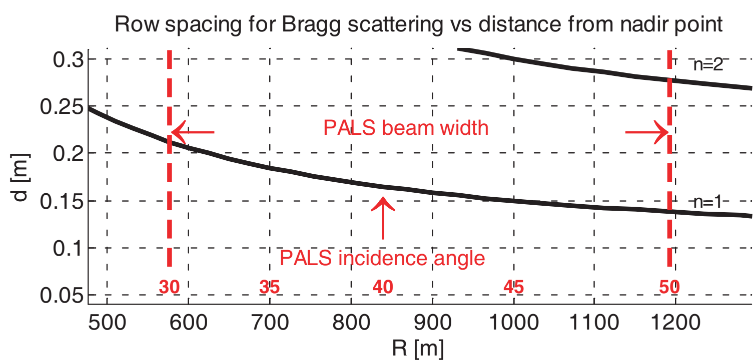



Figure 2: The row spacing allowing Bragg scattering as a function of the distance from the aircraft nadir point in case of $n = 1$, 2 and 3 (multiples of wavelength).

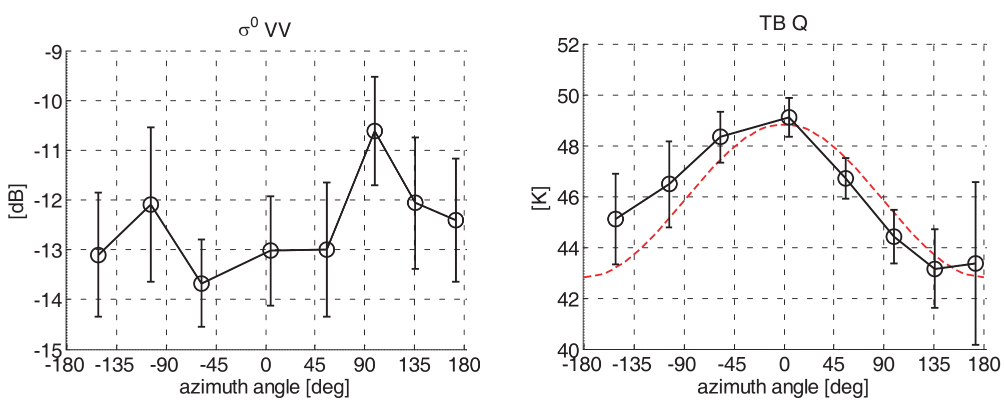



Figure 3: VV-polarized NRCS (left) and the second Stokes parameter (right) as a function of the azimuth angle. The error bars show the spread of the measurement made at that particular angle. The red dashed line on the TB plot shows a cosine curve fitted to the measurement points.

cross-sections in these bins were combined using the inverse distance squared weighting [9] to obtain representative values for TB and NRCS. Additionally, the standard deviations of the values in the bins were calculated to estimate the variance of the measurement at each angle.

Figure 3 shows the VV-polarized normalized radar cross-section (left) and the second Stokes parameter (which is defined as vertical polarization minus horizontal polarization) of brightness temperature (right) as a function of the azimuth angle ($0^\circ$ is parallel to the row structure). The error bars indicate the standard deviation of the measurement samples in the azimuth bins.

Figure 2 indicates that the row spacing of the field would cause Bragg scattering in one part of the footprint. This backscatter enhancement can be observed at $-90^\circ$ and $90^\circ$ in Figure 3. This is consistent with Bragg scattering effect which occurs in perpendicular orientation from the row direction. The Bragg scattering contribution is in the order of 1 to 3 dB. The fact that the whole footprint does not satisfy the condition for the Bragg scattering obviously lowers the contribution (see Figure 1). The scattering symmetry effect, however, cannot be detected from the radar cross-section measurements.

In the second Stokes parameter measurement the signature is dominated by the first harmonic coefficient of the Fourier series in Equation (1). A cosine curve (dashed red line in Figure 3) is fitted to the measurement with about 3 K magnitude. This would imply that the medium under investigation, defined by mature soy beans over moist soil surface, would be reflection symmetric.

## 5. CONCLUSIONS

Airborne measurements of brightness temperature and normalized radar cross-section of a crop field over the azimuth were presented. The measurement results suggest that the mature soy bean field with distinct row structure invokes the backscatter enhancement under Bragg scattering conditions. Furthermore, observations of the second Stokes parameter suggest that the soil/soy bean medium has reflection symmetric properties. The result is expected for a fine resolution scale measurement of a homogeneous field surface such as the one in question. The result is useful in assessing the contribution of periodic structures on the forward models on small resolution scale. These results can then be combined at coarser resolution while accounting for the natural distribution of targets different types of targets.

## ACKNOWLEDGMENT

The research described in this publication was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

## REFERENCES

1. "Soil moisture active and passive mission," http://smap.jpl.nasa.gov, Jet Propulsion Laboratory, California Institute of Technology.
2. Yueh, S., S. Dinardo, S. Chan, E. Njoku, T. Jackson, R. Bindlish, J. Johnson, J. Piepmeier, and C. Ruf, "PALS-ADD and airborne campaigns to support soil moisture and sea surface salinity missions," *Proc. IEEE IGARSS09*, July 12–17, 2009.
3. Yueh, S., R. Kwok, and S. Nghiem, "Polarimetric scattering and emission properties of targets with reflection symmetry," *Radio Sci.*, Vol. 29, No. 6, 1409–1420, November–December 1994.
4. Naderi, F. M., M. H. Freilich, and D. G. Long, "Spaceborne radar measurement of wind velocity over the ocean — An overview of the NSCAT scatterometer system," *Proc. IEEE*, Vol. 79, No. 6, 850–856, June 1991.
5. Gaiser, P., et al., "The windsat spaceborne polarimetric microwave radiometer: Sensor description and early orbit performance," *IEEE Trans. Geosci. Remote Sens.*, Vol. 42, No. 11, 2347–2361, November 2004.
6. Yueh, S., R. Kwok, F. Li, S. Nghiem, W. Wilson, and J. Kong, "Polarimetric passive remote sensing of ocean wind vectors," *Radio Sci.*, Vol. 29, No. 4, 799–814, July–August 1994.
7. Randa, J., et al., "Recommended terminology for microwave radiometry," NIST Technical Note 1551, National Institute of Standards and Technology, August 2008.
8. Wentz, F. J., S. Peteherych, and L. A. Thomas, "A model function for ocean radar cross sections at 14.6 GHz," *J. Geophys. Res.*, Vol. 89, No. C3, 3689–3704, May 20, 1984.
9. Shepherd, D., "A two-dimensional interpolation function for irregularly-spaced data," *Proc. ACM National Conference*, 1968.

# Wave Equations in Electromagnetic and Gravitational Fields

**Zi-Hua Weng**
School of Physics and Mechanical & Electrical Engineering
Xiamen University, Xiamen 361005, China

**Abstract**— The paper studies the wave equations for the electromagnetic field when there is the gravitational field at the same time. In the description with the algebra of octonions, the deductions of wave equations are identical with that in conventional electromagnetic theory with vector terminology. By means of the octonion exponential function, we can draw out that the electromagnetic waves are transverse waves in a vacuum. The inference can be used to rephrase the law of reflection, Snell's law, and Fresnel formula etc. The study claims that the deductions of wave equation for electromagnetic fields keep unchanged in the case for existence of gravitational fields. The electric components of electromagnetic waves can not be determined simultaneously with magnetic components in electromagnetic fields.

## 1. INTRODUCTION

In the electromagnetic field, the inferences of wave equation are believed to be correct. Until to now, this validity is confined only to one special condition that there exist only the electromagnetic fields. When there are gravitational fields, the people doubt whether the wave equation of electromagnetic field is still correct or not. The validity about wave equations is being questioned continuously, and it remains as puzzling as ever. The existing theories do not explain why the deductions of wave equation should keep unchanged, and then do not offer compelling reason for the unique situation. The paper tries to find out why the deductions of wave equation keep the same in most cases.

The algebra of quaternions was invented by W. R. Hamilton [1] in 1843, and then was first used by J. C. Maxwell [2] in 1861 to represent field equations of electromagnetic field. O. Heaviside in 1884 recast Maxwell equation in terms of vector terminology and electromagnetic forces [3], thereby reduced the original twenty equations down to the four differential equations.

In 1871, H. Helmholtz made clear the electromagnetic theory effectively [4]. Studying Helmholtz equation [5] and boundary conditions can deduce some wave features, including the law of reflection, Snell's law, Fresnel formula, and total internal reflection etc. In 1887, H. R. Hertz was the first to satisfactorily demonstrate the existence of electromagnetic radiation by constructing an apparatus to produce and detect the electromagnetic waves.

In the paper, rephrasing with the algebra of octonions, we will achieve Maxwell equation in the electromagnetic fields, and then some wave properties in a vacuum, including the electromagnetic wave equation and the transverse waves etc.

## 2. ELECTROMAGNETIC AND GRAVITATIONAL FIELDS

The wave feature of gravitational field and electromagnetic field can be described simultaneously by the octonion space, which consists of two quaternion spaces [6].

In the quaternion space, the coordinates are $r_0$, $r_1$, $r_2$, and $r_3$, with the basis vector $\mathbb{E}_g = (\mathbf{i}_0, \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3)$. The radius vector is $\mathbb{R}_g = \Sigma(r_i \mathbf{i}_i)$, and the velocity $\mathbb{V}_g = \Sigma(v_i \mathbf{i}_i)$, with $\mathbf{i}_0 = 1$. Herein $r_0 = v_0 t$. $v_0$ is the speed of gravitational intermediate boson, and $t$ is the time. $i = 0, 1, 2, 3$.

In the quaternion space for the gravitational field, the gravitational potential is $\mathbb{A}_g = \Sigma(a_i \mathbf{i}_i)$, and the gravitational strength is $\mathbb{B}_g = \Sigma(b_i \mathbf{i}_i) = \lozenge \circ \mathbb{A}_g$. While the gravitational strength $\mathbb{B}_g$ covers two parts, $\mathbf{g}/v_0 = \partial_0 \mathbf{a} + \nabla a_0$, and $\mathbf{b} = \nabla \times \mathbf{a}$. That is, $\mathbb{B}_g = b_0 + (\mathbf{g}/v_0 + \mathbf{b})$. The gauge equation is, $b_0 = \partial_0 a_0 + \nabla \cdot \mathbf{a} = 0$, with $\mathbf{a} = \Sigma(a_j \mathbf{i}_j)$, $\nabla = \Sigma(\mathbf{i}_j \partial_j)$. Herein the $\circ$ denotes the quaternion multiplication, the operator $\lozenge = \Sigma(\mathbf{i}_i \partial_i)$, with $\partial_i = \partial / \partial r_i$. $j = 1, 2, 3$.

In the quaternion space for the electromagnetic field, the basis vector is $\mathbb{E}_e = (\mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3)$, the radius vector is $\mathbb{R}_e = \Sigma(R_i \mathbf{I}_i)$, and the velocity is $\mathbb{V}_e = \Sigma(V_i \mathbf{I}_i)$. The $\mathbb{E}_e$ is independent of the $\mathbb{E}_g$, with $\mathbb{E}_e = \mathbb{E}_g \circ \mathbf{I}_0$ and $\mathbf{I}_0 \circ \mathbf{I}_0 = -1$. $V_0$ is the speed of electromagnetic intermediate boson.

These two quaternion spaces can be combined together to become an octonion space [7], with the octonion basis vector $\mathbb{E} = (\mathbf{i}_0, \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3)$. The radius vector in the octonion space is $\mathbb{R} = \Sigma(r_i \mathbf{i}_i + R_i \mathbf{I}_i)$. And that the octonion velocity is $\mathbb{V} = \Sigma(v_i \mathbf{i}_i + V_i \mathbf{I}_i)$. Herein the symbol $\circ$ denotes the octonion multiplication.

The potential of gravitational field and electromagnetic field are $\mathbb{A}_g = \Sigma(a_i \mathbf{i}_i)$ and $\mathbb{A}_e = \Sigma(A_i \mathbf{I}_i)$ respectively. They are combined together to become the field potential, $\mathbb{A} = \mathbb{A}_g + k_{eg} \mathbb{A}_e$, in the octonion space. The field strength $\mathbb{B}$ consists of the gravitational strength $\mathbb{B}_g$ and electromagnetic strength $\mathbb{B}_e$. That is, $\mathbb{B} = \lozenge \circ \mathbb{A} = \mathbb{B}_g + k_{eg}\mathbb{B}_e$, with the $k_{eg}$ being a coefficient. While the strength $\mathbb{B}_e = B_0\mathbf{I}_0 + (\mathbf{E}/V_0 + \mathbf{B})$, and covers two parts, $\mathbf{E}/V_0 = (\partial_0\mathbf{A} + \nabla A_0) \circ \mathbf{I}_0$ and $\mathbf{B} = (\nabla \times \mathbf{A}) \circ \mathbf{I}_0$. The gauge equations are $b_0 = 0$ and $B_0 = 0$. Herein, $B_0 = \partial_0 A_0 + \nabla \cdot \mathbf{A}$, with $\mathbf{A} = \Sigma(A_j\mathbf{i}_j)$.

The electric current density, $\mathbb{S}_e = q\mathbb{V}_e$, is the source of electromagnetic field in the octonion space. Meanwhile the linear momentum density is that of gravitational field. And the source $\mathbb{S}$ was devised to describe consistently the field source of electromagnetism and gravitation.

$$(\mathbb{B}/v_0 + \lozenge)^* \circ \mathbb{B} = -\mu\mathbb{S} = -(\mu_g\mathbb{S}_g + k_{eg}\mu_e\mathbb{S}_e), \tag{1}$$

where $\mu$ is a coefficient. $\mu_e$ and $\mu_g$ are the electromagnetic and gravitational constants respectively, with $k_{eg}^2 = \mu_g/\mu_e$. $q$ is the electric charge. And $*$ denotes the conjugate of octonion.

According to the basis vectors, the above can be decomposed further as follows,

$$\mathbb{B}^* \circ \mathbb{B}/v_0 + \lozenge^* \circ \mathbb{B}_g = -\mu_g\mathbb{S}_g, \tag{2}$$
$$\lozenge^* \circ \mathbb{B}_e = -\mu_e\mathbb{S}_e. \tag{3}$$

In the above, Eq. (2) is the field equation for the gravitational field, meanwhile Eq. (3) is for the electromagnetic field. Further, the latter one in the above can be rewritten as,

$$\nabla \cdot \mathbf{B} = 0, \tag{4}$$
$$\partial_0\mathbf{B} + \nabla^* \times \mathbf{E}/V_0 = 0, \tag{5}$$
$$\nabla^* \cdot \mathbf{E} = -(q/\varepsilon_e)\mathbf{I}_0, \tag{6}$$
$$\partial_0\mathbf{E}/V_0 + \nabla^* \times \mathbf{B} = -\mu_e\mathbf{S}, \tag{7}$$

where the coefficient $\varepsilon_e = 1/(\mu_e V_0^2)$. $\mathbf{S} = \Sigma(S_j\mathbf{I}_j)$, $S_j = qV_j$. $\mathbf{S}_0 = S_0\mathbf{I}_0$, $S_0 = qV_0$.

By contrast with Maxwell equation in vector terminology, it is found that the above is the same as that in conventional electromagnetic theory, except for the direction of displacement current. And the gauge equation $B_0 = 0$ is different either.

Table 1: The octonion multiplication table.

| | $1$ | $\mathbf{i}_1$ | $\mathbf{i}_2$ | $\mathbf{i}_3$ | $\mathbf{I}_0$ | $\mathbf{I}_1$ | $\mathbf{I}_2$ | $\mathbf{I}_3$ |
|---|---|---|---|---|---|---|---|---|
| $1$ | $1$ | $\mathbf{i}_1$ | $\mathbf{i}_2$ | $\mathbf{i}_3$ | $\mathbf{I}_0$ | $\mathbf{I}_1$ | $\mathbf{I}_2$ | $\mathbf{I}_3$ |
| $\mathbf{i}_1$ | $\mathbf{i}_1$ | $-1$ | $\mathbf{i}_3$ | $-\mathbf{i}_2$ | $\mathbf{I}_1$ | $-\mathbf{I}_0$ | $-\mathbf{I}_3$ | $\mathbf{I}_2$ |
| $\mathbf{i}_2$ | $\mathbf{i}_2$ | $-\mathbf{i}_3$ | $-1$ | $\mathbf{i}_1$ | $\mathbf{I}_2$ | $\mathbf{I}_3$ | $-\mathbf{I}_0$ | $-\mathbf{I}_1$ |
| $\mathbf{i}_3$ | $\mathbf{i}_3$ | $\mathbf{i}_2$ | $-\mathbf{i}_1$ | $-1$ | $\mathbf{I}_3$ | $-\mathbf{I}_2$ | $\mathbf{I}_1$ | $-\mathbf{I}_0$ |
| $\mathbf{I}_0$ | $\mathbf{I}_0$ | $-\mathbf{I}_1$ | $-\mathbf{I}_2$ | $-\mathbf{I}_3$ | $-1$ | $\mathbf{i}_1$ | $\mathbf{i}_2$ | $\mathbf{i}_3$ |
| $\mathbf{I}_1$ | $\mathbf{I}_1$ | $\mathbf{I}_0$ | $-\mathbf{I}_3$ | $\mathbf{I}_2$ | $-\mathbf{i}_1$ | $-1$ | $-\mathbf{i}_3$ | $\mathbf{i}_2$ |
| $\mathbf{I}_2$ | $\mathbf{I}_2$ | $\mathbf{I}_3$ | $\mathbf{I}_0$ | $-\mathbf{I}_1$ | $-\mathbf{i}_2$ | $\mathbf{i}_3$ | $-1$ | $-\mathbf{i}_1$ |
| $\mathbf{I}_3$ | $\mathbf{I}_3$ | $-\mathbf{I}_2$ | $\mathbf{I}_1$ | $\mathbf{I}_0$ | $-\mathbf{i}_3$ | $-\mathbf{i}_2$ | $\mathbf{i}_1$ | $-1$ |

## 3. OCTONION EXPONENTIAL FUNCTION

In the octonion space, defining the octonion exponential function as

$$\exp\{\mathbf{I}(a + ib)\} = \cos(a + ib) + \mathbf{I}\sin(a + ib), \tag{8}$$

where $a$ and $b$ are all real. $i$ is the imaginary unit. $\mathbf{I}$ is the octonion unit, with $\mathbf{I} \cdot \mathbf{I} = -1$.

When $b = 0$, the above reduces to one octonion, $\exp(\mathbf{I}a) = \cos(a) + \mathbf{I}\sin(a)$. When $a = 0$, Eq. (8) reduces to the exponential function,

$$\exp(\mathbf{I}ib) = \cos(ib) + \mathbf{I}\sin(ib). \tag{9}$$

The octonion exponential function $\exp(-i\mathbf{I}\alpha)$ satisfies,

$$\exp(-i\mathbf{I}\alpha) \circ \exp(i\mathbf{I}\alpha) = 1, \quad \exp(-i\mathbf{I}\alpha) \circ \exp(-i\mathbf{I}\alpha) = \exp(-i\mathbf{I}2\alpha), \tag{10}$$
$$\partial_j \exp(-i\mathbf{I}\alpha) = i\mathbf{I}K_j \circ \exp(-i\mathbf{I}\alpha), \quad \partial_j^2 \exp(-i\mathbf{I}\alpha) = K_j^2 \exp(-i\mathbf{I}\alpha), \tag{11}$$

where $\alpha = -(\Sigma K_j r_j)$, with $K_j$ being the coefficient.

In case there is only the quaternion components in the unit $\mathbf{I}$, Eq. (8) reduces to the quaternion exponential function, $\exp\{\mathbf{i}(a+ib)\} = \cos(a+ib) + \mathbf{i}\sin(a+ib)$. And that the octonion unit $\mathbf{I}$ is reduced to the quaternion unit $\mathbf{i}$, with $\mathbf{i} \circ \mathbf{i} = -1$.

## 4. WAVE EQUATION

Studying wave transmission of electromagnetic field need to solve Maxwell equation Eqs. (4)–(7). In a vacuum far away from the field sources, there does not exist free electric charge $S_0$ and electric current $\mathbf{S}$, and then field equation Eq. (3) can be reduced to,

$$\lozenge^* \circ \mathbb{B}_e = 0. \tag{12}$$

Applying the operator $\lozenge$ to the above, we have the wave equation

$$\lozenge \circ (\lozenge^* \circ \mathbb{B}_e) = 0, \tag{13}$$

or Laplace equation,

$$(\partial_0^2 + \partial_1^2 + \partial_2^2 + \partial_3^2)\mathbb{B}_e = 0. \tag{14}$$

The Laplace equation is also the wave equation in the octonion space. In a similar way, this wave equation can be obtain from Eq. (12) with the conjugate operator $\lozenge^*$ in a vacuum.

Proceeding with the operator $\partial_0$ and $\nabla$, we can obtain the wave equation about the components of field strength from the Maxwell equation directly,

$$(\partial_0^2 + \partial_1^2 + \partial_2^2 + \partial_3^2)\mathbf{E} = 0, \tag{15}$$
$$(\partial_0^2 + \partial_1^2 + \partial_2^2 + \partial_3^2)\mathbf{B} = 0. \tag{16}$$

The above means that the Eq. (14) can be separated into two parts in the above. And two strength components of electromagnetic field, $\mathbf{E}$ and $\mathbf{B}$, are both possessed of wave features, and can be detected at present.

Table 2: The operator and multiplication of the physical quantity in the quaternion space.

| definition | meaning |
|---|---|
| $\nabla \cdot \mathbf{a}$ | $-(\partial_1 a_1 + \partial_2 a_2 + \partial_3 a_3)$ |
| $\nabla \times \mathbf{a}$ | $\mathbf{i}_1(\partial_2 a_3 - \partial_3 a_2) + \mathbf{i}_2(\partial_3 a_1 - \partial_1 a_3) + \mathbf{i}_3(\partial_1 a_2 - \partial_2 a_1)$ |
| $\nabla a_0$ | $\mathbf{i}_1 \partial_1 a_0 + \mathbf{i}_2 \partial_2 a_0 + \mathbf{i}_3 \partial_3 a_0$ |
| $\partial_0 \mathbf{a}$ | $\mathbf{i}_1 \partial_0 a_1 + \mathbf{i}_2 \partial_0 a_2 + \mathbf{i}_3 \partial_0 a_3$ |

## 5. TRANSVERSE WAVE

In a similar way, for the electromagnetic wave with the angular frequency $\omega$, the field strength should be a harmonic function, $\cos \omega t$, and can be chosen as the function $\exp(-i\omega t)$.

The electromagnetic strength $\mathbf{E}$ and $\mathbf{B}$ can be written as follows,

$$\mathbf{E} = \mathbf{E}(r)\exp(-i\omega t), \quad \mathbf{B} = \mathbf{B}(r)\exp(-i\omega t), \tag{17}$$

substituting the above in Eqs. (15) and (16),

$$(-\omega^2 + \Sigma\partial_j^2)\mathbf{E}(r) = 0, \quad (-\omega^2 + \Sigma\partial_j^2)\mathbf{B}(r) = 0. \tag{18}$$

From Maxwell equations, Eqs. (5) and (7) in a vacuum, we find that the amplitude of electromagnetic strength will be increased steadily or deceased continuously. So the field strength is one hyperbolic cosine $\cos(i\alpha)$, and can be replaced by the $\exp(-i\mathbf{I}\alpha)$ further. Herein, $i$ is the imaginary unit. The wave vector is $\mathbf{K} = \Sigma(\mathbf{I}_j K_j)$, and vector radius $\mathbf{r} = \Sigma(\mathbf{i}_j r_j)$.

And then, the field strength $\mathbf{E}(r)$ and $\mathbf{B}(r)$ are

$$\mathbf{E}(r) = \mathbf{E}_0 \circ \exp(-i\mathbf{I}\alpha), \quad \mathbf{B}(r) = \mathbf{B}_0 \circ \exp(-i\mathbf{I}\alpha), \tag{19}$$

Table 3: The operator and multiplication of the physical quantity in the octonion space.

| definition | meaning |
|---|---|
| $\nabla \cdot \mathbf{S}$ | $-(\partial_1 S_1 + \partial_2 S_2 + \partial_3 S_3)\mathbf{I}_0$ |
| $\nabla \times \mathbf{S}$ | $-\mathbf{I}_1(\partial_2 S_3 - \partial_3 S_2) - \mathbf{I}_2(\partial_3 S_1 - \partial_1 S_3) - \mathbf{I}_3(\partial_1 S_2 - \partial_2 S_1)$ |
| $\nabla \mathbf{S}_0$ | $\mathbf{I}_1 \partial_1 S_0 + \mathbf{I}_2 \partial_2 S_0 + \mathbf{I}_3 \partial_3 S_0$ |
| $\partial_0 \mathbf{S}$ | $\mathbf{I}_1 \partial_0 S_1 + \mathbf{I}_2 \partial_0 S_2 + \mathbf{I}_3 \partial_0 S_3$ |

where $\mathbf{E}_0$ and $\mathbf{B}_0$ both are the constant vectors in the octonion space.

Further, substituting Eq. (19) in Eq. (18), we have the result,

$$-(\omega/V_0)^2 + \Sigma K_j^2 = 0. \tag{20}$$

In the electromagnetic field, we will find that the electromagnetic waves are the transverse waves in a vacuum. From the field equations

$$\nabla \cdot \mathbf{E} = 0, \quad \nabla \cdot \mathbf{B} = 0, \tag{21}$$

we obtain

$$\mathbf{K} \cdot \mathbf{E}_0 = 0, \quad \mathbf{K} \cdot \mathbf{E}'_0 = 0, \quad \mathbf{K} \cdot \mathbf{B}_0 = 0, \quad \mathbf{K} \cdot \mathbf{B}'_0 = 0, \tag{22}$$

where $\mathbf{E}'_0 = \mathbf{E}_0 \circ \mathbf{I}$, $\mathbf{B}'_0 = \mathbf{B}_0 \circ \mathbf{I}$.

The above states that the electromagnetic waves, $\mathbf{E}_0$ and $\mathbf{B}_0$, both belong to transverse waves. So do two new wave components, $\mathbf{E}'_0$ and $\mathbf{B}'_0$. Moreover, the amplitudes of wave components, $\mathbf{E}'_0$ and $\mathbf{B}'_0$, are the same as that of $\mathbf{E}_0$ and $\mathbf{B}_0$ respectively.

In the electromagnetic field, there exist some relationships among the wave components, $\mathbf{E}_0$ and $\mathbf{B}_0$, with two new kinds of wave components, $\mathbf{E}'_0$ and $\mathbf{B}'_0$. The field equations,

$$\nabla \times \mathbf{E} = \partial \mathbf{B}/\partial t, \quad \nabla \cdot \mathbf{E} = 0, \tag{23}$$

can be rewritten as,

$$\nabla \circ \mathbf{E} = \partial \mathbf{B}/\partial t. \tag{24}$$

Expanding the $\exp(-i\mathbf{I}\alpha)$, the above implies that

$$[\mathbf{K} \circ \mathbf{E}'_0 + \omega \mathbf{B}_0]\cosh(\alpha) + [\mathbf{K} \circ \mathbf{E}_0 - \omega \mathbf{B}'_0]\sinh(\alpha) = 0, \tag{25}$$

considering Eq. (22) and the $\alpha$ may be any value, and then

$$\mathbf{K} \times \mathbf{E}'_0 + \omega \mathbf{B}_0 = 0, \quad \mathbf{K} \times \mathbf{E}_0 - \omega \mathbf{B}'_0 = 0. \tag{26}$$

In the same way, from the field equations

$$V_0^2 \nabla \times \mathbf{B} = \partial \mathbf{E}/\partial t, \quad \nabla \cdot \mathbf{B} = 0, \tag{27}$$

we have

$$\mathbf{K} \times \mathbf{B}'_0 + \omega \mathbf{E}_0/V_0^2 = 0, \quad \mathbf{K} \times \mathbf{B}_0 - \omega \mathbf{E}'_0/V_0^2 = 0. \tag{28}$$

The above means that there are some relationships between the $\mathbb{B}_e$ with $\mathbb{B}_e \circ \mathbf{I}$. The $\mathbf{K}$ and $\mathbf{E}_0$ will yield a new wave component $\mathbf{B}'_0$, while the $\mathbf{K}$ and $\mathbf{B}_0$ produce a new wave component $\mathbf{E}'_0$. Associating with the $\mathbf{K} \times \mathbf{B}_0$ is only the $\mathbf{E}'_0$, and with the $\mathbf{K} \times \mathbf{E}_0$ is the $\mathbf{B}'_0$. This implies that it is impossible to determine simultaneously the $\mathbf{E}_0$ and $\mathbf{B}_0$ via the $\mathbf{K}$. However, we can measure the $\mathbf{E}_0$ and $\mathbf{B}'_0$ synchronously, or the $\mathbf{E}'_0$ and $\mathbf{B}_0$ at the same time. In the nature, there are many substances being able to slow down the speed of electromagnetic waves, so the electromagnetic strength may possess the wave features, including the reflection and refraction etc.

## 6. CONCLUSIONS

In the electromagnetic and gravitational fields, the field equations and their wave features can be rephrased with the algebra of quaternions and of octonions. The related conclusions include the wave equation and transverse wave etc.

In the electromagnetic field, making use of the algebra of octonions, we can deduce the wave equation, and find that electromagnetic waves belong to the transverse wave in a vacuum. Bringing in the octonion exponential function $\exp(i\mathbf{I}\alpha)$, the wave vector and electromagnetic wave components, $\mathbf{E}_0$ and $\mathbf{B}_0$, will produce two new wave components, $\mathbf{E}'_0$ and $\mathbf{B}'_0$, respectively. In contrast to conventional electromagnetic theory with vector terminology, the research points out that the $\mathbf{E}_0$ and $\mathbf{B}_0$ can not be determined simultaneously via the wave vector. But the $\mathbf{E}_0$ and $\mathbf{B}'_0$ can be measured at the same time, while the $\mathbf{E}'_0$ and $\mathbf{B}_0$ can also be.

It should be noted that the study for the deductions of wave equations in the electromagnetic fields examined only some simple cases in octonion spaces. Despite its preliminary characteristics, this study can clearly indicate that the wave equation of electromagnetic field can be deduced from the field equation with the algebra of octonions, and obtain some inferences in the vacuum far away from the field sources. For the future studies, the research will concentrate on only the electromagnetic waves transmitting in the conductors.

## ACKNOWLEDGMENT

This project was supported partially by the National Natural Science Foundation of China under grant number 60677039.

## REFERENCES


1. Hamilton, W. R., *Elements of Quaternions*, Longmans, Green & Co., London, 1866.
2. Maxwell, J. C., *A Treatise on Electricity and Magnetism*, Dover Publications Inc., New York, 1954.
3. Rothman, T. and S. Boughn, "The Lorentz force and the radiation pressure of light," *American Journal of Physics*, Vol. 77, No. 2, 122–127, 2009.
4. Moroz, A., "Quasi-periodic Green's functions of the Helmholtz and Laplace equations," *Journal of Physics A: Mathematical and Theoretical*, Vol. 39, No. 36, 11247–11282, 2006.
5. Ben-Avraham, D. and A. S. Fokas, "Solution of the modified Helmholtz equation in a triangular domain and an application to diffusion-limited coalescence," *Physical Review E*, Vol. 64, No. 1, 016114, 6 pages, 2001.
6. Gogberashvili, M., "Octonionic electrodynamics," *Journal of Physics A*, Vol. 39, No. 22, 7099–7104, 2006.
7. Weng, Z.-H., "Maxwell equation in electromagnetic and gravitational fields," *PIERS Proceedings*, Xi'an, China, March 22–26, 2010.

# Theory of the $\hat{L}(\hat{c}, \hat{n})$ Numbers and Its Application to the Slow Wave Propagation in the Circular Ferrite Waveguide

**Georgi Nikolov Georgiev**[1] **and Mariana Nikolova Georgieva-Grosse**[2]

[1]Faculty of Mathematics and Informatics, University of Veliko Tirnovo "St. St. Cyril and Methodius"
BG-5000 Veliko Tirnovo, Bulgaria
[2]Meterstrasse 4, D-70839 Gerlingen, Germany

**Abstract**— The theorem for existence and for the main properties of the $\hat{L}\,(\hat{c},\,\hat{n})$ numbers ($\hat{c}$ — real, $\hat{n}$ — restricted positive integer), is formulated with the help of three lemmas and proved numerically. Lemma 1 discloses the existence of quantities and determines them for $\hat{c} \neq \hat{l}$, ($\hat{l} = 0, -1, -2, \ldots$) as the common limits of some couples of infinite sequences of positive real numbers, constructed by means of the positive real zeros of a real Kummer confluent hypergeometric function of specially picked out parameters. Lemma 2 defines the same in case $\hat{c} = \hat{l}$ (when the function in question has simple poles) as the common limit of the sequences of $\hat{L}(\hat{l} - \hat{\varepsilon}, \hat{n})$ and $\hat{L}(\hat{l} + \hat{\varepsilon}, \hat{n} + 1)$ numbers in the sense of Lemma 1 attained, if the positive real number $\hat{\varepsilon}$ becomes vanishingly small and shows also that under the circumstance referred to $\hat{L}(\hat{l} + \hat{\varepsilon}, 1)$ approximates to zero. Lemma 3 states that for $\hat{c} = \hat{l}$ and $\hat{c} = 1 \pm \hat{l}$ it holds $\hat{L}(\hat{c}, \hat{n}) = \hat{L}(2 - \hat{l}, \hat{n})$ and $\hat{L}(1 + \hat{l}, \hat{n}) = \hat{L}(1 - \hat{l}, \hat{n})$, resp., and that $\hat{L}(0.5, \hat{n})$ and $\hat{L}(1.5, \hat{n})$ are related with the Ludolphian number $\pi$. The application of results obtained in the theory of waveguides is demonstrated.

## 1. INTRODUCTION

The name $L$ numbers has been given to the common limits of some infinite sequences of positive real numbers with terms, devised by means of the positive purely imaginary (real) zeros of definite functions, constructed with the help of complex (real) confluent hypergeometric ones [1] of specially chosen parameters and eventually also through cylindrical functions [2–6]. They appeared in the theory of azimuthally magnetized circular ferrite waveguides, sustaining normal (slow) $TE_{0n}$ ($\hat{TE}_{0\hat{n}}$) modes, based on the functions mentioned, studied for the development of microwave ferrite control components [2–6].

In this investigation the theorem for existence and for the basic features of the $\hat{L}(\hat{c}, \hat{n})$ numbers — an integral part of the aforesaid class, is advanced and substantiated numerically. Following the recently elaborated model [3, 5], it is structured as a composition of three lemmas. The first of them reveals the existence of numbers and defines them as the limits of the infinite sequences of numbers $\{\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-)\}$ and $\{\hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-)\}$ in which $\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-) = |\hat{k}_-|\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ and $\hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-) = |\hat{a}_-|\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ in case $\hat{k}_- \to -\infty$, ($\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ — $\hat{n}$th positive real zero of the Kummer function $\Phi(\hat{a}, \hat{c}; \hat{x})$ in $\hat{x}$ with $\hat{a}$, $\hat{c}$, $\hat{x}$ — real, $\hat{a} < 0$, $\hat{c} > 0$ or $\hat{a} < \hat{c} < 0$, $\hat{c} \neq \hat{l}$, $\hat{l} = 0, -1, -2, \ldots$, $\hat{k}_- = \hat{a} - \hat{c}/2$ — real, $\hat{a}_- \equiv \hat{a}$, $\hat{x} > 0$, $\hat{n}$ — positive integer, linked with $\hat{a}$ or with $\hat{a}$ and $\hat{c}$. Lemma 2 determines the quantities in case $\hat{c} = \hat{l}$ as the common limit of the sequences $\{\hat{L}(\hat{l} - \hat{\varepsilon}, \hat{n})\}$ and $\{\hat{L}(\hat{l} + \hat{\varepsilon}, \hat{n} + 1)\}$ for $\hat{\varepsilon} \to 0$ ($\hat{\varepsilon}$ — infinitesimal positive real number). Lemma 3 reveals some of their characteristics. (The hats "^" are used to designate real quantities.) The employment of quantity $\hat{L}(3, \hat{n})$ is commented.

## 2. THEOREM FOR EXISTENCE AND FOR THE MAIN PROPERTIES OF THE $\hat{L}(\hat{c}, \hat{n})$ NUMBERS

**Theorem 1:** The statement of the theorem is expressed by the following three Lemmas:

**Lemma 1:** *If $\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ is the $\hat{n}$th positive real zero of the Kummer confluent hypergeometric function $\Phi(\hat{a}, \hat{c}; \hat{x})$ in $\hat{x}$ provided $\hat{a}$, $\hat{c}$, $\hat{x}$ are real, $\hat{a} < 0$, $\hat{c}$ — restricted positive number, $\hat{c} > 0$ ($\hat{n} = 1, 2, \ldots, \hat{p}$, $\hat{p} = abs[\hat{a}]$) or $\hat{a} < 0$, $\hat{c}$ — restricted negative real number, different from zero or a negative integer, $\hat{c} < 0$ ($\hat{c} \neq \hat{l}$, $\hat{l} = 0, -1, -2, \ldots$), $\hat{a} < \hat{c} < 0$, ($\hat{n} = 1, 2, \ldots, \hat{s}$, $\hat{s} = \hat{p} - \hat{q}$, $\hat{p} = abs[\hat{a}]$, $\hat{q} = abs[\hat{c}]$, $\hat{q} = 1, 2, \ldots, \hat{p} - 1$), $\hat{k}_- = \hat{a} - \hat{c}/2$ — real, negative, $\hat{a} = \hat{c}/2 + \hat{k}_-$, $\hat{x} > 0$, ($[\hat{a}]$ denotes the largest integer less or equal to $\hat{a}$), $\hat{K}_-(\hat{c}, \hat{n}, \hat{k}_-) = |\hat{k}_-|\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$, $\hat{M}_-(\hat{c}, \hat{n}, \hat{k}_-) = |\hat{a}_-|\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ ($\hat{a}_- \equiv \hat{a}$),* *then the infinite sequences of positive real numbers* $\{\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}\}$, $\{\hat{K}_-(\hat{c},\hat{n},\hat{k}_-)\}$ *and* $\{\hat{M}_-(\hat{c},\hat{n},\hat{k}_-)\}$ *are convergent for* $\hat{k}_- \to -\infty$, $(\hat{c},\hat{n}$ *— fixed). The limit of the first sequence is zero and the limit of the second and third ones is the same. It equals the finite positive real number* $\hat{L}$ *where* $\hat{L} = \hat{L}(\hat{c},\hat{n})$. *It is valid:*

$$\lim_{\hat{k}_-\to-\infty} \hat{K}_-\left(\hat{c},\hat{n},\hat{k}_-\right) = \lim_{\hat{k}_-\to-\infty} \hat{M}_-\left(\hat{c},\hat{n},\hat{k}_-\right) = \hat{L}\left(\hat{c},\hat{n}\right). \tag{1}$$

**Lemma 2:** *If* $\hat{L}(\hat{l}-\hat{\varepsilon},\hat{n})$ *and* $\hat{L}(\hat{l}+\hat{\varepsilon},\hat{n}+1)$ *are finite positive real numbers in the sense of Lemma 1 in which* $\hat{l} = 0,-1,-2,\ldots$ *is zero or a negative integer,* $\hat{\varepsilon}$ *is a positive real number, less than unity,* $(0 < \hat{\varepsilon} < 1)$ *and* $\hat{n}$ *is a restricted positive integer, taking the same values as in Lemma 1, then the infinite sequences of positive real numbers* $\{\hat{L}(\hat{l}-\hat{\varepsilon},\hat{n})\}$ *and* $\{\hat{L}(\hat{l}+\hat{\varepsilon},\hat{n}+1)\}$ *are convergent for* $\hat{\varepsilon} \to 0$ *and possess a common limit. The sequence* $\{\hat{L}(\hat{l}-\hat{\varepsilon},\hat{n})\}$ $(\{\hat{L}(\hat{l}+\hat{\varepsilon},\hat{n}+1)\})$ *tends to it from the left (right). The limit mentioned is accepted as a value of the* $\hat{L}(\hat{c},\hat{n})$ *number*

Table 1: Zeros $\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ and numbers $\hat{K}_-(\hat{c},\hat{n},\hat{k}_-)$ and $\hat{M}_-(\hat{c},\hat{n},\hat{k}_-)$ for $\hat{c} = 1$ and 3, assuming $\hat{n} = 1$ and 2 in case of large negative $\hat{k}_-$.

| $\hat{c}$ / $\hat{k}_-$ | $\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ | $\hat{K}_-(\hat{c},\hat{n},\hat{k}_-)$ | $\hat{M}_-(\hat{c},\hat{n},\hat{k}_-)$ | $\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ | $\hat{K}_-(\hat{c},\hat{n},\hat{k}_-)$ | $\hat{M}_-(\hat{c},\hat{n},\hat{k}_-)$ |
|---|---|---|---|---|---|---|
| | 1 | | | 3 | | |
| | | | $\hat{n} = 1$ | | | |
| $-1.10^4$ | (–4) 1.44579 64919 | **1.44579 649**19 | **1.4457**2 42021 | (–4) 6.59365 41513 | **6.59365 415**13 | **6.592**66 51031 |
| $-2.10^4$ | (–5) 7.22898 24551 | **1.44579 649**10 | **1.4457**6 03461 | (–4) 3.29682 70590 | **6.59365 41**179 | **6.593**15 95939 |
| $-4.10^4$ | (–5) 3.61449 12270 | **1.44579 6490**8 | **1.4457**7 84184 | (–4) 1.64841 35274 | **6.59365 410**96 | **6.593**40 68475 |
| $-6.10^4$ | (–5) 2.40966 08179 | **1.44579 6490**8 | **1.4457**8 44425 | (–4) 1.09894 23513 | **6.59365 410**80 | **6.593**48 92667 |
| $-8.10^4$ | (–5) 1.80724 56134 | **1.44579 6490**8 | **1.4457**8 74545 | (–5) 8.24206 76344 | **6.59365 410**75 | **6.593**53 04765 |
| $-1.10^5$ | (–5) 1.44579 64907 | **1.44579 64907** | **1.4457**8 92618 | (–5) 6.59365 41072 | **6.59365 410**72 | **6.593**55 52024 |
| | | | $\hat{n} = 2$ | | | |
| $-1.10^4$ | (–4) 7.61781 56311 | **7.61781 5**6311 | **7.617**43 47403 | (–3) 1.77125 00013 | **17.712**50 00134 | **17.7**0984 31384 |
| $-2.10^4$ | (–4) 3.80890 77986 | **7.61781 55**972 | **7.617**62 51518 | (–4) 8.85624 99003 | **17.71249 9**8007 | **17.711**17 13632 |
| $-4.10^4$ | (–4) 1.90445 38972 | **7.61781 558**87 | **7.617**72 03660 | (–4) 4.42812 49369 | **17.71249 97**475 | **17.711**83 55288 |
| $-6.10^4$ | (–4) 1.26963 59312 | **7.61781 558**72 | **7.617**75 21054 | (–4) 2.95208 32896 | **17.71249 97**377 | **17.712**05 69252 |
| $-8.10^4$ | (–5) 9.52226 94833 | **7.61781 558**66 | **7.617**76 79753 | (–4) 2.21406 24668 | **17.71249 97**342 | **17.712**16 76248 |
| $-1.10^5$ | (–5) 7.61781 55864 | **7.61781 558**64 | **7.617**77 74973 | (–4) 1.77124 99733 | **17.71249 97**326 | **17.712**23 40451 |

Table 2: Zeros $\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ and numbers $\hat{K}_-(\hat{c},\hat{n},\hat{k}_-)$ and $\hat{M}_-(\hat{c},\hat{n},\hat{k}_-)$ for $\hat{c} = -1.03527\ 61948$ and $0.21846\ 97035$, assuming $\hat{n} = 1$ and 2 in case of large negative $\hat{k}_-$.

| $\hat{c}$ / $\hat{k}_-$ | $\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ | $\hat{K}_-(\hat{c},\hat{n},\hat{k}_-)$ | $\hat{M}_-(\hat{c},\hat{n},\hat{k}_-)$ | $\hat{\zeta}^{(\hat{c})}_{\hat{k}_-,\hat{n}}$ | $\hat{K}_-(\hat{c},\hat{n},\hat{k}_-)$ | $\hat{M}_-(\hat{c},\hat{n},\hat{k}_-)$ |
|---|---|---|---|---|---|---|
| | –1.03527 61948 | | | 0.21846 97035 | | |
| | | | $\hat{n} = 1$ | | | |
| $-1.10^4$ | (–4) 6.40310 64354 | **6.40310 6**4354 | **6.403**43 78846 | (–5) 2.41561 51178 | **0.24156 151**18 | **0.2415**5 88731 |
| $-2.10^4$ | (–4) 3.20155 32018 | **6.40310 6**4035 | **6.403**27 21281 | (–5) 1.20780 75589 | **0.24156 151**18 | **0.24156** 01924 |
| $-4.10^4$ | (–4) 1.60077 65989 | **6.40310 639**56 | **6.403**18 92578 | (–6) 6.03903 77944 | **0.24156 151**18 | **0.24156** 08521 |
| $-6.10^4$ | (–4) 1.06718 43990 | **6.40310 639**41 | **6.403**16 16356 | (–6) 4.02602 51963 | **0.24156 151**18 | **0.24156** 10720 |
| $-8.10^4$ | (–5) 8.00388 29919 | **6.40310 639**36 | **6.403**14 78247 | (–6) 3.01951 88972 | **0.24156 151**18 | **0.24156** 11819 |
| $-1.10^5$ | (–5) 6.40310 63933 | **6.40310 639**33 | **6.403**13 95382 | (–6) 2.41561 51177 | **0.24156 151**18 | **0.24156** 12479 |
| | | | $\hat{n} = 2$ | | | |
| $-1.10^4$ | (–3) 1.74371 44423 | **17.43714 4**4227 | **17.43**804 70358 | (–4) 4.47004 35373 | **4.47004 35**373 | **4.4**6999 47089 |
| $-2.10^4$ | (–4) 8.71857 21078 | **17.43714 4**2156 | **17.43**759 55221 | (–4) 2.23502 17627 | **4.47004 352**54 | **4.4700**1 91112 |
| $-4.10^4$ | (–4) 4.35928 60409 | **17.43714 41**638 | **17.43**736 98170 | (–4) 1.11751 08806 | **4.47004 352**24 | **4.4700**3 13153 |
| $-6.10^4$ | (–4) 2.90619 06924 | **17.43714 41**542 | **17.43**729 45897 | (–5) 7.45007 25364 | **4.47004 3521**9 | **4.4700**3 53838 |
| $-8.10^4$ | (–4) 2.17964 30189 | **17.43714 41**508 | **17.43**725 69775 | (–5) 5.58755 44021 | **4.47004 3521**7 | **4.4700**3 74181 |
| $-1.10^5$ | (–4) 1.74371 44149 | **17.43714 41**493 | **17.43**723 44106 | (–5) 4.47004 35216 | **4.47004 3521**6 | **4.4700**3 86387 |

*in case $\hat{c} = \hat{l}$ ($\hat{c}$ — zero or a negative integer). Thus, it is assumed that:*

$$\hat{L}\left(\hat{c}, \hat{n}\right) = \lim_{\hat{\varepsilon}\to 0} \hat{L}\left(\hat{l} - \hat{\varepsilon}, \hat{n}\right) = \lim_{\hat{\varepsilon}\to 0} \hat{L}\left(\hat{l} + \hat{\varepsilon}, \hat{n} + 1\right). \tag{2}$$

*Moreover, the sequence $\{\hat{L}(\hat{l} + \hat{\varepsilon}, \hat{n})\}$ in which $\hat{l} = 0, -1, -2, -3, \ldots$ is zero or a negative integer, $\hat{\varepsilon}$ is a positive real number, less than unity, ($0 < \hat{\varepsilon} < 1$) and $\hat{n} = 1$ is convergent, as well when $\hat{\varepsilon} \to 0$ with the zero as its left limiting point. Correspondingly, it is written:*

$$\lim_{\hat{\varepsilon}\to 0} \hat{L}\left(\hat{l} + \hat{\varepsilon}, 1\right) = 0. \tag{3}$$

**Lemma 3:** *Under the conditions of Lemmas 1 and 2, it is true: i) In case $\hat{c} = \hat{l}$, it holds: $\hat{L}(\hat{l}, \hat{n}) = \hat{L}(2 - \hat{l}, \hat{n})$; ii) If $\hat{c} = 1 \pm \hat{l}$, ($\hat{c} = 0, \pm 1, \pm 2, \pm 3, \ldots$), it is fulfilled: $\hat{L}(1 + \hat{l}, \hat{n}) = \hat{L}(1 - \hat{l}, \hat{n})$; Besides, it is valid:*

$$\begin{pmatrix} \hat{L}\left(0.5, \hat{n}\right) \\ \hat{L}\left(1.5, \hat{n}\right) \end{pmatrix} = \begin{pmatrix} \left(2\hat{n} - 1\right)^2 \\ \left(2\hat{n}\right)^2 \end{pmatrix} \left(\frac{\pi}{4}\right)^2 \tag{4}$$

**Numerical proof**: The proof of Lemma 1 is illustrated in Tables 1 and 2 for $\hat{c}$ — positive integers and arbitrary real numbers (positive and negative, save for $\hat{c} = \hat{l}$), and that of Lemmas

Table 3: Numbers $\hat{L}(\hat{c}, \hat{n})$ with $\hat{c} = \hat{l} - \hat{\varepsilon}$ and $\hat{L}(\hat{c}, \hat{n} + 1)$ with $\hat{c} = \hat{l} + \hat{\varepsilon}$ in case $\hat{l} = 0, -1, \ldots, -5$, $\hat{n} = 1$ and 2 for $\hat{\varepsilon} = 10^{-i}$, $i = 1, 2, \ldots, 5$.

| $\hat{\varepsilon}$ \ $\hat{l}$ | 0 | –1 | –2 | –3 | –4 | –5 |
|---|---|---|---|---|---|---|
| $\hat{n} = 1$ | | | | | | |
| $\hat{L}(\hat{l} - \hat{\varepsilon}, \hat{n})$ | | | | | | |
| $1.10^{-1}$ | **3.**31569 73770 | **6.**05485 68064 | 9.44333 92952 | **1**3.45838 85151 | **1**8.08452 98622 | **2**3.31039 98436 |
| $1.10^{-2}$ | **3.6**3470 41442 | **6.5**3958 16381 | **10.**10326 29760 | **14.3**0218 06384 | **19.**12012 98110 | **24.**54513 46009 |
| $1.10^{-3}$ | **3.66**691 06818 | **6.5**8824 48059 | **10.16**928 06088 | **14.3**8638 11950 | **19.2**2327 55144 | **24.6**6793 09171 |
| $1.10^{-4}$ | **3.670**13 44314 | **6.593**11 31561 | **10.17**588 28648 | **14.39**479 98367 | **19.23**358 64627 | **24.68**020 44600 |
| $1.10^{-5}$ | **3.6704**5 68374 | **6.5936**0 00120 | **10.176**54 30967 | **14.395**64 16897 | **19.234**61 75267 | **24.681**43 17627 |
| $\hat{L}(2 - \hat{l}, \hat{n})$ | | | | | | |
| | **3.67049 26605** | **6.59365 41068** | **10.17661 64546** | **14.39573 52258** | **19.23473 20834** | **24.68156 81193** |
| $\hat{L}(l + \hat{\varepsilon}, \hat{n} + 1)$ | | | | | | |
| $1.10^{-5}$ | **3.6705**2 84839 | **6.593**70 82026 | **10.1766**8 98148 | **14.395**82 87672 | **19.234**84 66501 | **24.681**70 44932 |
| $1.10^{-4}$ | **3.670**85 08967 | **6.59**419 50621 | **10.17**735 00455 | **14.39**667 06119 | **19.23**587 76963 | **24.68**293 17652 |
| $1.10^{-3}$ | **3.67**407 53350 | **6.59**906 38670 | **10.1**8395 24164 | **14.**40508 89478 | **19.2**4618 78515 | **24.6**9520 39706 |
| $1.10^{-2}$ | **3.**70635 07496 | **6.**64777 25059 | **10.**24998 15422 | **14.**48925 89278 | **19.**34925 42483 | **24.**81786 65343 |
| $1.10^{-1}$ | 4.03223 78730 | 7.13702 39958 | **10.**91101 76357 | **1**5.32993 74668 | 20.37684 63973 | **2**6.03912 57896 |
| $\hat{n} = 2$ | | | | | | |
| $\hat{L}(\hat{l} - \hat{\varepsilon}, \hat{n})$ | | | | | | |
| $1.10^{-1}$ | **1**1.70308 17308 | **1**6.93402 62374 | **2**2.85963 85091 | 29.46141 82532 | **3**6.72453 54054 | **4**4.63690 08475 |
| $1.10^{-2}$ | **12.**24413 74109 | **17.**63436 01501 | **23.**72317 15028 | **30.**49220 78437 | **3**7.92659 32481 | **46.**01412 90483 |
| $1.10^{-3}$ | **12.**29856 31700 | **17.7**0468 28252 | **23.8**0976 84684 | **30.**59547 29546 | **38.0**4691 75870 | **46.1**5189 40180 |
| $1.10^{-4}$ | **12.304**00 89570 | **17.71**171 80099 | **23.81**843 06442 | **30.60**580 13907 | **38.0**5895 12928 | **46.16**567 10478 |
| $1.10^{-5}$ | **12.3045**5 35690 | **17.7124**2 15575 | **23.819**29 68921 | **30.606**83 42536 | **38.060**15 46919 | **46.1670**4 87563 |
| $\hat{L}(2 - \hat{l}, \hat{n})$ | | | | | | |
| | **12.30461 40804** | **17.71249 9729**8 | **23.81939 31360** | **30.60694 90162** | **38.06028 83854** | **46.16720 1834**8 |
| $\hat{L}(l + \hat{\varepsilon}, \hat{n} + 1)$ | | | | | | |
| $1.10^{-5}$ | **12.3046**7 45944 | **17.712**57 79022 | **23.819**48 93910 | **30.60**706 37790 | **38.060**42 21108 | **46.167**35 49136 |
| $1.10^{-4}$ | **12.30**521 92111 | **17.71**328 14563 | **23.8**2035 56335 | **30.60**809 66462 | **38.06**162 54813 | **46.16**873 26232 |
| $1.10^{-3}$ | **12.3**1066 57118 | **17.7**2031 72898 | **23.8**2901 83612 | **30.6**1842 55115 | **38.0**7365 94717 | **46.1**8250 97741 |
| $1.10^{-2}$ | **12.3**6516 28481 | **17.**79070 48475 | **23.9**1567 05215 | **30.**72173 35481 | **38.**19401 22805 | **46.**32028 68487 |
| $1.10^{-1}$ | **12.**91335 50520 | **1**8.49752 46490 | **2**4.78471 92258 | **3**1.75681 03910 | **3**9.39891 01103 | **4**7.69871 66990 |

Table 4: Numbers $\hat{L}(\hat{c}, \hat{n})$ with $\hat{c} = \hat{l} + \hat{\varepsilon}$ in case $\hat{l} = 0, -1, \ldots, -5$, $\hat{n} = 1$ for $\hat{\varepsilon} = 10^{-i}$, $i = 1, 2, \ldots, 5$.

| $\hat{\varepsilon}$ \ $\hat{l}$ | 0 | –1 | –2 | –3 | –4 | –5 |
|---|---|---|---|---|---|---|
| | $\hat{L}(\hat{l}+\hat{\varepsilon},1)$ | | | | | |
| $1.10^{-5}$ | **0.0000**1 00000 | **0.00**448 54849 | **0.0**4994 90596 | **0.**20013 48158 | **0.**51378 35703 | 1.03676 73410 |
| $1.10^{-4}$ | **0.000**10 00050 | **0.0**1427 59657 | **0.**10925 61462 | **0.**36404 21445 | **0.**83761 28606 | 1.57171 80732 |
| $1.10^{-3}$ | **0.00**100 04999 | **0.0**4607 07880 | **0.**24374 70870 | **0.**67670 89436 | 1.39589 82715 | 2.43470 57174 |
| $1.10^{-2}$ | **0.0**1004 99171 | **0.**15530 16048 | **0.**57183 93596 | 1.32229 98657 | 2.44008 91907 | 3.94611 12078 |
| $1.10^{-1}$ | **0.**10492 12125 | **0.**60135 94011 | 1.54771 56128 | 2.95799 04285 | 4.83919 24333 | 7.19567 87737 |

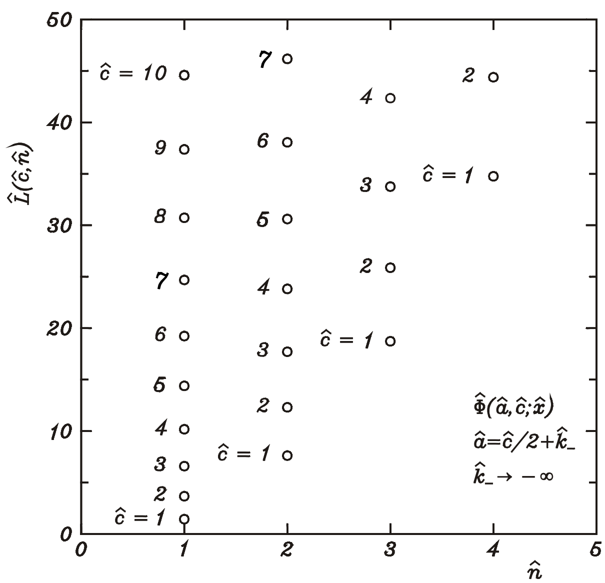



Figure 1: Dependence of $\hat{L}(\hat{c}, \hat{n})$ numbers on $\hat{n}$ with $\hat{c}$ as parameter.

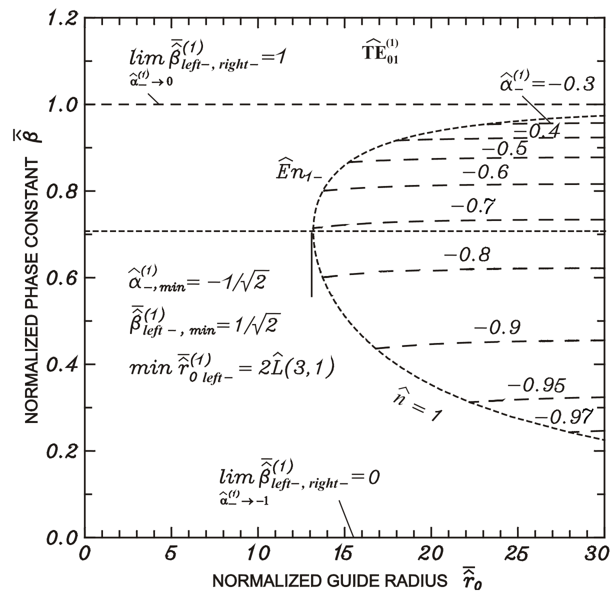



Figure 2: $\bar{\bar{\beta}}^{(1)}(\bar{\bar{r}}_0^{(1)})$ — curves of the slow $T\hat{E}_{01}^{(1)}$ mode in the ferrite guide for $-1 < \hat{\alpha}^{(1)} < 0$.

2 and 3 $i$) — in Tables 3 and 4 (cf. the digits marked by bold face type). The case $\hat{n} = 1$ and 2 is considered, except in Table 4, corresponding to $\hat{n} = 1$ only. Statement $ii$) of Lemma 3 is an obvious corollary of point $i$) of the same. The authentity of Eq. (4) is checked directly, (e.g., $\hat{L}(0.5, 1) = 0.61685\,02751$ and $\hat{L}(1.5, 1) = 2.46740\,11003$). The effect of $\hat{c}$ and $\hat{n}$ on $\hat{L}(\hat{c}, \hat{n})$ is presented in Fig. 1 and Table 5.

## 3. APPLICATION

The propagation of slow $\hat{\mathrm{TE}}_{0\hat{n}}$ modes of phase constant $\hat{\beta}$ is considered in a circular waveguide of radius $\hat{r}_0$, entirely filled with azimuthally magnetized ferrite, described by a permeability tensor of off-diagonal element $\hat{\alpha} = \gamma \hat{M}_r/\omega$, ($\gamma$ — gyromagnetic ratio, $\hat{M}_r$ — remanent magnetization, $\omega$ — angular frequency of the wave), and a scalar permittivity $\varepsilon = \varepsilon_0 \varepsilon_r$. The solution of Maxwell equations subject to the boundary condition at the wall $\bar{\bar{r}} = \bar{\bar{r}}_0$ reveals that they are governed by the characteristic equation:

$$\Phi\left(\hat{a}, \hat{c}; \hat{x}_0\right) = 0, \tag{5}$$

with $\hat{a} = \hat{c}/2 + \hat{k}$, $\hat{c} = 3$, $\hat{x}_0 = 2\bar{\bar{\beta}}_2 \bar{\bar{r}}_0$, $\hat{k} = \hat{\alpha}\bar{\bar{\beta}}/(2\bar{\bar{\beta}}_2)$, $\bar{\bar{\beta}}_2 = [\bar{\bar{\beta}}^2 - (1 - \hat{\alpha}^2)]^{1/2}$ — radial wavenumber. It is valid provided $\bar{\bar{\beta}}_2 = \hat{\zeta}^{(\hat{c})}_{\hat{k},\hat{n}}/(2\bar{\bar{r}}_0)$, giving the eigenvalue spectrum looked for. It may be shown that transmission of two slow waves $T\hat{E}_{0\hat{n}}^{(1)}$ and $T\hat{E}_{0\hat{n}}^{(2)}$ is possible in two areas for $|\hat{\alpha}_-^{(1)}| < 1$ and $|\hat{\alpha}_-^{(2)}| > 1$, resp. Fig. 2 presents with dashed lines the phase characteristics for $T\hat{E}_{01}^{(1)}$ mode, corresponding to the first case computed, following the procedure, developed earlier [2, 6]. The curves are restricted by an envelope line from the side of lower frequencies of equation: $\bar{\hat{r}}_{0en-} = \hat{L}(\hat{c},\hat{n})/\{|\hat{\alpha}^{(1)}_{en-}|[1-(\hat{\alpha}^{(1)}_{en-})^2]^{1/2}\}$, $\bar{\hat{\beta}}^{(1)}_{en-} = [1-(\hat{\alpha}^{(1)}_{en-})^2]^{1/2}$ ($\hat{\alpha}^{(1)}_{en-}$ is a parameter) where $\bar{\hat{r}}_0 = \beta_0\hat{r}_0\sqrt{\varepsilon_r}$, $\bar{\hat{\beta}} = \hat{\beta}/(\beta_0\sqrt{\varepsilon_r})$, $\bar{\hat{\beta}}_2 = \hat{\beta}_2/(\beta_0\sqrt{\varepsilon_r})$ and $\beta_0 = \omega\sqrt{\varepsilon_0\mu_0}$. (Throughout the paper the subscripts "$-$" and "$en-$" distinguish the quantities, relevant to negative ($\hat{\alpha}_- < 0$, $\hat{k}_- < 0$) magnetization, resp. to the envelope.)

Table 5: Values of $\hat{L}(\hat{c},\hat{n})$ as a function of $\hat{c}$ for $\hat{n} = 1$ and 2.

| $\hat{c}$ | $\hat{L}(\hat{c},\hat{n})$ | $\hat{c}$ | $\hat{L}(\hat{c},\hat{n})$ | $\hat{c}$ | $\hat{L}(\hat{c},\hat{n})$ | $\hat{c}$ | $\hat{L}(\hat{c},\hat{n})$ | $\hat{c}$ | $\hat{L}(\hat{c},\hat{n})$ | $\hat{c}$ | $\hat{L}(\hat{c},\hat{n})$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\hat{n} = 1$ | | | | | | | | | | | |
| –3.0 | | –2.0 | | –1.0 | | 0 | | 1.0 | 1.44579 64907 | 2.0 | 3.67049 26605 |
| –2.8 | 4.00209 21312 | –1.8 | 2.22586 15078 | –0.8 | 0.96338 27690 | 0.2 | 0.21940 16817 | 1.2 | 1.83206 07322 | 2.2 | 4.20080 97355 |
| –2.6 | 5.69683 88860 | –1.6 | 3.37709 13769 | –0.6 | 1.62888 38548 | 0.4 | 0.47564 39874 | 1.4 | 2.24827 40891 | 2.4 | 4.75859 34562 |
| –2.4 | 7.22882 15474 | –1.4 | 4.45444 70199 | –0.4 | 2.28943 38223 | 0.6 | 0.76647 86833 | 1.6 | 2.69377 63812 | 2.6 | 5.34351 81897 |
| –2.2 | 8.70942 88338 | –1.2 | 5.51970 18426 | –0.2 | 2.96757 32740 | 0.8 | 1.09027 62038 | 1.8 | 3.16800 91269 | 2.8 | 5.95529 25831 |
| –2.0 | | –1.0 | | 0.0 | | 1.0 | 1.44579 64907 | 2.0 | 3.67049 26605 | 3.0 | 6.59365 41068 |
| $\hat{n} = 2$ | | | | | | | | | | | |
| –3.0 | | –2.0 | | –1.0 | | 0 | | 1.0 | 7.61781 55859 | 2.0 | 12.30461 40804 |
| –2.8 | 16.26302 22910 | –1.8 | 11.64784 18558 | –0.8 | 7.68564 25102 | 0.2 | 4.40112 03549 | 1.2 | 8.49636 43640 | 2.2 | 13.32925 60604 |
| –2.6 | 18.13168 59216 | –1.6 | 13.13246 77929 | –0.6 | 8.80050 96937 | 0.4 | 5.16075 88218 | 1.4 | 9.40449 70823 | 2.4 | 14.38254 76497 |
| –2.4 | 20.01051 78375 | –1.4 | 14.63601 10992 | –0.4 | 9.94098 66095 | 0.6 | 5.94997 97114 | 1.6 | 10.34204 18046 | 2.6 | 15.46430 37705 |
| –2.2 | 21.90515 62688 | –1.2 | 16.16193 26198 | –0.2 | 11.10869 26401 | 0.8 | 6.76899 23948 | 1.8 | 11.30881 16816 | 2.8 | 16.57434 53219 |
| –2.0 | | –1.0 | | 0.0 | | 1.0 | 7.61781 55859 | 2.0 | 12.30461 40804 | 3.0 | 17.71249 97298 |

## 4. CONCLUSION

The positive real numbers $\hat{L}(\hat{c},\hat{n})$ are defined for $\hat{c}$ — arbitrary real and $\hat{n}$ — restricted natural number as limits of certain sequences of real ones. On condition that $\hat{c} \neq \hat{l}$, $\hat{l} = 0, -1, -2, -3, \ldots$, the terms of the latter are composed through definite zeros of a real Kummer function of suitably selected parameters, whereas if $\hat{c} = \hat{l}$ (for which the function mentioned does not exist), as such serve the quantities $\hat{L}(\hat{l} - \hat{\varepsilon}, \hat{n})$ and $\hat{L}(\hat{l} + \hat{\varepsilon}, \hat{n} + 1)$ with $\varepsilon \to 0$. Besides, a symmetry of the numbers toward the point $c = 1$ is found. The use of quantity $\hat{L}(3,\hat{n})$ in the theory of waveguides is shown.

## ACKNOWLEDGMENT

We express our gratitude to our mother Trifonka Romanova Popnikolova for her self-denial, patience and for the tremendous efforts she exerts to support all our undertakings.

### REFERENCES

1. Tricomi, F. G., *Funzioni Ipergeometriche Confluenti*, Edizioni Cremonese, Rome, Italy, 1954.
2. Georgiev, G. N. and M. N. Georgieva-Grosse, "A new property of the complex Kummer function and its application to waveguide propagation," *IEEE Antennas Wireless Propagat. Lett.*, Vol. 2, 306–309, December 2003.
3. Georgiev, G. N. and M. N. Georgieva-Grosse, "The $L(c, n)$ numbers and their application in the theory of waveguides," *Proc. Int. Conf. Days Diffr., DD'08*, 44–57, St. Petersburg, Russia, June 3–6, 2008.
4. Georgiev, G. N. and M. N. Georgieva-Grosse, "A property of the $L(c, \rho, n)$ numbers and its application to waveguide propagation," *Proc. XXIX URSI General Assembly*, BK.6(120), in CDROM, Chicago, IL, USA, August 7–16, 2008.
5. Georgiev, G. N. and M. N. Georgieva-Grosse, "Theorem for the $L(c, \rho, n)$ numbers," *PIERS Proceedings*, 1478–1482, Moscow, Russia, August 18–21, 2009.
6. Georgiev, G. N. and M. N. Georgieva-Grosse, "Effect of the dielectric filling on the phase behaviour of the circular waveguide with azimuthally magnetized ferrite toroid and dielectric cylinder," *Proc. Asia-Pacific Microwave Conf.*, WE4B-4(1680), in CDROM, Singapore, December 7–10, 2009.

# About the Specific Heat of Black Holes

**Antonio Puccini**
Department of Neurophysiology, Order of Malta, Naples, Italy

**Abstract**— If we give a certain thermic energy to a black hole (BH) it can happen that its temperature decreases and its specific heat (SH) has a negative number. At the same time its mass will increase, according to Hawking's equation: $T_{\mathrm{BH}} = 8\pi/m$.

The mathematical formalism is easy and elegant, however, we do not know the real physical checking which may explain this peculiar phenomenon. With this paper we try to describe how heat may give a mass to the BH.

## 1. INTRODUCTION

The specific heat (SH) of a body is measured by the increase of its temperature when a thermic energy is applied. In the case of an ordinary body we have a positive number, because the temperature usually increases when a heat is applied. This is related to the second law of thermodynamics according to which the heat passes from a warmer body to a colder one, until they have the same temperature. However, if we analyse the entropy and the SH of a Black Hole (BH), as well as its variations, the topic becomes more complex. As Hawking tells us "the second law of thermodynamics has got its own status quite different from the other scientific laws, since it doesn't seem to be always applicable, but only in most of the cases. Apparently in the case of a BH there is quite a simple way to violate the second law, such as throwing in BH some matter with a certain entropy, for example a container full of gas. In this way, the matter's total entropy would decrease and it is not possible to see, inside the BH, how much entropy the matter has. Hence the area of its event horizon (EH) increases every time matter falls in the BH" [1] (in Space-Time the EH is a 3-surface **H** where, of course, **H** stands from *Hamiltonian*).

It was Bekenstein to suggest that the area of the EH was a measure of the BH's entropy. Thus every time the matter (carrying some entropy) fell in a BH, the area of the EH would increase, so that the total entropy (that is the entropy inside and outside the BH) would not decrease. In this way the second law of thermodynamics was not violated. However, a BH having entropy implies a thermic radiation, an inside temperature, so it should behave as a *black body*. In fact "a body with a particular temperature must emit radiation with a certain rhythm. This radiation is required to prevent the violation of the second law. Thus, BH should emit a radiation too and, according to Heisenberg's Uncertainty Principle, BH (rotating and non rotating) should create and emit particles.

"Calculations confirmed this emission, and showed that the spectrum of the emitted particles was exactly the one which would have been emitted by a warm body, and that BH emitted particles exactly at the right rhythm to prevent the violation of the second law. Calculus, repeated by several researchers, have always confirmed that BH should always emit particles and radiations exactly as if it was a very hot body, with a temperature depending only on its mass: the higher its mass the lower the temperature" [1], and *vice versa*. In fact "a BH with a mass some times bigger than the sun's, would have a temperature one millionth degree bigger than Absolute Zero, that is much lower than the radiation of microwaves which is in the universe (about 2,7 degree higher than Absolute Zero), so that BHs emit even less than what they absorb" [1].

On the contrary "the lower the BH mass, the higher its temperature. Thus, as the BH loses its mass, its temperature, and its emission rhythms, will increase, as a consequence it loses mass even more quickly and it is doomed, probably, to disappear completely in a tremendous final explosion, as big as the explosion of millions of H bombs. It is likely there exist primordial BHs with a very small mass, produced by the collapse of *irregularities* emerged in the very first phases of the universe.

Apparently these BHs have a very high temperature and they emit radiations with a much higher rhythm. A primordial BH with a mass of a billiard of tons would live more or less as long as the present age of the universe. Primordial BHs with a smaller initial mass would be already completely *evaporated*, whereas those with a slightly bigger mass would be still emitting radiations as X and gamma rays: in fact we observe a bottom of gamma rays" [1], as well as a bottom X radiation, as detected by Giacconi and Tucker [2].

## 2. DISCUSSION

We need to consider that in the ordinary General Theory of the Relativity, where Space-Time is four dimensional, a *stationary*, isolated BH is described in *Kerr' metric* [3], characterized by the values of only two real parameters (not negative): $m$ and $a$, where $m$ is the total mass of BH and $a \cdot m$ its total *angular momentum*. "The unrestrainable nature of BHs, since they collect any sort of material, makes so BHs have very big entropies. BH's entropy has a geometrical interpretation: it is proportional to the EH' area of that Hole!" [4]. According to the well known formula of Beckenstein-Hawking, it is possible to attribute a well defined entropy:

$$S_{\mathrm{BH}} = \frac{Kc^3 A}{4G\hbar} \tag{1}$$

where $A$ is EH'superficial area of the BH, $K$ is Boltzmann' constant, $c$ is the speed of light, $G$ is the gravitational constant, $\hbar$ is Planck's constant written in Dirac's way, $S$ is the entropy. As Penrose suggests "it is often extremely convenient to adopt for all these *constants* the unitary value, i.e., 1:

$$G = c = \hbar = K = 1 \tag{2}$$

In this way also the temperature unit becomes an absolute thing" [4]. Hence, the Eq. (1) can be reformulated in following way:

$$S_{\mathrm{BH}} = \frac{1}{4}A \tag{3}$$

that is the entropy of a BH, according to Beckenstein-Hawking's formula, it will just be one fourth of EH's area of the BH we took in consideration. In Kerr's solution we find:

$$A = \frac{8\pi G^2}{c^4} m \left(m + \sqrt{m^2 - a^2}\right) \tag{4}$$

$$S_{\mathrm{BH}} = \frac{2\pi G K}{c\hbar} m \left(m + \sqrt{m^2 - a^2}\right) \tag{5}$$

expressed in general units" [4]. It is possible to infer that the higher the value of $A$, that is EH's surface area, the bigger BH's *evaporation* phenomenon. Hawking provides a relevant result "demonstrating that a BH must have also an its own *temperature*, which is proportional to the so called *superficial gravitation* of the Hole. Using Kerr's geometry, we get:

$$T_{\mathrm{BH}} = \frac{1}{4\pi m \left[1 + \left(1 - \frac{a^2}{m^2}\right)^{\frac{1}{2}}\right]} \tag{6}$$

where $T$ is the temperature of BH, $m$ its mass, $a \cdot m$ its angular momentum. This temperature can be obtained by the standard formula of Thermodynamics:

$$TdS = dE \tag{7}$$

where, varying the energy $E$, we keep the angular momentum constant. Hence, BH will emit photons, as it was a physical object in a thermal equilibrium, irradiating energy with the characteristic spectrum of (Planck's) black body, at a temperature $T_{\mathrm{BH}}$" [4]. We need to specify that in order that Kerr'Geometry describes really a BH, not a *naked singularity*, it is necessary that $m$ is bigger (or equal to) than $a$. "The limit case of an ordinary BH happens when $m = a$, which just justifies it as a *BH*. It is really a limit case (not reachable from an astrophysical point of view) of common BH, with a zero value for its *Hawking temperature*. It is also necessary to mention a peculiar thermodynamic property of normal BHs, which with small *angular momentum* have *negative* SH. If we give heat to a BH we notice that its temperature *decreases*" [4]. If we apply heat to an ordinary body, its temperature will increase and its SH will have a *positive* value. Whereas, if we apply a heath to a BH "the thermal energy provides mass to the BH (according to $E$=$mc^2$) and it becomes more massive. According to Hawking's relation:

$$T_{\mathrm{BH}} = \frac{8\pi}{m} \tag{8}$$

related to a Schwarzchild's BH, its temperature will decrease, so that its SH is really negative" [4].

## 3. CONCLUSION

We may wonder: how can the SH of a BH be negative? We may answer that the increase in its mass makes its SH value decrease! But then, which mechanism explains that? The application of heat to BH. Thus, it is the heat to give mass to the BH, to make it more massive. What is the heat made of? It is well known that heat is thermic energy, i.e. electromagnetic waves (EMWs), photons (Ps) after all. Thus, it is Ps which provide mass to the BH. However it has always been stated that P's mass is zero. This takes to an incongruence: mass less particles are able to increase the mass of a body, the BH in our case. This is a real and concrete fact, indeed its SH is modified. We get a confirmation of the phenomenon from Mathematics; in fact the increase of the mass (m), which in Hawking's relation is placed at the denominator — see Eq. (8) — will cause proportionally a decrease of BH's temperature ($T_{\mathrm{BH}}$) and of its SH, the latter will be negative.

Then, how does heat, that is thermic Ps — normally known as massless — to give mass to a BH? We need to keep in mind that any P, whatever its frequency, has a *momentum*, as stated in (9):

$$p = \frac{h}{\lambda} \tag{9}$$

where $h$ is Planck's constant, corresponding to $6,625 \cdot 10^{-27}\,[\mathrm{erg} \cdot \mathrm{s}]$, and $\lambda$ represents the wave length of the considered $P$. Since we are considering thermic Ps, that is infrared rays, their $\lambda$ can correspond to $5 \cdot 10^{-3}\,[\mathrm{cm}]$. If we develop the (9) we have:

$$p = \frac{6,625 \cdot 10^{-27}\,[\mathrm{erg} \cdot \mathrm{s}]}{5 \cdot 10^{-3}\,[\mathrm{cm}]} \tag{10}$$

Since erg can be expressed in g$\cdot$ cm/s$^2\cdot$ cm, that is erg = g$\cdot$ cm$^2$/s$^2$, we can write:

$$p = \frac{6,625 \cdot 10^{-27}\,[\mathrm{g} \cdot \mathrm{cm}^2/\mathrm{s}]}{5 \cdot 10^{-3}\,[\mathrm{cm}]} \tag{11}$$

we have:

$$p = 1,325 \cdot 10^{-24}\,[\mathrm{g} \cdot \mathrm{cm/s}] \tag{12}$$

This is the value of the *momentum* of a P in the infrared band. It shows that a common infrared particle is not so ethereal [5]. We can infer from (12) that the P with its *momentum* carries also a mass, or rather a *dynamic-mass*, since we are talking about a particle which is continuously in motion. We also know that as soon as the P interacts with another particle the P transfer to the latter its *momentum* [6] thus its *dynamic-mass.* If we consider that the most frequent hadronic particle, that is the proton, weighs $1{,}6272 \cdot 10^{-24}$ [g], we have that a thermic particle, that is an infrared P, carries a *dynamic-mass* more or less equal to the mass of a proton. Moreover, if we consider that a single EMW of the infrared band carries about $10^{13}$ Ps per second, then also the infinitesimal quantity of energy (and *hidden dynamic-mass)* will have a certain value, being able to make more massive our BH and make its SH become negative.

With this work we hope to give a small contribution to understand, in its intimate mechanism, a singular astrophysical phenomenon: the negativity of BHs' SH.

### REFERENCES


1. Hawking, S., *A Brief History of Time*, Rizzoli Ed. (Milano), 124–126, 129–130, 1993.
2. Giacconi, R. and W. G. Tucker, *The X Ray Universe. The Quest for Cosmic Fire from Black Holes to Intergalactic Space*, A. Mondadori Ed. (Milano), 56, 2003.
3. Kerr, R. P., "Gravitational field of a spinning mass as an example of algebraically special metrics," *Physical Review Letters*, Vol. 11, No. 5, 237–238, 1963.
4. Penrose, R., *The Road to Reality*, RCS Ed. (Milano), 714–716, 824, 917–920, 2005.
5. Puccini, A., "About the zero mass of the photon," *Progress In Electromagnetics Research*, PIER 55, 117–146, 2005.
6. Feynman, R. P., *The Feynman Lectures on Physics*, Addison Wesley Longman, 1989; Zanichelli Ed., Bologna, I, 34–15, 16; I, 39–8, 2001.

# Application of Analytical Method to Weak Global Positioning System Signal

**H. Babazadeh, S. Askari, A. Safaian, and M. Razfar**
Department of Electrical Engineering, California State University Los Angeles
Los Angeles, CA 90032, USA

**Abstract**— The aims of the study were to analytically detect and treat the nonlinear behavior of weak indoor Global Positioning System (GPS) signal. Since analyzing nonlinear problems are of great difficulty, scientists implement different numerical methods to treat such problems. Ignoring nonlinearity is inevitable in the actual world of communications; therefore, nonlinear analysis has been of great importance to the scientists in its field. The advantages of analytical methods, specially Homotopy Perturbation Method (HPM), are that these methods are capable of solving both regular and strong non-linear equations; plus, these methods are simple to apply and will not increasing complexity. The obtained results are compared simultaneously with numerical ones and as result shown in graphs and in tables; analytical solutions are in good agreement with those of the numerical method. As a strategy, chaotic oscillators, which are sensitive to periodic signal and inert to noise, possess huge advantages in weak signal acquisition. In this paper chaotic oscillator is employed in weak GPS signal acquisition. In the final section, the results achieved from computer simulation indicate that chaotic oscillator algorithm can acquire GPS signal at $-48\,\text{dB}/2\,\text{MHz}$ SNR [1].

## 1. INTRODUCTION

Up to this date, different numerical methods have been implemented to solve the problem of a nonlinear oscillators. This paper represents the nonlinear behavior of a Chaotic oscillator in weak GPS signal [1]. It has been attempted to propose an analytical method as a solution for this problem. The analytical method is simple and is also a strong method for engineers to use in their designs.

In the dynamic model of this problem, the chaotic oscillator, weak GPS signals has been modeled in the next part. The analytical methods, including Homotopy Perturbation Method (HPM), are used for solving many engineering problems by a variety of scientists in different fields [2–8]. This method is capable of solving highly nonlinear problems, while the constant coefficients of the equations in the problems are parametrically inserted into the equation. Therefore, the obtained results can be graphically shown and analyzed for different cases, and by inserting different values for these parameters regarding each single case of study.

Fading, refraction, reflection and multipath interference are the major causes for weakening GPS indoor signals [1]. Normally line-of-sight GPS signal is at 44 dBHz [9] while signal strength will degrade larger than 25 dB in bad case [10]. Common commercial GPS receivers, however, can only acquire GPS signal above 38 dBHz [1]. To achieve successful position and navigation in today's electromagnetic signal filled environment, there has been so many efforts towards the prolonging integration duration to achieve increase in the SNR [1]. Chaotic oscillator [11] is sensitive to periodic signal and inert to noise, which can be utilized to achieve successful acquisition in weak signal. The detection of weakened linear frequency modulation (LFM) signal after decrypting was first used by Bo [12]. Bo's work showed that chaotic oscillator detect extremely weak LFM signal, even at −27 dBHz. In the end, a comparative study is conducted to verify the accuracy of the analytical method as compared with the numerical solution. The study solves the GPS equation and spring equation which are basically the same equation and the results are shown in graphs and tables.

## 2. DYNAMIC AND MATHEMATICAL MODEL OF THE PROBLEM [13, 14]

This section, introduces the dynamic and mathematical approaches. Among the chaotic oscillators, Duffing oscillator has been studied [15]. Duffing Oscillator was brought into non-linear dynamics in 1918 by Duffing [16]. Holmes [17] modified the primal equation and reached a novel one which depicts forced double-well model, which can be the general equation like an oscillator, compared with a nonlinear spring, a linear spring and a damper under a harmonic load. This comparison is as follows [18]:

$$m\ddot{x} + c\dot{x} + k_1 x + k_2 x^3 = F_0 \cos(\omega t) \tag{1}$$

Subject to that following initial conditions:

$$x(0) = A, \quad \dot{x}(0) = 0 \tag{2}$$

where $m$ is the mass, $c$ is a viscous damping coefficient, $k_1$ is a linear stiffness coefficient, and $k_2$ is a nonlinear stiffness coefficient. The harmonic excitation force is characterized by the force amplitude, $F_0$, with excitation frequency of $\omega$. $A$ is the initial amplitude of displacement.

### 2.1. GPS Signal Model

Mark [19] proposed a typical received GPS signal model as: $r_k = Ad(t_k)C[(1+\eta)(t_k - t_s)\cos[\omega_{IF}t_k - (w_D t_k + \varphi_0)] + \delta(t_k)$ where $r_k$ is output of RF front end at sample time $t_k$ [1], $f_D = \frac{\omega_D}{2\pi}$ is Doppler shift, $\eta = \frac{\omega D}{2\pi * 1575.42 * 10^6}$ accounts for Doppler shift caused chip length distortion, $f_{IF} = \frac{\omega_{IF}}{2\pi}$ is carrier frequency, $\varphi_0$ is initial carrier phase, $\delta(t_k)$ is Gaussian band-limited white noise, constant $A$ is signal amplitude, and $d(t_k)$ stands for GPS data stream.

Front-end output is at frequency $f_{IF} - f_D$ sampling frequency $f_S$ is 5 MHz. After sampling, output from low pass filter with 2 MHz bandwidth is at frequency $f_G = f_{IF} - |f_D| - nf_s$, corresponding to angular frequency $\omega_G$. By deducing the original Duffing oscillator equation in the GPS signal acquisition, the following equation can be achieved [1]: $\ddot{x}(t) + \omega_G k\dot{x}(t) - \omega_G^2[x(t) - x^3(t)] = \omega_G^2 f\cos(\omega_G t)$ this equation is the application of Duffing oscillator in GPS signal acquisition at $\omega_G$. If the angular frequency $\omega$ of GPS signal is the same as the inherent angular frequency $\omega_G$ of oscillator system (previous equation which exactly the Duffing equation of spring), motion state of trajectories on phase plane will change largely, which indicates that the received data contains expected GPS signal.

As in [20], $\omega$ can be found easily by having the parameters, $A$, $c$, $m$, $k_1$ and $k_2$:

$$\left(\left(k_1 - m\omega^2\right)A + \frac{3}{4}k_2A^3\right)^2 + (c\omega A)^2 = F_0^2 \tag{3}$$

The stiffness coefficients of nonlinear and linear springs behave as well as the GPS's equation, where $f(x)$ is the spring force and both GPS and $x$ is the displacement:

In the following section, the basic concepts of the analytical and numerical methods are discussed and later applied to the nonlinear equation above.

## 3. THE BASIC CONCEPT OF THE SOLUTIONS HPM

To illustrate the basic ideas of this method, we consider the following equation:

$$A(x) - f(r) = 0 \quad r \in \Omega \tag{4}$$

with the boundary condition of:

$$B\left(x, \frac{\partial x}{\partial t}\right) = 0 \quad r \in \Gamma \tag{5}$$

where $A$ is a general differential operator, $B$ a boundary operator, $f(r)$ a known analytical function and $\Gamma$ is the boundary of the domain $\Omega$. $A$ can be divided into two parts of $L$ and $N$, where $L$ is linear and $N$ is nonlinear. Eq. (4) can therefore be rewritten as follows:

$$L(x) + N(x) - f(r) = 0 \quad r \in \Omega \tag{6}$$

Homotopy perturbation structure is shown as follows:

$$H(v, p) = (1-p)[L(v) - L(x_0)] + p[A(v) - f(r)] = 0 \tag{7}$$

where,

$$v(r, p) : \Omega \times [0, 1] \to R \tag{8}$$

In Eq. (7), $p \in [0, 1]$ is an embedding parameter and $x_0$ is the first approximation that satisfies the boundary condition. We can assume that the solution of Eq. (4) can be written as a power series in $p$, as follows:

$$\nu = \nu_0 + p\nu_1 + p^2\nu_2 + \ldots = \sum_{i=0}^{n} \nu_i p^i \tag{9}$$

and the best approximation for the solution is:

$$x = \lim_{p \to 1} v = v_0 + v_1 + v_2 + \ldots \tag{10}$$

### 3.1. Runge-Kutta

For the numerical approach to verify the analytic solution, the fourth RK (Runge-Kutta) method has been used. This iterative algorithm is written in the form of the following formulae for the second-order differential equation:

$$\begin{aligned} \dot{x}_{i+1} &= \dot{x}_i + \frac{\Delta t}{6}(h_1 + 2h_2 + 2h_3 + k_4) \\ x_{i+1} &= x_i + \Delta t \left( \dot{x} + \frac{\Delta t}{6}(h_1 + h_2 + k_3) \right) \end{aligned} \tag{11}$$

where, $\Delta t$ is the increment of the time and $h_1$, $h_2$, $h_3$, and $h_4$ are determined from the following formulae:

$$\begin{aligned} h_1 &= f(\dot{x}, x_i, \dot{x}_i)\, k \\ h_2 &= f\left(t_i, \frac{\Delta t}{2}, x_i + \frac{\Delta t}{2}\dot{x}_i, \dot{x}_i + \frac{\Delta t}{2} h_1\right) \\ h_3 &= f\left(t_i, \frac{\Delta t}{2}, x_i + \frac{\Delta t}{2}\dot{x}_i, \frac{1}{4}\Delta t^2 h_1, \dot{x}_i + \frac{\Delta t}{2} h_2\right) \\ h_4 &= f\left(t_i, \Delta t, x_i + \Delta t \dot{x}_i, \frac{1}{2}\Delta t^2 h_2, \dot{x}_i + \Delta t h_3\right) \end{aligned} \tag{12}$$

The numerical solution starts from the boundary at the initial time, where the first value of the displacement function and its first-order derivative are determined from initial condition (see Section 2). Then, with a small time increment $\Delta t$, the displacement function and its first-order derivative at the new position can be obtained using Eq. (11). This process continues to the end of the time limit.

## 4. THE SOLUTIONS

In this section, the applications of the two methods to the nonlinear equation of oscillator are discussed.

### 4.1. HPM (Analytic)

As the HPM was applied to the nonlinear equation of (1), we have:

$$(1-p)(m\ddot{x}_0 + c\dot{x}_0 + k_1 x_0) + p\left(m\ddot{x}_1 + c\dot{x}_1 + k_1 x_1 + k_2 x_0^3 - F\cos(\omega t)\right) \tag{13}$$

After expanding the equation and collecting it based on the coefficients of $p$terms, we have:

$$\left[\begin{array}{l} p^0 = m\ddot{x}_0 + c\dot{x}_0 + k_1 x_0 \\ p^1 = m\ddot{x}_1 + c\dot{x}_1 + k_1 x_1 + k_2 x_0^3 - F\cos(\omega t) \\ p^2 = m\ddot{x}_2 + c\dot{x}_2 + k_1 x_2 + 3k_2 x_0^2 x_1 \\ p^3 = m\ddot{x}_3 + c\dot{x}_3 + k_1 x_3 + 3k_2 x_0^2 x_2 + 3k_2 x_0 x_1^2 \end{array}\right. \tag{14}$$

One can now try to obtain the solution of different iterations (14), in the form of:

$$\begin{aligned} x_0(t) \;=\; & \frac{1}{2}\frac{1}{c^2 - 4k_1 m}\left(c\sqrt{c^2 - 4k_1 m} + c^2 - 4k_1 m\right) e^{\frac{1}{2}\frac{(-c+\sqrt{c^2-4k_1 m})t}{m}} \\ & + \frac{1}{2}\frac{1}{c^2 - 4k_1 m}\left(c^2 - c\sqrt{c^2 - 4k_1 m} - 4k_1 m\right) e^{\frac{\left(-c-\sqrt{c^2-4k_1 m}\right)t}{2m}} \end{aligned} \tag{15}$$

The obtained iteration is used to generate the equation for the next iteration, and therefore the second and third iterations are obtained. Since the two other ones and therefore the general solution are too long to be written in this article, we have shown them in graphs (see Section 5). In Table 1, the numerical values for $x$ and $\dot{x}$ for different points of time and for $f = 0.5$, $A = 0.06$, $\omega = 4.163379415$.

### 4.2. Runge-Kutta (Numerical)

The Maple Package has been utilized for the numerical analysis of the problem, in which the rkf45 is used to solve ODEs. The solution for the displacement and the velocity for eleven different points of time are shown in Table 2.

## 5. RESULTS AND DISCUSSIONS

In this section, the results for displacement and the velocity for different times are shown in Tables 3 and 4, for different $f$'s and $A$'s, in order to evaluate the accuracy of the analytic solution.

As it is obviously seen, the results of the analytic and numerical approaches have shown excellent compatibility. In order to have a better scheme of the problem, displacement $x$ is shown in Figure 1 based on time, for ten seconds (different $f$'s and $A$'s are assumed).

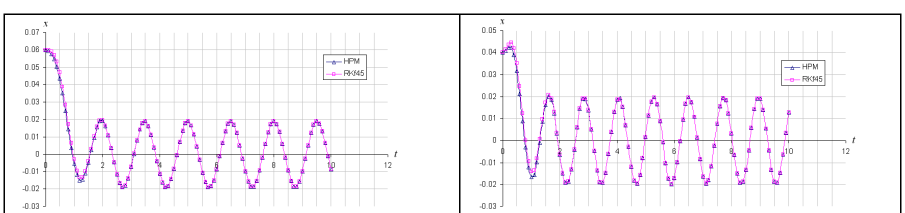
In the Figure 2, the velocity of each position is drawn versus its position; therefore, the velocity of any specific point $x$ can be easily read. This can only be done using the analytic method; since the equation of displacement is readily given by this method, the first and second differentiations

Table 1: The numerical values for $x$ and $\dot{x}$ for ten different points of time (analytical ) for $f = 0.5$, $A = 0.06$, $\omega = 4.163379415$.

| $t$ | $x$ | $\dot{x}$ |
|---|---|---|
| 0 | 0.06 | 0 |
| 1 | −0.005350926 | −0.078888606 |
| 2 | 0.019216633 | −0.013109041 |
| 3 | −0.007335627 | 0.072786956 |
| 4 | −0.011085293 | −0.064177133 |
| 5 | 0.018934134 | −0.005887804 |
| 6 | −0.008674275 | 0.070317298 |
| 7 | −0.009880823 | −0.067501783 |
| 8 | 0.018986752 | 0.000133187 |
| 9 | −0.009935365 | 0.067363257 |
| 10 | −0.00861736 | −0.070438539 |

Table 2: The numerical values for $x$ and $\dot{x}$ for ten different points of time (Numerical), for $f = 0.5$, $A = 0.06$, $\omega = 4.163379415$.

| $t$ | $x$ | $\dot{x}$ |
|---|---|---|
| 0 | 0.0599999999999999978 | 0 |
| 1 | −0.00281678838383509455 | −0.0825804788941202074 |
| 2 | 0.0195449692025107016 | −0.0138897877722475686 |
| 3 | −0.00730911369839314530 | 0.0727177020000000000 |
| 4 | −0.0110834704233601106 | −0.0641821083631454459 |
| 5 | 0.0189342438965698738 | −0.00588816238606704354 |
| 6 | −0.00867426974330831082 | 0.0703173716512146968 |
| 7 | −0.00988082636899610680 | −0.0675016034979940006 |
| 8 | 0.0189867450545043934 | 0.000133072037385784277 |
| 9 | −0.00993535776201434629 | 0.0673626924105089342 |
| 10 | −0.00861735199761238732 | −0.0704385763384659941 |



(a) $f$=0.5, $A$=0.06, ω=4.163379415 (b) $f$=0.7, $A$=0.04, ω=5.147879675

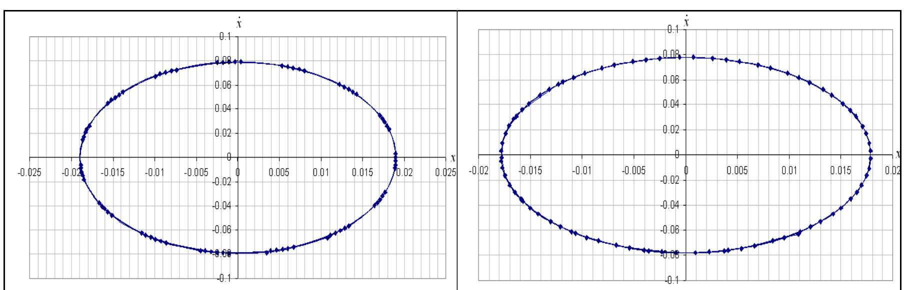
Figure 1: Displacement $x$ based on time $t$.

Table 3: A comparative table for error detection of the analytic method, for $f = 0.5$, $A = 0.06$, $\omega = 4.163379415$.

| | $x$ | | $\dot{x}$ | |
|---|---|---|---|---|
| $t$ | HPM | RKf45 | HPM | RKf45 |
| 0 | 0.06 | 0.06 | 0.00 | 0.00 |
| 1 | −0.00535092579 | −0.00281678838383509455 | −0.07888860559 | −0.0825804788941202074 |
| 2 | 0.01921663298 | 0.0195449692025107016 | −0.01310904065 | −0.0138897877722475686 |
| 3 | −0.007335626939 | −0.00730911369839314530 | 0.07278695633 | 0.0727177020766173188 |
| 4 | −0.01108529346 | −0.0110834704233601106 | −0.06417713286 | −0.0641821083631454459 |
| 5 | 0.01893413390 | 0.0189342438965698738 | −0.00588780420 | −0.00588816238606704354 |
| 6 | −0.008674275495 | −0.00867426974330831082 | 0.07031729844 | 0.0703173716512146968 |
| 7 | −0.009880823315 | −0.00988082636899610680 | −0.06750178333 | −0.0675016034979940006 |
| 8 | 0.01898675166 | 0.0189867450545043934 | 0.00013318678 | 0.000133072037385784277 |
| 9 | −0.009935365073 | −0.00993535776201434629 | 0.06736325729 | 0.0673626924105089342 |
| 10 | −0.008617359691 | −0.00861735199761238732 | −0.07043853899 | −0.0704385763384659941 |

Table 4: A comparative table for error detection of the analytic method, for $f = 0.7$, $A = 0.04$, $\omega = 5.147879675$.

| | $x$ | | $\dot{x}$ | |
|---|---|---|---|---|
| $t$ | HPM | RKf45 | HPM | RKf45 |
| 0 | 0.040000000 | 0.0400000000000000010 | 0.00 | 0.00 |
| 1 | −0.01668447952 | −0.0141498204358211938 | −0.01790373528 | −0.0215962732978834948 |
| 2 | −0.006766243615 | −0.00643783101998039124 | −0.09550085587 | −0.0962817733972154006 |
| 3 | 0.01386657809 | 0.0138931356222193238 | −0.07218316940 | −0.0722528614780738738 |
| 4 | 0.01856032332 | 0.0185621543419157848 | 0.03426369079 | 0.0342587020066141973 |
| 5 | 0.001795120001 | 0.00179523506518929110 | 0.1010828182 | 0.101082536865805456 |
| 6 | −0.01704578135 | −0.0170457707679198232 | 0.05102090371 | 0.0510209094083521886 |
| 7 | −0.01617683835 | −0.0161768394708507908 | −0.05803595515 | −0.0580359227265403871 |
| 8 | 0.003397207022 | 0.00339721173637169130 | −0.09998605091 | −0.0999864615455689084 |
| 9 | 0.01904310147 | 0.0190430958431267415 | −0.02632428341 | −0.0263237149352404214 |
| 10 | 0.01266966369 | 0.0126696551453435498 | 0.07777683300 | 0.0777768691824906740 |



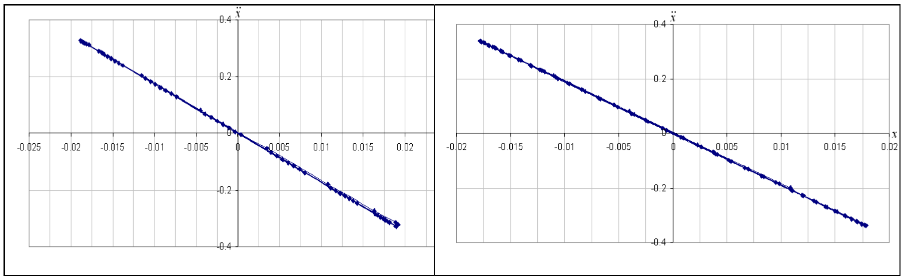
(a) $f$=0.5, $A$=0.06, ω=4.163379415 (b) $f$=0.6, $A$=0.05, ω=4.582639115

Figure 2: $\dot{x}$ based on $x$.

can be simply done by differentiating with respect to $t$.

Also using Figure 3, the acceleration of any specific point $x$ can be easily read. As mentioned earlier, this can only be done using the analytic approach.

The important point that cannot be seen on the figures of $(\dot{x} - x)$ and $(\ddot{x} - x)$ is that the starting



(a) $f$=0.5, $A$=0.06, ω=4.163379415 (b) $f$=0.6, $A$=0.05, ω=4.582639115

Figure 3: $\ddot{x}$ based on $x$.

part of these diagrams, which refers to the times between $t$'s from 0 to 1, is not drawn. The reason is that in this period of time, the behavior of the displacement equation has not yet become harmonic, and therefore, the velocity and acceleration is not in the range of the above diagrams.

## 6. CONCLUSION

In this paper, chaotic oscillator is used to acquire weak GPS signal using HPM, which is a strong analytical method applied to the nonlinear equation of an oscillator with damping term, and the results have been compared with that of the numerical solution. The advantage is driven from properties of non-linear dynamics.

Some other details in engineering have not been discussed in this paper. For example, in order to cover the possible Doppler shift range 20 kHz [21], oscillator acquisition system can be implemented by parallel structure with multiple oscillator acquisition channels. The main advantage of applying HPM is that the results are readily obtained and a few iterations are used. The significant merit of the analytic approach is to provide scientists with the general parametric relation between the dependent and independent variables, namely, displacement and time respectively. Therefore, the related equations can be simply obtained, giving one the opportunity for further studies, for different cases and thereby different parameters.

## REFERENCES


1. Huang, P., Y. Pi, and Z. Zhao, "Weak GPS signal acquisition algorithm based on chaotic oscillator," *Journal of Advances in Signal Processing, (EURASIP)*, in Press.
2. He, J. H., "Homotopy perturbation method for bifurcation of nonlinear problems," *International Journal of Nonlinear Science and Numerical Simulation*, Vol. 6, No. 2, 207–208, 2005.
3. He, J. H., "The homotopy perturbation method for nonlinear oscillators with discontinuities," *Appl. Math. Comput.*, Vol. 151, No. 1, 287–292, 2004.
4. He, J. H., "A coupling method of homotopy technique and perturbation technique for nonlinear problems," *Int. J. Nonlinear Mech.*, Vol. 35, 37–43, 2000.
5. Tari, H., D. D. Ganji, and M. Rostamian, "Approximate solutions of K(2,2), KdV and modified KdV equations by variational iteration method, homotopy perturbation method and homotopy analysis method," *International Journal of Nonlinear Science and Numerical Simulation*, Vol. 8, No. 2, 203–210, 2007.
6. Ganji, D. D. and A. Sadighi, "Application of He's homotopy-perturbation method to nonlinear coupled systems of reaction-diffusion," *International Journal of Nonlinear Science and Numerical Simulation*, Vol. 7, No. 4, 411–418, 2006.
7. Ganji, S. S., D. D. Ganji, H. Babazadeh, and S. Karimpour, "Variational approach method for nonlinear oscillations of the motion of a rigid rod rocking back and cubic-quintic duffing oscillators," *Progress In Electromagnetics Research M*, Vol. 4, 23–32, 2008.
8. Babazadeh, H., D. D. Ganji, and M. Akbarzade, "He's energy balance method to evaluate the effect of amplitude on the natural frequency in nonlinear vibration systems," *Progress In Electromagnetic Research M*, Vol. 4, 143–154, 2008.
9. Tsui, J. B.-Y., *Fundamental of Global Positioning System Receivers a Software Approach*, 136–145, John Wiley & Sons, Inc, New York, 2004.
10. Akos, D. M., P. L. Normark, J. T. Lee, K. G. Gromov, J. B. Y. Tsui, and J. Schamus, "Low power global navigation satellite system (GNSS) signal detection and processing," *Proceeding of the ION GPS 2000*, 784–791, Salt Lake City, UT, 2000.

11. Steven, H. S., *Nonlinear Dynamics and Chaos*, 366–369, Perseus Publish, LLC, Massachusetts, 2000.
12. Bo, L. E., et al., "Chaotic oscillator and other techniques for detection of weak signals," *IEICE — Tran. Fund. Elec., Comm. & Comp. Sci.*, Vol. E88-A, No. 10, 2699–2701, 2005.
13. Rao, S. S., *Mechanical Vibrations*, 3rd Edition, Addison-Wesley, 1995.
14. Rao, V. D. and J. Srinivas, *Textbook of Mechanical Vibrations*, Prentice Hall, 2004.
15. Wang, G., et al., "The application of chaotic oscillators to weak signal detection," *IEEE Transactions on Industrial Electronics*, Vol. 46, No. 2, 440–444, 1999.
16. Steven, H. S., *Nonlinear Dynamics and Chaos*, 366–369, Perseus Publish, LLC, Massachusetts, 2000.
17. Holmes, P., "A nonlinear oscillator with strange attractor," *Phil. Trans. Roy. Soc. A*, Vol. 292, 419, 1979.
18. Mann, B. P. and F. A. Khasawneh, "An energy-balance approach for oscillator parameter identification," *Journal of Sound and Vibration*, Vol. 321, No. 1–2, 65–78, Mar. 2009.
19. Mark, L. P., "Block acquisition of weak GPS signals in a software receiver [C]," *14th Int. Tech. Meeting of the Satellite Division of the Institute of Navigation (ION GPS 2001)*, 2838–2850, Salt Lake City, UT, 2001.
20. Rao, S. S., *Mechanical Vibrations*, 3rd Edition, Addison-Wesley, 1995.
21. Tsui, J. B.-Y., *Fundamental of Global Positioning System Receivers a Software Approach*, 136–145, John Wiley & Sons, Inc, New York, 2004.

# Electrodynamics in Expanding Cavities

**J. A. Grzesik**
Allwave Corporation, 3860 Del Amo Boulevard, Suite 404, Torrance, CA 90503, USA

**Abstract**— One embodiment of a plasma fusion reactor, in vogue a number of years ago at the Lawrence Livermore National Laboratory, employed a cylindrical chamber having its ends capped by massive, Yin-Yang (Y-Y) magnetic coils serving as barriers against charged particle escape.[1] Such Y-Y coils, by their very geometry, require opposed current flow in close proximity, circumstance which summons forth a dilatational magnetic pressure raising device disintegration to the level of a calamitous possibility. And, while such Y-Y fragmentation is surely not a welcome design outcome, nevertheless it does invite a preliminary analysis as to its potential violence, analysis which enjoys besides a modicum of theoretical interest by virtue of making relevant a scenario of electrodynamics in an expanding cavity.

With these dual aims in mind, we had many years ago undertaken the study of the very simplest of such expanding cavity situations, namely, the growing interstitial (vacuum) wafer separating two massive metallic plates undergoing a symmetric flight from one another. Quick penetration into the heart of this problem was provided by the observation that, on the one hand, a quasi-static (QS) field computation would surely suffice, while, on the other, that a moving boundary condition (MBC) could be fashioned in lowest relativistic order by combining laboratory-frame electric $\mathbf{E}$ and magnetic $\mathbf{B}$ fields, and the boundary velocity $\mathbf{v}$, and thence requiring that the effective tangential electric field $\sim \hat{\mathbf{v}}\times\{\mathbf{E}+\mathbf{v}\times\mathbf{B}\}$ vanish upon both plate boundaries. In this process, a secondary computation of $\mathbf{E}$ was bootstrapped upon a primary, QS one for $\mathbf{B}$ via Faraday's law, whereby the obligatory time derivative of the latter was implicitly tethered to the dynamic evolution of its underlying separation parameter $\eta(t)$.[2] Under this viewpoint there easily emerged the invariance against time of the product of $B$ by $\eta$ (or else of current $I$ by $\eta$) leading to a simple differential equation for the dynamical evolution of the net separation $\eta(t)$, and, in particular, to the identification of a characteristic time scale $\tau$ suggesting a most vigorous magnet disintegration. This aspect of the work has been previously reported in summary form,[3] and is set out anew here for the purpose of building an intuitive, heuristic base concerning field evolution within the primitive, expanding wafer cavity now at hand.

A heuristic base of this sort is far too coarse to account for field retardation effects due to signal transit at finite light speed $c$. We remove this defect by returning to the Faraday/Ampère equations in their primitive form and subjecting them first to Fourier transformation in coördinate $z$ along the direction of cavity expansion perpendicular to magnet walls. Such transformation embraces the entire interval $-\infty < z < \infty$ and, as such, submits to a null-field attitude which regards the field, in both its electric $\mathbf{E}$ and magnetic $\mathbf{B}$ manifestations, as being zero exterior to the expanding wafer, i.e., $\forall |z| > \eta(t)/2$. Due deference must of course be paid, in the form of Dirac delta function sources placed at $z = \pm\eta(t)/2$, to the radiation emanating from surface current density $\pm I(t)$ flowing on cavity walls. Elimination of either field transform leads then to a simple harmonic differential equation in time $t$ having a source gauged by $I(t)$. Its solution is readily gotten in a form that allows inverse Fourier transformation to proceed smoothly and, in particular, to identify a retarded signal emission time $t_* < t$ as gauged from either plate which obeys the intuitively pleasing condition $c(t - t_*) = \{\eta(t) + \eta(t_*)\}/2$. All in all one confronts at this point a relatively simple pattern of connections between the field and its source $I$ as reckoned at retarded times $t_*$ suitably structured so as to track upper/lower plate emissions, connections which succumb at length to an *a posteriori* enforcement of the non-relativistic limit $\dot{\eta}(t) \ll c$ so as to recover anew the key $I \times \eta$ invariance previously inferred during the prelude of approximate, QS/MBC analysis.

---

[1] A rather dramatic depiction of Y-Y coils, which serves to put their size into perspective against the scale of a human figure, perched astride, and masked against the glare of his welder's blowtorch, appears on the cover of *Physics Today*, Volume 34, Issue 9, September 1981, herewith reproduced. Confer also the accompanying article, "A tandem mirror in place at MFTF" by Bertram M. Schwartzschild on p. 22.

[2] Thus, as regards boundary velocities $v$ along the plate normal, $v = \pm\dot{\eta}(t)/2$, the overdot signifying a time derivative.

[3] Grzesik, J., "Kinetic energy release in missile ejection from confining magnets," *Transactions of the American Nuclear Society Annual Meeting*, 34, New York, June 12–16, 1977.

Some preliminary work from that same epoch, intended to account for a merely partial magnet disintegration in the form of a ribbon missile, parallel to the prevailing current flow and ejected from just one plate, is reported in:

Kastenberg, W. E. and D. Okrent, "Some safety studies for conceptual fusion-fission hybrid reactors," UCLA, *School of Engr. & Applied Science (Prepared for the Electric Power Research Institute, Palo Alto, California)*, I-1, I-32, July 1978.

This complementary work, however, awaits to this day both its amendment and completion.

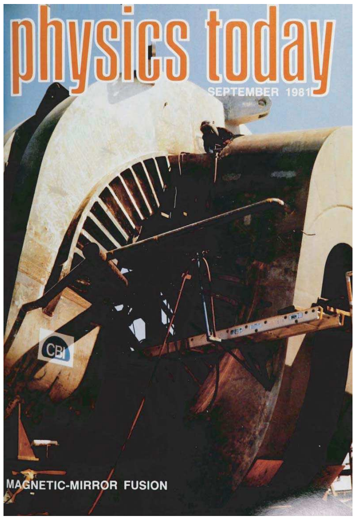



Figure 1: In medias res: Yin-Yang mirror magnet coils under construction.

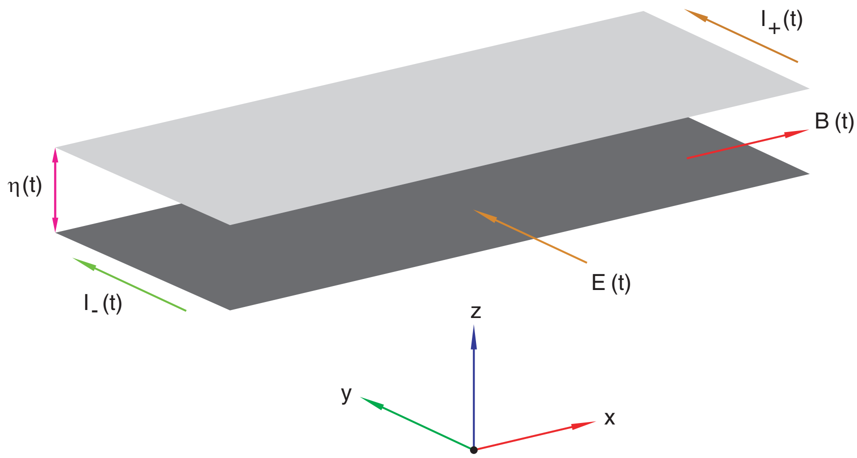



Figure 2: Expanding plate geometry.

## 1. EXPANDING CAVITY FIELDS: THE INTUITIVE, QS/MBC SCENARIO

We place the origin of a right-handed Cartesian coördinate system anywhere along the mid-plane between the two magnet faces, each one of them at right angles to the $z$-axis and situated, prior to the onset of cavity dilation, at $z = \pm a/2$. Before such dilation, steady current sheets of linear density $\pm I_0 \hat{\mathbf{e}}_y$ are assumed to flow upon these faces, and to sustain thus a sectionally uniform magnetic field, equal to[4]

$$\mathbf{B}_0 = -\mu_0 I_0 \hat{\mathbf{e}}_x \tag{1}$$

within the interstitial sandwich $-a/2 < z < a/2$, and zero otherwise (cf., Figure 2, which, for $t \geq 0$, depicts an evolved plate separation $\eta(t) \geq a$ and the associated field/current strengths $B(t)$ and $I_\pm(t)$ from Eqs. (6) and (7) below). The source of magnetic field (1) is equally apportioned among the two surface currents and because, as is easily shown, the steady-state magnetic field of any given current distribution is inherently incapable of exerting a net self-force, each of the two current sheets at $z = \pm a/2$ experiences, in the field of its partner, a repulsive pressure in an amount

$$p = \frac{\mu_0 I_0^2}{2} = \frac{B_0^2}{2\mu_0}, \tag{2}$$

with note taken of the fact that $B_0 = -\mu_0 I_0$.

We imagine next that the supporting structures, while they may indeed be designed to withstand such pressure when uncompromised, nevertheless somehow do fail at time $t = 0$, allowing the repulsion to initiate cavity dilation.[5] Regarding such dilation as a QS process, we arrive at a picture wherein, at any given moment $t > 0$, the magnetic field, with value $B(t)$, remains both uniform and $\hat{\mathbf{e}}_x$-directed, and is accompanied now by an $\hat{\mathbf{e}}_y$-directed electric field

$$\mathbf{E}(z,t) = E(z,t)\hat{\mathbf{e}}_y. \tag{3}$$

A simple symmetry argument, involving coördinate system rotation through angle $\pi$ around $\hat{\mathbf{e}}_x$, shows next that $E(z,t)$ is antisymmetric around cavity mid-plane $z = 0$, whereas its value is then fixed at

$$E(z,t) = z\dot{B}(t) \tag{4}$$

through recourse to Faraday's law of induction when applied to a rectangular circuit in a plane perpendicular to $\hat{\mathbf{e}}_x$, having its horizontal leg aligned along $\hat{\mathbf{e}}_y$, and the whole symmetrically disposed

[4]We describe field quantities in standard SI units. Vectors $\hat{\mathbf{e}}_x$, $\hat{\mathbf{e}}_y$, and $\hat{\mathbf{e}}_z$ comprise together the usual Cartesian orthonormal triad.

[5]We further idealize by ignoring the simultaneous presence of gravity pull, both as to its magnitude and its rôle in symmetry disruption. We focus instead exclusively upon the time scale of the dynamics engendered by purely electromagnetic forces.

with respect to vertical origin $z = 0$. At the evolving plate positions $\pm\eta(t)/2$, MBC further requires that

$$\eta\dot{B} + \dot{\eta}B = 0, \tag{5}$$

a perfect derivative which yields

$$B(t) = \frac{B_0 a}{\eta(t)} \tag{6}$$

so as to agree with the fact that $\eta(0) = a$. Accompanying (6), of course, are currents of strength

$$I_{\pm}(t) = \pm\frac{I_0 a}{\eta(t)} \tag{7}$$

flowing upon upper/lower plates.

Magnetic pressure $p$ retains its form on the right in (2) provided, of course, that $B_0$ is duly replaced by $B(t)$ as now discovered. If next one writes $\rho$ for the mass per unit area[6] of each plate, then the dynamics of their expansion is seen to be governed by

$$\frac{\rho\ddot{\eta}}{2} = \frac{B^2}{2\mu_0}, \tag{8}$$

or else

$$\frac{\rho}{4}\frac{d\dot{\eta}^2}{d\eta} = \frac{B_0^2 a^2}{2\mu_0 \eta^2}, \tag{9}$$

which latter admits a first quadrature in the form

$$\dot{\eta}^2 = \frac{2B_0^2 a^2}{\mu_0 \rho}\left\{\frac{1}{a} - \frac{1}{\eta}\right\}. \tag{10}$$

We proceed now to displace $B_0$ in favor of its more operationally accessible attribute $I_0$, so as to rewrite (10) as

$$\frac{d}{dt}\left(\frac{\eta}{a}\right) = \sqrt{\frac{2\mu_0 I_0^2}{a\rho}}\sqrt{\frac{\frac{\eta}{a} - 1}{\frac{\eta}{a}}}, \tag{11}$$

a gesture which identifies the dimensionless variables

$$Q = \frac{\eta}{a} \tag{12}$$

and

$$T = \frac{t}{\tau}, \tag{13}$$

with

$$\tau = \sqrt{\frac{a\rho}{2\mu_0 I_0^2}}. \tag{14}$$

In its newly acquired, dimensionless guise, Eq. (11) thus reëmerges in the form

$$\frac{dQ}{dT} = \sqrt{\frac{Q-1}{Q}} \tag{15}$$

which, when brought to quadrature as

$$T = \int_1^Q \sqrt{\frac{\zeta}{\zeta - 1}}d\zeta, \tag{16}$$

yields to an elementary substitution[7] so as to provide the explicit result

$$T = \sqrt{Q(Q-1)} + \ln\left\{\sqrt{Q} + \sqrt{Q-1}\right\}, \tag{17}$$

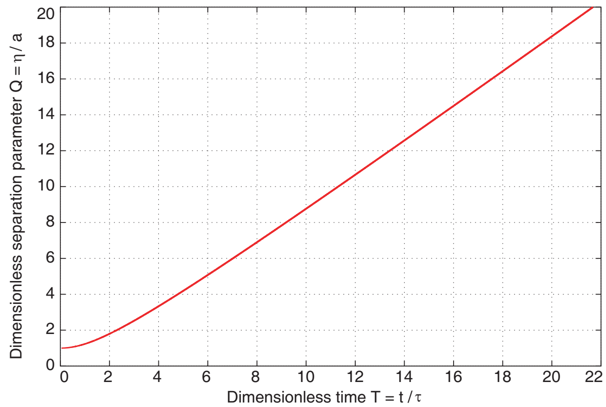



Figure 3: Universal expansion curve for separating magnet plates.

plotted above as Figure 3.

About this result, we note at once the seemingly innocuous asymptotic limit

$$T \approx Q \tag{18}$$

as both $T$ and $Q$ tend toward $\infty$, an approximate equality whose physical import is that of magnet faces receding at a steady, terminal velocity, evidently unaffected by any residual magnetic pressure. From (12) and (14), we then restore a physical terminal expansion rate in an amount

$$\dot{\eta}_\infty = \frac{a}{\tau} = \sqrt{\frac{2a\mu_0 I_0^2}{\rho}} \tag{19}$$

and, corresponding to it, a terminal kinetic energy $KE_\infty$ per unit magnet area

$$KE_\infty = \frac{\rho\dot{\eta}_\infty^2}{4}, \tag{20}$$

one which combines of course the contribution of both plates. When finally written as

$$KE_\infty = \frac{\mu_0 I_0^2 a}{2} = \frac{B_0^2 a}{2\mu_0}, \tag{21}$$

Eq. (21) conveys nothing other than the complete exchange into terminal kinetic energy of the magnetostatic energy initially resident between the undisturbed magnet faces. Our physical instincts would in fact be expected to recoil from any potential outcome at variance with this result.

In Figure 3, we plot the universal expansion profile (17) with dimensionless time $T$ as its abscissa and dimensionless separation $Q$ as ordinate. This curve is of necessity concave up,[8] with the asymptotic limit (18) approached from below, simply because of the inevitable time lag prior to full conversion of magnetic into kinetic energies. That conversion, however, is quite precipitous in physical terms. For instance, if we invoke the typical values $\rho \approx 10^3\,\mathrm{kgm}^{-2}$, $a \approx 1\,\mathrm{m}$, and $\mu_0 I_0 = 8.1\,\mathrm{Tesla}$, we find $\tau \approx 3.09\ldots$ milliseconds, a time short enough to be deemed as a token of a most violent expansion.

### 1.1. QS/MBC Appendix: Field Energetics

Inasmuch as all field attributes lie at hand in such simple form, it is useful to pause so as to reassure ourselves yet again of their self-consistency as regards energy conservation. All of our ensuing discussion unfolds now on a per unit area basis, without further comment.

The magnetic energy $W_m$ is easily disposed of, and reads

$$W_m = \frac{B^2\eta}{2\mu_0} = \frac{B_0^2 a^2}{2\mu_0\eta}. \tag{22}$$

[6]Units of $\mathrm{kgm}^{-2}$.

[7]Only the most routine integrals emerge from (16) after one sets $\zeta = \cosh^2\vartheta$.

[8]One easily verifies on the basis of Eq. (15) that such upward concavity persists unabated even as $T \to \infty$.

Somewhat more complicated is the accompanying electric energy $W_e$, for which we now find, from Eqs. (4) and (6),

$$W_e = \epsilon_0 \int_0^{\eta/2} E^2 dz = \frac{1}{12}\left(\frac{\dot{\eta}}{c}\right)^2 W_m, \tag{23}$$

with $c = (\epsilon_0\mu_0)^{-1/2}$ being the customary expression for the velocity of light in vacuum. And since, from (10), $\dot{\eta}$ rises monotonically only toward a finite, limiting value (19), $W_e$ from (23) ultimately declines together with $W_m$. On the other hand, inasmuch as the expression on the right in (23) is clearly a relativistic correction, and quadratic at that, we are, in our present, QS frame of mind, entirely justified in neglecting it.

Now the decline of $W_m$ is channeled via the Poynting vector $\mathbf{S}=\mathbf{E}\times\mathbf{B}/\mu_0$, here strictly parallel/antiparallel to $\hat{\mathbf{e}}_z$, and it would be pleasing indeed to witness the magnitude of $\mathbf{S}$ dissected into one part, $S_{fm}$, which accounts for (the now primarily) magnetic field penetration into a growing domain, and another, $S_{ke}$, which urges the plates toward ever higher, if necessarily bounded, kinetic energy. Evidently we have, from (6), with both plates taken into account,

$$S_{fm} = \frac{B_0^2 a^2 \dot{\eta}}{2\mu_0 \eta^2}. \tag{24}$$

At the same time, from (10), and again for both plates,

$$S_{ke} = \frac{d}{dt}\left(\frac{\rho\dot{\eta}^2}{4}\right) = \frac{B_0^2 a^2 \dot{\eta}}{2\mu_0 \eta^2}, \tag{25}$$

identical in every respect with $S_{fm}$. Adding the two we get

$$S_{fm} + S_{ke} = \frac{B_0^2 a^2 \dot{\eta}}{\mu_0 \eta^2}, \tag{26}$$

and this is precisely what emerges by forming the Poynting vector when (4) and (6) are applied, with component directions properly considered, at both upper and lower plates, $z = \pm\eta/2$. We have before us now in plain view both energy conservation and a neat equipartition of energy flow into its field and kinetic reservoirs.

## 2. EXPANDING CAVITY FIELDS: RIGOROUS MAXWELL EQUATIONS *cum* NULL EXTERIOR FIELD

The QS field description so far presented is necessarily coarse-grained on a time scale set by $\eta(t)/c$, simply because, by deliberate choice, it ignores the details of the cross-talk whereby the plates inform each other of their evolving separation. Simply put, it conveys in some sense an average over many cycles of $\eta(t)/c$. Thus, so as to assure its relevance on physical grounds, we should evidently prefer to have the underlying parameters $a$, $\rho$, and $I_0$ constrained in a way to assure that $\tau \gg \eta(t)/c$. Adequate as this picture may be, it is nevertheless desirable to refine its temporal resolution, a task to which we now turn by strengthening our adherence to the underlying Maxwell equations.

Adherence to Maxwell's equations raises the additional issue of just how much attention need be paid to Lorentz-transformed field components. The answer, as it happily turns out, is none, inasmuch as the entire discussion is free to unfold in the stationary frame of the pre-disruption magnet arrangement. Accordingly, the only nod to special relativity, and a passing one at that, is a recognition of signal retardation, a recognition which emerges in a natural way at a later point of Fourier transform inversion. All in all, we seek here not so much to gratify an intellectual indulgence but, rather, to establish an analytic basis from which practical approximations, such as (6) and (7), may emerge in a natural way.

### 2.1. Basic Relations

Maxwell's equations or not, the symmetries of the preceding discussion remain intact, and indicate the presence of just two field components[9]

$$\mathbf{E}(z,t) = E(z,t)\hat{\mathbf{e}}_y \tag{27}$$

[9]However, do confer the provisional material in Subsection 2.6 at paper's end.

and

$$\mathbf{B}(z,t) = B(z,t)\hat{\mathbf{e}}_x, \tag{28}$$

of which the first is odd against sign change in $z$, *viz.*,

$$E(-z,t) = -E(z,t), \tag{29}$$

while the second is even,

$$B(-z,t) = B(z,t). \tag{30}$$

On the plates themselves at $z = \pm\eta/2$ there flow in addition the time-dependent currents

$$\mathbf{I}_{\pm}(t) = \pm I(t)\hat{\mathbf{e}}_y \tag{31}$$

which are the source of fields (27)–(28) and respond in turn to the local values thereof.

These up/down sources immediately assert themselves on noting that projection of Ampère's equation onto $\hat{\mathbf{e}}_y$ takes the form

$$\frac{\partial B(z,t)}{\partial z} = \frac{1}{c^2}\frac{\partial E(z,t)}{\partial t} + \mu_0 I(t)\left\{\delta\left(z-\eta(t)/2\right) - \delta\left(z+\eta(t)/2\right)\right\} \tag{32}$$

involving on its right two instances of Dirac's delta function $\delta$. The corresponding projection of Faraday's connection upon $\hat{\mathbf{e}}_x$ is source-free, and simply reads

$$\frac{\partial E(z,t)}{\partial z} = \frac{\partial B(z,t)}{\partial t}. \tag{33}$$

We observe in passing that Eqs. (32)–(33) accommodate without conflict the field symmetries posited in (29)–(30).

We next subject Eqs. (32) and (33) to Fourier transformation against coördinate $z$, distinguishing the transforms from their parent quantities by a tilde placed atop. Thus

$$\tilde{E}(k,t) = \int_{-\infty}^{\infty} e^{-ikz} E(z,t)dz \tag{34}$$

and

$$\tilde{B}(k,t) = \int_{-\infty}^{\infty} e^{-ikz} B(z,t)dz. \tag{35}$$

Following such transformation, Eqs. (32) and (33) become

$$ik\tilde{B}(k,t) = \frac{1}{c^2}\frac{\partial \tilde{E}(k,t)}{\partial t} - 2i\mu_0 I(t)\sin\left(\frac{k\eta(t)}{2}\right) \tag{36}$$

and

$$ik\tilde{E}(k,t) = \frac{\partial \tilde{B}(k,t)}{\partial t}. \tag{37}$$

While our main objective is $\tilde{B}(k,t)$ (or, more precisely, its Fourier parent $B(z,t)$), we first eliminate it from Eqs. (36)–(37) in favor of $\tilde{E}(k,t)$, and then, with $\tilde{E}(k,t)$ thus in hand, return to Eq. (36) with full attention dedicated to that primary target. Such elimination results in

$$\frac{\partial^2 \tilde{E}(k,t)}{\partial t^2} + (kc)^2\tilde{E}(k,t) = \frac{2i}{\epsilon_0}\frac{d}{dt}\left\{I(t)\sin\left(\frac{k\eta(t)}{2}\right)\right\} \tag{38}$$

a structure whose solution can be exhibited as

$$\tilde{E}(k,t) = \frac{2i}{\epsilon_0}\int_{-\infty}^{t} G_{\mathrm{ret}}(t-t_1)\frac{d}{dt_1}\left\{I(t_1)\sin\left(\frac{k\eta(t_1)}{2}\right)\right\}dt_1 \tag{39}$$

in terms of the retarded Green's function $G_{\mathrm{ret}}(t-t_1)$, which satisfies the following analogue of (38)

$$\frac{d^2 G_{\mathrm{ret}}(t)}{dt^2} + (kc)^2 G_{\mathrm{ret}}(t) = \delta(t) \tag{40}$$

and the additional stipulation that $G_{\text{ret}}(t) = 0$ whenever its temporal argument $t$ becomes negative.

As the solution of (40) is standard, and reads[10]

$$G_{\text{ret}}(t) = U_+(t)\frac{\sin(kct)}{kc}, \tag{41}$$

with $U_+(t)$ being the Heaviside unit step, positive on its right, representation (39) is brought into the form

$$\tilde{E}(k,t) = \frac{2i}{k}\sqrt{\frac{\mu_0}{\epsilon_0}}\int_0^t \sin\left(kc(t-t_1)\right)\frac{d}{dt_1}\left\{I(t_1)\sin\left(\frac{k\eta(t_1)}{2}\right)\right\}dt_1 \tag{42}$$

once it is recognized that both current $I(t)$ and separation $\eta(t)$ are quiescent prior to the onset of magnet separation at $t = 0$. A routine integration by parts further reduces this into

$$\tilde{E}(k,t) = -\frac{2i}{\epsilon_0}\left[I_0\sin\left(\frac{ka}{2}\right)\frac{\sin(kct)}{kc} - \int_0^t I(t_1)\sin\left(\frac{k\eta(t_1)}{2}\right)\cos\left(kc(t-t_1)\right)dt_1\right]. \tag{43}$$

Collaboration between (36) and (43) then yields

$$\tilde{B}(k,t) \;=\; -2\sqrt{\frac{\mu_0}{\epsilon_0}}\left[I_0\sin\left(\frac{ka}{2}\right)\frac{\cos(kct)}{kc} + \int_0^t I(t_1)\sin\left(\frac{k\eta(t_1)}{2}\right)\sin\left(kc(t-t_1)\right)dt_1\right]. \tag{44}$$

### 2.2. Fourier Inversion

Recovery of the spatial field dependence from either (43) or (44) is quite similar and so, given the magnetic context of our discourse, we give initial pride of place to (44). When considering thus the Fourier inversion

$$B(z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikz}\tilde{B}(k,t)dk \tag{45}$$

we encounter two distinct categories of terms, those proportional to initial current density $I_0$ {Category I}, and those associated with the quadrature over time {Category II}. By virtue of the factor $\sin(ka/2)/k$, the Fourier inversions appropriate to Category I are, in their aggregate, free from singularities along the real $k$ axis, a feature which permits one to indent the inversion contour away from the axis of reals, say into $C_-$ running ever so slightly below the origin $k = 0$. The Fourier inversions appropriate to Category II are, by inspection, likewise free from singularities along the real $k$ axis. There emerge the following results.

**Category I**

Even though singularity-free at $k = 0$ in the aggregate, each individual exponential contribution into which the Category I term splinters does have a simple pole there. We find

$$\frac{1}{2\pi}\int_{C_-}\frac{dk}{k}\exp\left(ik\left\{z+\sigma_1 a/2+\sigma_2 ct\right\}\right) = iU_+\left(z+\sigma_1 a/2+\sigma_2 ct\right) \tag{46}$$

once again in terms of the Heaviside step as previously introduced, and with $\sigma_1 = \pm$ and $\sigma_2 = \pm$ being the independent signatures of all four contributing exponentials in their native states of occurrence.

**Category II**

In Category II, the Heaviside steps are replaced by Dirac deltas. We get, in the first place

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(ik\left\{z+\sigma_1\eta(t_1)/2+\sigma_2 c(t-t_1)\right\}\right)dk = \delta\left(z+\sigma_1\eta(t_1)/2+\sigma_2 c(t-t_1)\right). \tag{47}$$

And then, since

$$\delta\left(z+\sigma_1\eta(t_1)/2+\sigma_2 c(t-t_1)\right) = \frac{1}{c}\delta\left(t_1-\left\{t+\frac{\sigma_2 z+\sigma_1\sigma_2\eta(t_1)/2}{c}\right\}\right), \tag{48}$$

[10] That (41) is the correct Green's function retarded solution is seen by inspection, and can in any event be verified by solving Eq. (40) *ab initio* through yet another appeal to Fourier transformation, now against time $t$.

we are instructed to seek a retarded time $t_1 = t_*$, $0 < t_* < t$, in accordance with

$$t_* = t + \frac{\sigma_2 z + \sigma_1\sigma_2\eta(t_*)/2}{c}, \tag{49}$$

since only then will any one of the Dirac deltas be empowered to contribute to the temporal quadrature on the right in (44). Moreover, in the neighborhood of any such $t_*$ we have

$$\frac{1}{c}\delta\left(t_1 - \left\{t + \frac{\sigma_2 z + \sigma_1\sigma_2\eta(t_1)/2}{c}\right\}\right) = \frac{1}{c\left(1 - \sigma_1\sigma_2\dot{\eta}(t_*)/2c\right)}\delta(t_1 - t_*). \tag{50}$$

Of course, allowed solutions $t_*$ of (49) obeying $0 < t_* < t$, if any, depend not only upon current time $t$ and observation plane $z$ but also upon signatures $\sigma_1$ and $\sigma_2$. We make this explicit by attaching them as suffixes, *viz.*, $t_* \rightarrow t_{*\sigma_1\sigma_2}$.

On the basis of Eqs. (44) through (50), there emerges the expression

$$\begin{aligned} B(z,t) \;=\; -\frac{\mu_0}{2}\Bigg[ & I_0\left\{U_+(z+a/2+ct) + U_+(z+a/2-ct) - U_+\left(z-a/2+ct\right) - U_+\left(z-a/2-ct\right)\right\} \\ & - \left\{\frac{I(t_{*++})}{1-\dot{\eta}(t_{*++})/2c} - \frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c} - \frac{I(t_{*-+})}{1+\dot{\eta}(t_{*-+})/2c} + \frac{I(t_{*--})}{1-\dot{\eta}(t_{*--})/2c}\right\}\Bigg] \end{aligned} \tag{51}$$

which is subject to the further understanding that, in its last two lines, entry is denied to any term whose $t_{*\sigma_1\sigma_2}$ may fail to conform with $0 < t_* < t$. Thus, for instance, the term corresponding to $t_{*++}$ cannot possibly contribute unless $z < -\eta(t_{*++})/2$.

It follows in similar fashion from (43) that[11]

$$\begin{aligned} E(z,t) = -\frac{c\mu_0}{2}\Bigg[ & I_0\left\{U_+(z+a/2+ct) - U_+(z+a/2-ct) - U_+\left(z-a/2+ct\right) + U_+\left(z-a/2-ct\right)\right\} \\ & - \left\{\frac{I(t_{*++})}{1-\dot{\eta}(t_{*++})/2c} + \frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c} - \frac{I(t_{*-+})}{1+\dot{\eta}(t_{*-+})/2c} - \frac{I(t_{*--})}{1-\dot{\eta}(t_{*--})/2c}\right\}\Bigg]. \end{aligned} \tag{52}$$

It is comforting finally to note that the *a priori* symmetries (29)–(30) have survived all of the intervening analysis and are faithfully imprinted upon (52) and (51), in that order. A direct verification of this attribute follows from the fact that, for any real argument $\zeta$, $U_+(\zeta)+U_+(-\zeta) = 1$, whereas $t_{*\sigma_1\sigma_2}(-z,t) = t_{*(-\sigma_1)(-\sigma_2)}(z,t)$ as an easy consequence of (49). Somewhat more succinctly, one notes from Eqs. (43)–(44) that the symmetries (29)–(30) in physical space $z$ are echoed exactly,

$$\tilde{E}(-k,t) = -\tilde{E}(k,t) \tag{53}$$

and

$$\tilde{B}(-k,t) = \tilde{B}(k,t), \tag{54}$$

in reciprocal space $k$. Then, for instance,

$$B(-z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ikz}\tilde{B}(k,t)dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikz}\tilde{B}(-k,t)dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikz}\tilde{B}(k,t)dk = B(z,t), \tag{55}$$

and similarly for the electric antisymmetry $E(-z,t) = -E(-z,t)$. The upshot of such symmetry verification is the freedom which it bestows upon us to limit all further field discussion to just the upper, $z > 0$ cavity/magnet interval.

[11]By way of a quick sanity check, it is useful to remark that magnetic field $B(z,t)$ from (51) bears more than a casual similarity to its simple-minded forebear (1), whereas the electric field $E(z,t)$ as now assembled echoes the dimensions on view in its own precursor (4).

**2.3. Current/Separation Link: Null Exterior Field for $|z| > \eta(t)/2$**

Nothing has been said to this point regarding the time evolution of current $I(t)$ or the expanding cavity width $\eta(t)$. But, whatever may finally prove to be the temporal history of $\eta(t)$, current $I(t)$ must be adjudicated on the basis of purely electromagnetic considerations, applied to the geometric configuration prevailing at any given moment in time. One such consideration, easily expressed in global terms, is a simple requirement of null field penetration exterior to the magnetic cavity, which is to say, a requirement that $B(z,t) = 0$ whenever $z > \eta(t)/2$. For any $z > a/2$, the null field epoch has the temporal extent $0 < t < t_>$, with upper bound $t_>$ determined on the basis of $z = \eta(t_>)/2$. Evidently $t_> \gg (z - a/2)/c$.

On taking note thus of the declared membership caveats for terms involving the several $t_{*\sigma_1\sigma_2}$ on the right in (51), and of the underlying constraint $U_+(\zeta)+U_+(-\zeta) = 1$, we now find it advantageous to rewrite (51) as[12]

$$B(z,t) = -\frac{\mu_0}{2}\left[U_+\left(t - \frac{z-a/2}{c}\right)\left\{I_0 - \frac{I(t_{*--})}{1-\dot{\eta}(t_{*--})/2c}\right\}\right.$$
$$\left. -U_+\left(t - \frac{z+a/2}{c}\right)\left\{I_0 - \frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c}\right\}\right]. \tag{56}$$

We confront thus a propagation onset at $t_{\min} = (z - a/2)/c$ and its cessation at $t_{\max} = (z + a/2)/c$ of (just one half of) the spatially uniform field initially present within the undisturbed cavity and escaping now upward. Furthermore, there enter into competition at these times the fields subsequently radiated by the moving current sheets, first the radiation at time $t_{\min}$ released from the more proximate, approaching magnet plate, and then at time $t_{\max}$ from the more distant, receding one. Our global requirement that this composite exterior field vanish fragments thus into three distinct time regimes.

**Time Regime I: $0 < t < (z - a/2)/c$**

The expression on the right in (56) is inherently null without imposing any condition upon current $I(t)$.

**Time Regime II: $(z - a/2)/c < t < (z + a/2)/c$**

We witness here the arrival of the initially resident cavity field and that radiated from the closest, upper (approaching) magnet plate only. Radiation launched from the farther, lower (receding) magnet face has not yet had sufficient time to reach observation plane $z$. Under the competition of just the first two of these three candidates the null field requirement reads

$$I_0 = \frac{I(t_{*--})}{1-\dot{\eta}(t_{*--})/2c} \tag{57}$$

and is operative across the (radiation) time interval $0 < t_{*--} < t_{*--}^{\max}$, with $t_{*--}^{\max}$ determined from $t_{*--}^{\max} = \left(\eta(t_{*--}^{\max}) + a\right)/2c > a/c$. On this time interval the null field requirement provides that current $I(t_{*--})$ remain essentially constant at its initial value $I_0$, save for a small, first-order relativistic correction proportional to plate velocity $\dot{\eta}(t_{*--})/2c$ as scaled by light speed $c$, regardless of how that ratio may be determined.

**Time Regime III: $(z + a/2)/c < t < t_>$**

Radiation emanating from the lower, receding plate has by now been admitted into field competition, with the null-field requirement amended so as to read

$$\frac{I(t_{*--})}{1-\dot{\eta}(t_{*--})/2c} = \frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c} \tag{58}$$

and to offer thus a basis for a recursive, bootstrap current determination. Indeed, the retarded times $t_{*(\pm)(-)}$, both of which rise together with $t$, are here shifted with respect to one another, with $t_{*--}$ having an onset at $t_{*--} = t_{*--}^{\max}$, leading $t_{*+-}$, whose onset at $t_{*+-} = 0$ is assuredly lower. In particular, the $t_{*--}$ handoff at the close of Time Regime II dovetails smoothly onto its value at the onset of Time Regime III, as it evidently must. Furthermore, we have $\dot{t}_{*--} > \dot{t}_{*+-}$, as is easily seen on the basis of Eq. (49). And so in principle we can exploit the information already gleaned

[12]As they are deployed in Eq. (56) and in its subsequent partner (62), the overt Heaviside space/time cutoffs $U_+$ are clearly compatible with the constraints which we had previously imposed upon the admission on its right in Eq. (51) of $I(t_{*--})$ and $I(t_{*+-})$.

from Eq. (57) so as to advance across the entire positive time axis, simply by setting the (here strictly *gedanken*) observation planes $z$ as far to the right as we please. On the other hand, such a procedure, while logically defensible, suffers from what might best be described as a relativistic time granularity in the amount $a/c$, whose smallness renders this approach algorithmically unattractive.

Instead of all this, we can undertake a systematic, broad-brush rentrenchment into the nonrelativistic domain by noting, on the basis of (49), that the difference $t_{*\sigma_1\sigma_2} - t$ is, *eo ipso,* itself a relativistic correction, allowing one to approximate by setting[13]

$$t_{*\sigma_1\sigma_2} - t \approx \frac{\sigma_2 z + \sigma_1\sigma_2\eta(t)/2}{c} + \frac{\sigma_1\sigma_2\dot{\eta}(t)/2}{c}(t_{*\sigma_1\sigma_2} - t) \approx \frac{\sigma_2 z + \sigma_1\sigma_2\eta(t)/2}{c}. \tag{59}$$

The null boundary condition (58) thus becomes

$$\frac{I(t) - \dot{I}(t)\left(\frac{z+\eta(t)/2}{c}\right)}{1+\dot{\eta}(t)/2c} \approx \frac{I(t) - \dot{I}(t)\left(\frac{z-\eta(t)/2}{c}\right)}{1-\dot{\eta}(t)/2c}, \tag{60}$$

or else, once the algebra has been reduced to its simplest terms,

$$I(t)\dot{\eta}(t) + \dot{I}(t)\eta(t) \approx 0, \tag{61}$$

stated again, as was the case with (59), to lowest order in $1/c$. But this is merely to say that the product $I(t)\eta(t)$ is approximately constant, allowing Eq. (7) to remain in force (and this even in the face of an insistence that Eq. (57) be honored *"in the small"*).

Analogous considerations affect the electric field from (52). One emulates there the magnetic transition from (51) to (56) by writing

$$\begin{aligned} E(z,t) \;=\; & \frac{c\mu_0}{2}\left[U_+\left(t - \frac{z-a/2}{c}\right)\left\{I_0 - \frac{I(t_{*--})}{1-\dot{\eta}(t_{*--})/2c}\right\}\right. \\ & \left. -U_+\left(t - \frac{z+a/2}{c}\right)\left\{I_0 - \frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c}\right\}\right] \end{aligned} \tag{62}$$

so that, in particular, with a retrospective glance at (56),

$$E(z,t) = -cB(z,t) \tag{63}$$

as befits a planar field, null or otherwise, freely propagating along the direction of positive $z$.[14] And, happily, the magnetic null-field apparatus embodied in Eqs. (57)–(61) serves here just as well to ensure that the exterior electric field be likewise brought to zero.

### 2.4. *Bona Fide* Cavity Field: $|\mathbf{z}| < \boldsymbol{\eta}(\mathbf{t})/2$

The considerations foregoing of the putative exterior field, null though it may be, was no idle frivolity, simply because the physically accepted absence of penetration into ideal, perfectly conducting metal does not occur merely on the strength of fiat. On the contrary, such absence occurs only as a result of total cancellation among all available fields, those initially present and subsequently propagating outward, and those contributed by *bona fide* radiating current sources. An active demand for this very cancellation imposed upon these sources offers a convenient stepping-stone toward their determination, after the fashion of Eqs. (57)–(58) or else their simplified version (7). With this knowledge regarding the plate currents now taken for granted, we seek next to adapt the general field representations (51) and (52), unrestricted as they are with regard to observation plane location $z$, to the *bona fide* cavity interior, expanding in accordance with $|z| < \eta(t)/2$. And, as was the case in the discussion just now concluded, we can, by virtue of field symmetry/antisymmtry, limit attention to merely the positive values of $z$, $0 < z < \eta(t)/2$.

---

[13]Of course, while we do require that both $\eta(t)/2c$ and $\dot{\eta}(t)/2c$ remain small, this evidently imposes no practical restriction. One may also note parenthetically that our somewhat more heuristic QS/MBC treatment likewise insists upon the smallness of these two ratios, the second through an explicit mandate and the first implicitly, inasmuch as signal back and forth bounce averaging cannot really be accorded any credence in the event that plate-to-plate transit time $\approx 2\eta(t)/c$ grows excessively large. For that reason, our retention of the QS/MBC framework even when $\eta \to \infty$ (*cf.* Eq. (18) *et seq*) may, in a sense, be regarded as a lapse in modeling fidelity, a physical blemish which should nevertheless be taken in stride on practical grounds. Indeed, this entire *physical* caveat has already been hinted at within the introduction, wherein was suggested the desirability of having $\tau \gg \eta(t)/c$.

[14]The disposition of signs in (63) assures a Poynting vector that lies along the positive $z$-axis.

Operative now is a different mix of retarded times $t_{*\sigma_1\sigma_2}$ than that appropriate to the exterior field representations (56) and (62). We find

$$B(z,t) = -\frac{\mu_0 I_0}{2}\left\{1 - U_+\left(t + \frac{z-a/2}{c}\right) + U_+\left(t - \frac{z-a/2}{c}\right) - U_+\left(t - \frac{z+a/2}{c}\right)\right\} - \frac{\mu_0}{2}\left\{\frac{I(t_{*-+})}{1+\dot{\eta}(t_{*-+})/2c} + \frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c}\right\} \tag{64}$$

and

$$E(z,t) = -\frac{c\mu_0 I_0}{2}\left\{1 - U_+\left(t + \frac{z-a/2}{c}\right) - U_+\left(t - \frac{z-a/2}{c}\right) + U_+\left(t - \frac{z+a/2}{c}\right)\right\} - \frac{c\mu_0}{2}\left\{\frac{I(t_{*-+})}{1+\dot{\eta}(t_{*-+})/2c} - \frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c}\right\}. \tag{65}$$

In Eqs. (64) and (65), the terms proportional to initial plate current $I_0$ and involving several Heaviside steps $U_+$ convey a process wherein the field present at $t = 0$ undergoes a progressive dismantling, with due attention to retarded propagation, by signals in sign opposition emitted from both plates at the moment when disintegration begins. In conjunction with this there is set in motion, so to speak, a radiative field restoration attributable to time-varying currents $I(t_{*(\pm)(\mp)})$, both of their contributions being restricted as to their participation in the manner expressed following Eq. (51).[15]

Of course, any given station $z > a/2$ qualifies as a legitimate cavity point only when $t > t_<$, with $t_<$ gotten from $z = \eta(t_<)/2$. In particular, as one readily verifies, an adherence to this space/time constraint gives free rein to radiative participation on the part of the term proportional to $I(t_{*-+})$, which embodies radiation emitted from the upper plate, properly receding beyond $z$ when $t > t_<$.[16] By contrast, the radiation emanating from the lower plate, with current $I(t_{*+-})$, streams by uninterruptedly after a time delay $(a/2 + z)/c$.

As is easily seen, the terms proportional to $I_0$ in both (64) and (65) vanish following this lower-plate time delay $(a/2 + z)/c$. The field structures then default into

$$B(z,t) = -\frac{\mu_0}{2}\left\{\frac{I(t_{*-+})}{1+\dot{\eta}(t_{*-+})/2c} + \frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c}\right\} \tag{66}$$

and

$$E(z,t) = -\frac{c\mu_0}{2}\left\{\frac{I(t_{*-+})}{1+\dot{\eta}(t_{*-+})/2c} - \frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c}\right\}, \tag{67}$$

subject to the participation caveats as just now set out. And then, after the passage of several plate-to-plate bounce intervals on the order of $a/c$ past $t_<$, we can seek yet again to invoke approximations of the type illustrated in Eqs. (59) and (60). Thus

$$\frac{I(t_{*-+})}{1+\dot{\eta}(t_{*-+})/2c} \approx \frac{I(t) + \dot{I}(t)\left(\frac{z-\eta(t)/2}{c}\right)}{1+\dot{\eta}(t)/2c} \tag{68}$$

and

$$\frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c} \approx \frac{I(t) - \dot{I}(t)\left(\frac{z+\eta(t)/2}{c}\right)}{1+\dot{\eta}(t)/2c}, \tag{69}$$

with

$$B(z,t) \approx -\mu_0\left\{\frac{I(t) - \dot{I}(t)\eta(t)/2c}{1+\dot{\eta}(t)/2c}\right\} \approx -\mu_0 I(t) \tag{70}$$

and

$$E(z,t) \approx -z\mu_0\left\{\frac{\dot{I}(t)}{1+\dot{\eta}(t)/2c}\right\} \approx -z\mu_0\dot{I}(t) \tag{71}$$

[15] In Eqs. (56) and (62) we have found it opportune to give explicit voice to such restrictions through the medium of Heaviside unit steps $U_+$ with suitable arguments.

[16] For $0 < z < a/2$, the field radiated by $I(t_{*-+})$ enjoys a steady presence beyond its onset at time $(a/2 - z)/c$.

as the corresponding demotions for Eqs. (66) and (67). In conjunction with Eq. (61), we have now recovered the full trio (4), (6), and (7) earlier suggested on the basis of field heuristics which embrace from the outset a coarse-grained evolution in time.

### 2.5. A Postscript on Plate Dynamics

As Eqs. (68) and (69) suggest, it is the second, $t_{*+-}$ term in both field representations (66) and (67) which conveys the radiation propagating upward from the lower, receding plate at $z = -\eta(t)/2$. And so, if we continue to adhere to the premise that any notion of a magnetic self-force must be discounted, we retain on the right in (66) only its second term and hence arrive at the modified dynamical equation

$$\frac{d}{dt}\left\{\frac{\rho\dot{\eta}(t)}{2\sqrt{1-(\dot{\eta}(t)/2c)^2}}\right\} \approx \frac{\rho\ddot{\eta}(t)}{2} = \frac{\mu_0 I(t)}{2}\left\{\frac{I(t_{*+-})}{1+\dot{\eta}(t_{*+-})/2c}\right\} \tag{72}$$

with $t_{*+-}$ gotten from

$$t_{*+-} = t - \frac{\eta(t)+\eta(t_{*+-})}{2c}. \tag{73}$$

And if, sufficiently emboldened by our previous work, we proceed next to sweep aside *en masse* all relativistic corrections in both (69) and (73), we duly reduce (72) into

$$\frac{\rho\ddot{\eta}(t)}{2} \approx \frac{\mu_0 I(t)^2}{2}, \tag{74}$$

which is nothing other than (8).

### 2.6. Some Speculations on the Possible Presence of a Longitudinal, Electrostatic Field Component

In addition to the transverse electric component

$$\mathbf{E}(z,t) = E(z,t)\hat{\mathbf{e}}_y \tag{75}$$

as introduced in (27), one should in principle entertain the possibility of there being also a longitudinal

$$\mathbf{E}_{\|}^{\rm in}(z,t) = E_{\|}^{\rm in}(z,t)\hat{\mathbf{e}}_z, \tag{76}$$

essentially electrostatic one. To the extent that it may be deemed to be really operative at all, this component, on the one hand, is decoupled from our basic relations (32) and (33) while, on the other, the mutual influence which it is able to exert between the plates is such as to *counteract* the magnetic pressure from either (8) or (72). Its rôle, in fact, is identical to that of the electrostatic field in a parallel-plate capacitor, whose presence engenders a mutual attraction. Indeed, if we write $\pm\varsigma(t)$ for the electric charge density upon the plates at $z = \pm\eta(t)/2$, then we find

$$\frac{1}{c^2}\frac{\partial E_{\|}^{\rm in}(z,t)}{\partial t} + \frac{\mu_0\varsigma(t)\dot{\eta}(t)}{2}\left\{\delta\left(z-\eta(t)/2\right)+\delta\left(z+\eta(t)/2\right)\right\} = 0 \tag{77}$$

as the counterpart of (32). And then, inasmuch as the equation of charge continuity in either plate insists upon $\dot{\varsigma}(t) = 0$, this latter amounts to

$$E_{\|}^{\rm in}(\pm z,t) = -\frac{\varsigma_0}{\epsilon_0}\left\{U_+\left(z+\eta(t)/2\right) - U_+\left(z-\eta(t)/2\right)\right\}, \tag{78}$$

a classical capacitor field, beginning/ending on charges $\pm\varsigma_0$ at $z = \pm\eta(t)/2$ and leading to a mutual attraction in opposition to the magnetic repulsion from (8), (72), and (74). Such attraction has the curious feature of remaining undiminished no matter how large $\eta(t)$ may ultimately become. Be that as it may, the underlying message of (78) would seem to be that equations (8) and (72) should more properly be viewed as providing *conservative* overestimates of expansion severity. Moreover, we behold in (78) a null field component exterior to the cavity, $|z| > \eta(t)/2$, exhibited on the basis of a truly meager input of mathematics, and, in particular, no trace of signal retardation whatsoever.

Nevertheless, the fact remains that longitudinal electric component (78) injects a vexing aura of physical indeterminacy. We can evade it, perhaps, by reminding ourselves that our glib, implicit assumption, to the effect that magnet plate metal extends upward and downward without bound, is clearly a fiction. And, if that fiction be removed,[17] with magnet faces assigned a depth $d$, then we could just as well repeat the above argument with *opposed* surface charge densities $\mp\varsigma_0$ residing at $z = \pm(\eta(t)/2 + d)$ and the electric component from (78) correspondingly *canceled* throughout cavity interior $|z| < \eta(t)/2$ by

$$E_{\|}^{\text{out}}(\pm z, t) = \frac{\varsigma_0}{\epsilon_0} \left\{ U_+ \left( z + \{\eta(t)/2 + d\} \right) - U_+ \left( z - \{\eta(t)/2 + d\} \right) \right\}. \tag{79}$$

The additional charge densities $\mp\varsigma_0$ on outer magnet faces are taken in opposition to their inner-face counterparts if only because, in its quiescent state prior to disintegration, each plate within the Y-Y apparatus may reasonably be presumed to be charge-neutral. While certainly plausible, all of this seems tentative and provisional, and should be viewed *cum grano salis*, at the very least. In particular, we have simply exiled to infinity in the plane of $(x, y)$ any credible circuit mechanisms for maintaining charge neutrality, be the magnet plates bounded or not as to their thickness.

## 3. EPILOGUE

In principle, we should go on to tease out the fine structure in the temporal evolution of current $I(t)$, cavity field components (66) and (67), and the dynamical Equation (72). On the other hand, it is clear that any such program necessarily involves plate separation $\eta(t)$ in a nonlinear way, and is thus best consigned to a numerical inquiry complete unto itself. And in any event, a fine-structure study of this sort, while interesting perhaps on academic grounds, would clearly have little practical relevance.

At the same time we have succeeded, we would hope, in demonstrating that, one the one hand, an intuitive, coarse-grained QS/MBC framework can, with but a modest expenditure of analytic effort, delineate most of the field features to be encountered within an expanding cavity, while, on the other, that the QS/MBC program is in fact underwritten as a controlled, nonrelativistic limit of a more robust, fully Maxwell-compliant field formulation. This limit amounts indeed to a blurring of focus as to the fine features of temporal evolution.

In order to make tangible progress, we have of course pruned away much realism by adopting what, at first sight, must seem to be an overly simplistic, strictly one-dimensional model. But even here are surely to be found a few nuggets of genuine physical and mathematical interest. In particular, we have been able to show that the null exterior field condition, repeatedly exploited by us in several other problems, offers a most fruitful alternative to the traditional technique of tangential component matching along surfaces of material discontinuity. All in all we would hope that, someday, this work may encourage similar efforts addressed to more complex cavity geometries. Spherical and (infinitely long) cylindrical cavities with radial dependence only spring to mind at once as the next possible candidates.

[17]We shan't attempt to revise at this late juncture any of the foregoing, transverse field developments, all of which were premised on the idealized existence of magnet plates having infinite depth. Revisions of this sort, similar in spirit to that about to be proposed with Eq. (79) for the longitudinal electrostatic field, would bring into null field competition an additional pair of up-traveling/down-traveling field systems, considerably complicating the arrival/departure temporal bookkeeping at any intended null field location. Furthermore, the option so offered for radiation escape to the exterior of the cavity would, on physical grounds, presumably serve to soften the interior magnetic pressure heretofore considered. But to explore all of these effects, even if one were to continue to take them seriously, would require an additional essay unto itself. It would seem that surface current determination on the basis of a null field constraint has already been adequately illustrated in the magnet model as it now stands, with its plate material extending indefinitely up and down along coördinate $z$.

# Numerical Simulation of EM Environment and Human Exposure When Using RFID Devices

**Aránzazu Sanchis**[1], **Javier Espinosa-García**[2], **and Agustín Martín**[2]

[1]Centro Nacional de Sanidad Ambiental (ISCIII), Majadahonda, Madrid, Spain
[2]Instituto de Física Aplicada (CSIC), Madrid, Spain

**Abstract**— In this work, RFID readers and active tags, operating at 2.4 GHz and 5.8 GHz, are simulated to calculate the electromagnetic field distribution generated in their surroundings, considering the presence of the user's body and possible scattering obstacles, and including the analysis of the specific absorption rate (SAR) in the human models (which can be very close to the tag). The finite-difference time-domain (FDTD) method and the finite elements (FE) method are used. Results show that significant field level can be found in regions far from the tag-reader direction. This fact could constitute a risk because of the possible presence of general public and eventual eavesdroppers. Nevertheless, the calculated values of SAR are always below the basic restrictions.

## 1. INTRODUCTION

Radio frequency identification (RFID) systems have become a very important and raising researching area in the past few years. This wireless communication technology uses radio waves for automatic identification and information, and is being efficiently applied in a wide variety of activities, such as tracking and tracing materials, goods, or people, improving theft prevention, security, and medical care. Distinction is made between more or less high-performance systems (low-end through high-end), depending on the field of application and the task to be performed. Among the most highly growing uses are real time location systems (RTLS), which is currently being used in schools and hospitals and reach a range of up to 200 m, and the electronic passports issues in the last years by many countries, which can be read from more than 10 m away. Thus, it becomes essential that the implementation of RFID technology takes place under a legal framework that affords citizens effective safeguards for fundamental values, health, environment, data protection, privacy and security [1, 2]. Moreover, the fast development and the expanding applications of this electromagnetic (EM) field-producing technology is demanding further work to tackle potential associated risks in terms of privacy, security and health issues, specially in complex electromagnetic environments, where signal reflections and other related effects may arise in unintended ways in presence of metallic objects.

The RFID working principle is based on a signal sent by a reader — interrogator — via a radio frequency to a tag — transponder —, which receives the signal and reflects a unique coded signal back to the reader from certain distance range. The operating frequencies are generally defined in four ranges within the unlicensed part of the electromagnetic spectrum, known as ISM (industrial, scientific and medical): Low frequency (LF) in 125 and 134 kHz, high frequency (HF) in 13.56 MHz, ultra high frequency (UHF) in 868–915 MHz, and microwave in 2.45 and 5.8 GHz [3]. Normally, LF and HF tags are passive tags without battery, and are inductively powered through a coil in the reader and a tiny coil in the tag, while UHF and microwave tags are active tags with battery and on-chip tag, relying on backscattering to communicate. LF and HF tags may be read in a shorter radio — typically a few cm — whilst UHF and microwave tags may communicate through longer distances — up to 200 m —.

In this paper, the operation of RFID readers and tag antennas is numerically simulated in possible actual environments where a user is placed. The objective is to analyze how the presence of possible obstacles can influence the electric field levels in different regions around the RFID user, so that possible unexpected security risks could arise. In addition, health issues during the use of the RFID devices can also be considered by comparing the electric field values obtained in both situations with the limits established in the exposure guidelines (reference levels), and comparing the SAR values inside the human body model with basic restrictions. In Section 2, a description of the used methodology is made. We also describe the geometrical configuration which simulates an operating environment where a user, wearing a personnel tracking badge, is located in front of a metallic cabinet. This could be a common situation in an office, a school or a hospital. Examples of results are shown and discussed in Section 3. Finally, conclusions are presented.

## 2. METHODOLOGY AND MODELS

In order to carry out the analysis of the proposed problem, the finite-difference-time-domain (FDTD) method is used. The FDTD method is a well known computational technique in which the Maxwell differential equations are discretized by means of a finite differences scheme implemented in a mesh of cubic cells, named Yee cells, where the geometries under study are spatially approximated. The cell size must be small enough to permit accurate results at the highest frequency of interest, taking into account the effect of the different materials on the wavelength. A lattice of 10 cells per wavelength usually gives accurate results. Once the cell size is selected, the maximum time step is determined by the Courant stability condition. Then, absorbing boundary conditions are fixed at the limits of the space under study to avoid undesired reflections [4].

Among the frequency bands assigned to the RFID applications, we'll focus on the 2.45 and 5.8 GHz operating frequencies. Most of the antennas can be simulated as half wavelength dipole antennas. They have an omni-directional radiation pattern and represent a worst case for dosimetry studies. Nevertheless, as the operating frequency rises into the microwave region, the conformal structure and compact size are also main concerns within the design process. Thus, we have used $\lambda/2$ dipole antennas at 2.45 and 5.8 GHz, and have also simulated an actual miniature folded-slot antenna suitable to be included in a 5.8 GHz personnel tracking badge. The geometrical and electrical parameters of this antenna can be found in [5].

A FDTD problem space of $400 \times 260 \times 400$ cubic cells has been generated to include a human body model and a metallic plate simulating a perfectly conducting obstacle. The greater size of the cells is $5 \times 5 \times 5\,\mathrm{mm}^3$, imposed by the stability condition for 5.8 GHz, and an adaptive mesh with $0.3 \times 0.3 \times 0.3\,\mathrm{mm}^3$ cells has been used when needed to accurately simulate the folded-slot antenna. For simulating a free space situation, perfectly matched layer (PML) boundary conditions (8 layers) are introduced at the limits of the computational space. A vertical metallic flat plate, with a width of 1.9 m and a height of 1.5 m, is modelled at a distance of 80 cm in front of the user's chest, to simulate the existence of a metallic cabinet before the user of the RFID tag.

To simulate the user's body, a 3D high-resolution body mesh provided by REMCOM (State College, PA, USA) has been used. It also has $5\,\mathrm{mm}^3$ cubic cells, has been obtained from the Visible Human Project data in collaboration with the Hershey Medical Center (Hershey, PA, USA), and includes 23 different tissues. Their densities and dielectric properties at 5.8 GHz are obtained from Gabriel et al. [6]. The antennas are simulated in the middle of the chest — as a hanged badge —, at a distance of 1 mm from the body and 1.40 m above the floor. The maximum radiated power is 25 mW [7]. A schematic view of the problem configuration is shown in Figure 1, where the geometry of the folded-slot antenna can be appreciated in the inset.

## 3. RESULTS

The numerical computations have been made by using the widely used commercial program XFDTD — produced by REMCOM — which accurately implements the FDTD technique. Results obtained with this commercial program have previously been compared by the authors with those obtained by means of other numerical methods, with very good agreement [8]. In order to validate our simulations, we have compared the XFDTD results for the radiation patterns of the dipole antennas

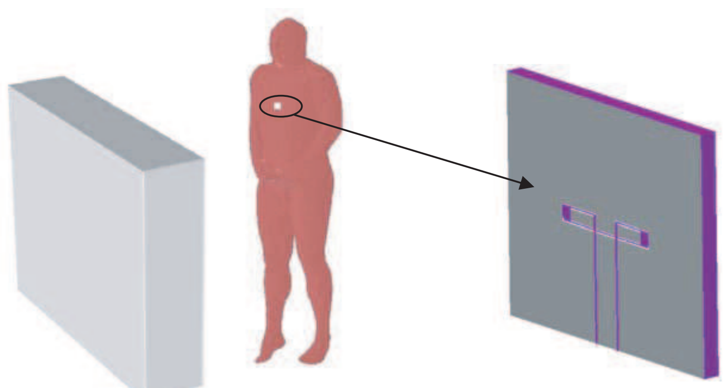


Figure 1: Squematic view of problem geometry.

with the results obtained after simulating the antennas using COMSOL, a well known simulating platform which is based in the finite-elements method. In this case, a tetrahedral mesh consisting of 193241 elements is employed to simulate the folded-slot antenna and a sphere to truncate the computational domain, using infinite elements available for COMSOL as boundary conditions at infinity. Calculations with both numerical techniques show a very good agreement.

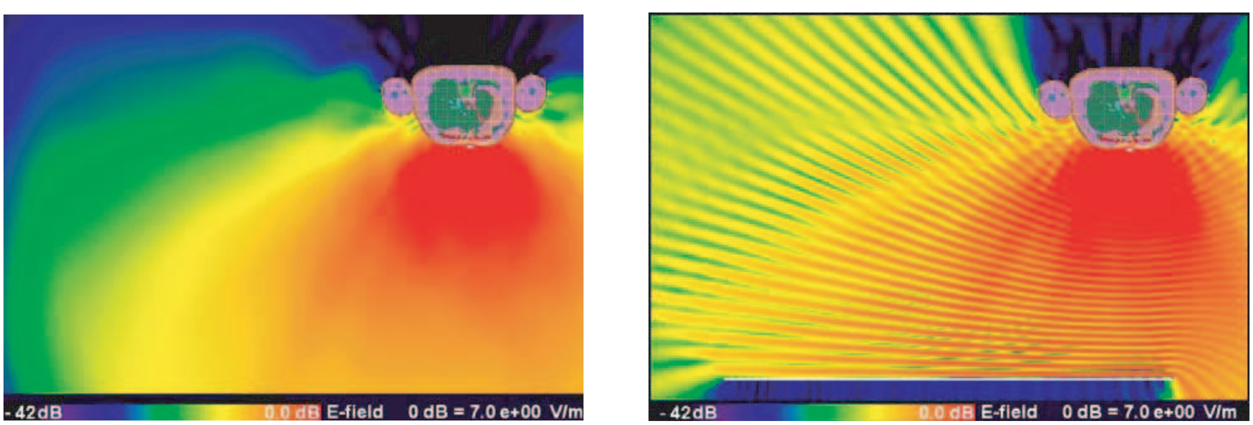
Firstly, values of the electric field, $E$, and SAR in the proximity of the user have been calculated for both studied frequencies when there are no obstacles in the surroundings, simulating a free space situation. Values for whole body averaged SAR ($\mathrm{SAR_{wb}}$) and local SAR averaged over 10 g of contiguous tissue ($\mathrm{SAR_{10g}}$) have been calculated. Then, the conducting flat plate is included in the model to simulate the existence of a metallic cabinet in front of the user of the RFID tag. The maximum memory required for the calculations is 1.81 GB, and the longer calculation time for a simulation, including SAR averaging, has been 5 h 33 min in a Pentium D 3.2 GHz processor with 2 GB of RAM memory.

As an example of the obtained results, contour plots of the calculated electric field generated by the tag antenna in the user's surroundings are shown in Figure 2. A comparison is made of the electric field $E$ obtained in both simulated configurations within the horizontal plane containing the antenna feed. As can be observed, the presence of the conducting obstacle clearly increases the field level in regions of the space where almost no direct field from the tag is found in free space situations (obviously, the antenna radiation pattern is greatly modified due to the presence of the user's body). This fact has to be taken into account when analyzing security risks during the operation of sensible identification devices.

The calculated values of $\mathrm{SAR_{wb}}$, so as the maximum $\mathrm{SAR_{10g}}$ values, in both configurations, are shown in Table 1. It can be appreciated that $\mathrm{SAR_{wb}}$ and maximum $\mathrm{SAR_{10g}}$ values are far below the corresponding basic restrictions for both geometries. Due to the large distance between the body and the obstacle, no difference is appreciated in the maximum averaged local SAR obtained in both situations. The whole body averaged SAR is only slightly higher when the obstacle is included, but the difference is insignificant.

Table 1: $\mathrm{SAR_{wb}}$ and max $\mathrm{SAR_{10g}}$ for the simulated exposure situations.

| | Free space | In presence of a metallic obstacle |
|---|---|---|
| $\mathrm{SAR_{wb}}$ (W/kg) | $2.985 \times 10^{-4}$ | $3.039 \times 10^{-4}$ |
| Max. $\mathrm{SAR_{10g}}$ (W/kg) | 0.748 | 0.748 |




(a) (b)

Figure 2: Contour plots of the calculated electric field $E$, generated by the active antenna of the badge tag within the horizontal plane containing the antenna feed (1.40 m above the floor) in the user's surroundings. (a) In a free space situation. (b) In presence of a metallic obstacle. Frequency: 5.8 GHz. Body-obstacle distance: 80 cm. Tag radiated power: 25 mW.

## 4. CONCLUSION

In this work, RFID readers and active tags, operating at 2.4 GHz and 5.8 GHz, are simulated to calculate the electromagnetic field distribution generated in their surroundings, considering the presence of possible scattering obstacles, and including the analysis of the SAR in human models which can be very close to the tag. The FDTD method has been used to simulate the antennas, using the FE method for validation. Results show that significant field level can be found in regions far from the tag-reader direction (which in an ideal case is expected to be that of the main lobes of the antennas). This fact could constitute a risk because of the possible presence of general public and eventual eavesdroppers. Nevertheless, the calculated values of SAR are always below the basic restrictions.

## ACKNOWLEDGMENT

This work has been supported by the Spanish Ministry of Science and Innovation, Projects SPY1416/09 and TEC2009-13964-C04-02.

## REFERENCES

1. European Commission, "Standarisation mandate to the european standarisation organisations CEN, CENELEC and ETSI in the field of information and communication technologies applied to radio frequency identification (RFID) and systems," M/436, December 2008.
2. IEC (International Electrotechnical Commission), "Evaluation of human exposure to electromagnetic fields from short range devices (SRDs) in various applications over the frequency range 0 GHz to −300 GHz — Part 1: Fields produced by devices used for electronic article surveillance, radio frequency identification and similar systems," IEC 62369-1, 1–149, Geneva, Switzerland, 2008.
3. IEE (International Electrical Engineers), "Radio frequency identification device technology (RFID) factile," The Institution of Electrical Engineers, 2005, available online at www.theiet.org/factfiles/it/rfid.cfm.
4. Taflove, A. and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method in Electromagnetics*, Artech House, Norwood, MA, 2000.
5. Chen, S.-Y. and P. Hsu, "CPW-fed folded-slot antenna for 5.8 GHz RFID tags," *Electron. Lett.*, Vol. 40, No. 24, 1516–1517, 2004.
6. Gabriel, C. and S. Gabriel, "Compilation of the dielectric properties of body tissues at RF and microwave frequencies," Technical Report AL/OE-TR-1996-0037, Physics Department, King's College, London, 1996, available at http://niremf.ifac.cnr.it/tissprop/.
7. Finkenzeller, K., *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification*, John Wiley & Sons, Ltd., 2003.
8. Martínez-Búrdalo, M., L. Nonídez, A. Martín, and R. Villar, "Near field time-domain physical optics and FDTD method for safety assessment near a base-station antenna," *Microw. Opt. Technol. Lett.*, Vol. 39, No. 5, 393–395, 2003.

# Distributed Transverse Orientation (DiTO) in Maxwell's Equations

**J. E. Carroll**
Centre for Advanced Photonics and Electronics, University of Cambridge, CB3 0FA, UK

**Abstract**— A remarkable, but obscure, property of classical Maxwell's equations is reviewed showing how to change rotational orientations about an axis $O_z$ equally for all transverse vector fields, $\mathbf{E_T}$ and $\mathbf{H_T}$, along with the transverse differential vector operator but *without* altering any axial fields. The theory may have implications for: (a) connecting classical and random electrodynamics; (b) the formation of plane wave-packets with finite energy; (c) a classical explanation of Schrödinger number states ($N + 1/2$); (d) reduced dispersion communications.

## 1. INTRODUCTION

The solutions of Maxwell's equations (MEs) can have additional vector solutions [1], without the sophistications that can encompass full vector forms [2, 3]. The theory of these lesser known solutions is reviewed in Section 2 with some potential applications outlined in Section 3. The general concept is that Transverse Electric (TE) and Transverse Magnetic (TM) waves, moving along an arbitrary axis $O_z$, can have their transverse vector fields, $\mathbf{E_T}$ and $\mathbf{H_T}$, along with the transverse gradient operator $\boldsymbol{\nabla}_\mathbf{T}$ rotated by arbitrary angles about the axis $O_z$ while leaving axial fields $E_z$ or $H_z$ unaltered. Although materials can be included, simplifications are made here by limiting the discussion to free space.

## 2. DISTRIBUTED TRANSVERSE ORIENTATIONS

TE and TM modes are 'dual' modes [4] so that from one set of calculations for TE fields one may find the results for TM fields. TE fields, with 'driving' field $H_z$ propagating along $O_z$ (with unit vector $\mathbf{n}$), can have their transverse vector fields $\mathbf{F_T}$ (i.e., $\mathbf{E_T}$ or $\mathbf{H_T}$ and gradient operator $\boldsymbol{\nabla}_\mathbf{T}$ written as:

$$\boldsymbol{\nabla}_\mathbf{T} = \begin{bmatrix} \partial_x \\ \partial_y \end{bmatrix}; \quad \mathbf{F_T} = \begin{bmatrix} F_x \\ F_y \end{bmatrix} \quad \text{with } (\mathbf{n} \times \mathbf{F_T}) \leftrightarrow \mathbf{j}\mathbf{F_T} \quad \text{where } \mathbf{j} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{1}$$

It can be shown that the TE equations from MEs using this matrix form are given by:

$$\boldsymbol{\nabla}^{/}_\mathbf{T}\mathbf{H_T} = -\partial_z H_z; \quad \boldsymbol{\nabla}^{/}_\mathbf{T}(\mathbf{j}Y_o\mathbf{E_T}) = \partial_{ct} H_z; \quad [\text{where } \partial_{ct} = (1/c)\partial/\partial t]. \tag{2}$$

$$\partial_z(\mathbf{H_T}) + \partial_{ct}(\mathbf{j}Y_o\mathbf{E_T}) = \boldsymbol{\nabla}_\mathbf{T} H_z; \quad Y_o = \sqrt{(\varepsilon_o/\mu_o)} = 1/Z_o. \tag{3}$$

Here $^{/}$ denotes transposition of rows and columns without conjugation. Given transverse modes $H_z(x, y)$ varying as $\exp[\mathrm{i}(kz - wt)]$, transverse field patterns are determined here from:

$$\boldsymbol{\nabla}^{/}_\mathbf{T}\boldsymbol{\nabla}_\mathbf{T} H_z(x, y) = -\kappa^2 H_z(x, y) : k^2 + \kappa^2 = (w/c)^2. \tag{4}$$

The 'cut-off' wave-vector $\kappa$ is taken here as a measure of diffraction. If $\kappa^2$ tends to zero, the waves tend to plane waves. Direct substitution shows permitted solutions of Eq. (2) as:

$$\mathbf{H_T} = (ik/\kappa^2)\boldsymbol{\nabla}_\mathbf{T} H_z; \quad (\mathbf{j}Y_o\mathbf{E_T}) = (iw/c\kappa^2)\boldsymbol{\nabla}_\mathbf{T} H_z. \tag{5}$$

The dual TM fields are found by replacing $(\mathbf{H_T},\ \mathbf{j}Y_o\mathbf{E_T}, H_z)$ with $(\mathbf{E_T}, -\mathbf{j}Z_o\mathbf{H_T}, E_z)$.

Just as a 90° change of *temporal* phase can be given by an operator i, with $\exp(\mathrm{i}q)$ giving a phase change $q$ radians, so here a 90° change of *rotational* orientation about $O_z$ is represented by $\mathbf{j}$ giving a unitary operator $\mathbf{R} = \exp(\mathbf{j}Q)$ rotating transverse fields and operator $\boldsymbol{\nabla}_\mathbf{T}$ through $Q$ radians:

$$\mathbf{R}\mathbf{F_T} = \cos Q\mathbf{F_T} + \sin Q\mathbf{j}\mathbf{F_T}; \ \mathbf{R} \ = (\cos Q)\boldsymbol{\nabla}_\mathbf{T} + (\sin Q)\mathbf{j}\boldsymbol{\nabla}_\mathbf{T}. \tag{6}$$

Here $\mathbf{j}\mathbf{F_T}$ is $\mathbf{F_T}$ rotated through 90° about $O_z$. Now for *all* $\mathbf{R}$, even when $Q$ is a function $Q(x, y)$:

$$\mathbf{R}^{/}\ \mathbf{R} \ = 1; \quad (\mathbf{R}\boldsymbol{\nabla}_\mathbf{T})^{/}\mathbf{R}\ \mathbf{H_T}) = \boldsymbol{\nabla}^{/}_\mathbf{T}(\mathbf{R}^{/}\mathbf{R})\ \mathbf{H_T} = \boldsymbol{\nabla}^{/}_\mathbf{T}\mathbf{H_T}. \tag{7}$$

Moreover, $\mathbf{j}\ \mathbf{R} = \mathbf{R}\ \mathbf{j}$ so that

$$(\mathbf{R}\boldsymbol{\nabla}_\mathbf{T})^{/}(\mathbf{j}Y_o\mathbf{R}\ \mathbf{E}_\mathbf{T}) = \boldsymbol{\nabla}^{/}_\mathbf{T}(\mathbf{j}Y_o\mathbf{E}_\mathbf{T}). \tag{8}$$

Provided that $Q$ is not a function of $z$ or $t$ then Eqs. (2) and (3) can be written as:

$$(\mathbf{R}\boldsymbol{\nabla}_\mathbf{T})^{/}(\mathbf{R}\mathbf{H}_\mathbf{T}) = -\partial_z H_z; \quad (\mathbf{R}\boldsymbol{\nabla}_\mathbf{T})^{/}(\mathbf{j}Y_0\mathbf{R}\mathbf{E}_\mathbf{T}) = \partial_{ct} H_z; \tag{9}$$

$$\partial_z(\mathbf{R}\mathbf{H}_\mathbf{T}) + \partial_{ct}(\mathbf{j}Y_0\mathbf{R}\mathbf{E}_\mathbf{T}) = \mathbf{R}\boldsymbol{\nabla}_\mathbf{T} H_z. \tag{10}$$

Arbitrary rotational patterns $Q(x, y)$ may then be applied to TE (or TM) modes replacing $\mathbf{E}_\mathbf{T}, \mathbf{H}_\mathbf{T}$ and $\boldsymbol{\nabla}_\mathbf{T}$ with $\mathbf{R}\mathbf{E}_\mathbf{T}$, $\mathbf{R}\mathbf{H}_\mathbf{T}$ and $\mathbf{R}\boldsymbol{\nabla}_\mathbf{T}$ without changing any modal patterns determined by $H_z(x, y)$ [or $E_z(x, y)$]. In the context of previous work such changes were called Distributed Spin Rotations (DSRs) [1] but are not always *spin* so are called here Distributed Transverse Orientations (DiTOs).

Of greater interest is the situation when $Q$ is a function of space-time: $Q(x, y, z, ct)$. Eqs. (7) and (8) still hold so that Eq. (9) also holds but Eq. (10) needs rewriting:

$$\partial_z(\mathbf{R}\mathbf{H}_\mathbf{T}) + \partial_{ct}(\mathbf{j}Y_0\mathbf{R}\mathbf{E}_\mathbf{T}) - \left[(\partial_z\mathbf{R})\mathbf{R}^{-1}\right](\mathbf{R}\mathbf{H}_\mathbf{T}) - \left\{[\partial_{ct}\mathbf{R}]\mathbf{R}^{-1}\right\}(\mathbf{j}Y_0\mathbf{R}\mathbf{E}_\mathbf{T}) = \mathbf{R}\boldsymbol{\nabla}_\mathbf{T} H_z. \tag{11}$$

From Eq. (11) it is seen that Eq. (10) is still valid provided that one can arrange for:

$$\left[(\partial_z\mathbf{R})\mathbf{R}^{-1}\right](\mathbf{R}\mathbf{H}_\mathbf{T}) + \left\{[\partial_{ct}\mathbf{R}]\mathbf{R}^{-1}\right\}(\mathbf{j}Y_0\mathbf{R}\mathbf{E}_\mathbf{T}) = 0. \tag{12}$$

Now set $\mathbf{R} = \exp[\mathbf{j}(Kz - Wt)]$, then Eq. (12) can be simplified to read

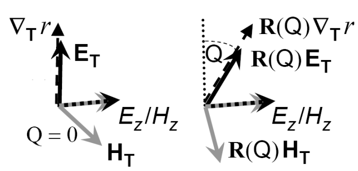
$$K\mathbf{H}_\mathbf{T} - (W/c)(\mathbf{j}Y_0\mathbf{E}_\mathbf{T}) = 0. \tag{13}$$

Using Eq. (5) with frequency $w$ and wave-vector $k$, it follows that Eq. (10) is still valid provided:

$$Kk - (W/c)(w/c) = 0. \tag{14}$$

From Eq. (4), the underlying classical mode has a *modal* group velocity $(v_g/c) = [(dw/c)/dk] = [k/(\omega/c)]$. The solution offered by Eqs. (12)–(14) requires a helical motion at frequency $W$ (independent of $w$) propagating at $W/Kc = [k/(\omega/c)] = (v_g/c)$, the modal *group* velocity of the underlying mode. This is unlike the vortex modes propagating at the *phase* velocity with a single modal frequency [5].

Figure 1(left) shows a key difference between a conventional rotation and a DiTO where the gradient differential operator is rotated *along with* all transverse field vectors. Figure 1(right) uses fluid flow as an analogy to partly illustrate the physics. Inviscid fluid with density fluctuations is interrupted by a rotating perforated disc. Fluid emerging from any one aperture has a helical path impressed on it but independent density fluctuations can be preserved in spite of the disc's rotation.



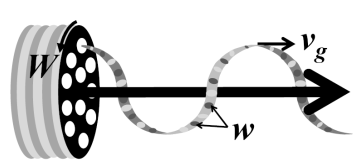




Figure 1: DiTO action. *Left* (a) The right handed sets of fields have no rotations. *Left* (b) $\mathbf{R}\mathbf{E}_\mathbf{T}$, $\mathbf{R}\mathbf{H}_\mathbf{T}$ and $\mathbf{R}\boldsymbol{\nabla}_\mathbf{T} r$ are all rotated through an angle $Q$ where $\boldsymbol{\nabla}_\mathbf{T} r$ with $r = [(x^2 + y^2)]^{1/2}$ as a reference. *Right.* A perforated disc rotating at frequency $W$ interrupts inviscid fluid flowing with velocity $v_g$. Helical path is impressed on the inviscid fluid preserving density fluctuations at $w$ independent of $W$.

## 3. POSSIBLE APPLICATIONS

### 3.1. Stochastic Electrodynamics

Stochastic or random electrodynamics [e.g., 6–8] assumes a background of random radiation with an energy per mode of $1/2\hbar w$ and from that assumption shows how features of quantum theory can be derived. The present analysis, using rotations and transverse fields, has a contribution to make by giving answers to the question 'What are the sources of randomness?'. The internal orientation change given by $Q(x,y)$ in Section 2 appears to be allowed to have a random structure by being totally undefined and of course the axis $O_z$ can be chosen randomly. The helical frequency $W$ can also be randomly chosen. These are then potential random 'sources' for stochastic electrodynamics. Of additional note, is that if one reverses $t$ and $\mathbf{j}$ then MEs are seen to be unaltered in Eqs. (2) and (3). Moreover, if one simultaneously reverses $w$ then $\exp[\mathrm{i}(kz-(-w)(-t)] = \exp[\mathrm{i}(kz-wt)]$. These reversals give a new type of wave that travels together with the conventional causal or 'reference' wave ('r' waves) but yet has time reversal. The source of these waves must come from the future. They are not the same as waves with advanced Green functions [9, 10] but more akin to the 'handshake' of $\Psi^*$ introduced by Cramer [9] but without complex conjugation. Elsewhere this class of solutions to classical MEs are referred to as 'adjoint' waves [1, 13] or 'a' waves also used in Section 3.5. The 'a' waves will have counter rotating helical motion compared to 'r' waves. Here is another potential cause for randomness in electrodynamic fields showing how DiTOs offer new insights.

### 3.2. Change of Amplitude by Interactions of DiTOs with Circularly Polarised Fields

One of the strangest features of DiTOs is their interaction with circularly polarised fields. Again the matrix formation is of considerable help in understanding this interaction. Given a field $\mathbf{F_T}$, then is it possible to project out the components of circular polarisation by using operators $\mathbf{P}_+$ and $\mathbf{P}_-$ where

$$\mathbf{P}_\pm = 1/2[1 \pm \mathrm{i}\mathbf{j}] = \begin{bmatrix} 1/2 & \mp 1/2\mathrm{i} \\ \pm 1/2\mathrm{i} & 1/2 \end{bmatrix}; \quad \mathbf{P}_+\mathbf{P}_+ = \mathbf{P}_+; \quad \mathbf{P}_-\mathbf{P}_- = \mathbf{P}_-; \quad \mathbf{P}_+\mathbf{P}_- = \mathbf{P}_-\mathbf{P}_+ = 0. \tag{15}$$

Consider these operators acting on a sinusoidal field $\mathbf{F_T}$:

$$\mathbf{F_T} = \exp(\mathrm{i}\Theta)\begin{bmatrix} F_x \\ F_y \end{bmatrix}; \quad \Theta = (kz - wt) \text{ noting that } \mathbf{P}_\pm\mathbf{j} = \mp\mathrm{i}\ \mathbf{P}_\pm. \tag{16}$$

Polarisation operators change $\mathbf{F_T}$ into having rotating equal $x/y$ amplitudes in phase-quadrature:

$$\mathbf{P}_\pm\mathbf{F_T} = \exp(\mp\mathbf{j}\Theta)\begin{bmatrix} (1/2F_x \mp 1/2\mathrm{i}F_y) \\ \pm\mathrm{i}(1/2F_x \mp 1/2\mathrm{i}F_y) \end{bmatrix}. \tag{17}$$

Consider a complex DiTO, $\mathbf{R}_\mathrm{A} = \exp(\mathrm{i}\mathbf{j}A)$ noting that transposition does not mean conjugation of i so that $\mathbf{R}'_\mathrm{A}\mathbf{R}_\mathrm{A} = 1$ : *essential* for any DiTO in Section 2. Now combine circular polarisation with this DiTO operation $\mathbf{R}_\mathrm{A}$. Noting $\mathbf{P}_\pm\mathbf{j} = \mp\mathrm{i}\ \mathbf{P}_\pm$; $\mathbf{P}_\pm\mathbf{R}_\mathrm{A} = \exp(\pm A)\mathbf{P}_\pm$ one sees that temporal and rotational phases interact to change the local *amplitudes* of circularly polarised fields:

$$\mathbf{P}_\pm\mathbf{R}_\mathrm{A}\mathbf{F_T} = \exp(\pm A)\exp(\mp\mathbf{j}\Theta)\begin{bmatrix} (1/2F_x \mp 1/2\mathrm{i}F_y) \\ \pm\mathrm{i}(1/2F_x \mp 1/2\mathrm{i}F_y) \end{bmatrix}. \tag{18}$$

Of even greater interest is when $A = 1/2(\xi x^2 + \zeta y^2))$ because that allows one to select an appropriate polarisation and an appropriate DiTO to describe convergent Gaussian fields, e.g.,

$$\mathbf{P}_-\mathbf{R}_\mathrm{A}\mathbf{F_T} = \exp[-1/2(\xi x^2 + \zeta y^2))]\exp(\mathbf{j}\Theta)\begin{bmatrix} (1/2F_x \mp 1/2\mathrm{i}F_y) \\ \pm\mathrm{i}(1/2F_x \mp 1/2\mathrm{i}F_y) \end{bmatrix}. \tag{19}$$

Gaussian patterns for transverse fields now allow convergence of transverse energy integrals.

### 3.3. Gaussian Circularly Polarised Plane-wave Packet Formation Using DiTOs

A classic wave-packet consists of plane waves with two different frequencies of equal amplitude:

$$E_x = E_{xo}\exp\mathrm{i}(k_r x - w_r t) + E_{xo}\exp\mathrm{i}(k_a x - w_a t) = 2E_{xo}\cos(k_D x - w_D t)\exp\mathrm{i}(k_S x - w_\Sigma t). \tag{20}$$

where $k_{S/D} = \frac{1}{2}(k_r + / - k_a)$; $w_{S/D}\frac{1}{2}(w_r + / - w_a)$. Creating a circularly polarised plane wave packet requires additional $E_y$ fields in the appropriate phase quadrature as in Eq. (17). The $H_x$ and $H_y$ fields must follow, also in phase quadrature, as required by the classic plane wave solutions for MEs. Here diffraction is zero and the fields are uniform over the whole space so that such circularly polarised plane-wave packets fail to have convergent energy integrals [4]. It is possible of course to consider specific paraxial approximations [12] and create Gaussian profiles but a different general approach is given by the DiTO of Eq. (19) that allows in principle a circularly polarised plane-wave-packet to have a convergent arbitrary Gausian profile. Because the centre of a plane wave is undefined, this Gaussian profile can be placed anywhere suggesting that plane waves might be made up from such random packets. However no 'number states' or internal structure can be suggested for this.

### 3.4. TM+TE Wave-packet Formation Using DiTOs

The next stage considers the equivalent of Section 3.3 for TM+TE wave-packets because it is argued that circularly polarised plane (TEM) waves are more correctly considered as a limiting case of TM+TE packets where diffraction in all axial fields tend to zero. More detailed accounts have been submitted elsewhere [1, 13]. TE and TM fields are combined in phase quadrature to give circularly polarised fields. This makes it possible to form collocated nodes in the $E$- and $H$-axial fields similar to Eq. (20). DiTOs are used now, as outlined in Eq. (18) to equalise the *transverse* fields to form collocated axial-field nodes similar to Eq. (20). Then, using another DiTO to create the Gaussian profile of Eq. (19), one is able to create a circularly polarised TE+TM wavepacket where all axial and transverse fields have collocated nodes that can trap the electromagnetic energy. The minimum duration of this wavepacket is determined by $2\pi/|w_r - w_a|$. As the axial fields are allowed to tend to zero, the TE+TM wavepacket tends to a TEM wave-packet that may appear to be the same as the TEM wavepacket in Section 3.3 but can have hidden structure as discussed in the next section.

### 3.5. Number States in TE+TM Packets Using Helical DiTOs

The TE+TM wave-packet discussed above can now be given internal structure using helical motion imparted by DiTOs (Section 2). About every local axis, the transverse gradient operator and field vectors are rotated helically at $W_r$ rotating clockwise (or anti-clockwise without restriction) with 'r' waves (frequency $w_r$) while the 'a' waves (with $t$, $\mathbf{j}$ and frequency $w_a$ all reversed) rotate at $W_a$ in the opposite direction (Figure 2(a)). These two helically changed waves, with the same circular polarisation, are superimposed to produce additional nodes in the transverse fields [13]. By appropriate choice of the helical frequencies in terms of 'multiples' of $(w_r - w_a)$, the ratio of axial-field-node-spacing to the transverse-field- node-spacing is given by by $L_{axial}/L_{transverse} = (2N + 1)$. Given $N$ is integer, the end nodes of all axial and transverse fields can collocate to form a complete packet.

While $E_z$ and $H_z$ are small but non-zero, the axial velocity of the TE+TM packet is fractionally below the velocity of light. This allows examination of packet's structure in its proper frame where its group velocity is zero [13], then the helical rotations change the effective optical length of an '$N$' state to $2N + 1$ times the effective length when $N = 0$. Power flows into the packet before any nodes form but ceases as soon as the nodes form. The enclosed energy then is a measure of the power flow times the effective packet temporal duration. It is for this reason that an '$N$' packet contains $2N + 1$ times the energy of any 'ground state' packet without helical changes ($N = 0$).

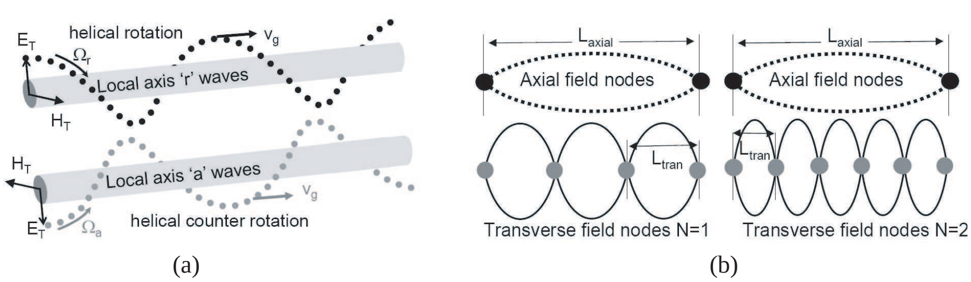



Figure 2: (a) Local helical rotations of transverse fields: 'a' waves rotate helically counter to 'r' waves though both waves propagate in the same direction. (b) Counter rotating helical DiTOs create additional nodes in the transverse fields. With appro-priate choice of the helical frequency the number of nodes can be increased by $2N$ where $N$ is integer. This quantized structure is independent of diffraction (i.e., $N$ remains unchanged as diffraction and $E_z/H_z$ tend to zero).

## 4. CONCLUSIONS

Distributed transverse orientations (DiTOs) of the transverse electromagnetic fields in TE/TM waves offer new, if speculative, insights into the working of classical MEs. Experimentally DiTOs pose a challenge over experimental accessibility. If ways of exciting these hidden internal changes can be devised, then new ways may be available for encoding information on electromagnetic beams. References [13] and [14] suggest that DiTOs are not subject to material dispersion so that encoding information by different helical frequencies or different DiTO phases might offer larger information bandwidths than at present obtainable. Future work will consider how waves, with helical DiTOs, react at beam splitters. If DiTOs were found to be accessible and able to construct photon like packets [1, 13] then such packets should emerge from one side of a beam splitter or the other but not both (similar to a photon). This might offer interesting ways of controlling optical paths.

## ACKNOWLEDGMENT

Special thanks are given to Ian White for valuable support at Cambridge and similarly to Joseph Beals IV who was a valued co-author of an earlier paper [1]. Grateful thanks are also given to Cambridge colleagues, Jonathan Ingham, Ruth Thompson, Jose Rosas-Fernandez, for their help and support.

## REFERENCES

1. Carroll, J. E. and J. Beals IV, "Photon-like solutions of Maxwell's equations in dispersive media," *The Nature of Light: What Are Photons*, C. Roychoudhuri, A. F. Kracklauer, & A. Y. Khrennikov (eds.), *Proc. Spie*, 0277-786X, Vol. 7421, 7421-17, 2009.
2. Davis, L. W., "Theory of electromagnetic beams," *Phys. Rev. A*, Vol. 19, 1177–1179, 1979.
3. Li, C.-F., "Representation theory for vector electromagnetic beams," *Phys. Rev. A*, Vol. 78, 063831, 2008.
4. Lekner, J., "TM, TE and 'TEM' modes: Exact solutions and their problems," *J. Opt. A: Pure Appl. Opt.*, Vol. 3, 407–412, 2001.
5. Berry, M. V., "Optical vortices evolving from helicoidal integer and fractional phase steps," *J. Opt. A: Pure Appl. Opt.*, Vol. 6, 259–268, 2004.
6. Marshall, T. W., "Random electrodynamics," *Proc. Roy. Soc. A*, Vol. 276, 475–491, 1963.
7. Boyer, T. H., "Random electrodynamics: The theory of classical electrodynamics with classical electromagnetic zero-point radiation," *Phys. Rev. D*, Vol. 11, 790–807, 1975.
8. Kracklauer, A. F., "Pilot wave steerage: A mechanism and test," *Found. Phys. Lett.*, Vol. 12, 441–453, 1999.
9. Jackson, J. D., *Classical Electrodynamics*, 2nd Edition, 226, Wiley, New York, 1975.
10. Wheeler, J. A. and R. P. Feynman, "Interaction of the absorber as the mechanism of radiation," *Rev. Mod. Phys.*, Vol. B7, 157–181, 1945.
11. Cramer, J. G., "The transactional interpretation of quantum mechanics," *Rev. Mod. Phys.*, Vol. 58, 647–687, 1986.
12. Nienhuis, G. and L. Allen, "Paraxial wave optics and harmonic oscillators," *Phys. Rev. A*, Vol. 48, 656–665, 1993.
13. Carroll, J. E., "Flying with 'Photons'?" submitted for publication, 2010.
14. Carroll, J. E., "Towards unlimited bandwidth?" *El. Lett.*, Vol. 43, 429–431, 2007.

# Detection of Partial Discharge inside of HV Transformer, Modeling, Sensors and Measurement

**P. Fiala**, **T. Jirku**, **P. Drexler**, **and P. Dohnal**
Brno, FEEC BUT, UTEE, Kolejní 2906/4, Brno 612 00, Czech Republic

**Abstract**— The aim of this paper is to present the particulars of result research in the HF measurement method and modeling of starting process partial discharge inside of high voltage transformer. The numerical analysis of the electromagnetic wave attenuation helped to set up conditions to decrease it and get of information for sensors conception preparing and detection apparatuses construction and measurement methods.

## 1. INTRODUCTION

One of the problematic conditions in the field of high-voltage technology, apparatuses and devices (machines) consists in the emergence of partial discharges [1]. At this point, let us also note that several other effects have combined with this notion over time [2–4]. In consequence of these effects there emerge short electromagnetic pulses with a defined and measurable spectrum in the characteristic frequency band [5]. The group of end products attributable to the emergence of interfering signals involves, for example, displacement current in a dielectric, pulse current on the interface between dielectrics, or the dielectric/metal interface owing to high electric field intensity and structure of the dielectric. In HV and VHV transformers the dielectric is mineral or synthetic oil.

Large distribution transformers are constructed in such a manner as to have structural measures facilitating oil purification. Also, these transformers are equipped with sensors indicating the initial stage of increase in pulse activity. In the course of this activity, as is well-known, there occurs an increase in the boundary value of the of the applied dielectric breakdown value. As referred to in the above text, oil is the dielectric. Under certain conditions, however, the separation of chemical compounds incurred by decomposition of the dielectric does not have to occur. Thus, free atoms of carbon, hydrogen and oxygen develop from hydrocarbons, and there also generates a certain percentage of water, other organic compounds, and semiconductive carbon. All of these elements decrease the quality of the dielectric; in addition to that, rapid increase in pulse activity may cause the formation of a hazardous explosive compound of oxygen and hydrogen. Then, this situation may result in a local explosion, damage to the device and reduction of its ability to perform the respective functions.

This work deals next steps after the analysis of electromagnetic field distribution in a transformer dielectric region. The structural parts enable the placement and choice of sensors, whose structure and concept must be adapted to the characteristics of the configuration in such a manner that, from all components of the device, there is a measurable (indicable) electromagnetic impulse signal.

The analysis will be realized for the minimum required level of an electromagnetic pulse for the discrete values of frequencies from the desired spectral interval. An example will be evaluated of electromagnetic field distribution in the region of critical parts of the device.

## 2. MATHEMATICAL MODEL

It is possible to carry out an analysis of an MG model as a numerical solution by means of the Finite element method (FEM). The electromagnetic part of the model is based on the solution of full Maxwell's equations

$$\nabla\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t},\quad \nabla\times\mathbf{H}=\sigma\mathbf{E}+\frac{\partial\mathbf{D}}{\partial t}+\mathbf{J}_s,\quad \nabla\cdot\mathbf{D}=\rho,\quad \nabla\cdot\mathbf{B}=0\ \text{in}\,\Omega. \tag{1}$$

where $\mathbf{E}$ and $\mathbf{H}$ are the electrical field intensity vector and the magnetic field intensity vector, $\mathbf{D}$ and $\mathbf{B}$ are the electrical field density vector and the magnetic flux density vector, $\mathbf{J}_s$ is the current density vector of the sources, $\rho$ is the density of free electrical charge, $\gamma$ is the conductivity of the material and $\Omega$ is the definition area of the model. The relationships between the electric and the magnetic field intensities and densities are given by material relationships

$$\mathbf{D}=\varepsilon\mathbf{E},\quad \mathbf{B}=\mu\mathbf{H}. \tag{2}$$

The numerical model analysis was described in paper [8].

## 3. TYPES OF SENSORS

The sensors of electromagnetic field were described in report [9]. Tests of antennas are prepared according Fig. 1, described in paper [1] and [5]. Results of numerical model analysis are presented in Fig. 3. The real position of antennas is shown in Fig. 2. There were tested three types of antennas. The first one was spiral types, Fig. 4. The frequency range for spiral antennas was from 500 MHz to 3.5 GHz. The SWR diagram was tested in anechoic chamber and it is shown in Fig. 5. The second type of antenna was based on cone design, Fig. 6, the SWR diagram is shown in Fig. 7. The last type of antenna was design like a Vivaldi. The used material of them was PCB of FR-4, Fig. 8.

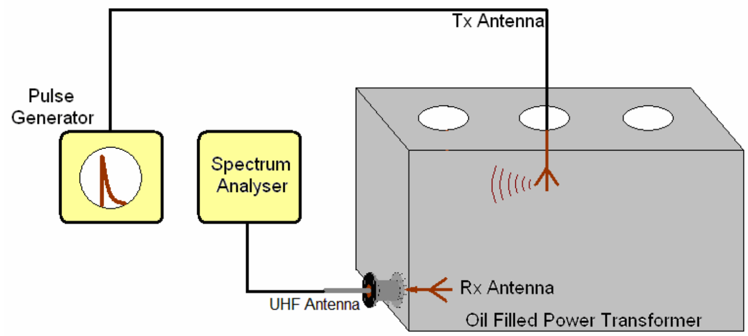



Figure 1: The sensor tests.

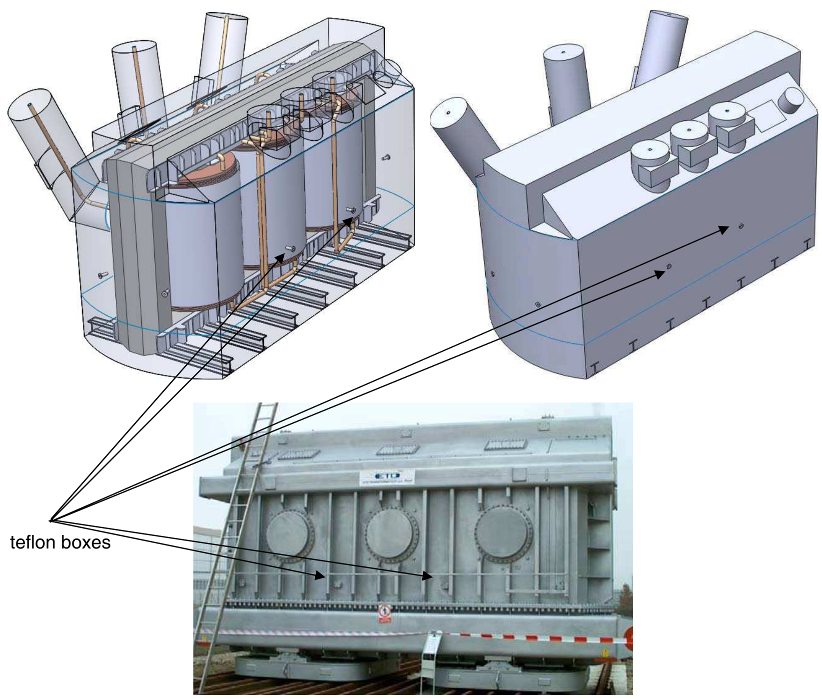



Figure 2: Detector position.

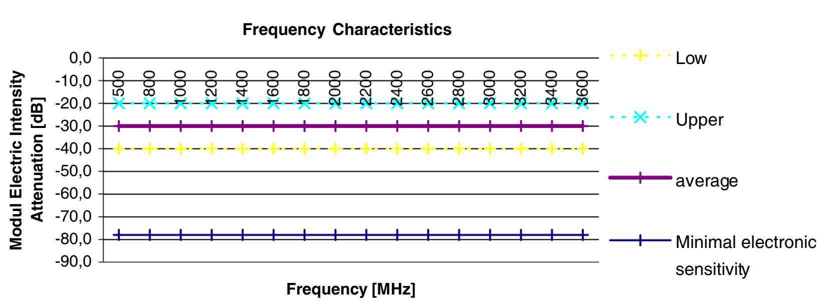



Figure 3: Numerical model analysis result of Electric intensity attenuation inside of HV transformer.

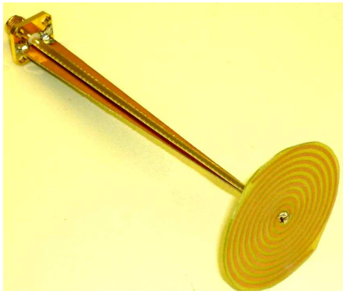


Figure 4: Spiral antenna for $f \in \langle 500; 3500 \rangle$ MHz.

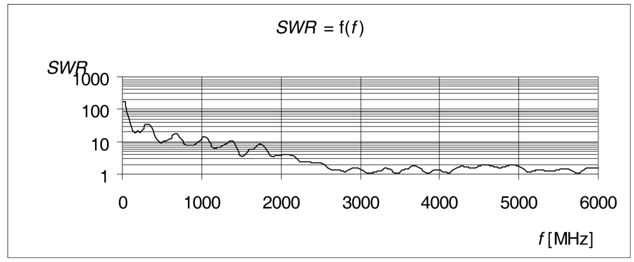



Figure 5: The SWR diagram of spiral antenna.

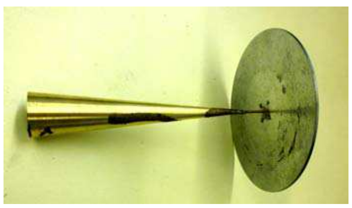


Figure 6: Cone type antenna design for $f \in \langle 500; 3500 \rangle$ MHz.

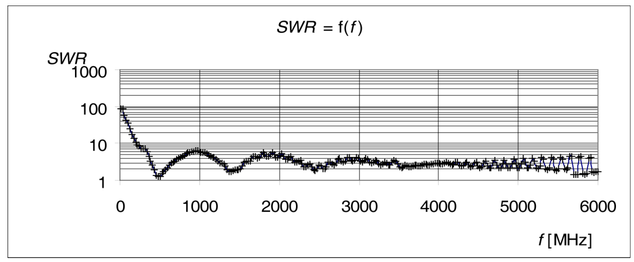



Figure 7: The SWR diagram of cone type antenna.

Figure 8: The Vivaldi antenna design for $f \in \langle 500; 3500 \rangle$ MHz.

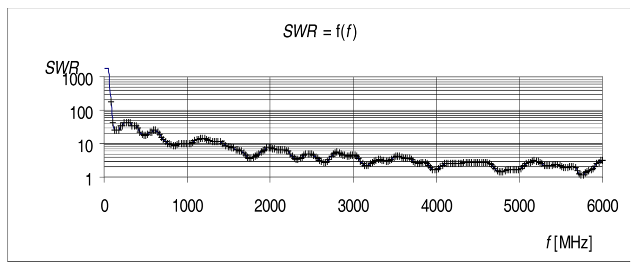



Figure 9: The SWR diagram of Vivaldi type antenna.

## 4. CONCLUSION

The basic research of the numerical model HF wide band signals inside of a VHV transformer has brought a considerable sum of experience in the field of signals and possibilities of their measurement and detection. It was used for sensor design conception and construction. There were tested three types of antenas and it was prepared next step of partial discharge detection.

## ACKNOWLEDGMENT

The research described in the paper was financially supported by FRVŠ (a fund of university development) by research plan No. MSM 0021630513 ELCOM, No. MSM 0021630516 and grant Czech ministry of industry and trade MPO No. FR-TI1/001, GACR 102/09/0314.

## REFERENCES

1. Kaneiwa, H., Y. Suzuoki, and T. Mizutan, "Characteristics of partial discharges, artificial simulated tree channels during in tree propagation," *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 8, No. 1, February 2001.
2. Fiala, P., "Secondary winding model of current transformer-switchable variant," Research report, Laboratory of Modeling and Optimization Field in Electromagnetic Systems, FEI VUT and ABB EJF a.s. Brno No. 1/99, 21.1.1999, Brno, Czech Republic, 1999.
3. Fiala, P., "Transformer partial discharge modeling, minimal breakdown value set in a critical parts of transformer design," Research report, Laboratory of Modeling and Optimization Field

in Electromagnetic Systems, FEI VUT and ABB EJF a.s. Brno No. 2/99, 18.3.1999, Brno, Czech Republic, 1999.

4. Fiala, P., "Optimization of two-pole voltage transoformer design $U_{\max} = 25\,\mathrm{kV}$, new variant 2000," Research report, Laboratory of Modeling and Optimization Field in Electromagnetic Systems, FEI VUT Brno No. 1/00, 20.2.2000, Brno, Czech Republic, 2000.
5. Sarathi, R., A. J. Reid, and M. D. Judd, "Partial discharge study in transformer oil due to particle movement under DC voltage using the UHF technique," *Electric Power Systems Research*, Vol. 78, 1819–1825, Elsevier, 2008.
6. *Manuals ANSYS*, v.8.1, USA.
7. Fiala, P., "Coupled electromagnetic model of smoothing choke," *AMTEE'01 Fifth International Conference*, Plzeň Str. C05, October 9–December 9, 2001.
8. Fiala, P., E. Gescheidtová, and T. Jirku, "Numerical modeling a microwave and detection of partial discharge inside of HV transformer," *PIERS Proceedings*, Xi'an, China, 2010.
9. Fiala, P., T. Jirku, and T. Drexler, "Discharge activities detection inside of power oil: Detection method, numerical model analysis transmiting, design," Research report, DTEEE FEEC BUT, Czech Republic, Novermber 20, 2009.

# EMHD Model Used for Linear Moving Objects Analysis

**Pavel Fiala, Zoltán Szabó, Tibor Bachorec, and Premysl Dohnal**
Brno, FEEC BUT, UTEE, Kolejní 2906/4, Brno 61200, Czech Republic

**Abstract**— The paper provides an insight into the issues of measurement methods for accuracy obtain information velocity objects. One of the measurement method used for its problems is based on electromagnetic field measurement. The accuracy measurement method is focused to use of EMHD effects analysis and results applied to the experimental measurement. This results and measurement method decrease lethality of wrong measurements of non-ballistics projectile velocity.

## 1. INTRODUCTION: MODELLING AND PHYSICAL MODELS

At the present time, the modelling and simulation of technical systems constitute a substantial part of the entire network of modelled problems. Among the modelled tasks there are various problems related to fields with the fastest pace of both the development and the application of modelling, namely to energetics, mechanical engineering, metallurgy, electrical engineering, space and nuclear technologies [1, 2].

Physical analogy embodies an extension of the notion of physical similarity to physically disparate phenomena having a similar mathematical description. Currently there exist a large number of various physical analogies used in disciplines such as hydromechanics, thermomechanics, elasticity, electrical engineering, applied mathematics, and non-technical subject fields. In addition to the required characteristics, the determining factors in the analogy selection process usually consist in the precision, accessibility, task attributes (character), and measurability of quantities.

From this angle of view we can mention, for example, a model described and analyzed by means of a mathematical apparatus which accepts different types of physical models (Poisson's equation for electrostatics, the diffusion equation for the propagation of heat) [3]. This model is solved as a coupled model.

## 2. HYBRID MODELS

In the making of hybrid simulation models, analogue and numerical (or analytical and experimental, numerical and experimental) models are combined to facilitate the utilisation of convenient characteristics of both the paired components. Here, information represented differently is processed. During the course of simulation there occurs unidirectional or bidirectional transmission of data; by the help of a connecting device, the information is converted from the numerical to the analogue form and vice versa.

Hybrid models provide an extensive potential applicable in the solution of tasks which are extreme, difficult to verify, or very intensive in terms of the analysis of results.

This paper examines the problem of utilisation of the coupled hybrid magnetic field model and the subject of the conductive area motion in a magnetic field. The model is dealt with *per partes*: the first aspect solved is a stationary model of a magnetic field, the following aspect consists in a quasi-stationary model for the assumed motion of elements of the task, and the last part constitutes a model respecting the relativistic principle in electromagnetic field modelling [4].

## 3. DESCRIPTION OF THE MODEL

The physical model is based on the solution of the reduced Maxwell equations [4]; the stationary magnetic field can be described as

$$\mathrm{rot}\mathbf{H} = \mathbf{J}, \tag{1}$$

$$\mathrm{rot}\mathbf{E} = 0, \tag{2}$$

$$\mathrm{div}\mathbf{B} = 0, \tag{3}$$

$$\mathrm{div}\mathbf{J} = 0, \tag{4}$$

$$\mathrm{div}\mathbf{D} = 0, \tag{5}$$

where **H, J, E, D, B** are the vectors of magnetic field intensity, current density, electric field intensity, electric flux density, magnetic flux density, respectively. The material relations are represented by the expressions

$$\mathbf{B} = \mu_0\mu_r\mathbf{H}, \tag{6}$$
$$\mathbf{J}_v = \gamma\mathbf{E}, \tag{7}$$
$$\mathbf{D} = \varepsilon_0\varepsilon_r\mathbf{E}. \tag{8}$$

where $\varepsilon$, $\mu$, $\gamma$ are the permittivity, permeability, conductivity of the environment. Vector functions of the electric and the magnetic fields **E, B** are expressed by the help of the scalar electric potential $\phi_e$ and the vector magnetic potential **A**; for the static task, in the relation for the electric field intensity, the time derivation of the vector magnetic potential is zero

$$\mathbf{E} = -\text{grad}\phi_e - \frac{\partial\mathbf{A}}{\partial t}, \tag{9}$$
$$\mathbf{B} = \text{rot}\mathbf{A}. \tag{10}$$

The resulting current density **J** from relation (4) is formed by the exciting current density $\mathbf{J}_s = \rho\mathbf{v}$ with the specific density of electric charge $\rho$ and the current density caused by eddy currents $\mathbf{J}_v$. Motion effect for the instantaneous velocity vector **v** is respected in the model by current density

$$\mathbf{J}_m = \gamma(\mathbf{v}\times\mathbf{B}). \tag{11}$$

Then, in respecting eddy currents (14),

$$\mathbf{J} = \mathbf{J}_v + \mathbf{J}_s + \mathbf{J}_m. \tag{12}$$

The electromagnetic field distribution is formulated using expressions (1) to (12).

$$\text{rot}\mathbf{H} = \mathbf{J}_v + \mathbf{J}_s + \mathbf{J}_m \text{ in the entire region of model } \Omega. \tag{13}$$
$$\text{rot}\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}. \tag{14}$$

For the individual parts of model $\Omega$ there holds $\Omega \subset \Omega_v \cup \Omega_s$, where $\Omega_v$ is the area with dominant eddy currents which behave according to relation (7), and $\Omega_s$ is the area with the known distribution of current density $\mathbf{J}_s$. In the assumed model there holds

$$\Omega_s \subset \Omega_v. \tag{15}$$

These basic relations were complemented with boundary and initial conditions and, using the finite element method, the physical model was transferred to the mathematical. The mathematical model is a system of non-linear, time-variable equations; reference [5] provides a complete derivation of the model. The ANSYS program system enabled us to perform a complete analysis; the ANSYS program was modified (a solution internally referred to as SOLID97) using suitably selected boundary conditions in order to respect the model characteristics described in relations (11) to (13), [6].

According to reference [7], in any investigation into electrodynamic systems it is necessary to respect the relativistic approach to electrodynamics. The entire problem begins at the moment when the vector of intensities of both the electric and the magnetic fields of the moving system is **relative**. In the static system, the model dynamics is shown in relations (11) and (14). In order to eliminate possible errors, it is suitable to include in relation (14) the term which respects Faraday's law of induction

$$\text{rot}\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t} + \text{rot}(\mathbf{v}\times\mathbf{B}). \tag{16}$$

and for the magnetic field relation (13)

$$\text{rot}\mathbf{H} = \mathbf{J} + \rho\mathbf{v} + \frac{\partial\mathbf{D}}{\partial t} + \text{rot}(\mathbf{v}\times\mathbf{D}). \tag{17}$$

where $\rho$ is the volume density of electric charge. The complete Maxwell equations are covariant in all the systems; therefore, it is not important to specify the system within which the observer moves as the described relations always hold true. After the derivation of the four-vector and respecting the Lorentz transformation, the current density is written in the form

$$\mathbf{J} = \rho \frac{\partial \mathbf{s}}{\partial t} + jc\rho \mathbf{u}_t. \tag{18}$$

where $j$ is the symbol of the imaginary component of the quantity complex shape, $c$ is the velocity of light module in the vacuum. For the continuity Equation (4) there holds

$$\operatorname{div}(\mathbf{J}) = 0. \tag{19}$$

In order to facilitate simplification, if we assume the motion of one reference system as ()' (for example, in the $x$ axis of the Cartesian coordinate system), the four-vector of current density can be written, after application of the Lorentz transformation, in the (invariant) form of

$$\mathbf{J}' = \frac{J_x + j\dfrac{v_x}{c} J_t}{\sqrt{1 - \left(\dfrac{\|\mathbf{v}\|}{c}\right)^2}} \mathbf{u}_x + J_y \mathbf{u}_y + J_z \mathbf{u}_z + j \frac{c\rho - \dfrac{v_x}{c} J_x}{\sqrt{1 - \left(\dfrac{\|\mathbf{v}\|}{c}\right)^2}} \mathbf{u}_t. \tag{20}$$

and the electric charge volume density in the reference system is, after transformation, written as

$$\rho' = \frac{\rho - \dfrac{v_x}{c^2} J_x}{\sqrt{1 - \left(\dfrac{\|\mathbf{v}\|}{c}\right)^2}}. \tag{21}$$

There is no problem obtaining, by means of reversal in the sign of velocity, transformed quantities of the opposite system (which is moving) for the above-expressed transformed quantities. Let us have a simple geometric task, Fig. 1.

## 4. SOLUTION

Several steps of the electromagnetic field analysis will be realized. Step 1 is the elementary magnetic field analysis without the strip movement, relations (1) to (5). Step 2 is the magnetic field analysis with a steady movement of the conductive strip and the eddy currents effect (16). Step 3 is the evaluation of current density in the moving strip with the electrodynamic effect according to relations (20) and (21). The SOLID97 element with the vector magnetic potential was used to

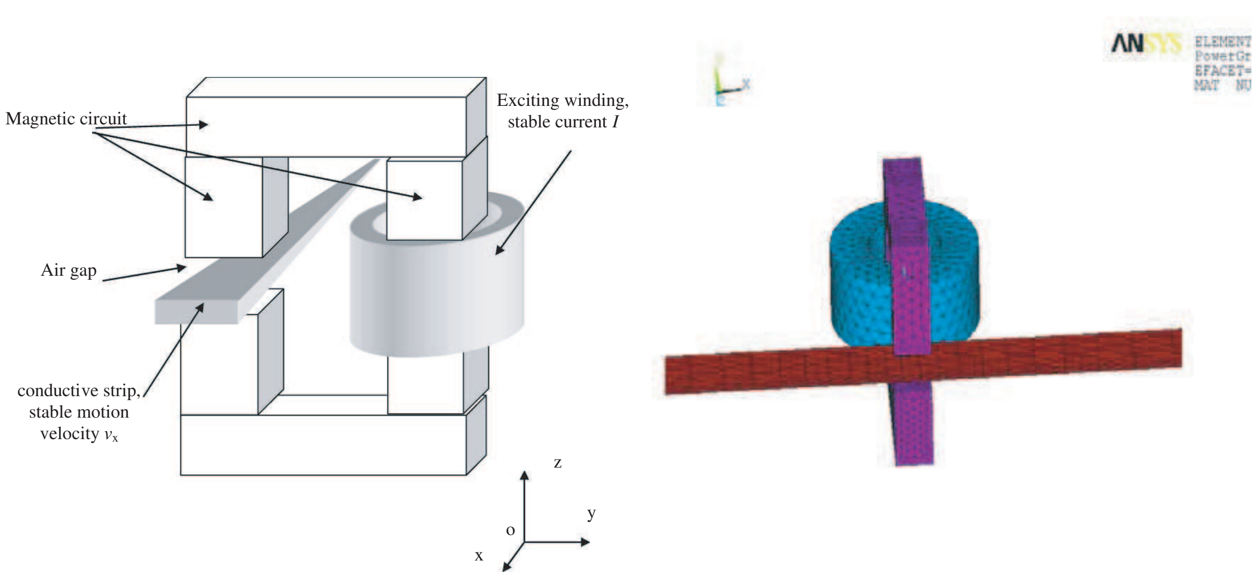



Figure 1: A simple geometry for the verification of the relativistic approach of electrodynamics, numerical model.

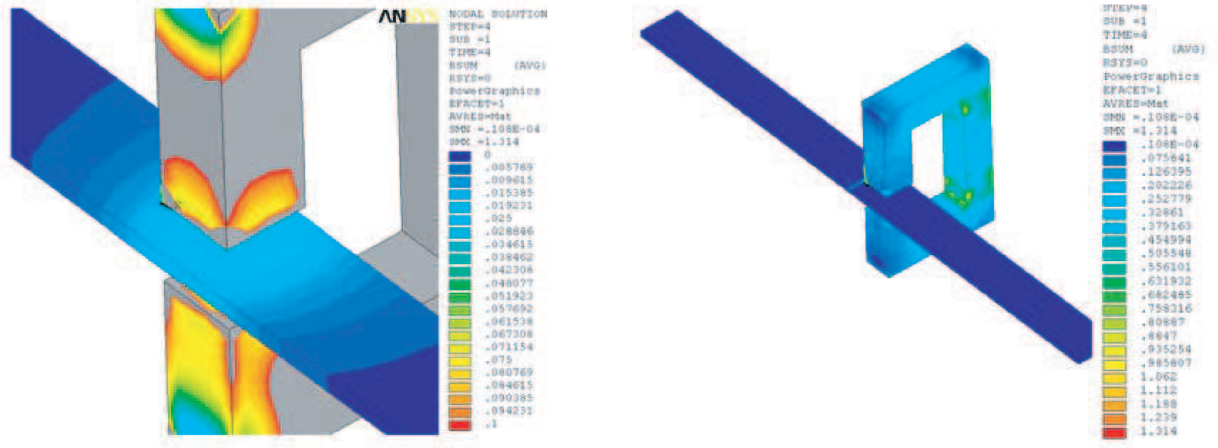
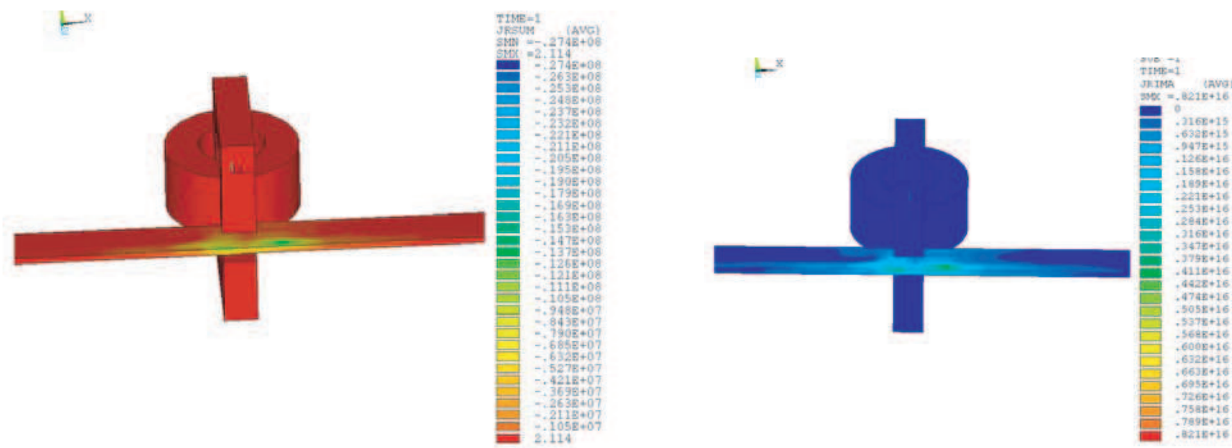


Figure 2: The magnetostatic task analysis: the the magnetic flux density vector module **B** distribution, the scalar magnetic potential.



(a) (b)

Figure 3: The electrodynamic magnetic field analysis: the current density $\mathbf{J}_{re}$ vector module distribution, the vector magnetic potential. (a) The real component. (b) The imaginary component.

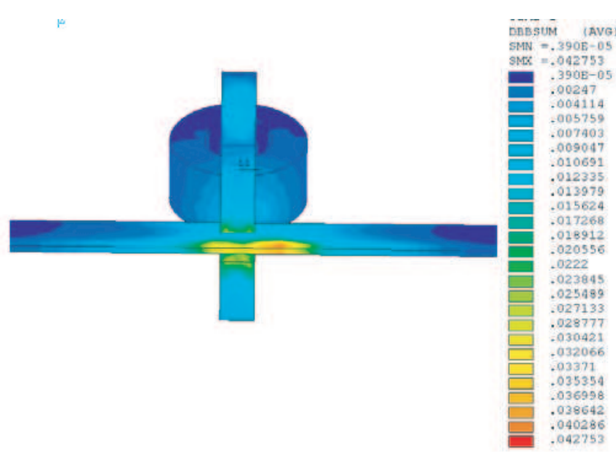


Figure 4: The electrodynamic magnetic field analysis: magnetic flux density vector module distribution $\mathbf{B}_{re}$ with an electrodynamics effect, the vector magnetic potential, excitation realized exclusively through the real component of current density $\mathbf{J}_{re,r}$.

facilitate the realization of these analyses; the element was checked by means of the model analysis using the SOLID98 component with scalar magnetic potential (the APDL language program for ANSYS in appendix 1). Step 1 was evaluated; the results of the magnetic the magnetic flux density vector module **B** agree in both models and correspond to the anticipated theoretical results for the stationary magnetic field, Fig. 2.

It was evaluated module of the current density vector in the task quasi-stationary model; the conductor moves in the $x$ axis direction at the speed of $v = 1\,\mathrm{m/s}$. According to relations (20) and (21) — step 3, the current density vector $\mathbf{J}_{re}$ was evaluated in the real and the imaginary components, Fig. 3. This value was superposed on the basic component of current density, and the magnetic field analysis for these conditions was performed. Magnetic field excitation was carried out using only the real component of the resulting current density **J**. The resulting distribution of the magnetic field and the magnetic flux density **B** vector module is shown in Fig. 3. Naturally, it is possible to perform an analysis for the imaginary part of the current density component.

## 5. CONCLUSION

The model makes apparent the distinction by order of the individual phenomena. It is evident that the electrodynamic principle is indispensable and introduces substantial changes into the original non-dynamic conception of this type of simple task analysis. In view of the numeric values, it is important $B_{\max} = 0.277 \times 10^{-3}\,\mathrm{T}$ for the quasi-stationary task and Fig. 4, $B_{\max re} = 0.427 \times 10^{-1}\,\mathrm{T}$ for the electrodynamic model. The effect is apparent of motion on the magnetic field imbalance. A higher value of an elementary magnetic field combines with a decrease in relative magnitude of the electrodynamic field influence.

## ACKNOWLEDGMENT

The research described in the paper was financially supported by FRVŠ (a fund of university development) from research plans No. MSM 0021630513 ELCOM, No. MSM 0021630516 and by a grant from the Ministry of Industry and Trade of the Czech Republic (MPO No. FR-TI1/368).

## REFERENCES

1. Kuneš, J., O. Vavroch, and V. Franta, *Základy Modelování*, ISBN 80-03-00147-1, SNTL, Praha, 1989.
2. ANSYS, *Theory Reference Manual*, Příloha Programu ANSYS.
3. Dědek, L. and J. Dědková, *Elektromagnetismus*, 234, VUTIUM, Brno, 2000.
4. Stratton, J., *Teorie Elektromagnetického Pole*, 51–60, Vyd., SNTL, Praha.
5. Fiala, P., "Secondary winding model of current transformer — switchable variant," Research Report, Laboratory of Modeling and Optimization Field in Electromagnetic Systems, FEI VUT and ABB EJF, a.s. Brno No. 1/99, Brno, Czech Republic, Jan. 21, 1999.
6. Fiala, P., "Transformer partial discharge modeling, minimal breakdown value set in critical parts of transformer design," Research Report, Laboratory of Modeling and Optimization Field in Electromagnetic Systems, FEI VUT and ABB EJF, a.s. Brno No. 2/99, Brno, Czech Republic, Mar. 18, 1999.
7. Haňka, L., *Teorie Elektromagnetického Pole*, Skripta, SNTL, Praha, 1971.

# Tuned Periodical Structures — Model, Experiments in THz Band Applied in Safety Application

**Pavel Fiala, Radim Kadlec, Petr Drexler, and Premysl Dohnal**
Brno, FEEC BUT, UTEE, Kolejní 2906/4, Brno 61200, Czech Republic

**Abstract**— The paper provides an insight into the issues of integration and application of non-lethal weapons and devices in the field of protection against special-type weapons. The structures like a metamaterials, left-handed type models were analyzed and prepared to the experimental measurements.

## 1. INTRODUCTION

Since 2001 discussions have been in full progress concerning legal adaptation to the newly emerged conditions or needs of finding methods and means that provide constrictive wounding effects [1]. Importantly, there have emerged several focus areas within the field. Firstly, let us turn our attention in this respect to legal conditions and ethics, where the problem is analyzed not only with a view to the question of how and when non-lethal weapons should be applied, but also from the perspective of the existing structures and regulations of international law as well as national legal systems. It is also important to note within this context that higher-order social functions have been taken into account, for example the maintenance of rights of individuals. In addition to this aspect, the overall discussion has involved the problem of an individual transgressing legal norms [2]. In the field of the operational-tactical outlook and tactical approaches to applying different classes of weapons, means and methods, the solution proposed in source [3] is of significant interest. Currently, psychological training or training conducted by a psychologist is being proven beneficial for various purposes in the province of tactics, where it has provided very good results. Yet, with a certain degree of contrast, the use of lethal means or methods can yield successful unravelling of local conflicts without causing major injuries to humans.

In the medicine-related sections of wounding effects and means evaluation [4, 5], methodology has been developed to evaluate the concrete effects on individual parts of the human body [6]. The ultimate province to be mentioned is the technical-technological field, where the potential and known physical principles are utilized to facilitate the designing of the means and principles of non-lethal effects [7–10].

One of the problematic aspects of various armed conflicts consists in the issue of snipers and their timely detection by protective forces. For this reason, full priority within the investigation into the problem of sniper position detection has been given to locating the sniper before the first shot; for the same reason, too, the applications of image recognition methods [11, 12] have emerged in the field of military tactics. Naturally, these methods are also utilized for several other purposes. Fig. 1 shows, according to [11], the analysis of the process of utilization of degrees of means and the application of image recognition as a sensor for the input evaluation of situation in the intervention process. Another sophisticated semiautomated system of intervention controlling consists in the procedure of identifying the physical characteristics of crowd behaviour [12] and, based on the evaluation of energy and behaviour of the crowd elements, the introduction of intervention with respect to the problem in the task. The task can be, for instance, loss minimization or the rapidity of crowd expansion inhibition. An example of the evaluation of necessary elements of image is shown in Fig. 2; here, the sources of crowd dynamics are found to provide the basis for evaluation of the crowd energy and for the identification of essential instruments to facilitate the ultimate controlling of the crowd, introduction of intervention, and intervention location recognition.

According to article [13], one of the options of sniper identification consists in the scanning of the required sector and the retrieval of the point that shows the corresponding optical characteristics. Other ways of solution can be found in the process of further investigation from the perspective of physics into the problem of snipers and their identification.

## 2. TARGET CHARACTERISTICS WITH RESPECT TO THE ELECTROMAGNETIC FIELD

The electromagnetic field enables us to see the problem from a different perspective. In an analysis of the sniper location task we may assume that the identification of the sniper will only be possible in the ex post mode after a single gunshot. In this respect, it is necessary to note that the sniper's activity is very precise in distances above the range of 50 m and that, after the shooting action, the sniper will not show elements of movement that can be easily discernible or identified. Then, it is possible to focus from the perspective of physics on the characteristics of the target. With respect to the instruments or means of the sniper, the target can show different characteristics; thus, for example, in certain physical conditions the target does not have to be clearly visible or can be visible in the light spectrum of white light with the wavelength of $\lambda \in \langle 400\,\text{nm}, 700\,\text{nm}\rangle$. Another possible solution consists in the situation when the image of the target from the observer-sniper shows, in the defined light spectrum, a position different from reality. Thus, there will not occur any fatal consequences in the absence of timely identification of the sniper position and, thanks to the fact, partial defensive actions to save the target will be made possible.

## 3. THE REFLECTION AND REFRACTION OF ELECTROMAGNETIC WAVES

In the DTEEE laboratories, an algorithm was derived and set up for the refraction and reflection of EMG wave on the interface by the help of the refraction coefficient the permeance coefficient. The algorithm was tested using the Matlab program, and an analysis of the issue was provided in [14]. The program generates a matrix of beams which propagate from the source. Intersections of the beams and objects in the model are evaluated and a new direction is determined of the reflected and the permeating beams. The program has been designed to facilitate the analysis and evaluation of any quantity of reflected beams. The number of reflections of each beam markedly increases the time exigence factor of the analysis. The reflections and permeation are solved on the basis of laws concerning electromagnetic waves. The reflection and permeation of a simple interface between two instances of the EMG wave environment is indicated in Fig. 1. It follows from the derivation of Snell's law [15] and [16] that, for the calculation of refraction angle, there holds the equation

$$\frac{\sin\theta_0}{\sin\theta_2} = \frac{\mathbf{k_2}}{\mathbf{k}_1}, \tag{1}$$

where $\mathbf{k}$ is the wave number with the wave propagation data, and its shape is:

$$\mathbf{k} = \sqrt{-j\omega\mu \cdot (\gamma + j\omega\varepsilon)}, \tag{2}$$

where $\varepsilon$ is the environment permittivity, $\mu$ is the environment permeability, and $\gamma$ the environment conductivity Relation (1) is formulated only for the interface between two dielectrics which are not subject to the occurence of total reflection. In general, $\mathbf{k}_1$ and $\mathbf{k}_2$ are complex, and then also angle $\boldsymbol{\theta}_2$ is complex. The propagation of light as an electromagnetic wave is understood as the propagation of the electric and the magnetic field intensity, the electric constituent of the incident constituent, according to Fig. 4. This can be written as:

$$\mathbf{E}_i = \mathbf{E}_0 e^{-j\mathbf{k}_1\mathbf{u}_{n0}\cdot\mathbf{r}}, \tag{3}$$

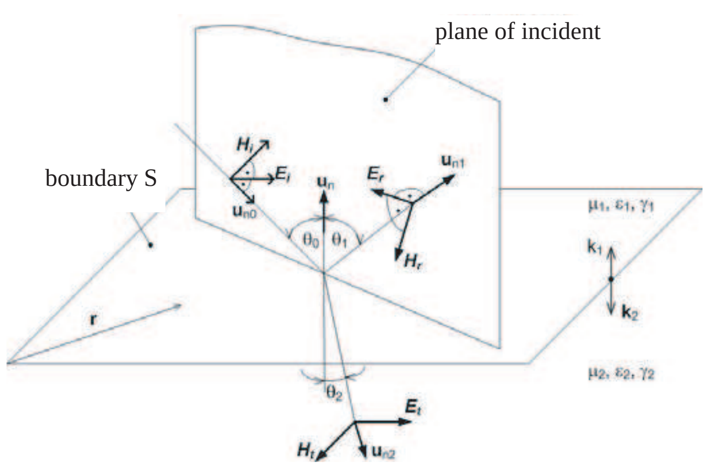



Figure 1: Reflection and refraction of a plane wave.

where $\mathbf{E}_0$ is the electric field intensity amplitude in the location of the interface, $\mathbf{r}$ is the position vector, and $\mathbf{u}_{n0}$ is the unit vector of the propagation direction. The reflected beams intensity and the permeated beams intensity are evaluated as:

$$\mathbf{E}_{\mathbf{r}} = \mathbf{E}_1 e^{-j\mathbf{k}_1 \mathbf{u}_{n1}\cdot\mathbf{r}}, \quad \mathbf{E}_t = \mathbf{E}_2 e^{-j\mathbf{k}_2 \mathbf{u}_{n2}\cdot\mathbf{r}}, \tag{4}$$

where $\mathbf{E}_1$ is determined from the amplitude in the interface location and reflection coefficient $\rho_E$, and $\mathbf{E}_2$ is determined from the amplitude in the location of the interface and transmission (permeation) factor $\tau_E$.

The procedure is adjusted for a multilayer heterogeneous material (layers, metamateials structure). The EMG wave reflections from the heterogeneous material and its permeation are solved by the help of numerical methods. The multilayer environment is schematically indicated in Fig. 2. The algorithm only processes the reflection starting at the number of 10 layers. The reflection from $n$ layers generates $n$ of primary (only once reflected) EMG waves (in Fig. 3, an EMG wave impinges upon 5 layers and there are 5 reflected EMG waves on the surface), which reflect in a multilayer environment.

The experimental model was based on metamaterials type of tuned structure shown in Fig. 4. It was designed for the frequency of 100 THz. The dimension is $x = 9\,\mu\text{m}$. The intensity module of an electromagnetic wave impinging on the structure is $E_{in} = 0.1\,\text{V/m}$, the wave orientation is in the direction of the $y$ axis, and the wave was modeled as linearly polarized. The solution and interpretation are presented in Fig. 5. Here, the basic charactericties of a periodic metamaterial structure are shown in a single element of the selected meta element.

The elements from Fig. 4 are prepared and tested on frequency spectrum $f \in \langle 400\,\text{nm}, 700\,\text{nm}\rangle$. There will be able to find beam deviation effect.

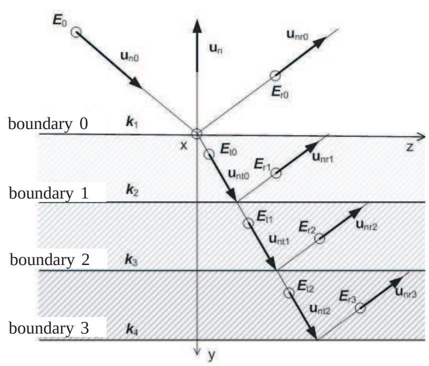



Figure 2: Multilayer enviornment of heterogeneous material.

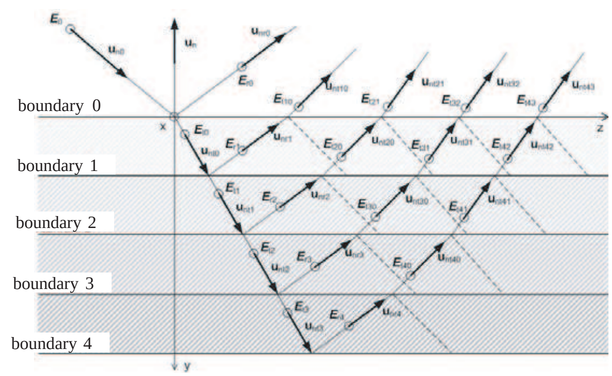



Figure 3: Waves on the surface of a heterogeneous material after the a reflection from several layers.

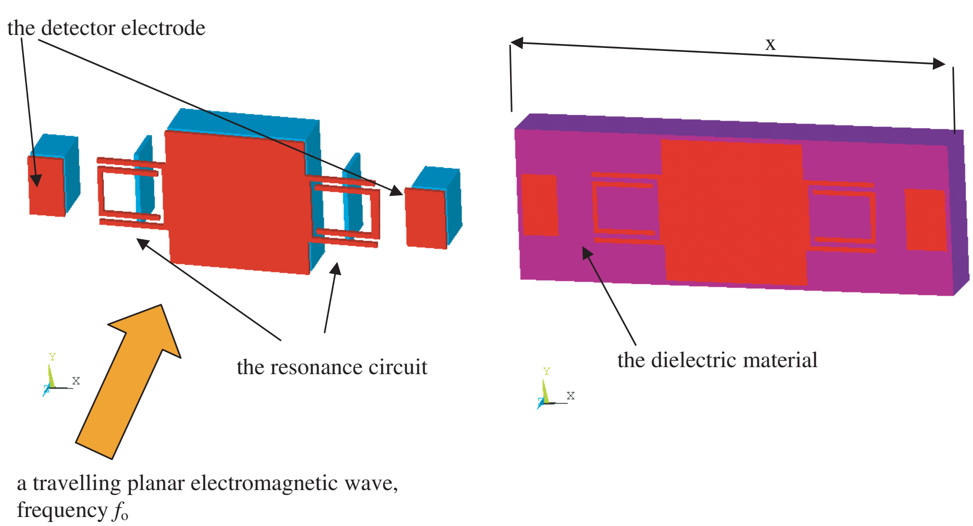



Figure 4: The geometrical model of the resonance circuit basic element.

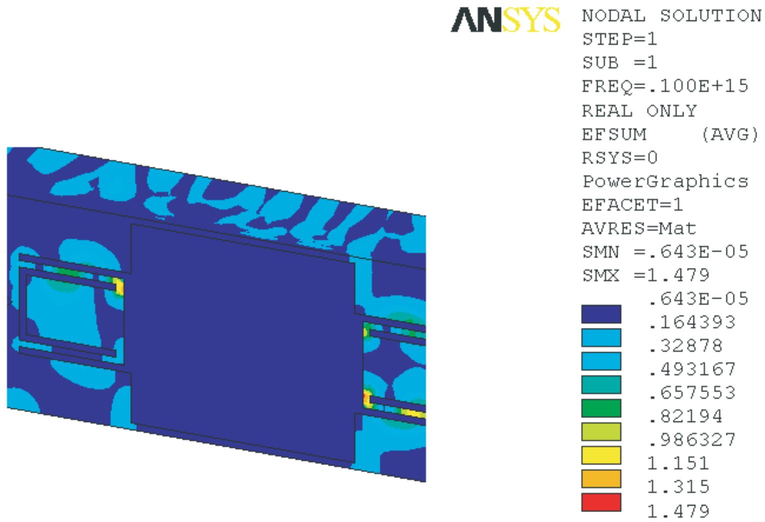



Figure 5: The interpretation of the solution, the electric field components of the resonant circuit basic element.

## 4. CONCLUSION

Both the basic and the applied types of research on optoelectronic systems and numerical modelling of wideband signals have led to conclusions in the field of multilayer and periodic structure optical materials.

The entire project was systematically guided by theoretical discussion and consideration, the results of modelling realized by the help of numerical models, and a large number of experiments. The main asset of the materialized work consists in the field of numerical modelling, in the proposed variants of models, and in the overall verification and calibration of the designed solutions using unique experiments. The activities have contributed significantly to the field of model design by proposing a combination of complementary yet different types of numerical models; these models then enabled marked acceleration of the calculation process while achieving and maintaining a satisfactory degree of accuracy. A valuable aspect consists in the methodology of a numerical model application and handling, where the correctness of the numerical analysis was verified by the experimental results.

It was prepared three metamaterials modification and prepared experimental verification its characteristics.

## ACKNOWLEDGMENT

The funding of the project was supported by the Ministry of Defence of the CR and Ministry of Industry and Trade of the CR (Diagnostics of superfast objects for safety testing, FR-TI1/368), Ministry of Education, Youth and Sports of the CR, and by insitutional resources from the Research Design — Electronic Communication Systems and New Generation Technologies (ELKOM) MSM0021630513, GACR 102/09/0314.

## REFERENCES

1. Krüger-Sprengel, F., "Legal adaptation of non-lethal capabilities in new conflict scenarious," *2nd European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 13–15, 2003.
2. Kim, B., "Between principles and absolutes: Non-lethal weapons and the law of armed conflict," *2nd European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 13–15, 2003.
3. Janssen, W. J. H. and E. J. M. Jansen, "Decision-making processes: The choice between lethal an non-lethal force," *2nd European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 13–15, 2003.
4. David, E., "A Effectiveness and risk during application of NLW from the medical point of view," *4th European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 10–12, 2005.
5. David, E., A. Fretz, and J. Reissenweber, "Mortality following taser exposure," *4th European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 21–23, 2007.

6. Wolf, F., "Multi-spectral measurement of NLW effects," *4th European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 21–23, 2007.
7. Fiala, P., "Finite element method analysis of a magnetic field inside a microwave pulsed generator," *2nd European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 13–15, 2003.
8. Steinbauer, M., P. Drexler, and P. Fiala, "Measurement of vircator ultra-short solitary electromagnetic pulses," *3rd European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 10–12, 2005.
9. Fiala, P., M. Steinbauer, and R. Kubasek, "High power generator test, PGP-II," *4th European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 21–23, 2007.
10. Fiala, P. and P. Drexler, "Measurement methods of pulsed power generators," *4th European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 21–23, 2007.
11. Groll, A., J. Haumann, and C. Pick, "Assessment of non-lethal-weapons (NLW)," *4th European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 21–23, 2007.
12. Kozyrev, V., V. V. Selivanov, V. V. Leonov, and A. A. Zverev, "Analysis of critical levels of physical effects on localized masses of people (crowd)," *4th European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 21–23, 2007.
13. Langhans, D. and D. Meisterhans, "Sniper locating system," *2nd European Symposium on Non-Lethal Weapons*, Ettlingen, SRN, May 13–15, 2003.
14. Kadlec, R., E. Kroutilova, and P. Fiala, "Improving of ray-tracing method for numerical modeling of lighting systems," *PIERS Proceedings*, 156–159, Beijing, China, March 23–27, 2009.
15. Dedek, L. and J. Dedková, *Elektromagnetismus*, Students Book, 232, ISBN 80-214-1548-7, VUTIUM, Brno 2. vyd., 2000.
16. Stratton, J. A., *Teorie Elektromagnetického Pole*, STNL, Praha, 1961.

# The Instruments for Noise Spectroscopy

**P. Drexler, P. Fiala, R. Kadlec, and R. Kubásek**
Dept. of Theoretical and Experimental Electrical Engineering, Brno University of Technology
Kolejní 2906/4, Brno 612 00, Czech Republic

**Abstract**— The article describes basic study of broadband noise signal application in the investigation of materials. The aim is find a metrology method utilizable for the research on metamaterials in the frequency range of about 100 MHz to 10 GHz. The instrumental equipment and other requirements are presented. This research report provides an overview of the current potentialities in the described field and summarizes the aspects necessary for noise spectroscopy.

## 1. INTRODUCTION

In the complex investigation of material structures for the micro-wave application (tensor and composite character), the properties of materials are studied by means of the classic single-frequency methods, which bring about certain difficulties in the process [1]. In boundary changes with a size close to the wave-length there can occur wrong information concerning the examined objects [2, 3]. One of the possible ways of suppressing the negative sources of signals consists in the use of wide-band signals like white noise, and in researching into the problem of absorption in the examined material [4]. These methods require a source of noise, a receiving and a transmitting antenna, and A/D conversion featuring a large bandwidth; for our purposes, the bandwidth ranged between 0 Hz and 10 GHz. Until recently it had not been possible to realize an A/D converter of the described speed, or devices with the above-mentioned bandwidth. Currently, high-end oscilloscopes are available with a sampling frequency of tens of Gsa/s.

## 2. NOISE SOURCE

For UWB systems, several methods of the generation of short pulses with large bandwidth have been developed to date [5]. However, these singly-iterative processes are not applicable for noise spectroscopy; in this respect, there is a need of a continuous source of noise signal (ideally of white noise) with the given bandwidth. The type of source referred to is currently being produced by certain manufacturers specialized in this field. Importantly, for the noise spectroscopy application we require a comparatively large output power of up to 0 dB/mW; the assumed bandwidth characteristics range up to 10 GHz. Nevertheless, at this point it is appropriate to mention the fact that there occurs the fundamental problem of finding active devices capable of performing signal amplification at this kind of high frequencies. As a matter of fact, our requirements are thus limited by the current status of technology used in the production of commercially available devices; the highest-ranking solution for the bandwidth of up to 10 GHz can be found only up to the maximum of 0 dB/mW.

Our response to the above-discussed problem consisted in an attempt to produce a noise generator in laboratory conditions as, in principle, this type of generator can be considered as sufficient for testing and basic measurement. In view of the price and availability of noise diodes we decided to apply thermal noise on electrical resistance as the basic source of noise. The specific connection is shown in Figure 1. The first transistor is in the CC configuration, where we require mainly a high input impedance of the amplifier. The thermal noise at the input is given by its input parameters. The generator could operate even without a resistor at the transistor input, yet the unconnected input would cause a substantial deterioration of the stability. The second and the third transistors form a cascade voltage amplifier in the CE configuration. The output impedance of the third amplifier is 50 Ω for its matching to coaxial line.

Figure 2 shows the realization of the tested noise generator; the BFP620 vf transistors were applied. This type of transistor features the characteristic of $f_t = 65$ GHz and the maximum stable amplification of 11 dB at the frequency of 6 GHz. The overall amplification of the two CE amplifiers in cascade for the output power of 0 dB/mW would have to approximate the value of 10000, and there is no hf transistor available for this kind of stable amplification. Therefore, for the stable noise generator we have to accept a lower output power.

Figure 3 shows the waveforms and the output voltage spectrum of the noise generator for three different working points. Figures 3(a) and 3(b) display the generator output voltage waveform and

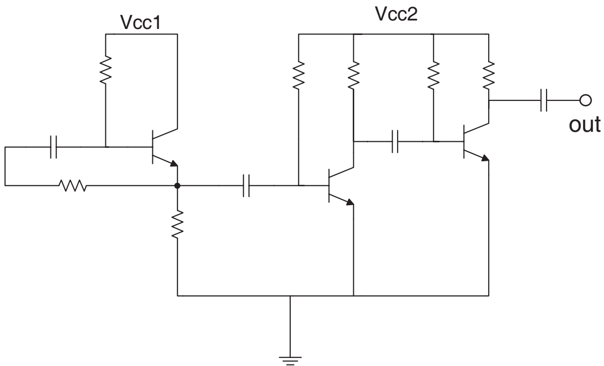



Figure 1: The noise generator configuration diagram.

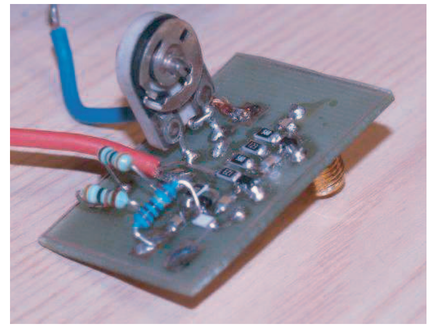


Figure 2: The noise generator realization.

spectrum for the situation when no large oscillations are yet measurable that normally cause the transistor closing and the subsequent transient process. There occur moderate oscillations, which are visible in the spectrum; their amplitude and frequency are subject to chaotic shifting and the output voltage can be considered as stochastic. However, we can not label this as white noise. Figures 3(c) and 3(d) show a situation when, owing to large amplification, there already occur oscillations causing the transistor closing that is well noticeable in the waveform in the time of 2 µs. These types of pulses are randomly repeated within the time range of 10 to 30 µs. In the spectrum, they will upgrade the frequency range in the field of 20 MHz. In applying the low-pass filter we would obtain a signal that could be regarded as stochastic.

The last one of the described situations is given in Figures 3(e) and 3(f); this situation is related to the operating field with the maximum amplification adjustable for the transistor safe operation region. The third transistor already randomly passes between the conditions of "closed" and "open" and the linear operating region. The closing interval ranges from units of ns to several tens of ns. No matter how stochastic this type of signal may be, still it rather models extreme shot noise combined with coloured noise. Its spectrum is the densest and importantly, the most stable of all the situations. For all measurement, spectrum analyzers noise background was −60 dB/mW.

## 3. ANTENNA

The purpose of the transmitting antenna connected to the noise generator output is to form an electromagnetic wave. As a matter of fact, in the field of noise we indeed have to consider a whole spectrum of electromagnetic waves, and it is not possible to define the antenna proximity area. In addition to this, most principles or rules related to the configuration of antennas have to be regarded as void here. The electromagnetic wave is let to impinge on the investigated material and the reflected or partially absorbed wave is then received through the receiving antenna, to the output of which an oscilloscope has been connected. This type of measurement configuration can be seen in Figure 4(a).

Both antennas ought to feature a large bandwidth with, if possible, constant amplitude and defined radiation pattern. In this respect, let us mention the fact that there exist approaches to the design of antennas that come close to the broadband requirements of noise spectroscopy. Suitable solutions include, for example, the spiral fractal antenna or the planar log-periodic antenna. The designed planar log-periodic antenna is applicable for transmission within the frequency range of between 100 MHz and 10 GHz; its real characteristics or qualities depend heavily on the quality of the design practical implementation. Figure 4(b) shows the realization of a planar log-periodic antenna. The numerical design was performed for currently available materials and its evaluation exhibited the undulating module frequency characteristics. The antenna realization experiments using the PCB showed, above all, troubled transmission at higher frequencies from 2 GHz and problematic modification of the feeder. Other antenna designs are directed towards applying the fractal spiral version on Figure 4(c).

The experiments and spectroscopy tests will be performed in an anechoic laboratory. We have selected a system of complementing the Faraday cage shielding with absorbers of electromagnetic waves. The absorbers were designed for the range of 100 MHz–10 GHz with dampening below 35 dB. Thanks to the shielded and separated chamber, the external environment should not affect

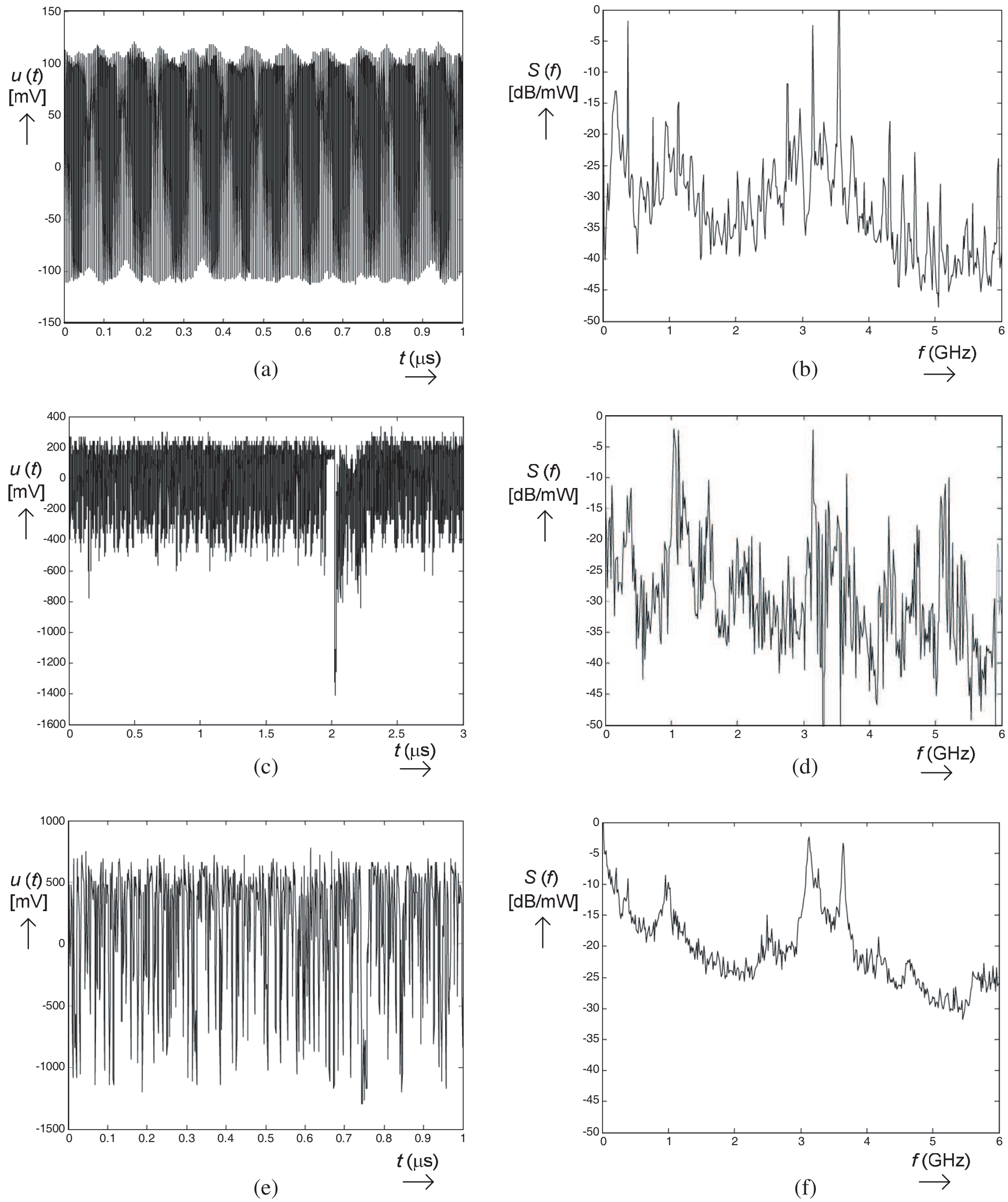



Figure 3: The waveform and spectrum of the noise generator output for three different working points.

the internal part of measurement, and the complemented absorbers will enable us to lower foreign signals to a level of below $-60\,\mathrm{dB/mW}$. Thus, the measurement will not be affected by the outside environment of external electromagnetic sources like mobile phone and Wi-Fi network signals or stationary waves and reflection within the Faraday cage.

## 4. DIGITAL RECIVER

The process of sensing (scanning) has to be necessarily supported by a costly high-end digital oscilloscope featuring a bandwidth of around 10 GHz. Here, the limiting factor is clearly represented by the final price, unfortunately, it is not possible to realize this type of fast A/D converter using the researchers' own means and facilities. In the assessment of spectral properties of materials, the use of a spectral analyzer is normally a fully sufficient and financially effective method. This method, however, can not be utilized for the assessment of a system via the input-output technique,

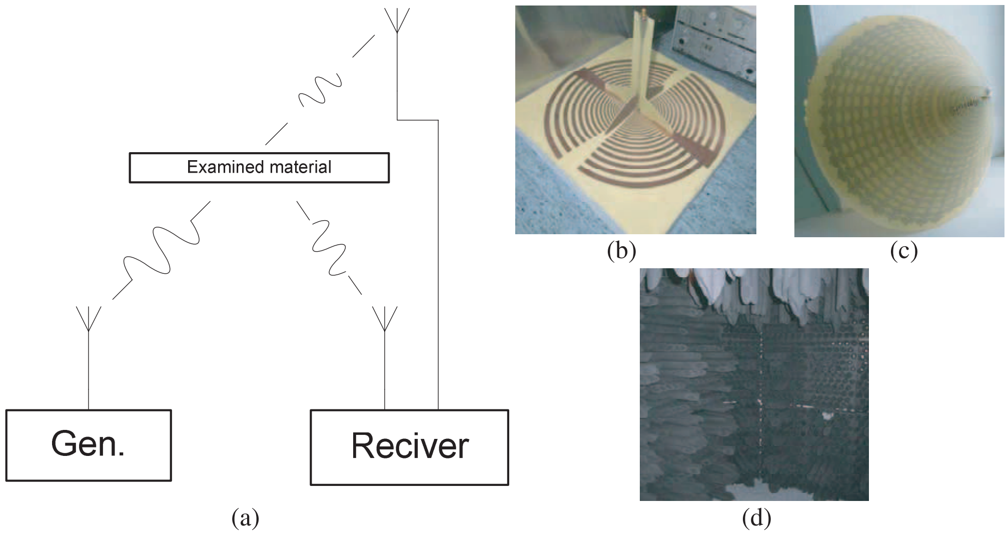



Figure 4: (a) The experiment configuration. (b) Planar log-periodic antenna. (c) Spiral fractal antenna. (d) Faraday cage.

for the correlation analysis, or for other techniques utilizing the measured signal waveform.

## 5. CONCLUSION

The research paper provides an elementary overview and description of laboratory equipment for the realization of noise spectroscopy measurement. In the text, it is generally noted that the search of an applicable source of noise will be markedly problematic in terms of the output power, which is normally low for broadband applications. Yet it is also noted that, for the laboratory tasks, considerably lower values of transmission power are sufficient for broadband systems in comparison with their "narrowband" counterparts. The designing of a broadband antenna is, to a great extent, an uneasy-to-solve problem, and the requirements placed on the width of the antenna transmitted band are very intensive. Another parameter to be mentioned in this respect is the linear transmission characteristics depending on the frequency. Noise spectroscopy for the frequency band of between 100 MHz and 10 GHz is realizable using the current means of technology.

## ACKNOWLEDGMENT

The research described in the paper was financially supported by the research programs MSM 0021630516 research plan MSM 0021630513, Ministry of Defence of the CR and Ministry of Industry and Trade of the CR (FR-TI1/001) and ACR 102/09/0314.

## REFERENCES

1. Maslovski, S., S. Tretyakov, and P. Alitalo, "Near-field enhancement and imaging in double planar polariton-resonant structures," *J. Appl. Phys.*, Vol. 96, 1293, 2004.
2. Freire, M. and R. Marques, "Near-field imaging in the megahertz range by strongly coupled magnetoinductive surfaces: Experiment and ab initio analysis," *J. Appl. Phys.*, Vol. 100, 063105, 2006.
3. Machac, J., P. Protiva, and J. Zehentner, "Isotropic epsilon-negative particles," *2007 IEEE MTT-S Int. Microwave Symp. Dig.*, Honolulu, USA TH4D-03, June 2007.
4. Fiala, P., E. Gescheidtova, and T. Jirku, "Tuned structures for special THz applications," *PIERS Proceedings*, 151–155, Beijing, China, March 23–27, 2009.
5. Protiva, P., J. Mrkvica, and J. Machác, "Universal generator of ultra-wideband pulses," *Radioengineering*, Vol. 17, No. 4, 74–79, 2008.

# Electromagnetic Wave Propagation in Heterogeneous Structures

**R. Kadlec**, **P. Fiala**, **and D. Nešpor**
Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejni 2906/4, Brno 612 00, Czech Republic

**Abstract**— The paper presents the problem of numerical modelling of high frequency electromagnetic waves propagation in inhomogeneous materials. For this method, a numerical model was prepared. The model was created in the MatLab and the COMSOL program environment For a layered heterogeneous medium, an algorithm was deduced for the reflection on several layers. The layers exist in the form of periodic structures which are composed of a homogeneous material. Reflection and refraction on heterogenous material are solved by means of the numerical method. Central in this respect are the refractions and reflections on the boundary of materials with different properties. This method is suitable for the design application of metamaterials. The deduced algorithm was projected for the visible spectrum.

## 1. INTRODUCTION

Generally, inhomogeneities and regions with different parameters appear even in the cleanest materials. During the electromagnetic wave passage through a material there occur, owing to the material characteristics such as conductivity, permittivity, or permeability, an amplitude decrease and a wave phase shift. If a wave impinges on an inhomogeneity, there occurs a change of its propagation. This change materializes in two forms, namely in reflection and refraction. In addition to this process, polarization and interference may appear in the waves.

For simple cases (such as a planar interface), the behaviour of an impinging wave can be calculated analytically by the help of Snell's refraction/reflection law and the Fresnel equations. However, in more complex structures it is difficult (and often infeasible) to perform an analytical calculation. Therefore, numerical methods are applied to facilitate the calculation process, and a wide range of programs like ANSYS, Comsol, or Matlab can be utilized in the realization of numerical modelling [1].

## 2. ELECTROMAGNETIC WAVES IN ISOTROPIC DIELECTRICS MATERIALS

In the Matlab program environment, algorithms were generated that simulate reflection and refraction in a lossy environment on the interface between two dielectrics. This section of the paper is linked to the previous modelling of light applying the related geometrical laws. The reflection and refraction is in accordance with Snell's law for electromagnetic waves as shown in Fig. 1 [2]. The form of Snell's law is

$$\frac{\sin\theta_0}{\sin\theta_2} = \frac{k_2}{k_1} = \frac{\sqrt{j\omega\mu_2 \cdot (\gamma_2 + j\omega\varepsilon_2)}}{\sqrt{j\omega\mu_1 \cdot (\gamma_1 + j\omega\varepsilon_1)}}, \tag{1}$$

where $k$ is the wave number, $\gamma$ is the conductivity, $\varepsilon$ the permittivity and $\mu$ the permeability. Relation (1) is defining for the boundary line between the dielectrics medium. Generally, $k_1$ and $k_2$ are complex; then angle $\theta_2$ is also complex. An electromagnetic wave is understood as the electric field strength and the magnetic field strength. The electric component incident wave according to Fig. 1 follows the formula

$$E_i = E_0 e^{-j\,k_1\,u_{n0}\cdot r}, \tag{2}$$

where $E_0$ is the amplitude electric field strength on the boundary line, $r$ is the positional vector, and $un_0$ is the unit vector of propagation direction.

The intensity of reflection beams and the intensity of refraction beams are expressed according to the formula

$$E_r = E_1 e^{-jk_1\vec{u}_{n1}\cdot\vec{r}}, \quad E_t = E_2 e^{-jk_2\vec{u}_{n2}\cdot\vec{r}}, \tag{3}$$

where $E_1$ is calculated from the intensity on boundary line $E_0$ and reflection coefficient $\rho_E$, and $E_2$ is calculated from the intensity on boundary line $E_0$ and transmission factor $\tau_E$:

$$E_1 = \rho_E \cdot E_0, \quad E_2 = \tau_E \cdot E_0. \tag{4}$$

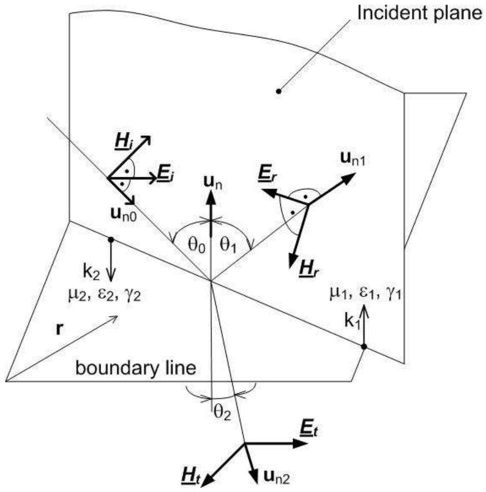



Figure 1: Reflection and refraction of light [2].

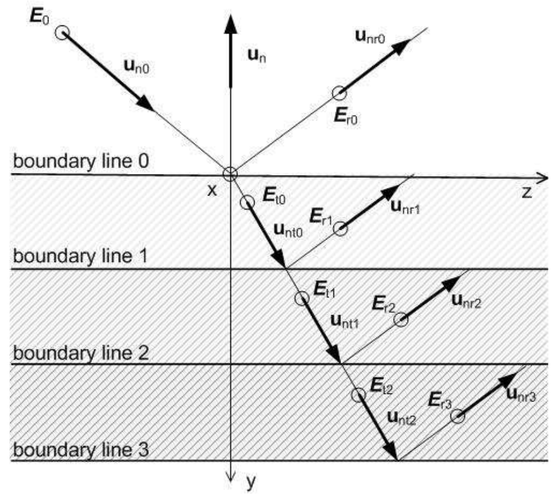



Figure 2: Reflection and refraction on a layered heterogenous material.

The calculation of reflection coefficient $\rho_E$ and transmission factor $\tau_E$ is according to these relations:

$$\rho_E = \frac{E_1}{E_0} = \frac{Z_{v2}\cos\theta_1 - Z_{v1}\cos\theta_2}{Z_{v2}\cos\theta_1 + Z_{v1}\cos\theta_2}, \quad \tau_E = \frac{E_2}{E_0} = \frac{2Z_{v2}\cos\theta_1}{Z_{v2}\cos\theta_1 + Z_{v1}\cos\theta_2}. \tag{5}$$

For numerical modelling, there is a suitable relation in the form of

$$E_r = \frac{\mu_2 k_1 \cos\theta_0 - \mu_1\sqrt{k_2^2 - k_1^2 \sin^2\theta_0}}{\mu_2 k_1 \cos\theta_0 + \mu_1\sqrt{k_2^2 - k_1^2 \sin^2\theta_0}} E_0 \cdot e^{-jk_1 u_{n1}\cdot r}, \quad E_t = \frac{2\mu_2 k_1 \cos\theta_0}{\mu_2 k_1 \cos\theta_0 + \mu_1\sqrt{k_2^2 - k_1^2 \sin^2\theta_0}} E_0 \cdot e^{-jk_2 u_{n2}\cdot r}. \tag{6}$$

These relations are calculated from the basic variable and they facilitate an acceleration of the calculation process.

Interpretation of the Fresnel equations and Snell's laws is simple in the case of the refraction on boundary line between the dielectrics medium. In case of refraction in a lossy medium, angle $\theta_2$ is complex. According to relation (1), angle $\theta_2$ depends on wave numbers $k_1$ a $k_2$, which are generally complex; then, in medium 2 an inhomogeneous wave is propagated.

For a layered heterogeneous medium, an algorithm is deduced for the reflection on several layers. The reflection and refraction on a heterogonous material is solved by the help of the numerical method. The reflection on a layered material on $n$ layers generates $n$ primary electromagnetic waves, according to Fig. 2. The interpretation of propagation of electromagnetic waves on a layered heterogeneous medium is according to relation

$$E_{rl} = E_{il}\rho_{El} \cdot e^{-j\,k_{(l+1)}\,u_{nrl}\cdot r_l}, \quad E_{tl} = E_{il}\tau_{El} \cdot e^{-j\,k_{(l+2)}\,u_{ntl}\cdot r_l}, \tag{7}$$

where $E_{rl}$ a $E_{tl}$ are the reflection and refraction electromagnetic waves on the boundary line $(l = 0, \ldots, \max)$ according to Fig. 2, $E_{il}$ is the amplitude electric field strength on boundary line $l$, $\rho_{El}$ a $\tau_{El}$ are the reflection coefficient and transmission factor on boundary line $l$, $k_l$ is the wave number of layer, $r_l$ is the electromagnetic wave positional vector on boundary line $l$, $u_{ntl}$ and $u_{nrl}$ are the unit vectors of propagation direction.

## 3. PLANE WAVE PROPAGATION IN METAMATERIALS

Metamaterials are artificial structures which show electrical and magnetic characteristics (permittivity, permeability) not present in the natural environment. The materials are composed of small segments such as thin conductive wires or planar coils; thus, as a matter of fact, these materials can be referred to as composites. However, owing to the fact that the electromagnetic wave of a given spectrum has a wavelength which is substantially longer in comparison with that of the individual components of a structure, the material can be regarded as homogeneous.

Generally, materials with negative parameters constitute a group of media that possesses a negative value of relative permittivity $\varepsilon_r$ or relative permeability $\mu_r$ [1]. Materials with negative permittivity are commonly found in the natural environment [1]. In this respect, the best known items are various types of low-lossy plasma, metals and semiconductors for electromagnetic waves in the optical and infrared spectra. Conversely, materials with negative permeability are less common in the natural environment. Only in ferromagnetics the mutual interaction of magnetic forces is sufficiently large and losses adequately small to facilitate the generation of negative permeability regions. Ferrites are magnetised up until saturation and, simultaneously, the permeability tensor is close to resonance. These materials are widely present in microwave applications.

By substituting wave equations into Maxwell's equations [1] we obtain, for the situation in Fig. 1, the relation

$$n \times E = \omega \mu H, \quad n \times E = -\omega \varepsilon E. \tag{8}$$

Equation (8) holds if the permittivity and permeability are positive. If these constants are negative (as shown in Fig. 3(c)), the equations can be written in the form

$$n \times E = -\omega |\mu| H, \quad n \times E = \omega |\varepsilon| E. \tag{9}$$

Among the most notable characteristics of materials with negative parameters is the negative refraction, namely the beam refraction to the other side from the perpendicular line during the passage through the interface. It is apparent from Fig. 3(c) that the direction of normal component $n_{2n}$ rotates, and therefore there also occurs rotation of the permeable wave direction. In view of the signs of permittivity and permeability, materials can be classified into four groups theoretically defined by Victor Veselago [1].

I. Both the permittivity and permeability are positive. The related items fall within the category of isotropic dielectrics; the behaviour of an electromagnetic wave on the interface of these materials corresponds to the description provided in Chapter 2.

II. Negative permittivity and positive permeability. The related items are metals and semiconductors for electromagnetic waves in the optical and infrared spectra.

III. Both the permittivity and permeability are negative. The related items are referred to as left-handed materials. At the incidence of a wave from the environment of group I into the environment of group III there occur a rotation of 180° of the Poyinting vector. The beam refracts to the other side from the perpendicular line as compared to the environment of group I.

IV. Positive permittivity and negative permeability; this combination applies to ferromagnetic materials.

Figure 3 shows numerical models in the Comsol program for the transition from non-lossy environment I. to non-lossy environment III. and vice versa. In Fig. 3, the following situations are described:

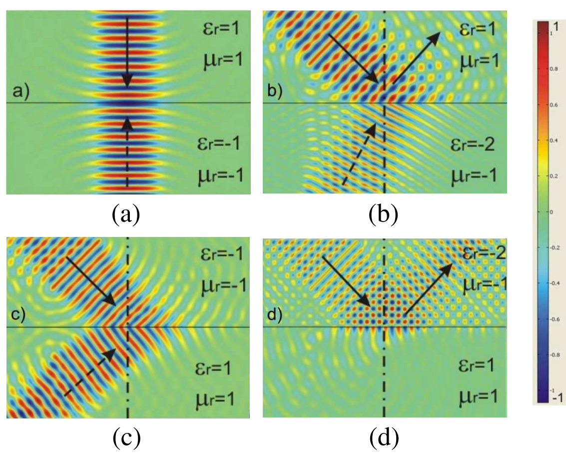



Figure 3: Calculations of refractions and reflections of waves on a planar interface for various material constants.

a) A wave during the passage from a material with parameters ($\varepsilon_r = 1, \mu_r = 1$) into a material with parameters ($\varepsilon_r = -1, \mu_r = -1$) at perpendicular incidence of the wave on the interface. A permeable wave has the same velocity. The Poynting vector rotates, and therefore there also occurs rotation of the wave propagation direction. During the passage into a material with parameters ($\varepsilon_r = -6, \mu_r = -1$), the permeable wave has a lower velocity.
b) A wave during the passage from a material with parameters ($\varepsilon_r = 1, \mu_r = 1$) into a material with parameters ($\varepsilon_r = -2, \mu_r = -1$) at the incidence of the wave on the interface at an angle of 45°. The permeable wave refracts to the other side from the perpendicular line (to the third quadrant). There forms a reflected wave; during the passage into a material with parameters ($\varepsilon_r = -1, \mu_r = -1$) there does not occur any reflection.
c) A wave during the passage from a material with parameters ($\varepsilon_r = -1, \mu_r = -1$) into a material with parameters ($\varepsilon_r = 1, \mu_r = 1$) at the incidence of the wave on the interface at an angle of 45°. The permeable wave refracts to the other side from the perpendicular line (to the third quadrant). There does not occur any reflection.
d) A wave during the passage from a material with parameters ($\varepsilon_r = -2, \mu_r = -1$) into a material with parameters ($\varepsilon_r = 1, \mu_r = 1$) at the incidence of the wave on the interface at an angle of 45°. There occurs total reflection.

## 4. CONCLUSIONS

Numerical modelling of wideband electromagnetic signals on field of multilayer and periodic structure optical materials in Matlab program is very time demanding. This method is suitable for specific purposes of detail analysis.

Algorithms created in the Matlab environment are verified by the help of programs based on the finite element method, namely programs such as Comsol and ANSYS. The paper includes a theoretical analysis and references to the generated algorithms, which are verified using numerical models.

Negative permittivity can be constructed on a periodic structure of electrically conductive wires, upon which a wave impinges vertically. The structure is commonly referred to as wire media. Another technique of constructing negative permeability is split ring resonators (SRRs). Several models were prepared in the Comsol program for negative permittivity materials.

## ACKNOWLEDGMENT

The research described in the paper was financially supported by the research program MSM 0021630516 and research plan MSM 0021630513, Ministry of Defence of the CR, Ministry of Industry and Trade of the CR (Diagnostics of Superfast Objects for Safety Testing, FR-TI1/368), and Czech Science Foundation (102/09/0314) and project of the BUT Grant Agency FEKT-S-10-13.

## REFERENCES

1. Nešpor, D., "Electromagnetic wave propagation study in heterogeneous structures," Ph.D. thesis, 2009.
2. Dedek, L. and J. Dedková, *Elektromagnetismus. 2*, 232, Brno, VITIUM, 2000.
3. Moss, C., "Numerical metods for electromagnetic wave propagation and scattering in comlex media," 240, 2004, available from http://portal.acm.org/citation.cfm?id=1023429.
4. Stratton, J. A., *Teorie Elektromagnetického Pole*, STNL, Praha, 1961.
5. Zhang, J. J., J. T. Huangfu, Y. Luo, H. S. Chen, J. A. Kong, and B.-I. Wu, "Cloak changing with background," *PIERS Online*, Vol. 4, No. 5, 596–600, 2008.
6. Fiala, P., "Finite element method analysis of a magnetic field inside a microwave pulsed generator," *2nd European Symposium on Non-lethal Weapons*, Ettlingen, SRN, May 13–15, 2003.

# Errors in Diffusion Coefficients Measurement

**P. Marcon**[1] **and K. Bartusek**[2]

[1]Department of Theoretical and Experimental Electrical Engineering, Brno University of Technology
Kolejni 2906/4, 612 00 Brno, Czech Republic
[2]Institute of Scientific Instruments, Academy of Sciences of the Czech Republic
Kralovopolska 147, 612 64 Brno, Czech Republic

**Abstract**— The magnetic resonance (MR) imaging techniques of tomography and spectroscopy are exploited in many applications. The quality of dynamic behavior of gradient magnetic fields is one of the important properties of devices exploiting the phenomenon of magnetic resonance (MR) for imaging or diffusion measurement [1]. A necessary precondition for the facilitation of MR instruments proper functioning consists in the maintenance of a high degree of homogeneity of the fundamental and gradient magnetic fields [2, 3]. Gradient changes induce eddy currents in the surrounding conducting arrangement, thus potentially causing image artifacts, localization errors, and signal distortion. While the use of actively shielded gradients has greatly reduced the magnitude of eddy currents, significant distortion often remains, mainly in the short time interval after the gradient is switched off. Residual eddy currents may require further reduction. This is frequently achieved by pre-emphasis correction in the relevant gradient channel and in the homogeneous $B_0$ shim or by active shielded gradient coils [4]. The rise time of the pulse of a magnetic field or its decrease to the level of non-homogeneity of the basic magnetic field should be as short as possible ($< 100\,\mu$s).

The experiments were carried out on an MR tomograph system 4.7 T/120 mm (i.e., 200 MHz for $^1$H nuclei). Actively shielded gradient coils yield a maximum gradient field magnitude of 180 mT/m. The data measured were processed in the MAREVISI and MATLAB programs.

The $b$-factor calculation software was created to characterize the device artifacts on the accuracy of measurement. It is possible to determine the diffusion measurement errors for changes of gradient field magnitude and inaccuracies of time parameters settings by this software. We characterized probable diffuse measurement errors and, primarily, the effect of gradient decay time on the accuracy of diffusion constants measurement of the chosen samples. The results of the simulation were experimentally tested and compared with the analytical model. The determination of the device artifacts effects with the diffusion constants measurement is useful mainly for biological samples with short relaxation times.

## 1. INTRODUCTION

Diffusion is the random translational motion of molecules given by their thermal energy. In the course of the motion there occur collisions between the molecules and microscopic obstacles generated by components such as cell structure. Thus, diffusion motion can be restricted to a certain (great or small) extent. In theory, if no obstacles are assumed and the molecular environment is considered isotropic, there holds the Einstein equation:

$$y = \sqrt{6 \cdot D \cdot t_D}, \tag{1}$$

where $y$ expresses the medium distance (in all three directions) covered by the molecules in time $t_D$. The constant of proportionality $D$ is the so-called diffusion coefficient, which indicates the rate of moleculear mobility in the given substance. During time $t_D$ the molecules can pass through regions with different $D$ coefficients, by means of which there originates the coefficient time dependence.

The diffusion measurement principle consists in a suitable application of gradients of the magnetic field (spatially dependent magnetic field) into the NMR measurement sequence. The moving spins (atoms) are, at the time of dephasing, in a different place (and, owing to the application of the magnetic field gradients, in a different magnetic field) as compared to the time of backphasing. Errors in the measurement of diffusion coefficients are dependent on several factors 5; this paper contains an analysis of two members of this group, namely the magnetic field inhomogeneity and the dependence of gradient characteristics on the diffusion measurement accuracy. The diffusion was measured and evaluated by the help of three techniques, whose evaluation is provided in the text below.

## 2. EXPERIMENT

The experiment was carried out using an MR tomograph available at the Institute of Scientific Instruments, Academy of Sciences of the Czech Republic (ISI ASCR). The MR tomograph provides static field flux density $B_0 = 4.7\,\mathrm{T}$; the nuclei resonance frequency is 200 MHz. We used a deionized water sample as the test medium, and the realization of the diffusion measurement took place at the constant temperature of 24.6°C. All calculations of relative errors are related to nominal value $D = 2.376 \cdot 10^{-9}\,\mathrm{m \cdot s^{-2}}$, which was published in source [6]. The utilized measurement technique consisted in the method of spin echo with the applied magnetic field gradients (the pulse field gradient spin echo — PFGSE). This method developed by Stejskal [7, 8] is described in Figure 1. In the process of measurement we observed the following time constants: $\delta = 3\,\mathrm{ms}$, $\Delta = 11.18\,\mathrm{ms}$, $T_E = 20\,\mathrm{ms}$. During the measurement, gradient magnitude $G_D$ was variable and determined by the magnitude of the Digital Analog Convertor (DAC). For the value of DAC = 27 k, the gradient magnitude equals to $G_D = 180\,\mathrm{mT/m}$.

In order to compensate the magnetic field inhomogenities, we performed also the measurement for both polarities of the gradient pulses [9, 10].

## 3. RESULTS AND DISCUSSION

First of all, at the temperature of 24.6°C, we measured image intensities in the deionised water sample at different magnitudes for the gradient. The magnitude is determined by the value of DAC (10 k, 15 k, 20 k, 25 k) at both polarities. From the thus measured values we evaluated the diffusion. Here, the first step consisted in examining the accuracy of results analyzed using two images in the Marevisi program. The diffusion was calculated from an image without the gradient as well as from another image that had been obtained at a different-sized gradient pulse in all coordinate directions. Relative measurement errors related to the nominal value of $D = 2.376 \cdot 10^{-9}\,\mathrm{m \cdot s^{-2}}$ are represented in Figure 2. At this point, measurement errors can be compensated for in such a manner that the measured diffusion value for the given absolute value of gradient is calculated as a diameter of the diffusion obtained from the two images for the positive gradient and the negative gradient.

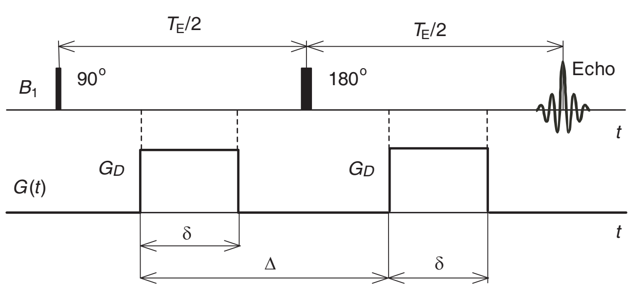



Figure 1: The MR spin-echo method of diffusion coefficients measurement utilizing gradient magnetic fields.

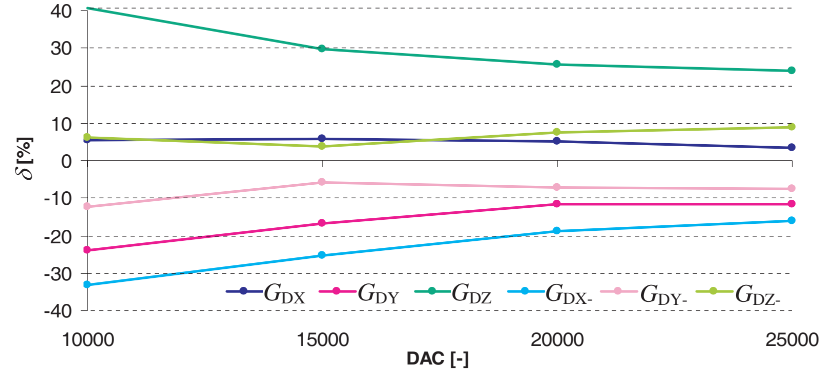



Figure 2: The dependence of relative error on the magnitude of gradients in the measurement of deionized water diffusion: evaluation from two images, temperature 24.6°C.

Another possibility of compensating for magnetic field inhomogenities consists in the evaluation of diffusion from three images. Here, the obtained results are more precise as, using the Matlab program, the first image is evaluated without the gradient field, the second one with the positive gradient pulse, and the third one with the negative gradient pulse. Thus, compensation is secured for the magnetic field inhomogeneities and non-linearities (Figure 3).

The procedure that showed the highest degree of accuracy in respect of the measurement and evaluation of diffusion was the measurement of a series of images at different gradient magnitudes and the subsequent approximation of these images in the Marevisi program [11]. An illustration of images obtained during the measurement with negative gradient pulses in the $x$ axis direction is provided in Figure 4. The series of images was obtained at gradient pulse magnitudes determined by DAC $= -10$ k, $-15$ k, $-20$ k, $-25$ k.

Figure 5 presents the obtained dependences of diffusion on the number of applied gradient pulses from which the diffusion constant was approximated. The measurement was performed for all the coordinate system directions and at both positive and negative polarities of the gradient pulse. The meaning of numbers on the $x$ axis is as follows: 5 = measured at |DAC| = 0, 10 k, 15 k, 20 k,

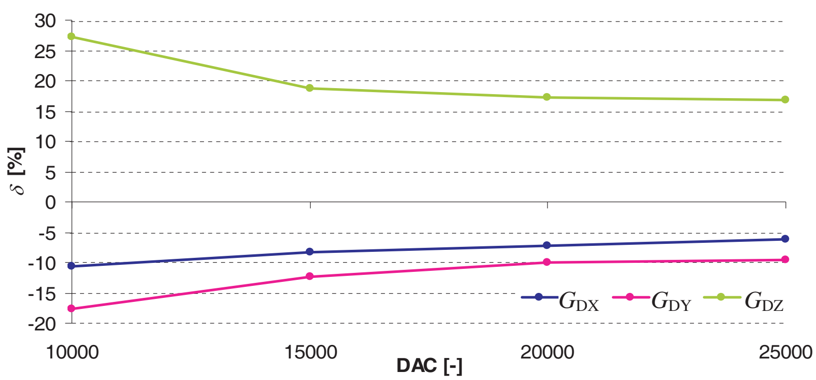



Figure 3: The dependence of relative error on the magnitude of gradients in the measurement of deionized water diffusion: evaluation from three images, temperature 24.6°C.

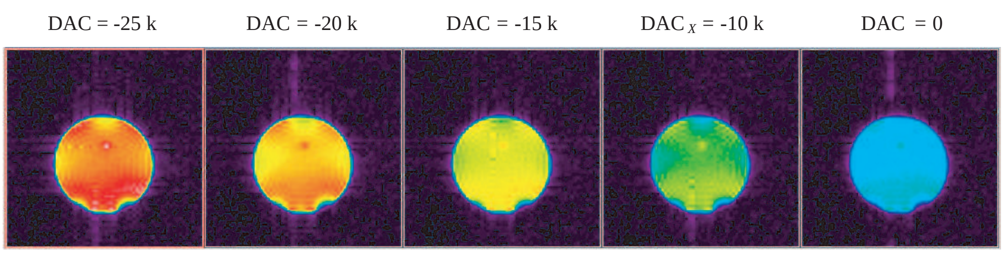



Figure 4: Diffusion images for the gradient $G_{DX}$ magnitudes.

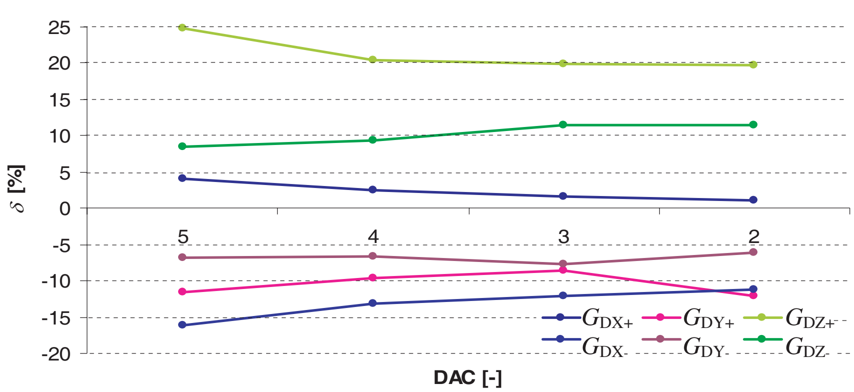



Figure 5: Diagram: diffusion evaluation from approximated images obtained at different DAC (positive and negative polarities evaluated separately), temperature of 24.6°C.

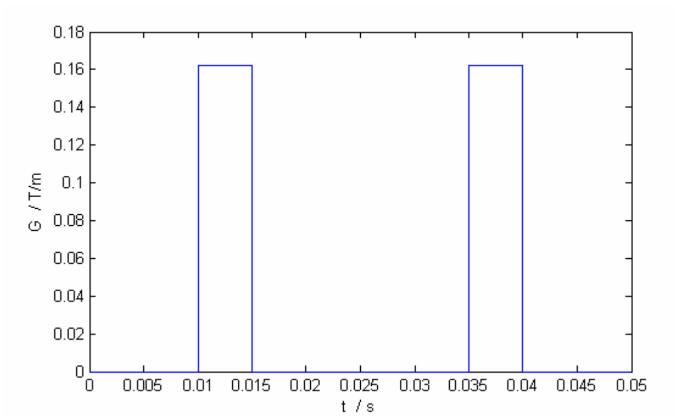



Figure 6: The course of gradient pulses.

Table 1: Comparison of the applied measurement methods and evaluation of diffusion.

| | $D$ [m$^2$/s] | | |
|---|---|---|---|
| | Evaluation from 2 images | Evaluation from 3 images | Approximation at different $G$ |
| $G_{DX}$ | $2.15880 \cdot 10^{-9}$ | $2.1845 \cdot 10^{-9}$ | $2.2510 \cdot 10^{-9}$ |
| $G_{DY}$ | $2.08756 \cdot 10^{-9}$ | $2.0815 \cdot 10^{-9}$ | $2.1820 \cdot 10^{-9}$ |
| $G_{DZ}$ | $2.81258 \cdot 10^{-9}$ | $2.8505 \cdot 10^{-9}$ | $2.7485 \cdot 10^{-9}$ |

25 k; 4 = 10 k, 15 k, 20 k, 25 k; 3 = 15 k, 20 k, 25 k; 2 = 10 k, 25 k. The diagram shows that it is sufficient to perform the measurement in gradients |DAC| = 20 k and 25 k and, then, approximate the obtained images. Thus, we obtain the diffusion value separately for the image measured at positive gradient values and, in the same manner, for the image measured at negative gradient values. The results can be made more precise if we average out the curve approximated for the positive gradient pulse with the corresponding curve approximated for the negative gradient pulse. Again, there will occur the compensation for the magnetic field inhomogenities.

Further, using the Matlab program, the program has been developed that facilitates investigation into the influence of gradient field shape on the measurement accuracy; thus, the compensation for diffusion measurement errors is also made possible. Real gradient pulses are not rectangular as is evidenced by Figure 6: They have entering and descending edges. The designed program realizes the numerical modelling also of these courses and calculates value $b$. Constant $b$ (referred to as the $b$-factor) indicates the pulse sequence sensitivity to the diffusion, and it is given by integral 5:

$$b = \gamma^2 \int_0^{T_E} \left[ \int_0^t G\left(t'\right) dt' \right]^2 dt. \tag{2}$$

Figure 6 shows the gradient simulated course in the Matlab program. By the help of the numerical model, the $b$-factor is calculated with great accuracy for the given course of the gradient. For the course provided in the image, the $b$-factor value is $b = 1.090 \cdot 10^9\,\mathrm{s \cdot m^{-2}}$. The theoretical value according to formula (2) yields an almost identical result: $b = 1.094 \cdot 10^9\,\mathrm{s \cdot m^{-2}}$.

## 4. CONCLUSION

The obtained diffusion values averaged out of the diffusion values measured at both polarities of the gradient pulse (Table 1) at the temperature of 24.6°C differ from the theoretical value $D = 2.376 \cdot 10^{-9}\,\mathrm{m \cdot s^{-2}}$ by a maximum of 18.3%. The diffusion values are not identical in all directions; especially in the $z$ axis there are differences as against the values of diffusion in axes $x$ and $y$. It follows from the values in Table 1 that the diffusion coefficients obtained by means of evaluation from two images (the mean of values measured at both positive and negative gradient pulses) show almost no difference from those obtained using three images. The most accurate values are yielded by means of approximation from images obtained at different gradient value. The related future research will focus on the achievement of a greater degree of accuracy in diffusion measurement and on the statistic processing of the results obtained.

## ACKNOWLEDGMENT

The research described in the paper was financially supported by research plans, No. MSM 00216305-16, and a grant from the Ministry of Industry and Trade of the CR, MPO No. FR-TI1/001 and project of the BUT Grant Agency FEKT-S-10-13.

## REFERENCES

1. Johansen-Berg, H. and T. E. J. Behrens, *Diffusion MRI: From Quantitative Measurement to in vivo Neuroanatomy*, 490, ISBN 978-0-12-374709-9, Elsevier, China, 2009.
2. Vlaardingerbroek, M., *Magnetic Resonance Imaging*, Springer-Verlag, 2000.
3. Bartusek, K. and E. Gescheidtova, "MR Measurement technique of the rapidly switched gradient magnetic fields in MR tomography," *Applied Magnetic Resonance*, (IF 0,665), Vol. 29, No. 12, 675–686, ISSN 0937-9347, 2006.
4. Mansfield, P. and B. Chapman, "Active magnetic screening of gradient coils in NMR imaging," *Journal of Magnetic Resonance*, Vol. 66, 573–576, 1986.
5. Bartušek, K., "Special methods of diffusion coefficients measurement by use of nuclear magnetic resonance," Inaugural Thesis, ISSN 1213-418X, Brno Univeristy of Technology, Brno, 2007.
6. Holz, M., S. R. Heil, and A. Sacco, "Temperature-dependent self-dfiffusion coefficients of water and six selected molecular liquids for calibration in accurate $^1$H NMR PFG measurements," *PCCP*, No. 2, 4740–4742, 2000.
7. Morris, P. G., *Nuclear Magnetic Resonance Imaging in Medicine and Biology*, Oxford University Press, Oxford, 1986.
8. Bartušek, K. and E. Gescheidtová, "MRI method of diffusion measurement in heterogeneous materials," *Measurement Science and Technology* (IF 1,297), Vol. 19, 1–8, ISSN 0957-0233, 2008.
9. Gescheidtová, E., R. Kubásek, Z. Smékal, and K. Bartušek, "Digital filter banks in MR measurement of gradient magnetic fields," *Applied Magnetic Resonance*, (IF 0,706), Vol. 33, No. 4, 399–417, ISSN 0937-9347, 2008.
10. Bartušek, K. and E. Gescheidtová, "Testing the quality of magnetic gradient fields for studying self-diffusion processes in biological specimens by magnetic resonance methods," *Measurement Science and Technology*, (IF 1,118), Vol. 17, 2256–2262, ISSN 0957-0233, 2006.
11. Mikulka, J., E. Gescheidtová, and K. Bartušek, "Perimeter measurement of spruce needles profile using MRI," *PIERS Proceedings*, 1128–1131, Beijing, China, March 23–27, 2009.

# Measuring of Temperature Fields Using MR Tomography

**M. Cap**[1], **P. Marcon**[1], **and K. Bartusek**[2]

[1]Dept. of Theoretical and Experimental Electrical Engineering, Brno University of Technology
Kolejní 2906/4, 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**— Magnetic resonance imaging is in medicine commonly used method for human body structures imaging. Progresses in imaging methods allow figure different parameters of the specimen then mere structural image. If we are able to image function of the human body then we can improve diagnosability of diseases. This article deal with noninvasive method for imaging of temperature distribution in required volume of specimen using MR tomography. Method is based on diffusion measurement [1, 4]. Temperature dependence of diffusion of materials makes possible measure temperature distribution in specimen [2].

As a specimen is used water solution of nickel sulfate ($NiSO_4$). This solution was chosen for similar properties to physiological solution. Pulsed Field Gradient Spin-Echo (PFG SE) method is used for measurement of self diffusion coefficients $D$ [3]. Results of measurement are compared with other materials measured in [2] to improve method for human body temperature fields measurement. The experiments were carried out on an MR tomograph system 4.7 T/120 mm (i.e., 200 MHz for $^1$H nuclei). Actively shielded gradient coils forms a maximum gradient field magnitude of 180 mT/m. The measured data were processed in the MAREVISI and MATLAB programs.

## 1. INTRODUCTION

Nuclear magnetic resonation is one of the most important imaging techniques for structural imaging of body structures. In many areas like MR imaging, MR spectroscopy or functional MRI is technique of magnetic resonance powerful diagnostic tool in medicine and biology. MR image can be T1, T2, proton density and diffusion weighted. Last named method, diffusion weighted imaging, is way how to measure distribution of temperature. Diffusion is a mass transport process arising in material, which results in molecular or particle mixing without requiring bulk motion. Diffusion should be not be confused with convection or dispersion, other transport mechanisms that require bulk motion to carry particles from one place to another [4]. Temperature makes difference in diffusion depending on different substances. Measuring this dependence makes possible temperature field mapping. This article deal with measuring temperature dependence for different substances.

## 2. EXPERIMENT

Experiment was accomplished on a MR tomograph at the Institute of Scientific Instruments, Academy of Sciences of the Czech Republic (ISI ASCR). The MR tomograph dispose of static field flux density $B_0 = 4.7\,\mathrm{T}$, $^1$H nuclei resonance frequency is 200 MHz.

Testing of diffusion temperature dependence was provided with three water solutions. All of used water solutions are presented in Table 1. Range of measured temperatures was set from 20°C (68°F) to 29.6°C (85.3°F). Temperature inside MR tomograph is regulated by temperature regulating system made by ISI ASCR Brno. Each measurement was provided after 30 minute temperature stabilization interval.

Table 1: Used water solutions.

| Sample Number | Sample |
|---|---|
| 1 | $H_2O + NiSO_4$ |
| 2 | $H_2O + 4 \times NiSO_4$ |
| 3 | Deionized watter |

Temperature fields mapping method use PFGSE (pulse field gradient spin echo) method for gradient measuring. This method is most used method for diffusion measurement was developed

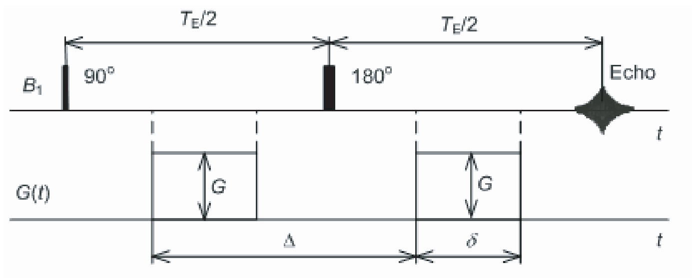



Figure 1: The RF pulse and gradient field waveforms of the pulsed field gradient spin-echo method.

by Stejskal [5]. Principles of this method are presented on Figure 1. Acquisition parameters are: $\delta = 3.5\,\text{ms}$, $\Delta = 16.8\,\text{ms}$, $T_E = 20\,\text{ms}$. $D$ is determined by digital-analog converter DAC = 25 000.

Value of $b$-factor is necessary for diffusion calculation. $b$-factor equation is equal to [6]

$$b = \gamma^2 G^2 \delta^2 \cdot \left(\Delta - \frac{\delta}{3}\right). \tag{1}$$

Diffusion was computed from three pictures for magnetic field inhomogeneity compensation.

These three pictures were measured without gradient pulse, with positive gradient pulse and with negative gradient pulse. After $b$-factor computing and all three data sets measuring is diffusion computed as:

$$D = \frac{1}{b} \cdot \ln\left(\frac{M_+ \cdot M_-}{M_0^2}\right), \tag{2}$$

where:

$b$ $b$-factor [s/m$^2$],
$M_0$ data measured without gradient pulse,
$M_+$ data measured with positive gradient pulse,
$M_-$ data measured with negative gradient pulse.

All data was processed in Marevisi and MATLAB program.

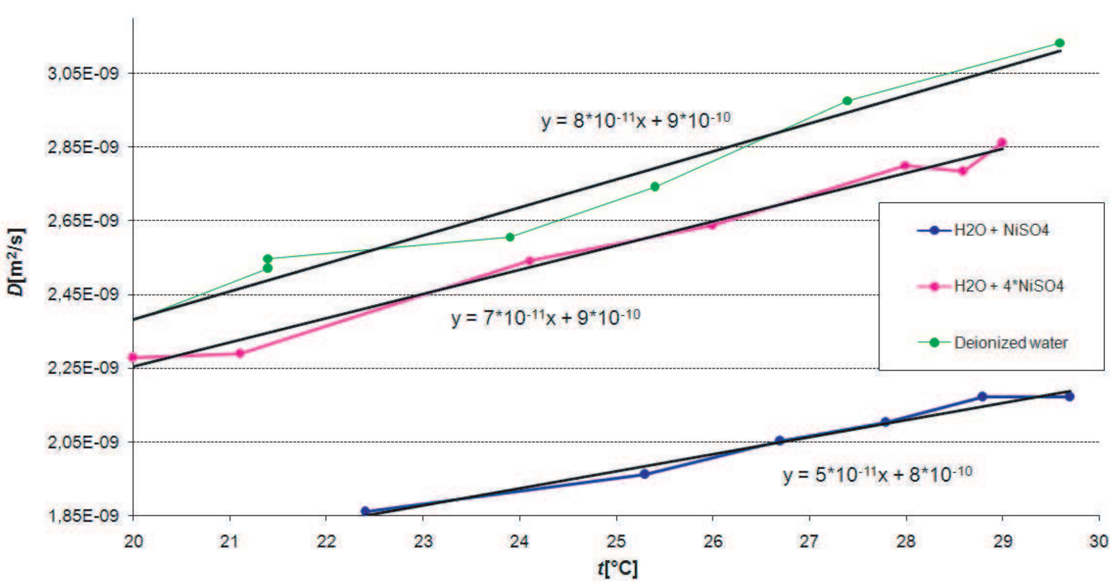
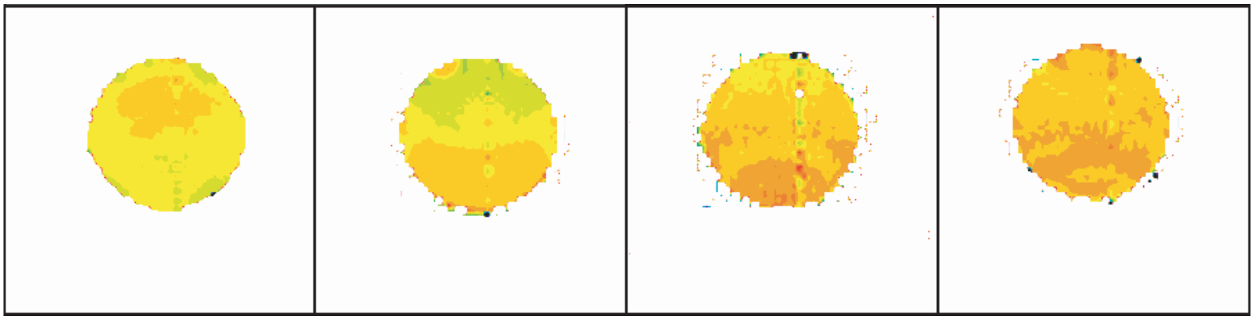



Figure 2: Temperature dependence of diffusion for deionized water and two water solutions.



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

Figure 3: Diffusion pictures of deionized water measured at temperatures. (a) 21.4°C, (b) 23.9°C, (c) 27.4°C, (d) 29.6°C.

## 3. RESULTS AND DISCUSSION

In Figure 2 are shown diffusion dependence on temperature for different water solutions presented in Table 1. Diffusion dependence on temperature for $H_2O+NiSO_4$, $H_2O+4\times NiSO_4$ and demonized water are almost parallel. All curves are line approximated with calculated equation of regression. Regression equation of ionized water was compared with regression equation of ionized water published in [2]

$$y = 6 \times 10^{-11} x + 9 \times 10^{-10}. \tag{3}$$

In Figure 3 diffusion pictures created in Marevisi are presented. These pictures show distribution of diffusion in sample slice. Diffusion is in the pictures modulated in the range of colors from yellow to red.

## 4. CONCLUSION

Diffusion weighted image acquired by PFGSE method gives possibility for temperature dependence of diffusion measurement. Measurements were provided for three water solutions and finally pictures computed from three pictures measured with different gradient. Temperature dependencies of diffusion are in selected temperature range linear and almost parallel for all three measured substances. With knowledge of this dependence for other substances is possible to measure temperature by NMR and map volume distribution of temperature.

## ACKNOWLEDGMENT

The research described in the paper was financially supported by FRVŠ (a fund of university development) by research plan No. MSM 0021630513 ELCOM, No. MSM 0021630516 and grant Czech ministry of industry and trade MPO No. FR-TI1/001, GACR 102/09/0314.

## REFERENCES

1. Vlaardingerbroek, M., *Magnetic Resonance Imaging*, Springer-Verlag, 2000.
2. Holz, M., S. R. Heil, and A. Sacco, "Temperature-dependent self-diffusion coefficients of water and six selected moleculer liquids for calibration in accure $^1$H NMR PFG measurement," *Chem. Phys.*, Vol. 2, 4740–4742, 2000.
3. Bartusek, K. and E. Gescheidtova, "Measurement technique of the rapidly switched gradient magnetic fields in MR tomography," *Acta Polytechnica*, Vol. 4, 30–35, 2003.
4. Johansen-Berg, H. and T. E. J. Behrens, "Diffusion MRI: From quantitative measurement to in vivo neuroanatomy," *Elsevier*, China, 2009.
5. Morris, P. G., *Nuclear Magnetic Resonance Imaging in Medicine and Biology*, Oxford Univ. Press, Oxford, 1986.
6. Bartusek, K., "Special methods of diffusion coefficients measurement by use of nuclear magnetic resonance," Inaugural Thesis, Brno University of Technology, Brno, 2007.

# 3D Reconstruction in Magnetic Resonance Imaging

**J. Mikulka[1] and K. Bartusek[2]**

[1]Department of Theoretical and Experimental Electrical Engineering
Brno University of Technology, Kolejni 2906/4, 612 00 Brno, Czech Republic
[2]Institute of Scientific Instruments, Academy of Sciences of the Czech Republic
Kralovopolska 147, 612 64 Brno, Czech Republic

**Abstract**— This article deals with three-dimensional reconstruction methods of nuclear magnetic resonance images. The testing images were observed by tomography with basic magnetic field of 4.7 T at the Institute of Scientific Instruments (Academy of Sciences of the Czech Republic). 10 slices of the testing phantom were acquired. The methods were proposed with the aim of getting utmost information about the shape of the testing phantom. One possible way is to increase the number of the sensed slices, but it implies decreasing of the signal to noise ratio. The second approach is finding the compromise between the effective count of slices and the following interpolation of other slices between the sensed ones. The both approaches were compared. The resultant images were segmented by the active contour methods which are based on partial differential equations solution [1, 2]. The advantage of these methods is that the images may not be preprocessed before segmentation. The appropriate segmented slices were compared. The following image processing leading to 3D model creation is proposed.

## 1. INTRODUCTION

The advantages of the nuclear magnetic resonance (NMR) were described in many publications. It is approach to acquisition of spatial data of soft tissues. The main advantage is absolutely the fact of unproved negative effects of the electromagnetic radiation to human organism subject to prescribed hygienic regulations. Against to other tomographic approaches the magnetic resonance images are with the higher resolution and quality. The observed images of sensed object can be used for three-dimensional model creation after the application of suitable preprocessing methods. The reconstructed object can be useful for example to the better diagnosis in medical sciences, for quantitative or qualitative description of tissues, tumors etc. The requirement is to have utmost slices for the three-dimensional model of sensed tissue. The signal-to-noise ratio (SNR) decreases with increasing the number of slices [1]. We have to specify compromises between the number of observed slices and satisfying SNR in images individually. However, if we actually need more slices for the 3D model creation, we can calculate the other images between observed ones. Two basic approaches are described for images interpolation of MR spatial data in this paper. The first method is based on ordinary averaging of neighboring pixels intensities with the same position in the frequency domain. We can get the images describing the real scene by Fourier transformation of each slices in the $k$-space, which we observed by tomography. We can get $(2n-1)$ images from original $n$ images. This algorithm is repeatable. By the second approach we process directly the data in $k$-space. There is created a vector of complex numbers by data observing with the same positions in the $k$-space. Then the frequency domain of this vector is obtained by Fourier transformation [2, 3]. The frequency domain is extended by zero values for higher frequency part. This extended spectrum is transformed back to the time domain by the inverse Fourier transformation. The length of the obtained vector depends on number of zeros included to frequency spectrum of the original vector. The slices of 4 flasks with water were inserted to work space of tomography with the magnetic field 4.7 T (200 MHz for $^1$H nuclei) for the testing. This phantom is shown in the Figure 1.

## 2. OBTAINING OF THE MR IMAGES

The images obtained by MR tomography in the time domain ($k$-space) are then transformed by Fourier transformation to the frequency domain. These images in the frequency domain represent the real scenes. The first method interpolates the three-dimensional data in the transformed frequency domain. The second method shows the possibility of data processing and interpolating the three-dimensional data in the original time domain — $k$-space. We can see in Figure 2 the procedure of scanned spatio-temporal slices processing and the transformation of the slices to the frequency domain, which represents the real scene.

Figure 1: The phantom for testing process of proposed interpolation methods of spatial MR data.

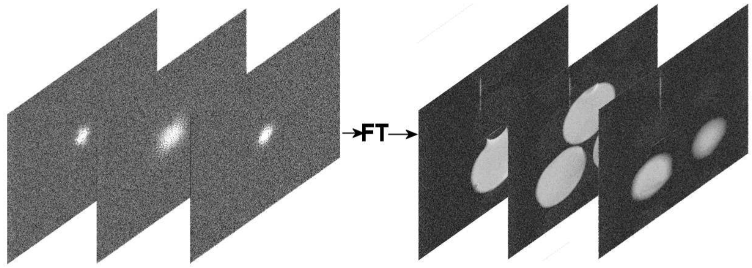



Figure 2: The procedure of obtaining the images by MR tomograph. Three selected slices in the $k$-space on the left, three selected slices of phantom in the frequency domain on the right.

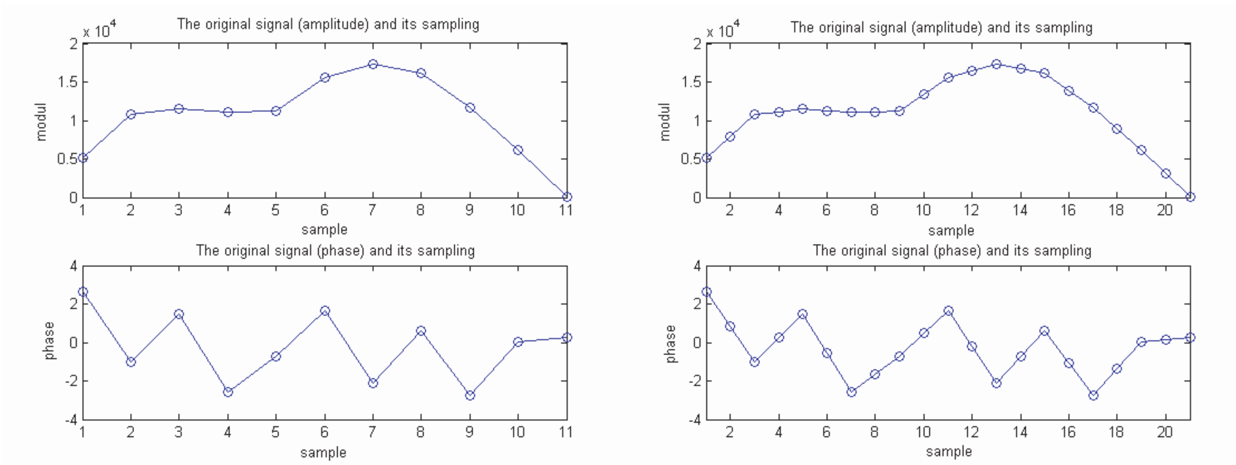



Figure 3: Interpolation of pixel intensity in the image between original slices by averaging. The pixel intensities vector (module, phase) is on the left, the pixel intensities resampled vector (module, phase) is on the right.

## 3. AVERAGING

The ordinary averaging method is used for nearby pixel intensities in the frequency domain for comparison of the proposed method of signal interpolation in the $k$-space, which are generated between two existing images obtained by MR tomography. The intensity of a new pixel is then given by the equation:

$$I_{\text{new}}\left(x,y\right)=\frac{I_{\text{new}-1}\left(x,y\right)+I_{\text{new}-1}\left(x,y\right)}{2}. \tag{1}$$

## 4. INTERPOLATION IN $k$-SPACE

This approach is a little bit complicated but it gives better results. The basic principle is extending of spectrum of the obtained signal in the $k$-space, which arises by values of the nearby pixel intensities with the same positions. In Figure 4, we can see the original complex signal generated by nearby pixel intensities. The same resampled signal transformed to frequency domain with including the zero values for higher frequencies and transformed back to the $k$-space is shown on the right. It is clear that by this approach the signal was supplemented by new samples, witch interpolates the $k$-space in the space where no slices were obtained by the tomography. This interpolated $k$-space is then transformed by FT to frequency domain with $(2n-1)$ images.

## 5. EXPERIMENTAL RESULTS

The both methods (averaging in the frequency domain, extending the spectrum of $k$-space by zero values) were tested on 11 original slices of phantom (Figure 1) obtained by MR tomography at the Institute of Scientific Instruments of the ASCR in Brno. The algorithms were implemented in the Matlab software [4]. The area of 11 slices was interpolated to 21 slices. The 10 new images were computed between the 11 original slices. The two selected slices (5th and 6th) were shown in Figure 5 and one interpolated slice with use the method of averaging in the frequency domain. In

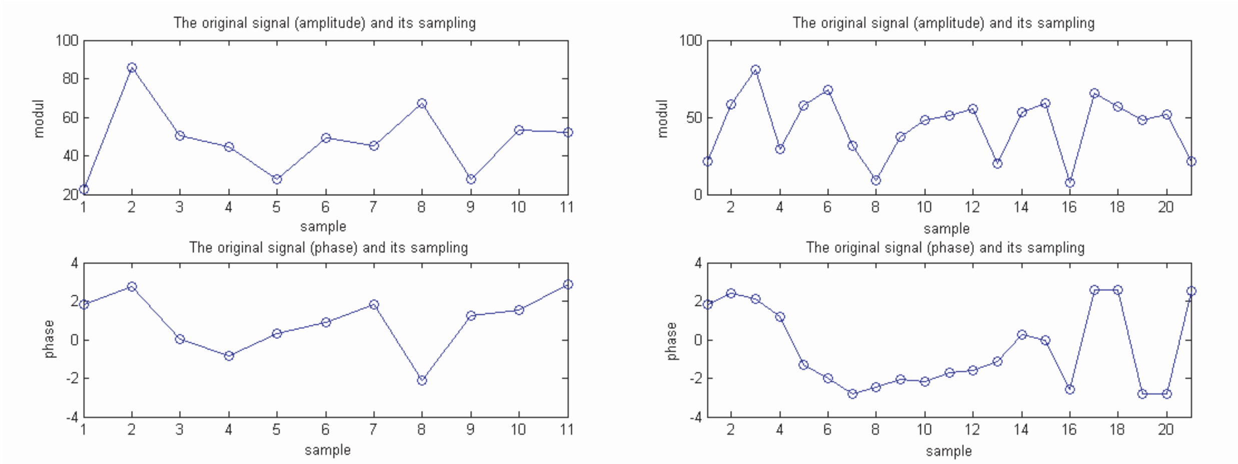



Figure 4: Interpolation of images between obtained slices by the spectrum adjustment in $k$-space. The original vector of $k$-space values (module, phase) is on the left, the resampled vector (module, phase) is on the right.

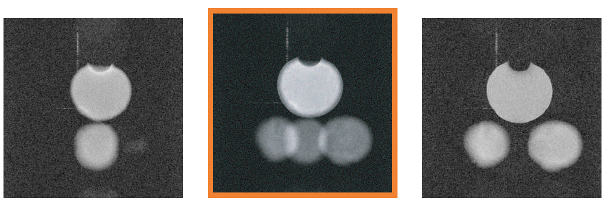


Figure 5: Interpolation of image between the obtained ones by averaging. On the left and on the right are two original images, in the middle is the interpolated image.

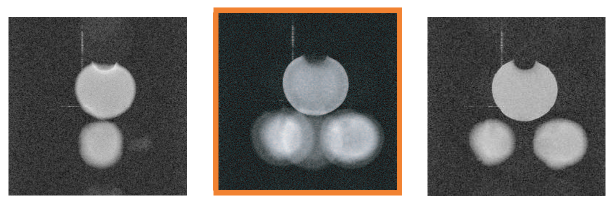


Figure 6: Interpolation of image between the obtained ones by extending the spectrum of signals in the $k$-space. On the left and right are the original images, in the middle is the interpolated image.

the Figure 6, we can see the result — the same slice — interpolated by the method of the spectrum in $k$-space extending.

## 6. CONCLUSIONS

We can compare both methods. The first one — the averaging method of the neighbor pixel intensities with the same position in the frequency domain gives the good results. The intensity of a new pixel in the interpolated image always reflects the values of obtained pixel intensities. There is a problem, after the segmentation of this image, the size of the segmented object will be always the same as in the neighbor image. But there is no other error brought to the image. It is clear that this method don't give any new spatial information, only about the average number of hydrogen nucleus in the area of interpolation.

By the second method — interpolation in $k$-space — there is brought an error due to resampling of the signal by extending its spectrum by zero values. This signal is devaluated by harmonic spatial signal. There are spatial artifacts in more images. The processed images will be segmented, whereas the suitable level of intensity for the segmentation will be found. The resultant contours will be used for accurate three-dimensional model creation and it will be made the comparison of the both methods.

## ACKNOWLEDGMENT

This work was supported within the framework of the research plan MSM 0021630516, MSM 0021630513 and project of the Grant Agency of the Czech Republic GA 102/09/0314 and project of the BUT Grant Agency FEKT-S-10-13.

## REFERENCES

1. Gescheidtova, E. and K. Bartusek, "Kriteria pro vyber vlnek pri zpracování MR obrazu," *Elektrorevue*, [online], 13.12.2009, 2009, [cit. 2010-02-14]. ISSN 1213-1539. Available

at www: http://elektrorevue.cz/cz/clanky/zpracovani-signalu/0/kriteria-pro-vyber-vlnek-pri-zpracovani-mr-obrazu/.
2. Fourier transformation [online], 2010 [cit. 2010-20-01], available at www: http://cs.wikipedia.org/wiki/Fourierova_transformace.
3. Vich, R. and Z. Smekal, "Císlicove filtry," *Academia*, Praha, 2000.
4. Matlab [online], [2009-] [cit. 2010-20-01], available at www: http://www.mathworks.com/products/matlab/.

# Magnetoinductive Lens for Experimental Mid-field MR Tomograph

**K. Bartusek, P. Drexler, P. Fiala, R. Kadlec, and R. Kubasek**
Department of Theoretical and Experimental Electrical Engineering, Brno University of Technology
Kolejni 2906/4, Brno 612 00, Czech Republic

**Abstract**— The growing interest of scientific community in the field of synthetic single/double negative materials has brought the effort to exploit their unique properties in various technical areas. One of the possible applications of these metamaterials is their utilization in subwavelength imaging. The recent work of some research teams deals with the application of single negative metamaterials with negative permeability for magnetic resonance imaging [1]. The concepts of the magnetoinductive lens have been proposed and realized [2]. The lens components are based on non-ferromagnetic materials. In spite of their insignificant response in the DC magnetic field, they exhibit magnetic response at certain resonant frequency. The resonant frequency approaches to the frequency of the excitation RF pulse [3]. By suitable configuration of the components the imaging effect can be observed and the device behaves like a magnetic lens. Most of the published papers describe the analysis of magnetoinductive lenses and the experiments in relatively low field medical MRI devices with the excitation frequency of tens of MHz [4]. This paper presents the design and development of magnetoinductive lens for application in the experimental mid-field MRI device which has a smaller field aperture and high resolution narrowband detection system.

## 1. INTRODUCTION

The possibilities of subwavelength imaging have been introduced by the recent progress in metamaterial science. There have been evaluated approaches for metamaterials structure analysis, design and development by means of analytical and numerical computing tools. Subwavelength imaging is one of the possible applications of metamaterial slabs with defined negative permitivity and negative permeability (double negative metamaterial). The subwavelength imaging is allowed through the process of evanescent field components amplification. The evanescent components are amplified in the slab by a factor which is inverse to the one which causes the decay in free space.

The recent experiments and theoretical findings have shown that the imaging can be observed in case of single negative metamaterials. This means by exploitation of metamaterial slab with only negative permitivity or only negative permeability. The typical characteristic of metamaterials structures is the dimension relation between the metamaterial slab components and the wavelength of the wave which is to carry through the slab. The dimensions of metamaterial components are much smaller than the wavelength of the interest. In spite of this dimension disproportion the metamaterial structures show response at these wavelengths. Regarding to these dimension relations it can be found that the metamaterial structure-wave interaction is in the realm of quasi-static fields. This phenomenon allows consider only electric or magnetic part for imaging purposes. The imaging utilizing electric component of the field has been demonstrated in case of excitation of surface plasmon polariton structures [5]. In the case of magnetic imaging huge progress has been made by Freire and Marques [6]. The promising application of magnetic imaging devices is in the improvement of signal in magnetic resonance tomography. The magnetic imaging devices have to be designed in respect to MR excitation frequency, detection bandwidth and the tomograph working space volume.

## 2. MAGNETOINDUCTIVE LENS

Authors in [4, 6] have studied in theoretical and experimental way the effect of two coupled resonator arrays on the process of image formation. The coupled resonator arrays are of planar type and consist of periodically arranged LC resonators on dielectric board. The typical realization for megahertz range consists of loop patterns made by etching of cooper foil on FR-4 board. The gaps in the loops are assembled with capacitors of suitable capacity. The section of resonator array is shown in Fig. 1.

The array is characterized by the resonators periodicity $a$. It should be noted that the resolution is equal the double of the lens periodicity and this resolution is far below the wavelength of the Larmor frequency in MR systems. The typical property of the resonator field is the inductive mutual coupling between the resonators. This coupling allows the propagation of magnetoinductive waves along the array's surface. The excitation frequency of the surface waves is equal to the resonant frequency of LC resonators.

The existence of the magnetoinductive waves on the resonator arrays has given the origin of the term magnetoinductive lens. The common magnetoinductive lens consists of two parallel resonator arrays at the distance $d$. The inductive coupling between the arrays causes the detuning of the equivalent circuits and the whole resonant characteristic has two peaks symmetrically placed around the original resonant frequency as can be shown in Fig. 2. The characteristic in the right part of Fig. 2 represents the excitation of two magnetoinductive waves at different frequencies around $f_r$.

The imaging effect of the lens has been observed when the frequency of the MR system has been located in relatively flat regime between the two peaks pertaining to the two magnetoinductive waves [6]. The arrays distance $d$ influences the distance between the two magnetoinductive peaks. Simultaneously the best imaging properties have been proved when the distance between the field source and the lens is the same as the lens width [7]. For imaging of objects which are beyond the original distance $d$ we need to increase the width of the lens. But, this causes the both magnetoinductive wave frequencies approaching and the imaging effect can disappear. For this reason, it is desirable to have a narrow peaks of both magnetoinductive waves.

In order to ensure the maximum narrow peak for each magnetoinductive wave the resonant frequencies of the LC resonators have to be equal. This requirement places demand on the accuracy of capacitor values and loop inductances. The selection of chip capacitor in the view of resultant resonant frequency accuracy is one of the issue of the discussion presented in this paper. The next issue is the possibility of loop dimensions calculation in order to have an accurate original resonant frequency value.

## 3. RESONATOR COMPONENTS SELECTION AND CALCULATIONS

The application of magnetoinductive lenses in MR systems specifies their dimensions to be able to operate in the tomograph space. Dimension constraints lead to resonators realization in form of planar cooper loops which are loaded by capacitors. The accuracy of required resonant frequency is dependent on values of inductance and capacity. Since the capacitor values are given by the production series the only way how to reach the desired resonant frequency is to vary the loop dimensions.

Suppressed temperature dependence of chip capacitors is desired, which requires specific dielectric material. According to magnetoinductive lens dimension the resonator loops outer diameter ranges around the value of $r_{\text{out}} = 10\,\text{mm}$. Corresponding loop inductance and the desired resonant frequency determines the capacitors value in tens of pF. In this range the chip capacitors based on NP0 material can be utilized. They offer the lowest temperature dependency and highest value accuracy in compare to other common capacitor materials (X7R,Y5V). However, the accuracy of this capacitors is $\pm 5\%$, so it is necessary to narrow down the capacitor selection by measurement.

By varying the outer diameter and width of the loop width (Fig. 1) the desired resonant frequency can be approached. For loop dimension calculation a common relations for inductance can be utilized. It has to be fulfilled the prerequisite of planar loop with thickness much smaller than the width. For this loop calculation a relations for planar coil can be exploited. On of the possibilities is to use a modified Wheeler formula which is in [8]. This relation has been derived for planar spiral coils in RF integrated circuits. With the presumption of the similar width to thickness ratio it can be modified for single loop

$$L_{mw} = \mu_0 K_1^{mw} \frac{d_{\text{out}} - w}{1 + K_2^{mw} \dfrac{w}{d_{\text{out}} - w}} = \mu_0 K_1^{mw} \frac{(d_{\text{out}} - w)^2}{d_{\text{out}} - w + K_2^{mw} w}, \tag{1}$$

where $K_1^{mw}$, $K_2^{mw}$ are constants for loop geometry approximation. However, this relation is usable in case of octagonal approximation of the loop where $K_1^{mw} = 2.25$ and $K_2^{mw} = 3.55$ [8]. For circular geometry, it is alternatively possible to use the current sheet approximation

$$L_{cs} = \frac{\mu_0 K_1^{cs} (d_{\text{out}} - w)}{2} = \left( \ln \frac{K_2^{cs} (d_{\text{out}} - w)}{w} + K_3^{cs} \frac{w}{d_{\text{out}} - w} + K_4^{cs} \left( \frac{w}{d_{\text{out}} - w} \right)^2 \right), \tag{2}$$

where $K_1^{mw} = 1$, $K_2^{mw} = 2.46$, $K_3^{mw} = 0$, $K_4^{mw} = 0.2$ are constants for loop geometry approximation [8].

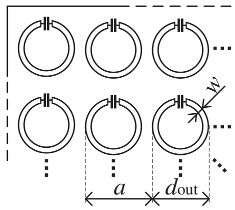



Figure 1: Section of planar resonator array.

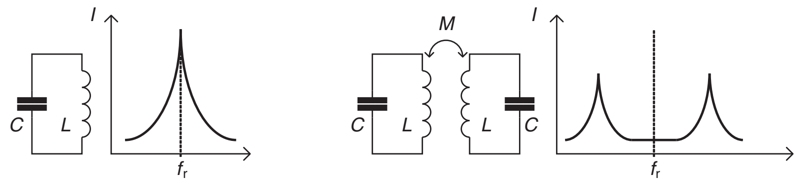



Figure 2: Equivalent LC circuit detuning by the mutual coupling.

## 4. SMALL MAGNETOINDUCTIVE LENS REALIZATION

Authors in [4] have build and verified a magnetoinductive lens for application in medical type MR system with large tomograph working space. There has been proved an imaging effect and MR signal improvements. Attempt to verify the imaging properties in MR system with small working space requires small magnetoinductive lens. The destination MR system is 4.7 T tomograph with excitation frequency close to 200 MHz. In order to achieve certain resolution and because of the higher frequency the loops have to be made smaller. A two planar resonator arrays on 2 mm thick FR-4 board has been realized with the total count of 25 resonators per layer. The resonator loop outer diameter was $d_{out} = 5\,\mathrm{mm}$ and the loop width $w = 1\,\mathrm{mm}$. The calculated inductance according to (2) was $L = 5.778\,\mathrm{nH}$. The loop has been assembled with parallel combination of three capacitors $C_1 = 100\,\mathrm{pF}$, $C_2 = 10\,\mathrm{pF}$ and $C_3 = 1.2\,\mathrm{pF}$ in attempt to approach the excitation frequency. The realized coupled resonators arrays are shown in Fig. 3 together with the experimental MR system.

The resonant frequencies of both arrays have been measured. The results are shown in Fig. 4. The resonant frequency of the first array was 199 MHz and 198.7 MHz for the second. The frequencies has been determined by spectrum analyzer with a loop on its input. The arrays has been excited by inductive coupling with dip meter. In the right part of Fig. 4 are shown peaks which correspond to the resonant frequencies of coupled arrays with distance $d = 3.3\,\mathrm{mm}$. The peak's frequencies are 191.85 MHz and 195.75 MHz. It is apparent that the expectation of symmetrical split of the resonant peaks around the original resonant frequency hasn't occurred. The centre frequency is shifted to value of approx. 193.8 MHz. There arises an opinion that the unwanted shift is caused by detuning of the resonator arrays through the coupling with the coil of the dipmeter.

Figure 3: Realized resonators array (left) and experimental MR system (right).

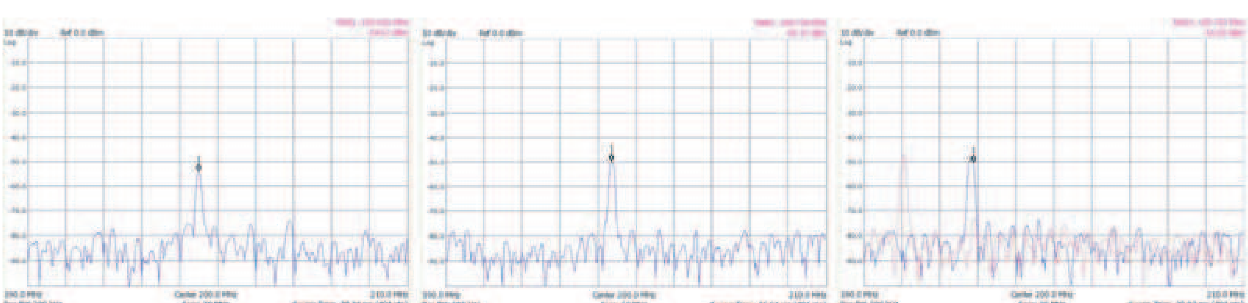


Figure 4: Peaks pertaining to resonant frequencies of single arrays (left and center) and peaks pertaining to resonant frequencies of coupled arrays (right).

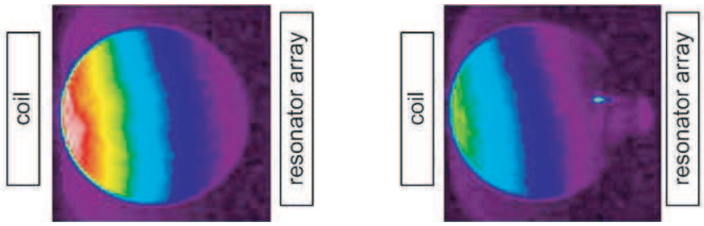



Figure 5: Measured field map in phantom with (left) and without (right) the presence of resonator arrays.

The loop areas of the resonators are relatively small. Because of the limited dip meter sensitivity the dip meter's coil had to be placed close to arrays and unwanted coupling might occured.

The effect on field distribution in phantom has been investigated in MR system. A small coil besides the phantom has been used for excitation and acquiring the signal of the slice of water volume together with resonator arrays placed on the other side. The results are shown in Fig. 5. It is apparent that no distinctive effect on the field distribution is noticeable. The RF field evacuation from the phantom volume is observable and we can resolve a small bright area of increased field intensity in the area of one loop resonator (Fig. 5). The excitation of the magnetoinductive surface wave is conceivable, which is undesirable for imaging ability. This confirms further that wrong arrays configuration might be chosen. Reported disappointing results place a demand for redesign of the resonator arrays.

## 5. CONCLUSION

In order to ensure the sufficient separation of the resonant peaks the coupling between the two resonator arrays has to be increased by using of thinner FR-4 boards and by using only one capacitor per loop. One capacitor per loop allows easier capacitor selection in view of accuracy. Subsequently the dimensions of the loop have to be calculated to obtain a right inductance for fixed capacitance and fixed resonant frequency. Next development will be focused on the redesign of the arrays and on the accuracy of arrays resonant frequency adjustment.

## ACKNOWLEDGMENT

The work described in the paper was financially supported by the research project GA102/09/0314 and research plan MSM 0021630513.

## REFERENCES

1. Wiltshire, M. C. K., J. B. Pendry, and J. V. Hajnal, "Sub-wavelength imaging at radio frequency," *Journal of Physics: Condensed Matter*, Vol. 18, No. 22, 2006.
2. Wiltshire, M. C. K., J. V. Hajnal, J. B. Pendry, D. J. Edwards, and C. J. Stevens, "Metamaterial endoscope for magnetic field transfer: Near field imaging with magnetic wires," *Optic Express*, Vol. 11, No. 7, 2003.
3. Solymar, L., O. Zhuromskyy, O. Sydoruk, E. Shamonina, I. R. Young, and R. R. A. Syms, "Rotational resonance of magnetoinductive waves: Basic concept and application to nuclear magnetic resonance," *Journal of Applied Physics*, Vol. 99, No. 12, 2006.
4. Freire, M. J., R. Marques, and L. Jelinek, "Experimental demonstration of a $\mu = -1$ metamaterial lens for magnetic resonance imaging," *Applied Physics Letters*, Vol. 93, No. 23, 2008.
5. Smolyaninov, I. I., Y. J. Hung, and C. C. Davis, "Super-resolution optics using short-wavelength surface plasmon polaritons," *Journal of Modern Optics*, Vol. 53, No. 16-17, 2006.
6. Freire, M. J. and R. Marques, "Near-field imaging in the megahertz range by strongly coupled magnetoinductive surfaces: Experiment and ab initio analysis," *Journal of Applied Physics*, Vol. 100, No. 6, 2006.
7. Freire, M. J. and R. Marques, "Optimizing the magnetoinductive lens: Improvements, limits and possible applications," *Journal of Applied Physics*, Vol. 103, No. 1, 2008.
8. Mohan, S. S., M. M. Hershenson, S. P. Boyd, and T. H. Lee, "Simple accurate expressions for planar spiral inductances," *IEEE Journal of Solid-State Circuits*, Vol. 34, No. 10, 1999.

# Homogenization of Arrays of Nanorods

**D. Felbacq**[1], **G. Bouchitté**[2], **A. I. Cabuz**[1], **F. Zolla**[3], **and A. Nicolet**[3]

[1]University of Montpellier 2, 34090 Montpellier, France
[2]University of Toulon, La Garde Cedex, France
[3]University of Aix-Marseille 3, Marseille, France

**Abstract**— The perfectly conducting infinite wire structure has been shown to exhibit negative effective permittivity (incidence normal to the wires), and strong spatial dispersion (in conical incidence) for wavelengths much larger than the wire separation. In this work we make use of a two-scale renormalization technique to treat the more general, realistic and practically relevant finite conductivity finite wire structure. The effective medium obeys a nonlocal constitutive relation which is verified by 3D numerical simulations.

## 1. INTRODUCTION

The effective medium theory of wire structures has been an active direction of research in recent years, due to the technological simplicity of the structures, and their interesting properties, particularly their effective negative permittivity, which has found use in the construction of negative index metamaterials [1–3]. However, the main thrust of investigations to date have focused on infinite wire media, and perfectly conducting wires. In this work we go further, and offer a rigorous treatment of the more realistic scenario of finite wires with finite conductivity.

The structure under study is a square biperiodic array of thin wires, of length $L$, radius $r$ and conductivity $\sigma$. We note the period $d$ and the wavelength $\lambda$. The renormalization (depicted in Fig. 1) involves a limiting process whereby the three quantities: $r$, $d$ and $1/\sigma$ tend simultaneously to zero. The parameter governing the limiting process is noted $\eta = d$, the period. The asymptotics of the other two parameters, $\sigma$ and $r$, with respect to $\eta$ are described by *fixed* parameters $\kappa$ and $\gamma$ according to the following relations:

$$\kappa = \frac{\pi r_\eta^2 \sigma_\eta}{\varepsilon_0 \omega \eta^2} \qquad \frac{1}{\gamma} = \eta^2 \log\left(\frac{r_\eta}{\eta}\right) \tag{1}$$

where $\omega$ is the angular frequency of the electromagnetic field. In other words the conductivity is renormalized inversely to the fill factor $\theta_\eta = \frac{\pi r_\eta}{\eta^2}$, while the radius is renormalized such that the expression $\eta^2 \log(\frac{r_\eta}{\eta})$ remains constant.

While these expressions may at first seem obscure, they have simple intuitive interpretations. The first requires the current density to remain constant during the renormalization. The second

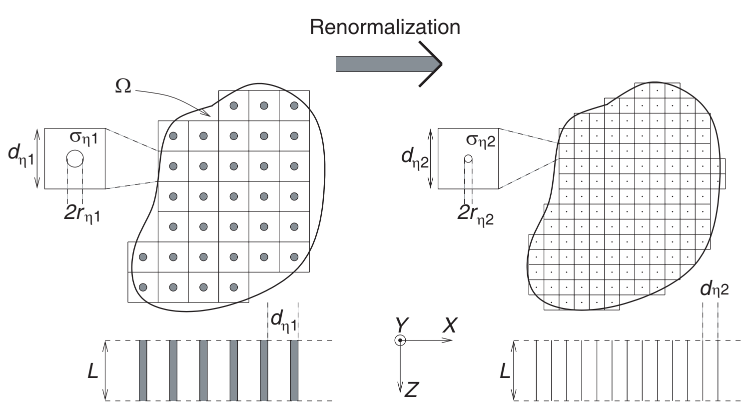



Figure 1: The bed-of-nails structure and the renormalization process. The conducting fibers occupy a region $\Omega \subsetneqq \mathbb{R}^2$, are oriented in the $z$ direction, and the structure is periodic in the $xy$ plane. Two renormalized structures are shown, corresponding to $\eta_1$ and $\eta_2$ respectively, with $\eta_1 > \eta_2$, $d_{\eta_1} > d_{\eta_2}$, $\sigma_{\eta_1} < \sigma_{\eta_2}$ and $r_{\eta_1}/d_{\eta_1} > r_{\eta_2}/d_{\eta_2}$ (see Eq. (1)). The real physical structure corresponds by definition to $\eta = 1$: $d_{\eta=1} = d$. The length $L$ and the wavelength $\lambda$ remain fixed, i.e., *we are homogenizing in the xy plane only*.

expression requires the average internal capacitance of the wires to remain constant during renormalization. This feature is known to be essential for their asymptotic behavior (see, for instance Ref. [4]).

The question to be answered now becomes: What happens in the limit $\eta \to 0$? The answer is that the fields converge (in a precise sense described in Ref. [5]) to the unique solution of the following system:

$$\begin{cases} \nabla \times E = i\omega\mu_0 H \\ \nabla \times H = -i\omega\varepsilon_0 \left(E + \frac{P}{\varepsilon_0}\hat{z}\right) \\ \frac{\partial^2 P_z}{\partial z^2} + \left(k_0^2 + \frac{2i\pi\gamma}{\kappa}\right) P_z = -2\pi\gamma\varepsilon_0 E_z, \quad z \in [-L/2, L/2] \\ \frac{\partial P_z}{\partial z} = 0, \quad z \in \{-L/2, L/2\} \end{cases} \tag{2}$$

where $P$ is the polarization field.

As a quick check of the reasonableness of this limit system, we consider the limit of infinite conductivity and infinitely long wires. It can be seen immediately that the material is non-local, even in the long-wavelength regime. By doing a Fourier transform on the third equation of system (2) (with $\kappa \to \infty$) we obtain: $\widehat{P_z} = -2\pi\gamma\varepsilon_0/(k_z^2 - k_0^2)\widehat{E_z}$ which gives $\varepsilon = 1 + 2\pi\gamma/(k_0^2 - k_z^2)$. This is fully consistent with existing results [6].

## 2. NUMERICAL RESULTS

We now proceed to test the homogeneous model by comparing it with 3D full vector simulations of the structure, i.e., we must compare the reflection, transmission and absorption coefficients and the current distribution of the homogeneous problem with those of the original bed-of-nails metamaterial. The solution to the homogeneous problem is obtained by integrating system (2) as described in Căbuz et al. (to be published).

The 3D full vector simulations of the bed-of-nails metamaterial were done using the Comsol Multiphysics finite element method software package. The periodicity was implemented using Floquet-Bloch conditions [7] in the two periodic directions ($x$ and $y$), and absorbing Perfectly Matched Layers [8] in the positive and negative $z$ directions. The linearity of the materials in the structure was used to treat the incident field as a localised source within the obstacle, as detailed in Ref. [9].

Figures 2 and 3 show the results of calculations for a structure of Toray T300® carbon fibers [10] with a conductivity of: $\sigma = 5.89 \cdot 10^4 (\Omega\text{m})^{-1}$ and a radius of 3.5 microns. The wires have an aspect ratio $L/r = 2.28 \times 10^5$, and they were modeled using an effective impedance technique [11] which eliminates the need to mesh the interior of the wires. A finite element model in which the interior

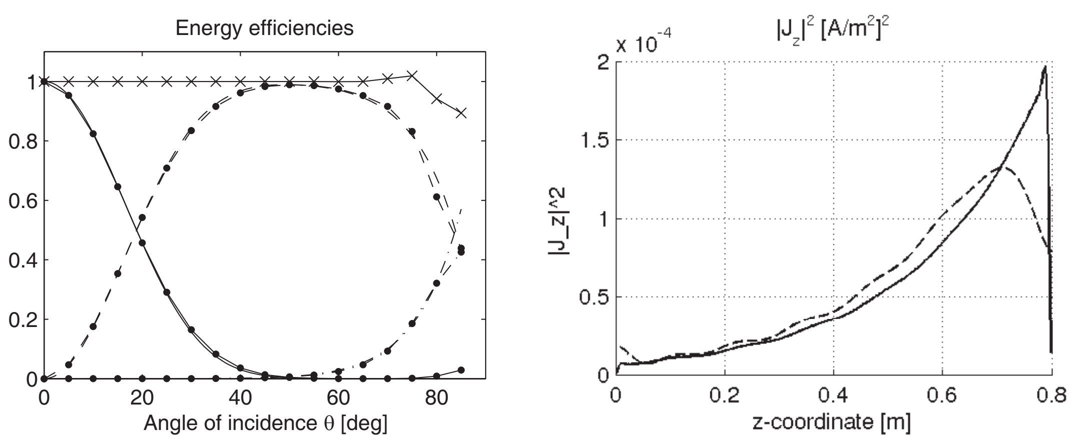



Figure 2: Transmission (solid), reflection (dot-dashed) and absorption (dashed) efficiency curves comparing the finite element solution (dot markers) and the effective medium (no markers) as a function of angle of incidence. The wire conductivity is that of Toray T300® carbon fibers $\sigma = 5.89 \cdot 10^4$ $(\Omega\text{m})^{-1}$. The structure has period $d = 0.01$ m, and dimensionless parameters $L/d = 80$, $\lambda/d = 20$, $r/d = 3.5 \cdot 10^{-4}$, and $\delta/r = 15$. Energy conservation of the finite element model ($\times$ markers) is respected to within better than one percent for most angles of incidence. The departure around 80° is explained by the poor performance of the PML absorbing layers when close to grazing incidence.

of the wires is meshed is a problem with around 3 million degrees of freedom, requiring several tens of Gigabytes of available RAM to solve. By comparison, using the impedance technique, the model of Fig. 2 (curves with markers), is a problem of approx. 62 thousand degrees of freedom, which requires less than one Gigabyte of available RAM and can therefore be solved on any sufficiently recent desktop or laptop computer.

Figures 2 and 3 illustrate the behavior which is typical of the model. The agreement remains good up to high incidence angles, and over a large wavelength domain (Fig. 5). The structure is transparent in normal incidence. For increasingly oblique angles of incidence the absorption increases more or less gradually, depending on the thickness $L$. The reflection is generally low,

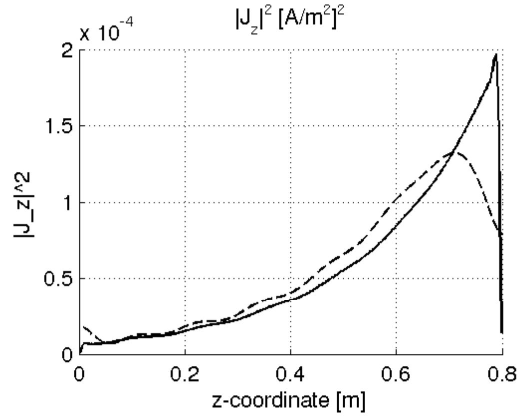



Figure 3: Square of the current density for the effective medium model (dashed) and the finite element solution (solid) as a function of position within the slab (which is positioned in $z \in (0, L)$). The structure is the same as in Fig. 2, illuminated at an angle of incidence $\theta = 40$ from the top. Note that the surface areas under the two curves are the same because they are proportional to the Joule dissipation rates, which are seen to be equal from Fig. 2 at the given angle of incidence.

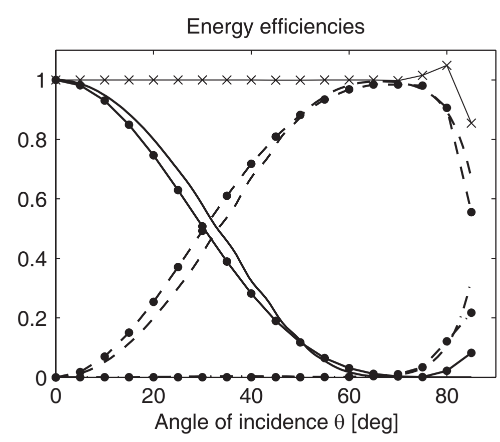



Figure 4: Transmission (solid), reflection (dot-dashed) and absorption (dashed) efficiency curves comparing the finite element solution (dot markers) and the effective medium (no markers) as a function of angle of incidence. The structure has a conductivity $\sigma = 1000\ (\Omega\text{m})^{-1}$, period $d = 0.01\,\text{m}$, and dimensionless parameters $L/d = 50$, $\lambda/d = 8$, $r/d = 0.002$, and $\delta/d = 13$. The reflection remains low for angles of incidence of up to 80° even as the Joule absorption reaches almost 100% for $\theta > 60$. Energy conservation is indicated by the $\times$ markers.

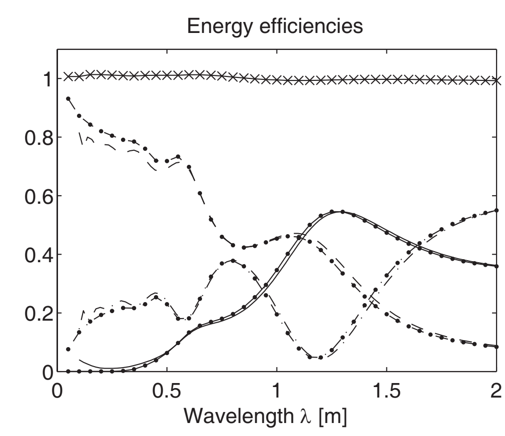



Figure 5: Transmission (solid), reflection (dot-dashed) and absorption (dashed) efficiency curves comparing the finite element solution (dot markers) and the effective medium (no markers) as a function of wavelength. Energy conservation for the finite element model is labeled with $\times$ markers. The structure has a conductivity $\sigma = 3000\ (\Omega\text{m})^{-1}$ (in the semiconductor domain), period $d = 0.01\,\text{m}$, and dimensionless parameters $L/d = 60$, $r/d = 0.003$, and the angle of incidence is $\theta = 70$. $\delta/r$ runs approximately from 4 to 25 from left to right over the domain of the plot. The model fails around $\lambda \lessapprox 0.1\,\text{m} = 10d$.

though it increases when approaching grazing incidence. The low reflection may be explained by the small radii of the wires: Their extremities have low capacitance, hence they exhibit very little charge accumulation, leading to an almost continuous normal component of the electric field. Certain configurations exhibit very low reflection for almost all angles of incidence, see Figs. 4 and 5 around $\lambda = 1.2\,\mathrm{m}$. The current density decreases roughly exponentially within the structure due to absorption.

## 3. DOMAIN OF VALIDITY

The boundaries of the domain of validity of the model are given by four dimensionless parameters: The ratio of the skin depth to the radius in the wires $\delta/r$, the ratio of the wire length to the period $L/d$, the ratio of the wavelength to the period $\lambda/d$ and the ratio of the wire radius to the period $r/d$.

The skin depth must be larger than the radius, due to the fact that the impedance used in defining $\kappa$ is the *static* impedance which differs from the *quasistatic* value by an imaginary inductive term $i\omega\mu/8\pi$ (see, for instance, Ref. [12]). Requiring this term to be negligible is equivalent to requiring that $\delta^2/r^2 \gg 1$. Moreover, in the rescaling process the skindepth/radius ratio is given by

$$\frac{\delta_\eta}{r_\eta} = \frac{\lambda}{\eta}\sqrt{\frac{1}{2\pi\kappa}}.$$

Since $\eta$ approaches zero in the rescaling process, it is natural to expect the homogeneous model to be valid when the skindepth is large compared to the radius.

In addition, recall that the definition of $\gamma$ in Eq. (1) fixes the capacitance of the wires to the value for thin, long wires. Consequently, we expect the model to hold for large $L/d$ and for small $r/d$. To these, we must add the general requirement for all effective medium models: the wavelength must be large compared to the period.

Due to the large (four dimensional) parameter space, an *exhaustive* numerical exploration of the bed-of-nails structure is not feasible in a reasonable timeframe. Still, our study has made it possible to broadly determine the boundaries of the domain of applicability of the effective medium model. Roughly, one must have $\lambda/d \gtrapprox 7-12$, $\delta/r \gtrapprox 4-8$, $L/d \gtrapprox 20-30$, $r/d \lessapprox 10$. Our (*a fortiori* limited) numerical exploration of the parameter space suggests that the skindepth-to-radius ratio is often the main limiting factor, particularly when considering highly conducting wires.

## 4. CONCLUSION

We have tested numerically the effective medium theory of the bed-of-nails structure, whose rigorous mathematical foundation is described in Ref. [5]. We have found good agreement between the transmission, reflection and absorption efficiencies between the effective medium model and a 3D finite element model, for a broad range of angles of incidence and wavelengths. The current density in the real structure corresponds to the polarization current density of the effective medium model. The medium is nonlocal, meaning that the polarization field depends on the electric field over a region of finite size. This nonlocal behavior also means that the permittivity depends on the wavevector, so it can no longer be seen, strictly, as a property of the *medium*, but rather, as a property of a given *wave* propagating in the structure [13, 14].

The bed-of-nails structure is a medium exhibiting high absorption with low reflection. It requires a very low filling fraction of conducting material, but exhibits near perfect absorption over a wide range of angles of incidence, for sufficiently large thicknesses. The low filling fraction is useful because it allows the engineer to fill the space between the wires with materials satisfying other design constraints, such as mass density, or mechanical, chemical or thermal robustness.

### REFERENCES

1. Shelby, R. A., D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science*, Vol. 292, No. 5514, 77, 2001.
2. Simovski, C. R. and B. Sauviac, "Role of wave interaction of wires and split-ring resonators for the losses in a left-handed composite," *Physical Review E*, Vol. 70, No. 4, 046607, 2004.
3. Woodley, J. F., M. S. Wheeler, and M. Mojahedi, "Left-handed and right-handed metamaterials composed of split ring resonators and strip wires," *Phys. Rev. E*, Vol. 71, 066605, 2005.
4. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, "Low frequency plasmons in thin-wire structures," *J. of Phys. — Cond. Matt.*, Vol. 10, No. 22, 4785, 1998.

5. Bouchitté, G. and D. Felbacq, "Homogenization of a wire photonic crystal: The case of small volume fraction," *SIAM Journal on Applied Mathematics*, Vol. 66, No. 6, 2061–2084, 2006.
6. Simovski, C. R. and P. A. Belov, "Low-frequency spatial dispersion in wire media," *Physical Review E*, Vol. 70, No. 4, 046616, 2004.
7. Nicolet, A., S. Guenneau, C. Geuzaine, and F. Zolla, "Modelling of electromagnetic waves in periodic media with finite elements," *Journal of Computational and Applied Mathematics*, Vol. 168, No. 1–2, 321–329, 2004.
8. Agha, Y. O., F. Zolla, A. Nicolet, and S. Guenneau, "On the use of PML for the computation of leaky modes — An application to microstructured optical fibres," *COMPEL — The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, Vol. 27, No. 1, 95–109, 2008.
9. Demésy, G., F. Zolla, A. Nicolet, M. Commandré, and C. Fossati, "The finite element method as applied tothe diffraction by an anisotropic grating," *Opt. Express*, Vol. 15, No. 26, 18089–18102, 2007.
10. Toray Carbon Fibers America Inc.
11. Carpes, Jr., W., L. Pichon, and A. Razek, "Analysis of the coupling of an incident wave with a wire inside a cavity using fem in frequency and time domains," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 44, 470, 2002.
12. Ramo, S., J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communication Electronics*, 3rd Edition, John Wiley and Sons, 1994.
13. Căbuz, A. I., "Electromagnetic metamaterials — From photonic crystals to negative index composites," Ph.D. Thesis, University of Montpellier II, 2007.
14. Căbuz, A. I., D. Felbacq, and D. Cassagne, "Spatial dispersion in negative-index composite metamaterials," *Phys. Rev. A*, Vol. 77, No. 1, 013807, 2008.

# Extended Krylov Subspace Methods for Transient Wavefield Problems

**R. F. Remis**
Laboratory of Electromagnetic Research
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands

**Abstract**— We present an extended Krylov subspace method to solve multiscale transient electromagnetic wavefield problems. A basis of an extended Krylov subspace is generated by iterating with the system matrix and its inverse. We show that such a basis can be computed very efficiently via three-term recurrence and CG-type updating formulas by exploiting specific symmetry properties of the system matrix, which are related to energy conservation and reciprocity. Multiscale transmission line and full electromagnetic wavefield problems are considered, and numerical experiments illustrate the performance of the method.

## 1. INTRODUCTION

In this paper we present an Extended Krylov Subspace method (EKS method, see [3]) to efficiently simulate electromagnetic wavefield propagation. In particular, we apply the EKS method to time-domain transmission line problems and transient electromagnetic wavefield problems are considered as well.

Starting point for both types of problems is a first-order finite-difference state-space representation obtained after discretizing the transmission line equations or Maxwell's equations in space using standard two-point finite-differences on a nonuniform staggered grid. The finite-difference representation is written in terms of a so-called system matrix and the field solution for transmission lines or Maxwell's equations is essentially given by a temporal convolution of the source vector and the system matrix evolution operator (matrix exponent).

For problems encountered in practice, the order of the system matrix can be very large and direct evaluation of the evolution operator is simply not feasible. Fortunately, we do not need the evolution operator on its own; only its action on the source vector is required. We therefore approximate the wavefield quantities by elements taken from an extended Krylov subspace. Such a space is generated by the source vector, the system matrix, and an inverse of the system matrix. Loosely speaking, the idea is that by appending a standard Krylov basis with vectors consisting of powers of an inverse of the system matrix, we approximate early- and late-time responses simultaneously.

To generate a basis of the extended Krylov space, we obviously need an inverse of the system matrix. For transmission line problems we can show that this inverse exists if (part of) the transmission line contains some loss mechanism. Moreover, we have found an explicit expression for this inverse and there is no need to numerically solve a system of equations at every step. For electromagnetic wavefield problems the inverse of the system matrix does not exist in general. Here, however, we can work with its so-called group inverse. We have found an explicit expression for this inverse as well.

Having the necessary inverses available, we still need to generate a basis of the extended Krylov space. We show that this can be done very efficiently by exploiting the symmetry properties of the system matrix and its (group) inverse. In particular, a (J-)orthogonal basis of the extended Krylov subspace can be generated via short Lanczos/CG-type recurrence relations. The expansion coefficients follow from a Galerkin procedure and having found these coefficients we can construct the EKS field approximations. Numerical experiments illustrate the performance of the method.

## 2. BASIC EQUATIONS

In this section, we briefly review the semidiscrete finite-difference systems for (multiconductor) transmission lines and $E$-polarized electromagnetic wavefields in two dimensions. In both cases, we arrive at these systems by discretizing the spatial coordinate(s) on a possibly nonuniform Yee-grid and using two-point finite-difference formulas for the spatial derivatives.

### 2.1. The Semidiscrete Transmission Line System

To fix ideas, we consider a transmission line of length $L$ with a position dependent per-unit-length (p.u.l.) capacitance $c$, inductance $\ell$, and resistance $r$. The $z$-direction is taken as the direction of propagation and the line is terminated at $z = L$ by a load resistance $R_{\mathrm{ld}}$ in series with a load inductance $L_{\mathrm{ld}}$. A voltage generator is present at the near end of the line ($z = 0$) and it consists of a voltage source $v_{\mathrm{s}}(t)$ in series with a source resistance $R_{\mathrm{s}}$.

For lines of the above type, the semidiscrete transmission line system can be written in the form

$$(\mathbf{A} + \mathbf{I}\partial_t)\mathbf{f} = v_{\mathrm{s}}(t)\mathbf{s}, \tag{1}$$

where $\mathbf{f} = [\mathbf{v}^T, \mathbf{i}^T]^T$ is the field vector and $\mathbf{v}$ and $\mathbf{i}$ are column vectors containing the finite-difference approximations of the voltage and the current, respectively. The total number of unknown time-dependent finite-difference approximations is denoted by $n$. Furthermore, $\mathbf{s}$ is the source vector and $\mathbf{A}$ is the so-called system matrix. It can be shown that if $R_{\mathrm{s}} > 0$, or $R_{\mathrm{ld}} > 0$, or the p.u.l. resistance $r > 0$ at least on some part of the transmission line, then the system matrix is nonsingular and an explicit expression for its inverse can be given as well (see [6]). There is no need to numerically solve a system of equations to compute the matrix-vector product $\mathbf{A}^{-1}\mathbf{x}$ for a given vector $\mathbf{x}$. Finally, the system matrix is J-symmetric because of reciprocity. More precisely, matrix $\mathbf{A}$ satisfies the symmetry relation

$$\mathbf{A}^T\mathbf{J} = \mathbf{J}\mathbf{A} \quad \text{with } \mathbf{J}^T = \mathbf{J}. \tag{2}$$

For the transmission line problem considered here, matrix $\mathbf{J}$ is diagonal and not positive definite. An explicit expression for matrix $\mathbf{J}$ can be found in [6]. The inverse of the system matrix is also J-symmetric, since J-symmetry is closed under inversion. Finally, we mention that multiconductor transmission lines can be handled essentially by tensoring the equations for a single transmission line.

### 2.2. The Semidiscrete Maxwell System

We consider $E$-polarized fields in a two-dimensional configuration that is invariant in the $z$-direction. The media in this configuration are characterized by the position dependent permittivity $\varepsilon$ and permeability $\mu$ and an external electric-current source $J_z^{\mathrm{ext}} = w(t)J_z^{\mathrm{sp}}(\mathbf{x})$ generates the electromagnetic field. The scalar function $w$ is called the source signature and it vanishes for $t < 0$. The electric and magnetic field strength vanish for $t < 0$ as well.

The semidiscrete Maxwell system can also be written in a form as given by Eq. (1). For electromagnetic wavefields, the field vector is given by $\mathbf{f} = [\mathbf{h}_x^T, \mathbf{e}_z^T, \mathbf{h}_y^T]^T$, where $\mathbf{h}_x$ contains all the time-dependent finite-difference approximations of the $x$-component of the magnetic field strength, and the finite-difference approximations of the $z$-component of the electric field strength and the $y$-component of the magnetic field strength are stored in the vectors $\mathbf{e}_z$ and $\mathbf{h}_y$, respectively. The total number of unknowns is again denoted by $n$.

The source vector $\mathbf{s}$ is given by $\mathbf{s} = [\mathbf{0}^T, (\mathbf{j}^{\mathrm{sp}})^T, \mathbf{0}]^T$, where $\mathbf{j}^{\mathrm{sp}}$ is a finite-difference approximation of $J_z^{\mathrm{sp}}$. Finally, for Maxwell's equations we have $v_{\mathrm{s}}(t) = -w(t)$ and $\mathbf{A}$ is the Maxwell system matrix. This matrix satisfies the symmetry relations

$$\mathbf{A}^T\mathbf{W}^{\mathrm{en}} = -\mathbf{A}\mathbf{W}^{\mathrm{en}} \quad \text{(energy conservation)} \quad \text{and} \quad \mathbf{A}^T\mathbf{J} = \mathbf{J}\mathbf{A} \quad \text{(reciprocity)}. \tag{3}$$

The first relation states that the system matrix is skew-symmetric with respect to the diagonal and positive definite energy matrix $\mathbf{W}^{\mathrm{en}}$. We refer to $\mathbf{W}^{\mathrm{en}}$ as an energy matrix, since $\frac{1}{2}\mathbf{f}^T\mathbf{W}^{\mathrm{en}}\mathbf{f}$ is a finite-difference approximation of the stored electromagnetic energy in the configuration. Using the skew-symmetry of matrix $\mathbf{A}$, we can show that energy is conserved.

The second relation in Eq. (3) states that the Maxwell system matrix is J-symmetric. Again, this symmetry relation holds because of reciprocity and here too matrix $\mathbf{J}$ is diagonal and not positive definite. Furthermore, $\frac{1}{2}\mathbf{f}^T\mathbf{J}\mathbf{f}$ is a finite-difference approximation of the field Lagrangian (see [2]). Explicit expressions for matrix $\mathbf{A}$, $\mathbf{W}^{\mathrm{en}}$, and $\mathbf{J}$ can be found in [7].

To construct a basis of a standard extended Krylov subspace, the inverse of the system matrix is required. The problem with the system matrix for Maxwell's equations is that its inverse does not exist. The system matrix is singular and we therefore resort to the Drazin (group) inverse of the system matrix. This inverse, denoted by $\mathbf{A}^{\mathrm{D}}$, is uniquely defined by the following conditions (see [1]):

$$\mathbf{A}^{\mathrm{D}}\mathbf{A}\mathbf{A}^{\mathrm{D}} = \mathbf{A}^{\mathrm{D}}, \quad \mathbf{A}\mathbf{A}^{\mathrm{D}} = \mathbf{A}^{\mathrm{D}}\mathbf{A}, \quad \text{and} \quad \mathbf{A}^{p+1}\mathbf{A}^{\mathrm{D}} = \mathbf{A}^p, \tag{4}$$

where $p$ is the index of matrix $\mathbf{A}$. Recall that the index of matrix $\mathbf{A}$ is defined as the smallest nonnegative integer $p$ such that $\mathrm{rank}(\mathbf{A}^p) = \mathrm{rank}(\mathbf{A}^{p+1})$. If the index of a matrix is equal to one, then its Drazin inverse is called the group inverse and it is denoted by $\mathbf{A}^{\#}$. The Maxwell system matrix has an index $p = 1$ and therefore it has a group inverse $\mathbf{A}^{\#}$. Since the source vector is in the range of the system matrix, we also have $\mathbf{A}\mathbf{A}^{\#}\mathbf{s} = \mathbf{A}^{\#}\mathbf{A}\mathbf{s} = \mathbf{s}$.

The group inverse of the system matrix is skew-symmetric with respect to $\mathbf{W}^{\text{en}}$ and it is J-symmetric as well. An explicit expression for the group inverse can be given and it shows that the solution of Poisson's equation is required to evaluate its action on a given vector. Specifically, evaluating the action of the group inverse on the source vector requires the solution of a single Poisson equation and $\mathbf{A}^{\#}\mathbf{s}$ is a finite-difference approximation of the Biot-Savart law [5].

## 3. EXTENDED KRYLOV SUBSPACE APPROXIMATIONS

In the extended Krylov method, the field vector $\mathbf{f}$ is approximated by elements drawn from the Krylov subspace

$$\mathbb{K}^{k,m} = \text{span}\{\mathbf{A}^{-k+1}\mathbf{s}, \mathbf{A}^{-k+2}\mathbf{s}, \ldots, \mathbf{A}^{-1}\mathbf{s}, \mathbf{s}, \mathbf{A}\mathbf{s}, \ldots, \mathbf{A}^{m-1}\mathbf{s}\}, \tag{5}$$

where $\mathbf{A}^{-1}$ has to be replaced by $\mathbf{A}^{\#}$ if electromagnetic wavefield problems are considered. Notice that the maximal dimension of $\mathbb{K}^{k,m}$ is $d = k + m - 1$ and we are interested in extended Krylov approximations for which $d \ll n$.

Constructing the approximations is a two-step procedure. We first generate a basis for $\mathbb{K}^{k,m}$ and expand the Krylov approximations in terms of the basis vectors. The second step consists of determining the time-dependent expansion coefficients.

By exploiting the symmetry properties of the system matrix (matrix $\mathbf{A}$ is J-symmetric or skew-symmetric or both), we can efficiently construct the basis vectors of $\mathbb{K}^{k,m}$. The skew-symmetry of the system matrix can be used to construct an orthogonal basis for $\mathbb{K}^{k,m}$ via three-term recurrence and CG-type update equations, while if we exploit the J-symmetry of the system matrix then a J-orthogonal basis for $\mathbb{K}^{k,m}$ can be constructed in much the same way as is done for the skew-symmetric case. We note that the algorithm based on J-symmetry may suffer from breakdown, since matrix $\mathbf{J}$ is not positive definite. In our numerical work we have never detected a breakdown of the algorithm, but we do not have a proof that no breakdown can occur either (in exact or finite-precision arithmetic).

Let us now give a brief description of how we generate a (J-)orthogonal basis for $\mathbb{K}^{k,m}$. We follow the approach outlined in [3]. A different (perhaps more flexible) approach is presented in [4].

Let the columns $\mathbf{w}_i$ of the $n$-by-$d$ matrix $\mathbf{W}_d$ form a (J-)orthogonal basis of $\mathbb{K}^{k,m}$. We expand the Krylov approximations in terms of these basis vectors. Specifically, the extended Krylov approximation of order $d$ is written as

$$\mathbf{f}_d(t) = \mathbf{w}_1 c_1(t) + \mathbf{w}_2 c_2(t) + \ldots + \mathbf{w}_d c_d(t) = \mathbf{W}_d \mathbf{c}(t), \tag{6}$$

where $\mathbf{c} = [c_1(t), c_2(t), \ldots, c_d(t)]^T$ is the vector of expansion coefficients. We now partition matrix $\mathbf{W}_d$ as

$$\mathbf{W}_d = (\mathbf{Q}_k, \mathbf{V}_{m-1}) \quad \text{where} \quad \mathbf{Q}_k = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k) \quad \text{and} \quad \mathbf{V}_{m-1} = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{m-1}). \tag{7}$$

The vectors $\mathbf{q}_i$, $i = 1, 2, \ldots, k$, are computed with $\mathbf{A}^{-1}$ (or $\mathbf{A}^{\#}$) as iteration matrix, while the vectors $\mathbf{v}_i$, $i = 1, 2, \ldots, m-1$, are generated with the system matrix $\mathbf{A}$. Specifically, the basis vectors are constructed as follows:

1. Run $k$ steps of the Lanczos algorithm with matrix $\mathbf{A}^{-1}$ (or $\mathbf{A}^{\#}$) and take the source vector $\mathbf{s}$ as a starting vector to obtain $\mathbf{q}_1$, $\mathbf{q}_2$, $\ldots$, $\mathbf{q}_k$;
2. Determine $\mathbf{v}_1$ and $\mathbf{v}_2$ by orthogonalizing $\mathbf{A}\mathbf{s}$ against $\mathbf{Q}_k$ and $\mathbf{A}\mathbf{v}_1$ against $\mathbf{v}_1$ and $\mathbf{Q}_k$;
3. Given the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, the remaining vectors $\mathbf{v}_3$, $\mathbf{v}_4$,$\ldots$,$\mathbf{v}_{m-1}$ can be computed via a three-term recurrence relation by exploiting the J-symmetry or skew-symmetry of the system matrix.

Notice that in step 2 all basis vectors $\mathbf{q}_i$ are required. Storage of $\mathbf{Q}_k$ can be avoided, however, by rewriting $\mathbf{v}_1$ and $\mathbf{v}_2$ in terms of an auxiliary vector (cf. [3]). This auxiliary vector can computed recursively by including a CG-type update formula in step 1. Storage of $\mathbf{Q}_k$ is avoided in this way.

After completion of the above algorithm, we have the summarizing equation

$$\mathbf{A}\mathbf{W}_d = \mathbf{W}_d \mathbf{R}_d + t_{m,m-1} \mathbf{v}_m \mathbf{e}_d^T, \tag{8}$$

where $t_{m,m-1}$ is a recurrence coefficient from step 3 and $\mathbf{e}_d$ is the $d$th column of the $d$-by-$d$ identity matrix. Furthermore, matrix $\mathbf{R}_d$ is a matrix of order $d$ with a block structure

$$\mathbf{R}_d = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{pmatrix}, \tag{9}$$

where $\mathbf{R}_{11}$ is a dense square matrix of order $k$, $\mathbf{R}_{22}$ is a square and tridiagonal matrix of order $m-1$, and all elements of $\mathbf{R}_{12}$ ($\mathbf{R}_{21}$) are zero, except for the elements in its first column (row).

The second step consists of determining the expansion coefficients. To this end, we substitute the Krylov approximation of Eq. (6) in the semidiscrete system of Eq. (1), use Eq. (8), and apply a (pseudo-)Galerkin procedure. The expansion coefficients are found as

$$\mathbf{c}(t) = \|\mathbf{s}\| \int_{\tau=0}^{t} v_{\mathrm{s}}(\tau) \exp[-\mathbf{R}_d(t-\tau)]\mathbf{e}_1 \, \mathrm{d}\tau \quad \text{for } t > 0, \tag{10}$$

where $\|\mathbf{s}\|$ is the 2-norm of the source vector (we normalize each basis vector such that it has unit 2-norm). With the expansion coefficients at our disposal, the extended Krylov approximation becomes

$$\mathbf{f}_d(t) = \|\mathbf{s}\|\mathbf{W}_d \int_{\tau=0}^{t} v_{\mathrm{s}}(\tau) \exp[-\mathbf{R}_d(t-\tau)]\mathbf{e}_1 \, \mathrm{d}\tau \quad \text{for } t > 0.$$

This expression is evaluated at specified receiver positions and therefore we only need to store rows of matrix $\mathbf{W}_d$ that correspond to these receiver locations.

## 4. NUMERICAL EXPERIMENTS

To illustrate the performance of the extended Krylov method, we solve a multiscale transmission line problem for which an exact solution is known. The transmission line has a length $L = 0.1\,\mathrm{m}$ and consists of a single active conductor only. The p.u.l. capacitance and inductance are given by $c = 0.2 \cdot 10^{-7}\,\mathrm{F/m}$ and $\ell = 0.05 \cdot 10^{-7}\,\mathrm{H/m}$, respectively, and the source signature is given by the derivative of a shifted Gaussian with a center frequency of 1.5 GHz (a 1.5 GHz monocycle). This pulse has a pulse width of about one nanosecond. Since we are interested in the late-time behavior of the Krylov approximations, we set the p.u.l. resistance to zero ($r = 0$) and we short-circuit the line at its far end ($R_{\mathrm{ld}} = 0$ and $L_{\mathrm{ld}} = 0$). Finally, we take $R_{\mathrm{s}} = 10\,\mathrm{m}\Omega$. The Gaussian pulse will now bounce back and forth and decays very slowly. The source is switched on at $t = 0$ and the length of the time interval of observation is one microsecond. The order of the system matrix is $n = 20000$.

Since the observation interval is 1000 times the pulse width of the source, it is difficult to illustrate how the Krylov method approximates the exact response on the complete time interval of observation. We therefore show the computed and exact response on a subinterval only running from $t = 245\,\mathrm{ns}$ to $t = 265\,\mathrm{ns}$. In Figure 1, the exact response (solid line) and the Krylov approximation of order $d = 80$ with $k = 40$ and $m = 41$ are shown. We observe that the arrival times are already captured, but the amplitudes are not correct. This is observed over the complete interval of observation. Increasing the order of the approximation to $d = 120$ using $k = 60$ and $m = 61$, we obtain the Krylov approximation as shown in Figure 2 (dashed line). The solid line signifies the

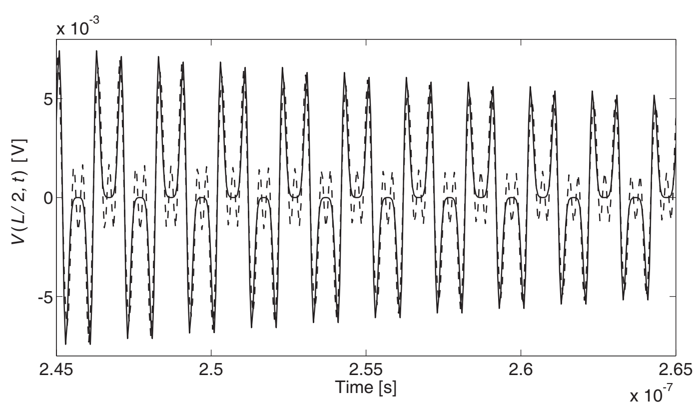



Figure 1: The voltage $V$ at $z = L/2$ from $t = 245\,\mathrm{ns}$ to $t = 265\,\mathrm{ns}$. The solid line signifies the exact result, the dashed line is the Krylov approximation of order $d = 80$ with $k = 40$ and $m = 41$.

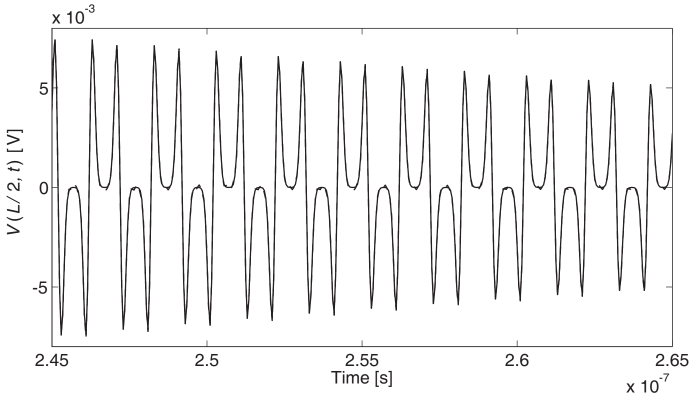



Figure 2: The voltage $V$ at $z = L/2$ from $t = 245\,\text{ns}$ to $t = 265\,\text{ns}$. The solid line signifies the exact result, the dashed line is the Krylov approximation of order $d = 120$ with $k = 60$ and $m = 61$.

exact response and we observe that now both the arrival times and the amplitudes are modeled correctly. This is also observed on the remaining part of the observation interval. This convergence behavior turns out to be typical. All numerical experiments that we have carried out so far indicate that arrival times are approximated very well on a large time interval of observation already after a few iterations of the extended Krylov method (using $m = k + 1$). Subsequent iterations correct for the amplitudes and the resulting Krylov approximation is accurate on large time intervals of interest after only a small number of iterations.

## 5. CONCLUSIONS

In this paper we have presented an extended Krylov subspace method for multiscale transient electromagnetic wavefield problems. A basis of the Krylov space $\mathbb{K}^{k,m}$ is generated by iterating with the system matrix and its (group) inverse. More precisely, first $k$ iterations are carried out with the inverse of the system matrix to obtain the first $k$ basis vectors, and subsequently the basis vectors related to powers of the system matrix are computed. Having this basis available, we obtain a basis for $\mathbb{K}^{k,m+1}$ by simply carrying out an additional iteration with matrix $\mathbf{A}$. However, it seems that it is not possible to efficiently construct a basis for $\mathbb{K}^{k+1,m}$ starting from an already computed basis of $\mathbb{K}^{k,m}$. A different approach in which this problem is avoided is presented in [4]. Furthermore, in all our numerical work we take $m = k+1$, but this is just a choice. Other relations between $k$ and $m$ may lead to faster convergence and in future work we plan to investigate this issue further.

## ACKNOWLEDGMENT

The research reported in this paper is supported by the Dutch Technology Foundation (STW). This support is gratefully acknowledged.

## REFERENCES

1. Campbell, S. L. and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*, Dover Publications, Inc., New York, 1991.
2. Cohen-Tannoudji, C., J. Dupont-Roc, and G. Grynberg, *Photons and Atoms — Introduction to Quantum Electrodynamics*, Wiley-Interscience, New York, 1997.
3. Druskin, V. and L. Knizhnerman, "Extended Krylov subspaces: Approximation of the matrix square root and related functions," *SIAM J. Matrix Anal. Appl.*, Vol. 19, No. 3, 755–771, 1998.
4. Jagels, J. and L. Reichel, "The extended krylov subspace method and orthogonal laurent polynomials," *Linear Algebra and Its Applications*, Vol. 431, 441–458, 2009.
5. Krauss, J. D., *Electromagnetics*, 4th Edition, McGraw-Hill, Inc., New York, 1992.
6. Remis, R. F., "An effcient model-order reduction approach to low-frequency transmission line modeling," *Progress In Electromagnetics Research*, PIER 101, No. 2, 139–155, 2010.
7. Remis, R. F., "Global form of the 2D finite-difference Maxwell system and its symmetry properties," *IEEE Antennas and Propagation Society International Symposium*, 624–627, San Antonio, USA, 2002.

# Interaction Dynamics of Solitons in a Linearly Coupled Ginzburg-Landau Equation with Cubic-quintic Nonlinearity

**Daniel R. Neill and Javid Atai**
School of Electrical and Information Engineering, The University of Sydney, NSW 2006, Australia

**Abstract**— We investigate the interactions of solitons in a model based on an asymmetric linearly coupled Ginzburg-Landau equations with a cubic-quintic nonlinearity. It is found that the outcomes of interactions are dependent upon the initial separation of the pulses, and the initial phase difference between them.

## 1. INTRODUCTION

Solitary pulses play an important role in the physical systems that are described by various forms of the Ginzburg-Landau equation. Models based on complex Ginzburg-Landau (CGL) equations have been used extensively to study pattern formation and solitary pulse dynamics in various physical systems [1–3]. In nonlinear optics, the CGL has been used to model pulses generated by fiber lasers [4, 5].

In the theoretical modeling of such pulses, the issues involved are the stability of the pulses, the interactions between them and the physical applications of the model. In Ref. [6], a model based on linearly coupled CGL equations was put forward that admitted exact stable soliton pulses. The model describes a doped dual core nonlinear optical fiber, with one core being passive and lossy and the other being active with dispersive losses, dispersion, and cubic nonlinearity.

In this paper we investigate the interaction of solitons in a model similar to that of Ref. [6] but with cubic-quintic nonlinearity. It should be noted that cubic-quintic nonlinear response has been experimentally observed in chalcogenide glasses [7] and some organic transparent materials [8]. Addition of quintic term to the model of Ref. [6] results in the following system of equations:

$$iu_z + \left(\frac{1}{2} - i\gamma_1\right) u_{\tau\tau} + \sigma\left(|u|^2 - \delta\,|u|^4\right) u - i\gamma_0 u + v = 0 \quad iv_z + i\Gamma_0 v + u = 0 \tag{1}$$

where $u$ and $v$ are the amplitudes of the pulses traveling in the active and lossy cores and $z$ and $\tau$ correspond to the propagation distance and retarded time, respectively. Dispersion is accounted for by $\sigma$, with normal dispersion being $\sigma = -1$ and anomalous dispersion being $\sigma = +1$. $\gamma_1$, $\gamma_0$ and $\Gamma_0$ are coefficients of the dispersive losses, the gain in the active core and the loss in the passive core, respectively. $\delta$ is the strength of the quintic nonlinearity.

## 2. SOLITON SOLUTIONS AND THEIR STABILITY

A key finding of Ref. [6] is that in the purely cubic model (i.e., when $\delta = 0$ in Eq. (1) *exact stable* solitary pulses exist. Also, it was demonstrated that there exist vast stability regions in different dispersion regimes on the model's phase planes.

Unlike the model of Ref. [6], the model of Eq. (1) does not admit exact analytical solutions. The only analytical results that can be obtained in the cubic-quintic model are the following two conditions that guarantee the stability of zero solutions (i.e., $u = v = 0$):

$$\gamma_0 < \Gamma_0 \quad \gamma_0\Gamma_0 < 1 \tag{2}$$

Inequalities (2) are the necessary conditions for the stability of solitary pulses of Eq. (1).

We have investigated the stability of solitary pulses of the system (1) by numerically solving (1). Different initial conditions were used to ensure that the results were not dependent on the choice of the initial pulse. Also, various asymmetric perturbations were initially added to the solitons to expedite the onset of any inherent instability. The results of the numerical stability analysis for anomalous dispersion regime is shown in Fig. 1.

As is shown in Fig. 1, a small increase in the strength of quintic nonlinearity results in a significant reduction in the size of stability region. Similar behavior is observed in the normal dispersion regime.

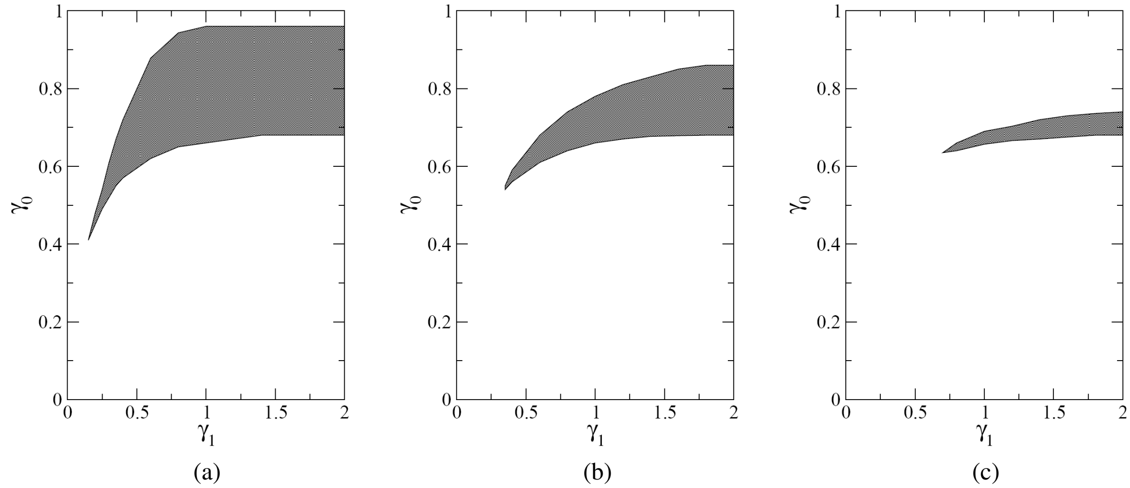



Figure 1: Stability regions (shaded) on the parametric plane $(\gamma_1, \gamma_0)$ for $\sigma = +1$ and $\gamma_0\Gamma_0 = 0.9$. (a) $\delta = 0.04$; (b) $\delta = 0.06$; (c) $\delta = 0.08$.

## 3. INTERACTION OF SOLITONS

We have investigated the interaction of solitons by numerically solving Eq. (1) subject to the following initial conditions:

$$u(\tau,0)=u\left(\tau-\frac{\Delta\tau}{2},0\right)+u\left(\tau+\frac{\Delta\tau}{2},0\right)\exp(i\Delta\phi)\,,\quad v(\tau,0)=v\left(\tau-\frac{\Delta\tau}{2},0\right)+v\left(\tau+\frac{\Delta\tau}{2},0\right)\exp(i\Delta\phi)\,,\quad (3)$$

where $\Delta\tau$ and $\Delta\phi$ are the initial separation and phase difference between the solitons and $u$ and $v$ belong to the stable region.

### 3.1. Interaction of Solitons in the Cubic Model

First, we consider the interaction of solitons for $\delta = 0$ (i.e., purely cubic nonlinearity). In this case, Eq. (1) reduce to the model of Ref. [6].

In the case of in-phase solitons (i.e., $\Delta\phi = 0$), it is found that in both anomalous and normal dispersion regimes, there exists a critical separation below which the interaction of solitons result in their merger into a single soliton. Additionally, in the normal dispersion regime, when the separation of solitons is greater than a critical value they initially repel each other until they reach a certain separation and then they propagate without any interaction.

When $\Delta\phi = \pi$, depending on the initial separation, the outcome of interactions in the anomalous dispersion regime may be (a) the decay of solitons, (b) their fusion into a single soliton, or (c) initial repulsion and subsequent parallel propagation. On the other hand, in the normal dispersion regime, only outcomes (a) and (c) are observed. Examples of interaction of solitons in anomalous and normal dispersion regimes are shown in Fig. 2.

### 3.2. Interaction of Solitons in the Cubic-quintic Model

The presence of quintic nonlinearity (i.e., $\delta \neq 0$) drastically alters the interaction dynamics of solitons in Eq. (1). The simulations demonstrate that the outcome of the interactions depend upon the initial separation and the phase difference between solitons. The interplay of these parameters results in a variety of outcomes. In particular, an outcome that is observed in this case and is absent in the cubic model is that in the anomalous dispersion regime, for both $\Delta\phi = 0$ and $\Delta\phi = \pi$, below a certain initial separation, the solitons attract each other and very quickly are destabilized and destroyed. Figs. 3(a) and 3(b) show examples of such an interaction for $\delta = 0.04$. A noteworthy feature shown in Fig. 3(c) is that increasing the initial separation from 8 to 9 results in the repulsion of solitons.

In the normal dispersion regime, when $\Delta\phi = 0$, the interaction of solitons with the same parameters as Fig. 3, results in the merger of solitons into a single soliton. As is shown in Fig. 4, depending on the initial separation the emerging soliton may be substantially larger than the original solitons. This outcome is due to the fact that the solitons in the model of Eq. (1) belong to the class of dissipative solitons. In such systems the interplay of gain, loss and nonlinearity may

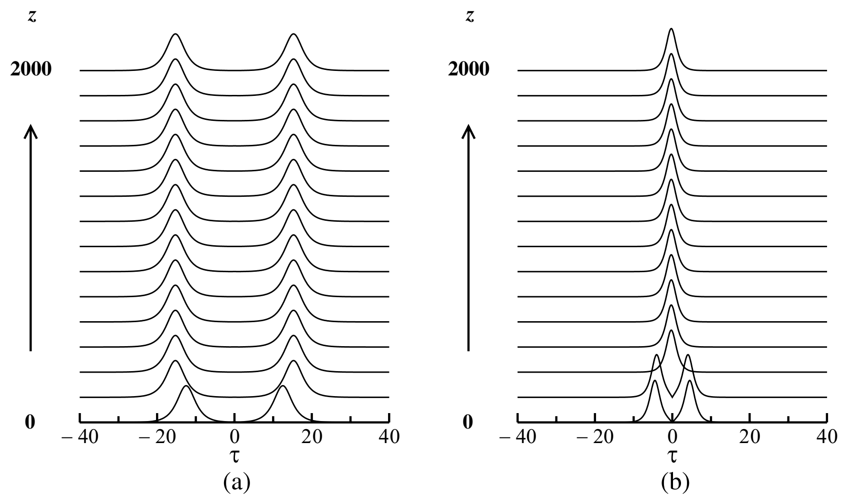



Figure 2: Examples of interaction of solitons in the model of Eq. (1) for $\delta = 0$ (i.e., the cubic model). The other parameters are $\gamma_1 = 1.2, \gamma_0 = 0.75$ and $\Gamma_0 = 1.2$. (a) Anomalous dispersion ($\sigma = +1$), $\Delta\phi = 0$, $\Delta\tau = 25$; (b) normal dispersion ($\sigma = -1$), $\Delta\phi = \pi$, $\Delta\tau = 9$.

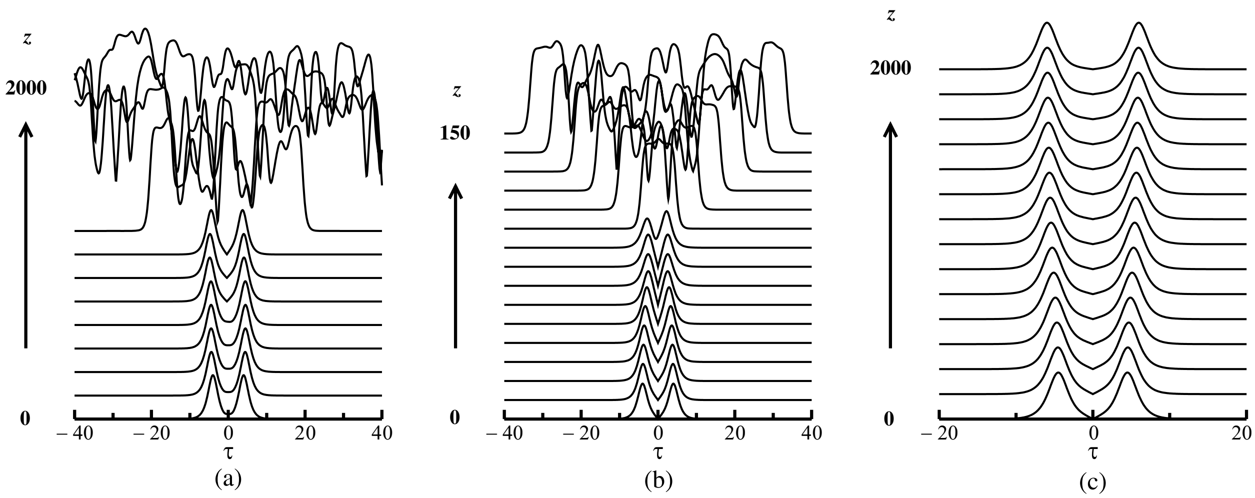



Figure 3: Examples of interaction of solitons in the model of Eq. (1) for $\sigma = +1$ and $\delta = 0.04$. The other parameters are $\gamma_1 = 1.2, \gamma_0 = 0.75$ and $\Gamma_0 = 1.2$. (a) $\Delta\phi = 0$ and $\Delta\tau = 8$; (b) $\Delta\phi = \pi$ and $\Delta\tau = 8$; (c) $\Delta\phi = \pi$ and $\Delta\tau = 9$.

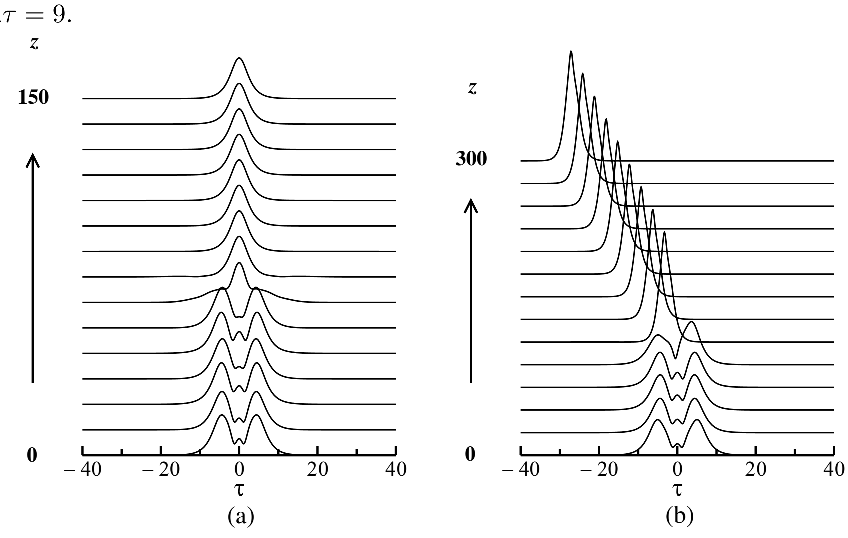



Figure 4: Examples of interaction of solitons in the model of Eq. (1) for $\sigma = -1$ and $\delta = 0.04$. The other parameters are $\gamma_1 = 1.2, \gamma_0 = 0.75$ and $\Gamma_0 = 1.2$. (a) $\Delta\phi = 0$ and $\Delta\tau = 9$; (b) $\Delta\phi = 0$ and $\Delta\tau = 10$.

lead to amplification of solitons [9]. In the case of $\Delta\phi = \pi$, when the initial separation of solitons is in the range $3 < \Delta\tau < 11$ they attract and form a single soliton of similar size. On the other hand, when $11 \leq \Delta\tau \leq 15$, similar to the case of $\Delta\phi = 0$ (i.e., Fig. 4(b)), the solitons merge into a larger soliton that moves with a certain velocity.

## 4. CONCLUSION

We have investigated the interaction of solitons in an asymmetric linearly coupled Ginzburg-Landau equations. It is found that the outcomes of interactions of solitons are dependent upon the initial separation of solitons and their initial phase difference. In particular, it is found that for certain parameters the interactions may result in destabilization and subsequent destruction of both solitons. Such an outcome does not exist in the purely cubic version of the model. In addition, in the normal dispersion regime, when $\delta = 0.04$, there exists a range of initial separations for which the solitons fuse into a larger soliton that moves with a certain velocity.

## REFERENCES

1. Malomed, B. A., "Complex Ginzburg-Landau equation," *Encyclopedia of Nonlinear Science*, A. Scott (Ed.), 157–160, Routledge, New York, 2005.
2. Arecchi, F. T., S. Boccaletti, and P. Ramazza, *Phys. Rep.*, Vol. 318, No. 1, 1999.
3. Aranson, I. S. and L. Kramer, *Rev. Modern Phys.*, Vol. 74, No. 99, 2002.
4. Fermann, M. E., A. Galvanauskas, G. Sucha, and D. Harter, *Appl. Phys. B*, Vol. 65, No. 259, 1997.
5. Ilday, F. O. and F. W. Wise, *J. Opt. Soc. Amer. B*, Vol. 19, No. 470, 2002.
6. Atai, J. and B. A. Malomed, *Phys. Lett. A.*, Vol. 246, No. 412, 1998.
7. Smektala, F., C. Quemard, V. Couderc, and A. Barthelemy, *J. Non-cryst. Solids*, Vol. 274, No. 232, 2000.
8. Zhan, C., D. Zhang, D. Zhu, D. Wang, Y. Li, D. Li, Z. Lu, L. Zhao, and Y. Nie, *J. Opt. Soc. Am. B*, Vol. 19, No. 369, 2002.
9. Akhmediev, N. N. and A. Ankiewicz, "Three sources and three component parts of the concept of dissipative solitons," *Dissipative Solitons from Optics to Biology and Medicine*, N. N. Akhmediev and A. Ankiewicz (Eds.), 1–28, Springer, Heidelberg, 2008.

# Modal Dispersion Characteristics of Different Cross Sectional Optical Waveguides

**Y. K. Prajapati**[1], **Vivek Singh**[2], **J. P. Saini**[1], **and Alka Verma**[1]

[1]Dept. of E & C, BIET, Jhansi 284128, U.P., India
[2]Banaras Hindu University, Varanasi 221005, U.P., India

**Abstract**— In this article, an analytical treatment of electromagnetic wave propagation having two different cross-sections based on a boundary matching condition and cutoff values ($V_c$) of a waveguides are computed and compare with standard circular optical waveguide. The boundary of waveguides can be described by equations $r = \xi \exp(1+\sin\theta)$ for optical waveguide with a core slightly flattened on one side, where $\xi$ is a size parameter and $x^N + y^N = a^N$, where $N$ takes the values 1 and 4. For $N = 1$, we have a circular boundary (standard fiber), for $N = 4$, the boundary is a Piet-Hein curve. At the core-cladding boundary, we will put its value equal to $a$, where $a$ is a fixed constant. We show that propagation characteristics of the guided wave are affected when the circular waveguide is deformed and when the circular symmetry no longer exists. And also we found that when the value of the $V$-parameter increases, the number of sustained modes also increases. For practical use the Piet Hein waveguide is therefore more acceptable than the slightly flattened waveguide.

## 1. INTRODUCTION

The invention of optical communication through dielectric waveguides [1, 2], the progress made in the field of fiber optics and transmission technology has been phenomenal. The impact of fiber optics on communication technology is evident in the widespread use of the dielectric waveguide in communication technology and various other technical applications. Technologists suggested that, not only are the transmission properties of an optical waveguide dictated by its refractive index profiles, but structural differences are also useful for various applications in integrated optics and optical communications. Consequently, various noncircular waveguides, like those having elliptical, rectangular, triangular, annular, Piet-Hein, cardiodic, and other cross sections [1–5], were researched by many investigators. Currently, the study of such noncircular waveguides has assumed great importance both theoretically and experimentally. we may point out that the problems are very difficult to analyse under the strong guidance condition [6, 7]. Here the difference between the refractive indices of the core material and the cladding material is not small. But in the case of weak guidance $(n_{\text{core}} - n_{\text{clad}}) \ll 1$, the partial differential equations are easier to solve by the method of separation of variables under some suitable approximation. Now differential equations appear when unfamiliar orthogonal coordinate systems are used and it is interesting to deal with the solutions of these new equations. Many new geometrical shapes have been analyzed using this simple method [3].

## 2. THEORY

We consider the circular core cross-section and two distorted shapes, the Piet Hein curve and slightly flattened waveguide shown in Fig. 1 respectively. The two distorted shapes in Figs. 1(a) and 1(b) can also be looked upon as distortions of a circle. We consider the two new shapes in the following sections.

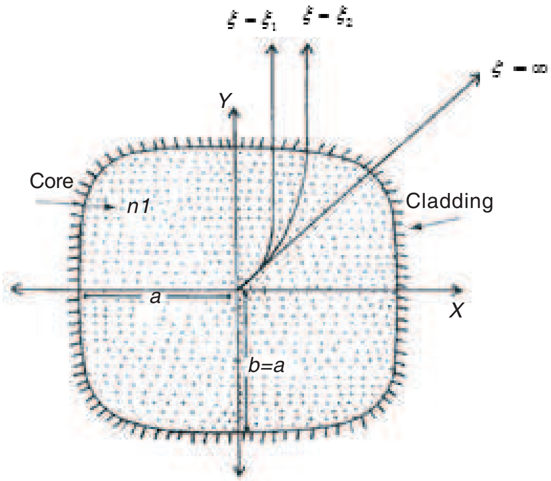
### 2.1. The Characteristic Equations

*2.1.1. The Waveguide with a Piet Hein Core Cross Section*

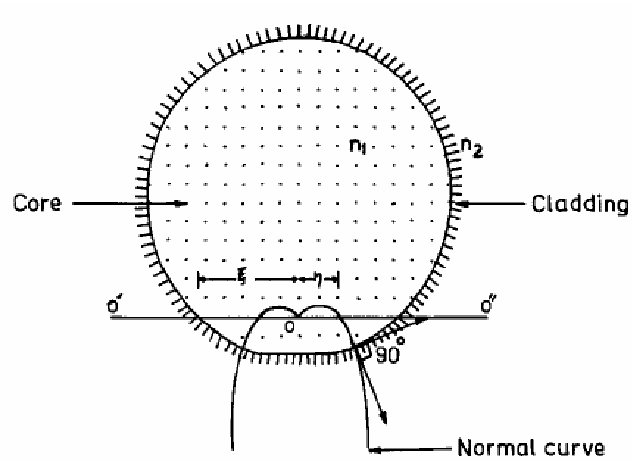
A schematic cross sectional view of the waveguide with a Piet Hein core cross section having core refractive index $n_1$ and cladding refractive index $n_2 \cdot (\frac{n_1 - n_2}{n_1} \ll 1)$ is shown in Fig. 1(a). This shape of the cross section is represented by the equation

$$x^N + y^N = a^N \tag{1}$$

where $N = 4$ and $a$ is a parameter related to the size of the proposed core cross section. We now want to choose an appropriate coordinate system $(\rho,\ \xi,\ z)$ suitable for the analysis of the proposed




(a)




(b)

Figure 1: (a) The crossCsectional view of one side slightly flat waveguide. (b) The cross-sectional view of the Piet-Hein waveguide.

cross section. If we take a variable size parameter $\rho$ instead of $a$ in Eq. (1) we obtain a set of parametric curves

$$x^4 + y^4 = \xi^4 \tag{2}$$

The set of curves normal to above set may be written as

$$\frac{1}{x^2} - \frac{1}{y^2} = \frac{1}{\xi^2} \tag{3}$$

where $\xi$ is to be treated as a new size coordinate variable. If we fix the value of $\rho$ we have a fixed curve. We may now make a coordinate transformation from ($x$, $y$, $z$) to ($\rho$, $\xi$, $z$). After straight forward steps we obtain the scale factors $h_1$, $h_2$ and $h_3$ as follows,

$$\begin{aligned} h_1 &= \frac{\rho^3}{A^{1/2}\left[(\xi^2+\rho^2) - \xi^2\sqrt{\xi^2+\rho^2}\right]^{1/2}} \\ h_2 &= \frac{\sqrt{\xi^2+\rho^2} - \xi^2}{A^{1/2}\left(\xi^2+\rho^2\right)^{1/4}} \\ h_3 &= 1 \end{aligned} \tag{4}$$

where $A = [\{\left(\xi^4+\rho^4\right)^{1/2} + \xi^2\}^2 - 4\xi^4]^{1/2}$.

Now the scalar Helmholtz equation can be written as

$$\nabla^2 E_z + \omega^2 \mu_0 \in E_z = 0 \tag{5}$$

where

$$\nabla^2 = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial\rho}\left(\frac{h_2 h_3}{h_1}\frac{\partial}{\partial\rho}\right) + \frac{\partial}{\partial\xi}\left(\frac{h_1 h_3}{h_2}\frac{\partial}{\partial\xi}\right) + \frac{\partial}{\partial z}\left(\frac{h_1 h_2}{h_3}\frac{\partial}{\partial z}\right)\right] \tag{6}$$

Here $E_2$ is the $z$-component of the electric field, $\omega$ is angular frequency of the electromagnetic wave, $\mu_0$ is the permeability of free space and $\mu$ is the permittivity of the medium.

Now using Eqs. (4) and (6) we deduce

$$\begin{aligned} \nabla^2 \;=\; & \frac{3A\xi^2}{\rho^7}\left(\sqrt{\xi^4+\rho^4} - \xi^2\right)\frac{\partial}{\partial\rho} + \frac{A\sqrt{\xi^4+\rho^4}}{\rho^6}\left(\sqrt{\xi^4+\rho^4} - \xi^2\right)\frac{\partial^2}{\partial\rho^2} \\ & + \frac{3A\xi}{\left(\sqrt{\xi^4+\rho^4} - \xi^2\right)^2}\frac{\partial}{\partial\xi} + \frac{A\sqrt{\xi^4+\rho^4}}{\left(\sqrt{\xi^4+\rho^4} - \xi^2\right)^2}\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial z^2} \end{aligned} \tag{7}$$

The wave equation for a structure having Piet-Hein geometry in terms of the new coordinates can be written by using Eqs. (5) and (7) as

$$\left[\frac{3A\xi^2}{\rho^7}\left(\sqrt{\xi^4+\rho^4}-\xi^2\right)\frac{\partial}{\partial\rho}+\frac{A\sqrt{\xi^4+\rho^4}}{\rho^6}\left(\sqrt{\xi^4+\rho^4}-\xi^2\right)\frac{\partial^2}{\partial\rho^2}\right.$$
$$\left.+\frac{3A\xi^0}{\left(\sqrt{\xi^4+\rho^4}+\xi^2\right)^2}\frac{\partial}{\partial\xi}+\frac{A\sqrt{\xi^4+\rho^4}}{\left(\sqrt{\xi^4+\rho^4}-\xi^2\right)^2}\frac{\partial^2}{\partial\xi^2}\right]E_z+\left(n^2k_0^2-\beta^2\right)E_z=0 \tag{8}$$

where $n$, $\beta$ and $k_0$ are the refractive index of the material, $z$-component of the propagation vector and the free space propagation constant respectively. The boundary conditions demand that the tangential components of the electric field $E$ and the magnetic field $H$ at the interface should be continuous. The boundary conditions can be written as

$$E(\rho)|_r = E(\rho)|_{ri} \quad \text{and} \quad \left.\frac{\partial E(\rho)}{\partial\rho}\right|_r = \left.\frac{\partial E(\rho)}{\partial\rho}\right|_{ri}$$

Thus we get a set of equations having twenty two unknown constants. The nontrivial solution will exist only when the determinant formed by the coefficients of the unknown constants is equal to zero. Calling this $4\times 4$ determinant $\Delta$, we have

$$\Delta = 0 \tag{9}$$

*2.1.2. The Waveguide with a One Side Slightly Flattened Waveguide*

The transverse cross sectional view of the proposed waveguide is shown in Fig. 1(b). The core region of the waveguide having refractive index $n_1$ and surrounded by the cladding layer of dielectric medium having refractive index $n_2$. Where $(n_1-n_2)/n_1 \ll n_1$ in the weak guidance approximation.

The polar equation of the curve in polar coordinates is written as

$$r = \xi\exp(1+\sin\theta) \tag{10}$$

where $\xi$ is a size parameter. At the core-cladding boundary, we will put its value equal to $a$, where $a$ is a fixed constant. The equation of the normal curve to the curve represented by Eq. (9). in polar coordinates is written as

$$r_\perp = \eta\frac{cos\theta}{(1+sin\theta)} \tag{11}$$

For our proposed waveguide structure, a new coordinate system ($\xi$, $\eta$, $z$) will be suitable. The scalar Helmholtz equation of electromagnetic wave propagation along the core is

$$\nabla^2\psi+\omega^2\varepsilon\mu\psi=0 \tag{12}$$

where $\nabla^2$ is a two-dimensional scalar Laplacian operator, the wave function $\psi$ may represent one of the orthogonal components of the electric/magnetic field, $\omega$ is the angular frequency of the field, $\mu$ is the permeability of the nonmagnetic dielectric material of the core region, $\varepsilon$ is the permittivity of the guiding or core region, and $\beta$ is the propagation constant of the guided mode [8]. Now, the slightly flattened part of our proposed waveguide is of main importance and, because the flattened portion lies within the normal curve, we make an approximation $\xi=\eta$ in order to facilitate our calculation.

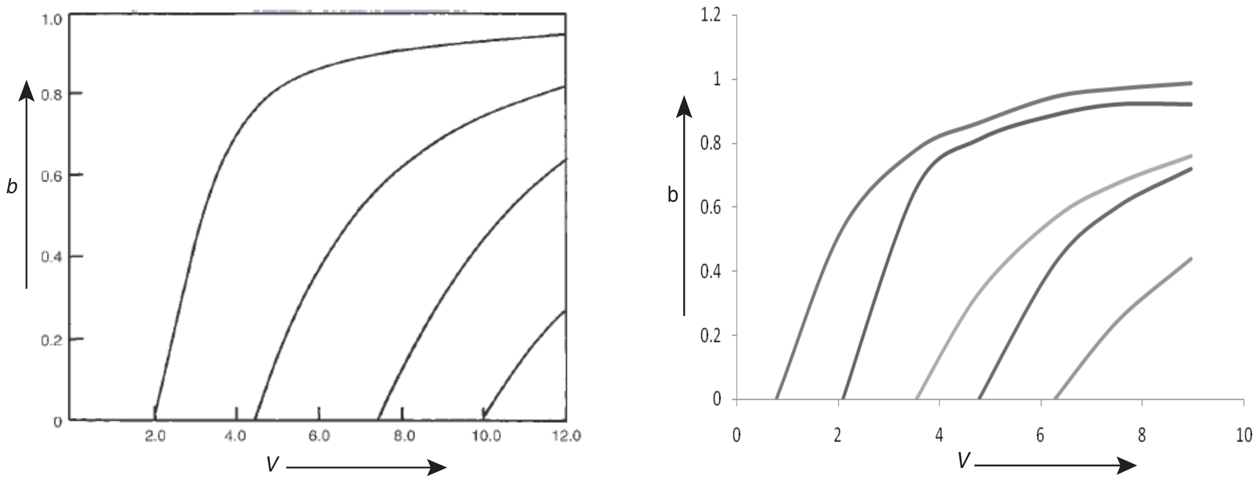
$$\frac{\xi}{\eta}\frac{\partial^2\psi}{\partial^2\xi}+\frac{1}{\eta}\frac{\partial\psi}{\partial\xi}+\frac{\eta}{\xi}\frac{\partial^2\psi}{\partial\eta^2}+\frac{1}{\xi}\frac{\partial\psi}{\partial\eta}+d^2\frac{\xi}{\eta}\frac{\partial^2\psi}{\partial z^2}+p^2\omega^2\mu\varepsilon\psi=0 \tag{13}$$

where $p=\sqrt{e/2}$ and $e=2.718$.

Now, for the core and cladding regions, which have permittivity's $\varepsilon_1$ and $\varepsilon_2$ respectively. After applying the boundary conditions, we obtain the following eigenvalue equation:

$$U\frac{J_0'(pUa)}{J_0(pUa)}-W\frac{K_0'(pWa)}{K_0(pWa)}=0 \tag{14}$$

After solving Eq. (14), we will get an idea of the modal properties of the light guide of the proposed cross section under the weak guidance condition.




(a) (b)

Figure 2: (a) Dispersion curves of normalized frequency $V$ versus normalized propagation constant $b$ for the Piet Hein waveguide. (b) Dispersion curves of normalized frequency $V$ versus normalized propagation constant $b$ for the one side slightly flat waveguide.

## 3. NUMERICAL RESULT AND DISCUSSION

The most important equation of our analysis is the characteristics Eqs. (9) and (14) which gives us all information regarding the modal characteristics of the proposed waveguides analytical method. We choose an operating wavelength $\lambda = 1.55\,\mu\text{m}$ and various values of dimensional parameter a in a regular increasing order. For each value of $a$ we compute the $V$-parameter and also obtain the values of $\beta$ from the characteristic Eq. (9) and Eq. (14) by graphical method. This means that the left hand side of the characteristic equation is plotted against $\beta$ for the assumed value of $a$ and the zero crossing of the graph with the $\beta$ axis are noted. These values are the solutions of the characteristic equation for different modes. For example the lowest zero crossing value of $\beta$ corresponds to the lowest order mode. From these values of $\beta$ we can find out the values of $b'$ and then plot the dispersion curves for the different modes. These graphs are shown in Fig. 2(a) and Fig. 2(b) for the Piet Hein waveguide and the one side slightly flat waveguide respectively. In the case of standard circular fiber, the first cutoff occurs at $V = 0$, and successive cutoffs occur at $V = 2.09$ etc. We observe that the first cutoff shifts from $V = 0$ to $V = 0.785$ approximately when we have a slightly flattened waveguide. The second and third shifts in the flattened waveguide follow closely. Thus, the first cutoff is shifted slightly toward higher values of $V$, whereas the next cutoff $V$-values occur at slightly lower values than those of the standard $V$-values for a circular waveguide. Finally, we get five modes when $V = 10.5$. And in the second equation, it is found that the Piet-Hein shape combines the desirable characteristics of the circular and the rectangular waveguide. Piet Hein cross section, we find that these curves have the standard shapes and for $V = 10.5$ there are four modes. The lowest order mode in this case has a cutoff $V$-value nearly $V = 2.0$. For practical use the Piet Hein waveguide is therefore more acceptable than the slightly flattened waveguide.

## REFERENCES

1. Snitzer, E., "Cylindrical dielectric waveguide modes," *J. Opt. Soc. Amer.*, Vol. 51, 491–498, 1961.
2. Yeh, C., "Modes in weakly guiding elliptical optical fibers," *Opt. Quantum Electron.*, Vol. 8, 43–47, 1976.
3. Singh, V., B. Prasad, and S. P. Ojha, "Weak guidance modal analysis and dispersion curves of an infrared-lightguide having a core cross-section with a new type of asymmetric loop boundary," *Opt. Fiber Technol.*, Vol. 6, 290–298, 2000.
4. Singh, V., B. Prasad, and S. P. Ojha, "Theoretical analysis and dispersion curves of an annular lightguide with a cross-section bounded by two Piet-Hein curves," *Journal of Electromagnetic Waves and Applications*, Vol. 17, No. 7, 1025–1036, 2003.

5. Rao, M. P. S., B. Prasad, P. Khastgir, and S. P. Ojha, "Modal cutoff conditions for an optical waveguide with a hypocycloidal cross section," *Microw. Opt. Technol. Lett.*, Vol. 14, 177–180, 1997.
6. Rahman, B. M. A. and J. B. Davies, "Finite element analysis of optical and microwave waveguide problems," *IEEE Trans. Microw. Theory Tech.*, Vol. 32, 20–28, 1984.
7. Lui, W. W., C. L. Xu, W. P. Huang, K. Yokoyama, and S. Seke, "Full vector mode analysis with considerations of field singularities at corners of optical waveguides," *J. Lightwave Technol.*, Vol. 17, 1509–1513, 1999.
8. Snyder, W. and J. D. Love, *Optical Waveguide Theory*, Chapman & Hall, New York, 1983.

# The Transition between Superluminal and Subluminal for Optical Resonant Cavity

**Yundong Zhang**[1], **Jing Zhang**[1], **Xiangchun Ju**[1], **Ping Yuan**[1], **Yuhua Zhang**[2], **and Sheng Qiang**[1]

[1]Harbin Institute of Technology, Harbin 150080, China
[2]Art College, Harbin Normal University, Harbin 150080, China

**Abstract**— The transmission extremum dots and dispersion reversal dots are analyzed to interpret the substance of CRIT and CRIA in the complex plane. The sensitivity of dispersion response is affected by the outmost sphere's loss coefficient and coupler parameter. Normal dispersion converts to abnormal dispersion for even number of active resonators, on the contrary, for odd numbers the dispersion changes from anomalous to normal on the under coupled condition. Meanwhile, a modified spectral intensity and a controllable light speed are of application of optical sensors, switcher and micro-lasers.

## 1. INTRODUCTION

Optical micro-cavity structures have more broad space for application than atomic systems because of miniaturization, integrating and easily control. The photonics resonance gradually displaces electronic resonance to control the propagation speed of light. Many researches focus on the basic characteristics of resonant systems without gain [1]. Generally, two or even numbers of micro-spheres resonant structures which cause a cancellation of absorption as a result of classic destructive interference and mode splitting generate slow light and odd numbers of resonators produce fast light lacking of gain media. The transition between superluminal and subluminal are investigated by means of adjusting coupler parameters and loss coefficients in coupled resonance micro-cavities. The effective phase shift is achieved by measurements of optical power originating from detector from the view of experiment [2]. The combination of fast/slow light mechanism and novel microstructures can significantly impact on the development of optical sensors [3] and interferometers [4]. In this letter, the gain-assistance system which is composed of micro-spheres coupled with a fiber taper has been considered. Different concentrations of $Er^{3+}$ are doped into the outmost sphere that contributes to reduction of inherent loss and amplication of input light intensity. Besides it is paid special attentions in the even numbers of resonators which could produce the transition from slow light to fast light and odd numbers of resonator is the other way round in under-coupled circumstance. The complex transfer function which links the transmissivity to the effective phase shift is decomposed into real component and imaginary part in the complex plane. The significant points in the complex plane are in depth analysis in order to boost the comprehension of dynamic characteristics in resonant micro-cavities. It's necessary to investigate the essential propagation properties of the gain-assistance system for the application as optical amplifiers, sensors and integrated micro-cavity lasers.

## 2. THE MODEL

We investigate a configuration of miniaturization structure consisting of two and three active micro-spheres coupled with a taper. Fig. 1 depicts the schematic of the gain-assistance structure. The pump light is originated from 980 nm laser diode to excite energy level of erbium ion in the outmost micro-sphere so as to compensate for probe signal. The probe light is a tunable 1550 nm laser diode with narrow linewidth in order to scan the transmittance spectrum. Two beams are coupled into the fiber taper by dichroic mirror and the straight light along the fiber taper partly permeates into the passive micropheres as an evanescent wave. Transfer function is the crucial parameter to describe the ratio of the output to the input optical field amplitude. Moreover, those based on the directional coupled-mode theory involving weakly guiding approximation are presented for analysis of the coupled resonant microcavity induced dispersion. For two and three passive spheres with equal length transfer constants are as follow:

$$\tau_n = \frac{E_{4n+3}}{E_{4n}} = \frac{b_n \cos(\varepsilon_n) - a_n b_n^2 \tau_{n-1} \exp(i\phi_n)}{1 - a_n b_n \cos(\varepsilon_n)\tau_{n-1}\exp(i\phi_n)} = \sqrt{T_{on}}\exp\left(i\phi_n^{eff}\right) \tag{1}$$

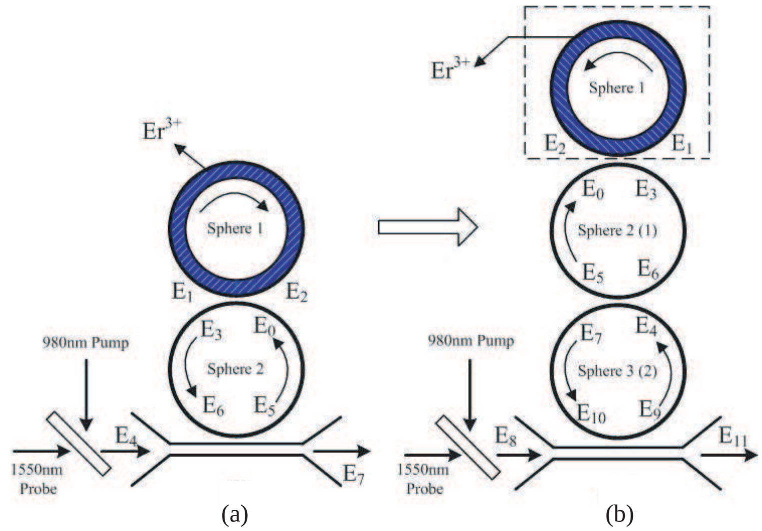



Figure 1: The configuration of the gain-assistance system.

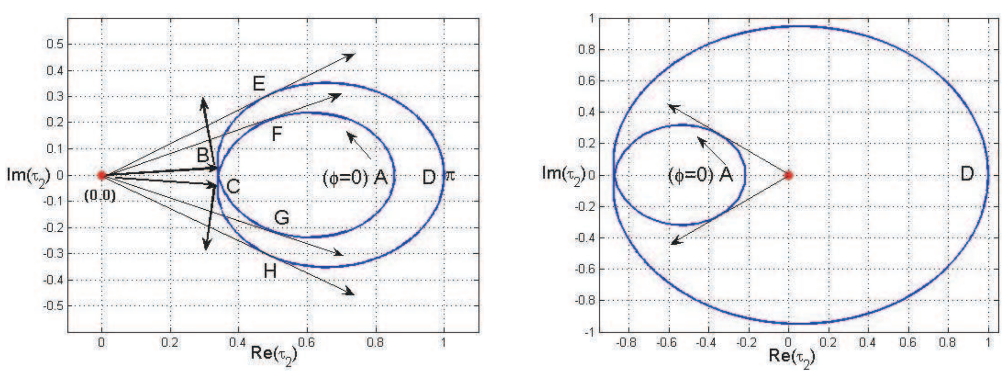



Figure 2: The images show CRIT ($m_1 = 0.999975$, $n_1 = 0.999999986$; $m_2 = 0.999820$, $n_2 = 0.999899$) on the left and CRIA ($m_1 = 0.999975$, $n_1 = 0.999999986$; $m_2 = 0.998120$, $n_2 = 0.999899$) on the right in the complex plane for conventional double microspheres resonators.

where the transfer function for single sphere structure is given by:

$$\tau_1 = \frac{E_3}{E_0} = \frac{b_1 \cos(\varepsilon_1) - b_1^2 a_1 \tau_1 \exp(i\phi_1)}{1 - a_1 b_1 \cos(\varepsilon_1)\,\tau_1 \exp(i\phi_1)} \tag{2}$$

$\varepsilon_n$ is the coupling strength among the spheres and the fiber taper; The taper loss is expressed as $b_n = 10^{-\frac{\gamma_n}{20}}$; $\gamma$ is the insertion loss. In each sphere the loss factor is $a_n = e^{(-\frac{\alpha_n L}{2})}$; $\alpha_n$ is single ring-trip loss. On the condition of gain it's set as $a_n = e^{(-\frac{(\alpha_n - g_n)L}{2})}$; $g_n$ is the gain factor. The loss coefficient is described as $m_n = a_n \cdot b_n$ and coupled parameter is set as $n_n = \cos_n(\varepsilon_n)$. $\tau_n(\tau_0 = 1)$ is transition function and $T_{on} = |\tau_2|^2$ ($T_{out}$ is output light intensity).The single-pass phase shift is shown as $\phi_n = \frac{nwL_n}{c}$; $\phi_n^{eff}$ is effective phase shift. For simplicity, we assume the taper loss as $b_n = 1$, $\phi = \phi_1 = \phi_2$ and then the structural properties of dispersion are discussed by different matching with respect to couple parameters and loss coefficients.

Figure 2 shows the complex transfer function which consists of the real part and imaginary components in the complex plane for two spheres resonant structure in the case of CRIT and CRIA. The electromagnetic wave going though the outmost sphere crosses the coupling region twice, and thus the phase of this optical field is $\pi$ rad shifted. This phase shift induces destructive interference between sphere 1 & 2 and the induced transparency window could be observed when the frequency of input electric field approximate whispering gallery mode (WGM) of microspheres at the equatorial surface. Therefore this effect is regarded as CRIT in artificial photonic structure analogous to EIT in atomic system. The appearance of mode splitting is in the form of double spiral curve in the complex plane as far as CRIT in two spheres coupled with a fiber taper. In

Fig. 2 double spiral curve which stars from point A ($\phi = 0$) moves clockwise passing dot D ($\phi = \pi$) and then go back to A over a free spectral range of $2\pi$ to describe the possible track of the complex transfer function $\tau_2(\phi)$ ($0 \le \phi \le 2\pi$). The horizontal and vertical axes are taken as its real part $\mathrm{Re}(\tau_2(\phi))$ and imaginary part $\mathrm{Im}(\tau_2(\phi))$ respectively. Furthermore, they represent dispersive and absorptive behavior accordingly for the under-coupled case in the complex plane.

The derivative of the effective phase shift is expressed as:

$$\frac{d\phi_2^{eff}}{d\phi_2} = \frac{1}{|\tau_2|^2}\left(\mathrm{Re}(\tau_2)\frac{d\,\mathrm{Im}(\tau_2)}{d\phi_2} - \mathrm{Im}(\tau_2)\frac{d\,\mathrm{Re}(\tau_2)}{d\phi_2}\right) = \frac{1}{|\tau_2|}\left|\vec{\tau}_2(\phi_2)\right| \cdot \left|\frac{d\vec{\tau}_2(\phi_2)}{d\phi_2}\right| \cdot \sin\theta \tag{3}$$

To acquire inflexions of the effective phase shift, the transmission phasor needs to parallel to its gradient. Obviously, four points, such as E, F, G and H, are marked which are called dispersion reversal dots corresponding to the four extreme points of dispersion curve. Besides they divide the complex plane into five parts over a free spectral range and near the resonant area the normal dispersion occurs in Fig. 2 on the left. The effective phase shift (the property of dispersion) curve alternates between abnormal and normal dispersion for CRIT. The fact that the effective phase shift is increasing or decreasing as the single pass phase shift augments would determine whether dispersion is normal or anomalous. However, as Fig. 2 shows on the right there only are two dispersion reversal dots describing the dispersion for CRIA.

Figure 3 shows the transmissivity changes from dip to peak as a function of the frequency detuning and loss coefficient near resonance for two passive resonators. The phenomenon of CRIT emerges as the amplitude loss ratio gradually approach the amplitude coupling ratio(the coupling ratio is kept constant $n_1 = 0.99999998$) and both are less than one in the case of under-coupled without gain. Especially, the maximum of peak transmittance corresponding to the maximum of dispersive sensitivity is obtained at $m_1 = 0.999975$ varying from 0.999 to 0.999975 because of intra-cavity coherence enhancement in Fig. 3. The loss coefficient of outmost sphere decides the magnitude of peak transmittance however the value of group delay dispersion related to low distortion pulse depends on the loss coefficient in lower sphere. In view of normal or abnormal dispersion which is determined by the derivative of the effective phase shift, optical pulse propagating in two spheres-taper system is numerically simulated to describe the delay and advance at the time domain in Fig. 4. Doping $Er^{3+}$ into the sphere 1 could give rise to the severe reduction of transmissivity at the resonant frequency yet the reversal of the effective phase shift curve occurred at the same time.

Two following resonant systems are in the case of under coupled that means $n_1 > m_1$, $n_2 > m_2|\tau_1|$ and $n_3 > m_3|\tau_2|$ at first, however, $n_1 < m_1$ is called over coupled. In Fig. 5(a) the red dotted line represents normal dispersion ($m_1 = 0.999975$, $n_1 = 0.999999986$; $m_2 = 0.999820$, $n_2 = 0.999899$) before outmost sphere is doped gain medium. And the blue line signifies it converts to anomalous dispersion as $m_1 = 1.01$ for two spheres linking with a fiber taper. The numbers of the dispersion reversal points change from four to two in the complex plane corresponding to the inflexions of effective phase shift as Fig. 5(a) illustrate. The original parameters are shown as: $m_2 = 0.999975$, $n_2 = 0.999999986$; $m_1 = 0.999975$, $n_1 = 0.999999986$; $m_3 = 0.999820$,

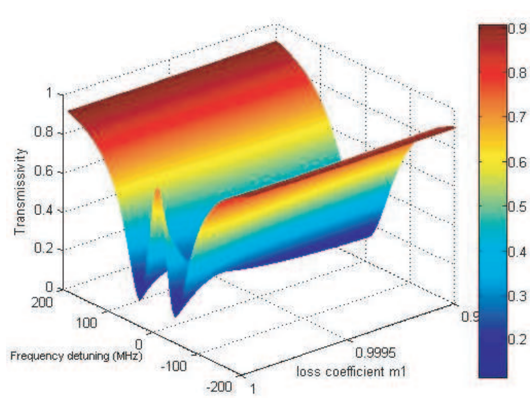



Figure 3: The transmissivity for CRIT.

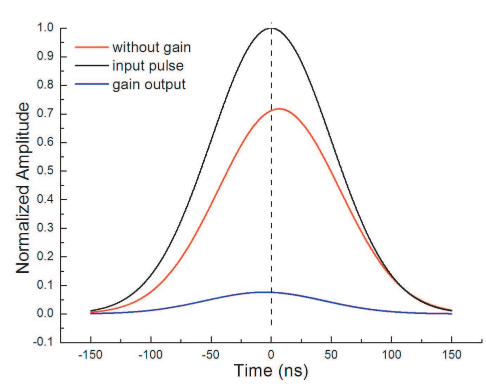



Figure 4: The propagation properties of pulse with gain.

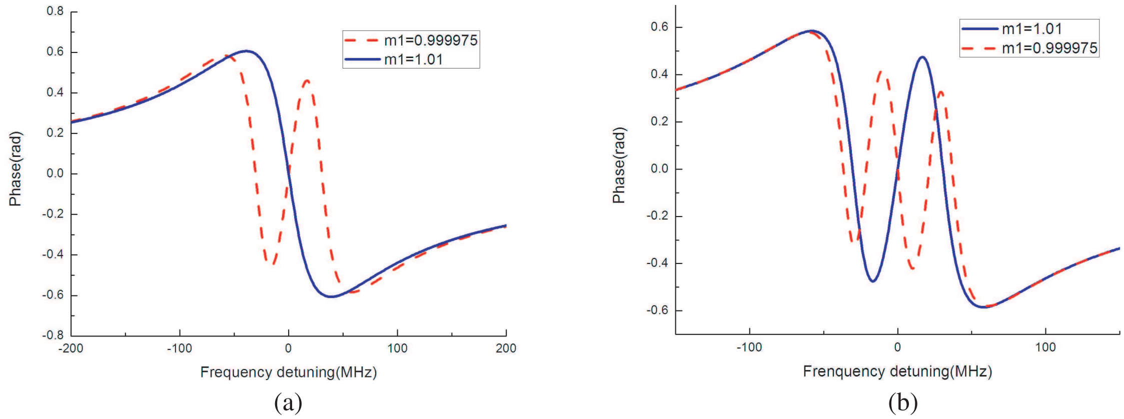



Figure 5: The dispersion reversal varying from normal to anomalous. (a) Two resonators. (b) Three resonators.

$n_3 = 0.999899$ for three coupled active resonators in Fig. 1(b). On the contrary, the transition between anomalous dispersion and normal dispersion in Fig. 5(b) due to the appearance of gain as $m_1 = 1.01$.

## 3. CONCLUSION

The extreme values of transmissivity occur as the transmission phasor is perpendicular to its derivative, however, the appearance of dispersion reversal is in the case that transmission phasor needs to parallel to its gradient in the complex plane for two passive resonators. The loss coefficient and coupler parameter of outmost sphere are a significant impact on spectral characteristics near the resonant frequency under this circumstance. The dispersion response becomes more sensitive the closer loss coefficient gets to coupler parameter. On the under coupled condition normal dispersion converts to anomalous dispersion accompanying with the reduction of the spectral intensity for even numbers of active resonators. On the contrary, for odd numbers the transmissivity enhancement could follow with the dispersion reversal from anomalous to normal. By means of adjusting the concentration of erbium ion the transmissivity would be modified instantly in order to change spectral intensity and control the velocity of light for the application of micro-cavity lasers, optical filters , optical switches and so on.

## ACKNOWLEDGMENT

The research is supported by the National Natural Science Foundation of China (Contracts No. 6027-2075, 60478014 and 60878006), and the National High Technology Research and Development Program (863 Program) of China (2007AA12Z112).

## REFERENCES


1. Smith, D. D., H. Chang, K. A. Fuller, A. T. Rosenberger, and R. W. Boyd, "Coupled-resonator-induced transparency," *Phy. Rev. A*, Vol. 69, 063804-1-6, 2004.
2. Dumeige, Y., T. K. N. Nguyên, L. Ghia, S. Trebaol, and P. Féron, "Measurement of the dispersion induced by a slow-light system based on coupled active-resonator-induced transparency," *Phy. Rev. A*, Vol. 78, 013818-1-5, 2008.
3. Smith, D. D., H. Chang, L. Arissian, and J. C. Diels, "Dispersion-enhanced laser gyroscope," *Phy. Rev. A*, Vol. 78, 053824-1-9, 2008.
4. Terrel, M., M. J. F. Digonnet, and S. H. Fan, "Ring-coupled Mach-Zehnder interferometer optimized for sensing," *Appl. Opt.*, Vol. 48, No. 26, 4874–4879, 2009.

# Observation of the Phase Shift and Group Delay in Nested Optical Fiber Ring Resonator

**Yundong Zhang**[1], **Jinfang Wang**[1], **Xiangchun Ju**[1], **Ping Yuan**[1], **Yuhua Zhang**[2], **and Sheng Qiang**[1]

[1]Harbin Institute of Technology, Harbin 150080, China
[2]Art College, Harbin Normal University, Harbin 150080, China

**Abstract**— We theoretically and experimentally demonstrate the transmission characteristics and time delay performance in a nested fiber ring resonator. The nested ring resonator is composed with a U-shaped fiber loop and a fiber ring nested together. With proper design of the coupling coefficient between the ring and the loop, the device is capable of generating a double-Fano spectral response. And the maximal group delay that it can be introduced depends on the value of the coupling coefficient. In this paper we measure the transmission properties of our nested fiber ring resonators with different coupling coefficients, using a Mach-Zehnder fiber interferometer. By inserting the nested fiber ring resonator into one arm of the fiber Mach-Zehnder interferometer we are able to optically measure the intensity transmission factor and the phase shift produced by the coupled resonators as a function of the optical frequency. The group delay can be deduced from this information. This theoretical and experimental demonstration shows the potential of nested ring resonators for dispersion tailoring and slow-light applications.

## 1. INTRODUCTION

Optical ring resonator have attracted much attention in resent years for their potential use as novel optical components, such as optical filters [1], switches [2], delay lines [3], buffers and interferometers [4]. The advantage of a fiber-ring resonator is that it can be constructed in a variety of ring sizes ranging from several meters to micrometers, with spectral characteristics that scale with the coupling coefficient, attenuation factor and size of the device. Furthermore, fiber-ring resonator can be readily constructed from standard optical fiber components, which allows their properties to be studied in a systematic manner.

In this paper we demonstrate the phase shift and timedelay of a nested fiber-ring resonator theoretically and experimentally. The schematic of the considered structure is shown in Fig. 1, which consists of a fiber ring coupled to a U-shaped fiber loop with coupling coefficient $\kappa_1$ and $\kappa_2$.

The paper is organized as follows. In Section 2 we present a theoretical formulation of the nested optical fiber ring resonator. In Section 3 we show experimental demonstration of this structure. And finally Section 4 shows the conclusion of the paper.

## 2. THEORETICAL FORMULATION OF THE NESTED FIBER-RING RESONATOR

For simplicity, it is assumed that there is no backward reflection in the coupling region, and that material dispersion is negligible. The coupling is treated as being lumped and localized only at

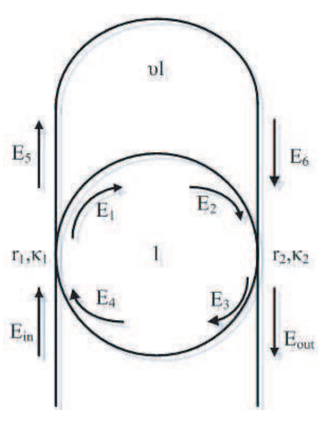



Figure 1: The nested optical fiber ring resonator studied in this paper.

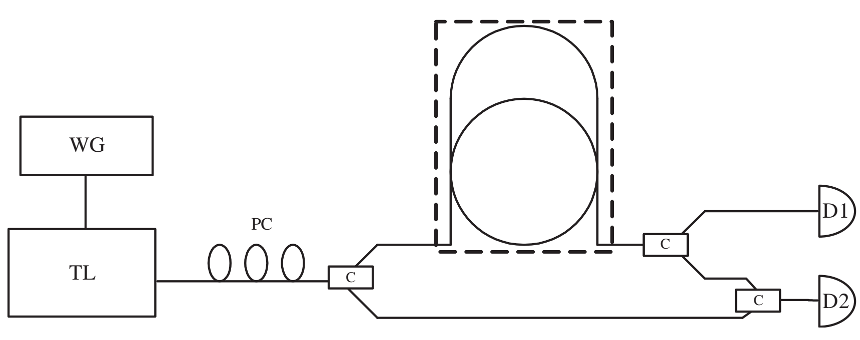



Figure 2: The experimental setups as the nested fiber ring resonator in one arm of a Mach-Zehnder fiber interferometer. There are wave generator, tunable laser, polarization controller, 3 dB couplers and detectors.

the point of tangent. We describe the coupling of light into and out of the resonator in terms of generalized splitter relations of the form. Our mathematical analysis of this structure proceeds as follows.

The transfer equations for the structures are given by

$$\begin{aligned} E_5 &= \sqrt{1-r_0}\,(r_1 E_{in} + i\kappa_1 E_4) \\ E_1 &= \sqrt{1-r_0}\,(i\kappa_1 E_{in} + r_1 E_4) \\ E_{out} &= \sqrt{1-r_0}\,(r_2 E_6 + i\kappa_2 E_2) \\ E_3 &= \sqrt{1-r_0}\,(i\kappa_2 E_2 + r_2 E_2) \end{aligned} \tag{1}$$

where $r_j$ and $\kappa_j$ ($j = 1, 2$) are the reflection and coupling coefficients of the two region, respectively, satisfying the relation $r_j^2 + \kappa_j^2 = 1 - r_0$; $i$ denotes the phase coupling which is a $\pi/2$ phase shift that occurs upon transmission across the coupling region, $r_0$ is the fractional coupler intensity loss.

The propagation between the adjacent coupling points is represented by

$$\begin{aligned} E_2 &= a_{12} e^{i\phi_{12}} E_1 \\ E_4 &= a_{34} e^{i\phi_{34}} E_3 \\ E_6 &= a_{56} e^{i\phi_{56}} E_5 \end{aligned} \tag{2}$$

where $a = e^{-\frac{\alpha l}{2}}$ is the attenuation factor and $\phi = \frac{n_1 l}{c}\omega$ is the single-pass shift, and $a_{12} = a_{34} = e^{-\frac{\alpha \frac{l}{2}}{2}} = \sqrt{a}$, $a_{56} = e^{-\frac{\alpha \nu l}{2}}$, $\phi_{12} = \phi_{34} = \frac{n_1 \frac{l}{2}}{c}\omega = \frac{\phi}{2}$, $\phi_{56} = \frac{n_1 \nu l}{c}\omega$, $\alpha$, $n$, $c$ and $l$ are the loss coefficient, the refractive index, the velocity of light in the vacuum and the length of the ring, respectively. The U-loop connecting the two coupler has a length of $\nu l$.

Combine Equations (1) and (2) to obtain the complex transmission coefficient for the structure.

$$\tau = \frac{(1-r_0)\left[-\kappa^2\sqrt{a}e^{\frac{i\phi}{2}} + r^2 a^v e^{iv\phi} - a^{v+1} e^{i(v+1)\phi(1-r_0)^3}\right]}{1-(1-r_0)\left(r^2 a e^{i\phi} + \kappa^2 a^{v+\frac{1}{2}} e^{i(v+\frac{1}{2}\phi)}\right)} \tag{3}$$

The transmission factors $T$ are defined as the squared modulus of $\tau$.

$$T = |\tau|^2 \tag{4}$$

The effective phase shift of the transmitted light is given by the argument of the complex transmission coefficient as follows:

$$\Phi^{eff} = \arg(\tau) \tag{5}$$

And the group delay is defined as the radian-frequency derivation of the phase shift $T_d = \frac{d\Phi^{eff}}{d\omega}$. Noting that the frequency $\omega$ is related to the single pass phase shift according to $\varphi_j = \omega t_j$, where $t$ is the round-trip time of the resonator. Hence $\varphi_j$ is indicative of the detuning of the input frequency from the resonance frequency.

## 3. EXPERIMENTAL RESULTS

As it is clear from the former analysis, the transmission characteristics of the optical fiber ring resonator would be influenced by the coupling coefficient. In this section we show experimental demonstration of this property. The experiment setup is shown in Fig. 2, it has the nested fiber ring resonator in one arm of a Mach-Zehnder fiber interferometer. The lengths of the fiber ring: $L = 1.4\,\mathrm{m}$, $\nu = 1.43$, and we assumed that $n\,(\delta)$ is the effective index of standard single mode fiber. The couplers marked as C in Fig. 2 is 3 dB couplers.

The intensity of the light which though the MZ interferometer is defined as

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\Delta\varphi \tag{6}$$

where $I$ is the interfere intensity, $I_1$ and $I_2$ respect light intensity of two arms, $\Delta\phi$ is the phase difference of the interferometers respectively. In our experiment the interfere intensity $I$ is measured though the coupler C2 by detector D2, and $I_1$ is equal to the intensity measurement of detector D1, $I_2 = I_0 r_0^2$. The expressions of optical powers inside the upper arm can be written as

$$I_1 = I_0 \kappa_0^2 T \kappa_1^2 \tag{7}$$

And the phase difference of the interferometers is

$$\Delta\varphi = \arccos\left(\frac{I - I_1 - I_2}{2\sqrt{I_1 I_2}}\right) = \arccos\left(\frac{I - I_0\kappa_0^2 T\kappa_1^2 - I_0 r_0^2}{2I_0\kappa_0 r_0 \kappa_1 \sqrt{T}}\right) \tag{8}$$

Then the transmittance and the whole phase shift $\phi(\delta)$ of the NFRR is deduced as

$$T = I_1 / I_0 \kappa_0^2 \kappa_1^2 \tag{9}$$

$$\Phi^{eff} = \varphi_0 + \Delta\varphi \tag{10}$$

Then the transmittance and the whole phase shift $\phi(\delta)$ of the NFRR can be get by scanning the probe wavelength and by simultaneously measuring the optical power values $I_1$, $I'$ ($I'$ is the intensity measured by detector D2, where $I' = I\kappa_2^2$) and the input power $I_0$.

The timedelay $T_0$ of the NFRR will be deduced by differentiating the phase shift $\phi(\delta)$. In order to reduce the noise impact on the numerical evaluation we smoothed these curves using successively Adjacent-Averaging filters.

Figure 3 is the transmission $T$ and phase shift $\phi(\delta)$ as a function of the detuning $\delta$ for different coupling coefficients. The solid curves represent experimental transmissions and the dashed lines are theoretical ones. From the figures we can find that the experimental data and theoretical calculations show a good agreement for several coupling coefficients. And the transmission at zero detuning increase with the coupling coefficients.

Figure 4 is the normalized input and output signal intensities for different coupling coefficients. The timedelay of the nested fiber ring resonator are 50 ns, 65 ns, 76 ns when the coupling coefficients are 0.9, 0.95 and 0.99.

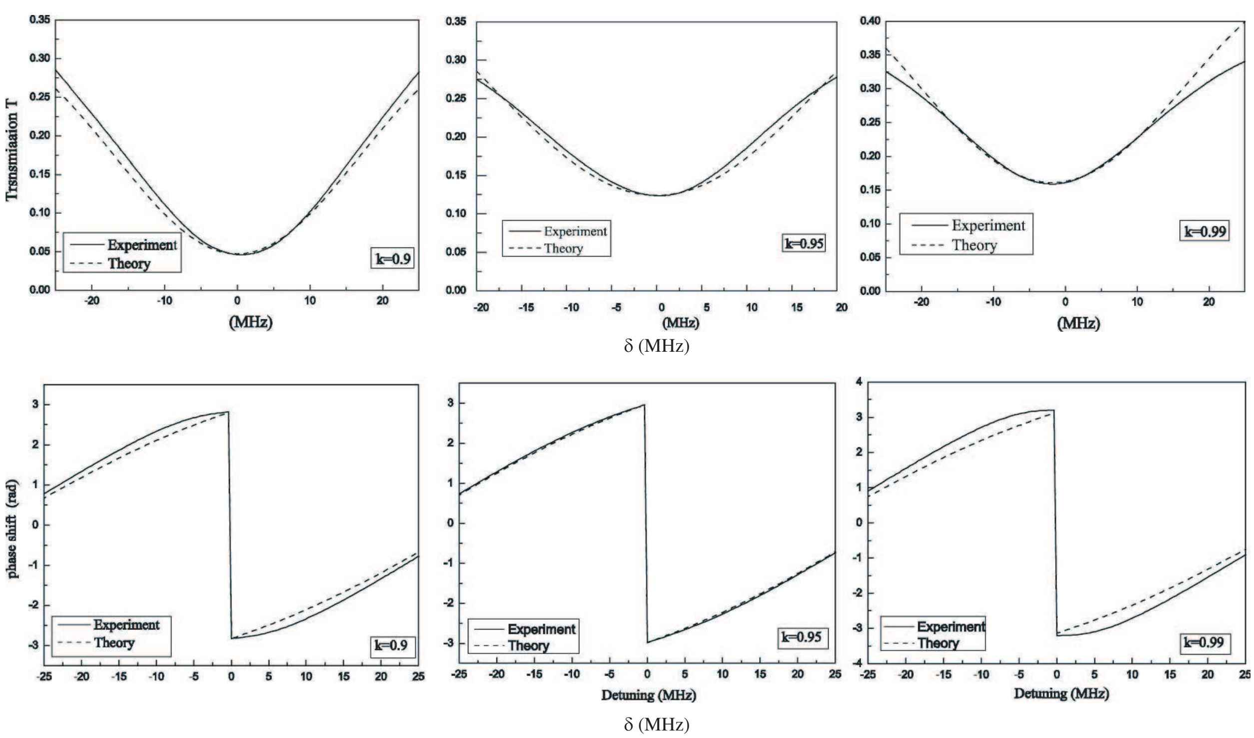



Figure 3: The transmission $T$ and phase shift as a function of the detuning $\delta$ for different coupling coefficients. The solid curves are experimental data and the dashed lines are theoretical fittings. The only parameter which is changed is the value of the coupling coefficient.

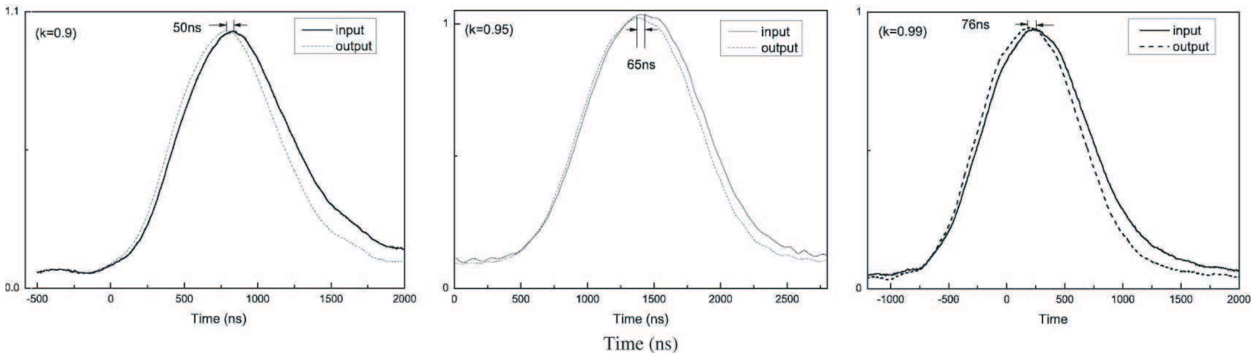



Figure 4: Normalized input and output intensities of different coupling coefficients.

## 4. CONCLUSION

In conclusion we demonstrate theoretically and experimentally the transmission characteristics and time delay performance in a nested fiber ring resonator. We have shown that the transmission and group delay are influenced by the coupling coefficient between the ring and the loop and other structure parameters. We measured the transmission properties and the phase shift of our nested fiber ring resonators with different coupling coefficients, by inserting the nested fiber ring resonator into one arm of the fiber Mach-Zehnder interferometer. Tens of nanosecond timedelay can be get using the nested fiber ring resonator.

## ACKNOWLEDGMENT

The research is supported by the National Natural Science Foundation of China under Grant No. 60878006, the National High Technology Research and Development Program (“863” Program) of China under Grant No. 2007AA12Z112.

## REFERENCES


1. Lenz, G., B. J. Eggleton, C. K. Madsen, et al., “Optical delay lines based on optical filters,” *IEEE J. Quant. Electron.*, Vol. 37, No. 4, 525–532, 2001.
2. Heebner, J. E. and R. W. Boyd, “Enhanced all-optical switching by use of a nonlinear fiber ring resonator,” *Opt. Lett.*, Vol. 24, No. 12, 847–849, 1999.
3. Maleki, L., A. B. Matsko, A. A. Savchenkov, et al., “Tunable delay line with interacting whispering-gallery-mode resonators,” *Opt. Lett.*, Vol. 29, No. 6, 626–628, 2004.
4. Terrel, M., M. J. F. Digonnet, et al., “Ring-coupled Mach–Zehnder interferometer optimized for sensing,” *Appl. Opt.*, Vol. 48, No. 26, 4874–4879, 2009.

# Differentiation of Human LAN-5 Neuroblastoma Cells by Electronically Trasmitted Retinoic Acid (RA)

**Alberto Foletti**[1] **and Settimio Grimaldi**[2]

[1]Independent Researcher, Lugano, Switzerland
[2]Institute of Neurobiology and Molecolar Medicine, CNR, Rome, Italy

**Abstract**— We report the transfer of the activity of Retinoic acid (RA) by electronic means.

Retinoic acid were placed at room temperature on one coil attached to an oscillator (VEGA select 719), while LAN-5 neuroblastoma cells were placed on another coil and incubate under controlled condition. The oscillator was then turned on for 12 hrs a day for 5 days, after which cells were counted and morphology studied by contrast microscopy. In a control experiment the effect of the differentiating agent directly added to the cell culture could be observed by a decrease in cell growth, metabolic activity and the protrusion of neurite like structure typical of the differentiated cells. These preliminary results suggest that retinoic acid molecules emit signals that can be transferred to LAN-5 neuroblastoma cells by artificial physical means in a manner that seems related to the chemical structure of the source molecules.

These results suggest that RA molecules emit signals that can be transferred to LAN-5 cells by artificial physical means in a manner that seems specific to the source molecules.

## 1. INTRODUCTION

Previously Benveniste et al. suggests that the electromagnetic molecular signal (EMS) from pharmacological active molecules can be digitally recorded and replayed [1]. In the same work, they show that normal human neutrophils reacted to PMA transmitted via an electronic oscillator by reducing cytochrome c as though they had been directly exposed to PMA. The Benveniste finding of the electromagnetic molecular signal, was recently confirmed by the nobel laureate Luc Montagnier [2] in a work were he discovered a novel property of DNA, that is the capacity of some sequences to emit electromagnetic waves in resonance after excitation by the ambient electromagnetic background.

In the present study we electronically captured, and transmitted the specific electromagnetic signal (EMS) of Retinoic acid (a potent chemical molecule acting on human cells as a differentiating agent) to a biological system constituted by human neuroblastama cells line (LAN-5).

LAN-5 neuroblastoma cells represent one of the most common paediatric solid tumors originating from the sympathoadrenal lineage of neural crest. This tumor shows extremely different clinical phenotypes such as spontaneous regression on one hand and aggressive growth on the other hand. Undifferentiated neuroblastoma cell line (Lan-5) represent a good model to study neuronal differentiation induced by a variety of stimuli such as retinoic acid treatment.

## 2. MATERIAL AND METHODS

### 2.1. Cell Cultures

LAN-5 cells were grown in RPMI (Gibco Laboratories, Scotland) supplemented with 10% Fetal Calf Serum (Gibco Laboratories, Scotland) and antibiotics (110 IU/ml of penicillin and 0.1 mg/ml of streptomycin) at $37 \pm 0.3^\circ$C, and 5% $CO_2$ as carbon source and sub-cultured twice a week at a 1 : 5 ratio. For every experiment, control and exposed cells were taken from the same flask.

### 2.2. Transmission Apparatus

For transmission experiments to cells, the input coil coupled to wave generator VEGA select 719 was operated at room temperature, while the output coil was placed in cell incubator. The source tube containing 5 µM RA and target coil containing LAN-5 cells. The electronic signal corresponding to RA was superimposed to both a 7 Hz sinusoidal frequency carrier modulated at 3 kHz.

The oscillator was then turned on for the 12 hrs a day for 5 days. During the experimental procedure, the various parameters such as power, voltage, capacitance and impedance remained constant.

### 2.3. Cellular Metabolic Activity and Proliferation by WST Assay

LAN-5 cells were exposed to the electronically transmitted RA EMS by Vega select 719. For each experiment LAN-5 cells were plated into 25 ml 4.2 × 5.2 cm base Corning flasks ($2.0 \times 10^5$/ml cells in a total volume of 5 ml). The flasks were kept in the exposure system continuously for up to 5 days with or without RA-EMS. Cells were then counted and metabolism determined by WST-1 method. The experiment was repeated three times.

The quantification of LAN-5 metabolic activity, as an index of cellular proliferation, was performed by a colorimetric assay based on oxidation of tetrazolium salts (Cell Proliferation Reagent water-soluble tetrazolium salt (WST)-1; Roche Diagnostics, Basel, Switzerland). Cells were cultured for up to 5 days in a normal humidified incubator (control) or in the presence of the RA-EMS (exposed), and they were analysed by means of the formazan dye every 24 h. WST reagent diluted to 1 : 10 was added in the wells at 4 h, 1, 2, 3 and 6 days after plating, and then incubated for 2 h in humidified atmosphere (37 8°C, 5% $CO_2$). Quantification of the formazan dye produced was performed by absorbance measurement at 450 nm with a scanning multiwell spectrophotometer (Biotrack II; Amersham Biosciences, Little Chalfont, UK).

### 2.4. Immunofluorescence

For immunofluorescence staining the cells were grown in Labtek chamber slides. The cells were then washed with PBS with Ca/Mg and fixed with absolute ethanol for 5 minutes, then incubated with the specific monoclonal antibodies, anti-200 KDa neurofilaments (Sigma) appropiately diluted for 1 hour at room temperature. Cells were then washed three times with PBS and incubated with fluoresceinated anti-mouse IgGF(ab')2 fragment (Sigma), appropriately diluted for 1 hour at room temperature.

### 2.5. RT-PCR Analysis

Total RNA was extracted from cells using TRIzol Reagent (Life Technologies, Merelbeke, Belgium) according to the manufacturers' instructions. Typically 5–10 μg total RNA per 10 cm$^2$ dish of cell culture was obtained. Reverse transcription-polymerase chain reaction (RT-PCR) was used to evaluate relative mRNA levels of neurofilament protein (NF-200) in control and RA-EMRexposed LAN-5 cells. One microgram of total RNA was used to synthesize first-strand cDNA with random primers using 100 U of ImProm-II$^{TM}$ RT-PCR kit (Promega, Madison, WI, USA) according to the manufacturer. The reaction was also carried out in the absence of reverse transcriptase (RT) to check for genomic DNA amplification. The NF-200 subunit-specific primers used for PCR were: 5'-aagtgaacacagatgctatgcg-3' 5'-ctgtcactccttccgtcacc-3' We used the 18 s as internal controls, because these genes are uniformly expressed during development. The subunit-specific primers used for PCR were: 5'-tttcggaactgaggccatgattaag-3' 5'-agtttcagctttgcaaccatactcc-3'. An aliquot (2 μL) of RT reaction was PCR-amplified in a final volume of 50 μL, by using 20 pmol of each primer, 200 μM of each dNTP, and 0.5 U of Taq DNA Polymerase (T. Aquaticus, Amersham-Pharmacia). PCR was carried out in a Bio-Rad I Cycle instrument. The thermo cycling conditions for each pair of primers were as follows: denaturation at 95°C for 3 min. followed by 30 cycles of denaturation at 95°C for 30 sec, annealing for 45 sec at 62°C, elongation at 72°C for 1 min and a final polymerization step at 72°C for 5 min for NF-200 and 20 cycles for 18 s. The amount of template and the number of amplification cycles were preliminarily optimized for each PCR reaction to avoid conditions of saturation. Aliquots (5 μL) of the reaction products were run on 1% agarose gels containing ethidium bromide (0.5 μ g/ml) to mark and visualise the PCR products. Gels were then photographed under UV light with Versadoc (Bio-rad) instrument. These experiments were replicated three different times.

### 2.6. Statistical Analysis

Statistics was performed with Student's t-test with $P < 0.05$ as the minimum level of significance.

## 3. RESULTS

### 3.1. Electronically Transmitted RA Effect on LAN-5 Cell Metabolism

The cell growth rate was analyzed by the WST-1 both in LAN-5 cells as control (not exposed electronically transmitted RA) or exposed to the field. An inhibition in the cell metabolism in the electronically transmitted RA exposed was statistically ($p < 0.01$) significant after 5 days exposure (Fig. 1).

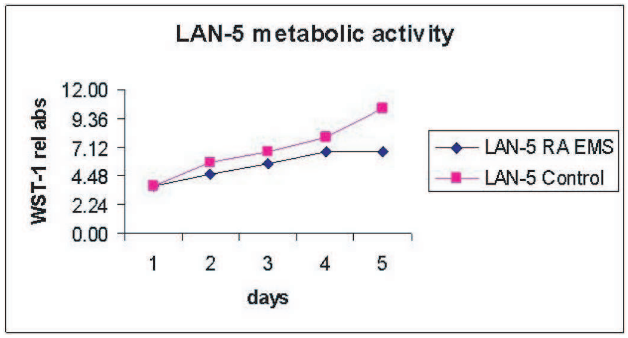



Figure 1: Cellular metabolic activity and proliferation by WST assay. LAN-5 metabolic activity by WST-1 analysis in presence (▶) or absence (■) on RA-EMS.

### 3.2. Electronically Transmitted RA Effect on LAN-5 Cell Morphology

By phase contrast and scanning electron microscopy LAN-5 control cells appeared small, poligonal, without neurite-like structures. The exposure to electronically transmitted RA induced morphological changes toward a more neuronal phenotype: the cells were stretched out and rich of neurite-like structures with blebs, mimicking the same effect induced by retinoic acid treatment (Fig. 2).

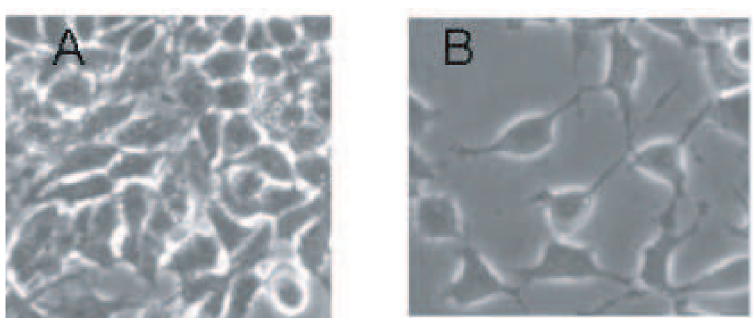



Figure 2: Electronically transmitted RA effect on LAN-5 cell morphology by contrast microscopy. Contrast microscopy of LAN-5 cells in absence (A) or presence of electronically transmitted RA on LAN-5 (B).

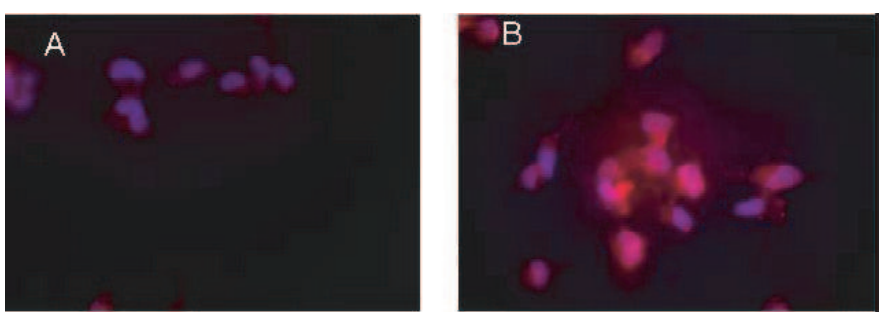



Figure 3: Electronically transmitted RA effect on LAN-5 cell by NF-200 indirect immunofluorescence. NF-200 indirect immunofluorescence of LAN-5 cells in absence (A) or presence of electronically transmitted RA on LAN-5 (B).

### 3.3. EMF Effect on Neurofilament Expression

Figure 3 shows the indirect immunofluorescent analysis of control and exposed Lan-5 cells with anti 200 neurofilaments. While control cell were little or not positive for NF 200 (CTR) the neurofilament protein become more fluorescent after exposure to the RA EMF (EXP). The same results were achieved by RT-PCR analysis for mRNA expression coding for NF-200. Fig. 4.

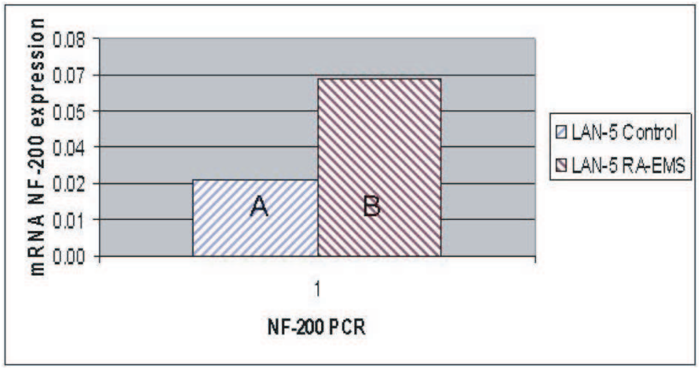



Figure 4: NF-200 mRNA RT PCR analysis on electronically transmitted RA effect on LAN-5 cell. Left bar (A) NF-200 mRNA RT PCR analysis on control LAN-5 cells, right bar (B) NF-200 mRNA RT PCR analysis on electronically transmitted RA effect on LAN-5 cell.

## 4. DISCUSSION

Low frequency electromagnetic fields at 50 or 60 Hz indeed are reported to stimulate nerve regeneration [3], alter gene transcription [4] and they may also play a synergistic role in cellular processes that are already activated, such as cell proliferation [5]. Despite an increasing number of publications demonstrate an effect of very low frequencies EM field on biological systems, other in vivo and in vitro studies suggest opposite results; in addition the possible interaction mechanism is not yet completely understood.

A possible mechanisms evoked to explain the mechanism of EM field action to biological system is involving $Ca^{2+}$ transport across cell membrane, to trigger the signal transduction cascade [6].

Electromagnetic therapeutic potential can be seen in the proven efficacy of low-energy pulsed magnetic fields in non-union bone fracture healing, confirming that under certain conditions non-ionising electro-magnetic energy can influence physiological processes in organisms. Physiological paradigms for non-ionising radiation effects are required. Clues may be found in the mechanisms by which EM field interacts with cultured cells under controlled laboratory conditions and by correlating in vivo evidence with in vitro data [7]. Brain maturation depends on a sequence of postnatal events [8]. Brushart et al. [9] found that electrical stimulation at 20 Hz, promote motoneuron regeneration, confirming previous finding of the use of electric field for the orientation and growth of neurite [10].

12 hrs a day for five days exposure to RA-EMS field has significant effects on cells proliferation leading to a 30% inhibition of cell metabolism (Fig. 1). In addition with the impairment in cell metabolism RA-EMS electromagnetic field exposure, generated a morphological change as reported by the contrast microscopy study showed in Fig. 2.

In particular contrast microscopy analysis indirect immunofluorescence (Fig. 3) and RT-NF-200 PCR (Fig. 4) showed a more neuronal morphology characterized by the development of neuritic like processes in the exposed cells compared to control. Immunofluorescence using monoclonal antibodies for the major neurofilament proteins NF 200 unequivocally demonstrated an increase in synthesis and accumulation of these neuronal proteins in the RA-EMS exposed cells. Taken together all these data support an evident effect of the electronically transmitted retinoic acid (RA-EMS) electromagnetic field of driving neuroblastoma cells toward a neuronal differentiation, which resembles the effect determined by morphogens, such as retinoic acid in its chemical form.

## ACKNOWLEDGMENT

This work has been partially supported by a grant from NAMED.

## REFERENCES

1. Thomas, Y., M. Schiff, L. Belkadi, P. Jurgens, L. Kahhak, and J. Benveniste, *Med. Hypothesys*, Vol. 54, No. 1. 33–39, 2000.
2. Montagnier, L., J. A. Aissa, C. Lavallee, M. M. Bamy, J. Varon, and H. Chenal, *Interdiscip Sci. Comput. Life Sci.*, Vol. 1, 245–253, 2009.
3. Rusovan, A. and M. Kanje, *Neuroreport*, Vol. 3, No. 12, 1039–1041, 1992.

4. Phillips, J. L., W. Haggren, W. J. Thomas, T. I. Jones, and W. Adey, *Biochimica Biophysica Acta*, Vol. 1132, 140–144, 1992.
5. Manni, V., A. Lisi, D. Pozzi, S. Rieti, A. L. Serafino, M. Ledda, L. Giuliani, and S. Grimaldi, *Bioelectromagnetics*, Vol. 23, 298–305, 2002.
6. Karabakhtsian, R., N. Bronde, N. Shalts, S. Kochlatyi, R. Goodman, and A. S. Henderson, *FEBS Letters*, Vol. 301, 53–59, 1994.
7. Glaser, R., *Bioelectrochem. Bioener*, Vol. 27, 255–268, 1992.
8. Vignes, M. and G. L. Collingridge, *Nature*, Vol. 388, 179–182, 1997.
9. Brushart, M. T., P. N. Hoffman, R. M. Royall, B. B. Murinson, C. Witzel, and T. Gordon, *J. Neurosci*, Vol. 22, 6631–6638, 2002.
10. Patel, N. and M. M. Poo, *J. Neurosc*, Vol. 2, 483–496, 1982.

# Effects of Extremely Low Frequency Electromagnetic Fields on the Antioxidant Enzymes Activity of C3 and C4 Plants

**Azita Shabrangi**[2], **Ahmad Majd**[1], **Masoud Sheidai**[3], **Mohammad Nabyouni**[1], **and Davoud Dorranian**[4]

[1]Department of Biological Science, Faculty of Science, Tarbiat Moalem University, Tehran, Iran
[2]Department of Biology, Tehran North Branch, Islamic Azad University, Tehran, Iran
[3]Biology Faculty, Shahid Beheshti University, Tehran, Iran
[4]Physic Plasma Research Centre, Science and Research Branch, Islamic Azad university, Tehran, Iran

**Abstract**— This experimental study was dedicated to the possible oxidative stress induced by electromagnetic field accumulation in the living tissues. In this research two states of seeds (wet, dry) of *Brassica Napus L* (canola) and *Zea mays L* (maize) were exposed to pulsed EMFs (15 min on, 15 min off) by magnitude of 1 to 7 mT in steps of 2 mT and the highest intensity was 10 mT for 1 to 4 hours in steps of 1 h. Activity of stress enzymes were carried out by spectrophotometer. Our investigation was focused on root and shoot of 7 days seedlings, which showed the most and the least growth comparing to corresponding control. Activity of all stress enzymes in root tissue of canola increased and in shoot tissue of this species decreased. In contrary, activity of the same stress enzymes in the same tissue of maize showed opposite results, except APX activity in the root tissue, which showed similar result in both species ($P < 0.05$).

All results suggested that antioxidant system of canola as a C3 plant and maize as a C4 plant react differently against EMFs.

## 1. INTRODUCTION

Several reports relate to the effects of strong magnetic fields on a variety of agriculturally important plants [5]. For example exposure of maize for 2–10 min to MF ranging from 0.06 to 0.2 T stimulate germination and increased the harvest by 29.5% [1]. Roots of *Zea mays* and hairy roots of *Dacus carota* grow faster at 0.5 T than they do in the geomagnetic field [10]. Numerous biological effects of extremely low frequency electromagnetic fields were recorded in the last decades [2]. They were observed at various cellular or molecular levels in living tissues. However no clear interaction mechanisms yet proved [4] Electric and magnetic treatments are assumed to enhance seed vigor by inducing the biochemical processes that involve free radicals and by stimulating the activity of proteins and enzymes [3, 11, 12]. Also, it is assumed that Electromagnetic fields (EMF), in ELF (extremely low frequency) activate the cellular stress response, a protective mechanism that induces the expression of stress response genes. Plant cells possess versatile antioxidant systems to protect organelles, membranes, and enzymes from the damaging action of toxic activated oxygen species [6, 15]. In chloroplasts, $H_2O_2$ is produced in the Mehler reaction and also during superoxide dismutase-catalyzed disproportionation of the superoxide radical anion. The first step of H20, degradation is performed by Catalysis. Then the rest of it, degraded by ascorbate peroxidase which reduces $H_2O_2$ to water [8].

In a previous study we investigated the effect of MFs exposure at different intensities (0.06 to 0.36 T) in *Lens culinaris L*. We also reported that MF increased seed germination, seedlings growth and biomass in lentil. we observed that stress enzymes especially APX increased by the increase of MFs intensity [14].

In the present study, the activities of antioxidative enzymes were investigated in order to find out if exposure to ELF-EMFs induces unspecific stress responses similar to those of other environmental stresses. To clarify the possible role of ROS causing EMFs effects as parameter of oxidative stress in *Zea mays L* and *Brassica napus L* were analyzed.

## 2. MATERIAL AND METHOD

Exposure to EMFs was performed by a locally designed EMF generator. The electrical power was provided by a 220 V, AC power supply with variable voltages and currents. This system consisted of one coil, cylindrical in form, made of polyethylene with 12 cm in diameter and 50 cm in length. The number of turns is 1000 of 0.5 mm copper wire, which were in two layers. A fan was employed to avoid the increase of temperature ($22 + 1°C$). Calibration of the system as well as tests for the accuracy and uniformity of EMFs (60 Hz) were performed by a tesla meter with a probe type of hall sound.

Three replicates, with 30 seeds in each one were used. They were spread in moist filter paper (for wet seeds) on Petri dishes. They were placed in the coil. The wet and dry seeds of two species were exposed to pulsed EMFs (15 min on, 15 min off) by magnitude of 1 to 7 mT in steps of 2 mT and the highest intensity was 10 mT for 1 to 4 hours in steps of 1 h. Then dry seeds in Petri dishes were moistened and all Petri dishes were placed in germinator with 23 °C temperature. Our investigation were focused on seedlings, which showed the most and the least growth parameters. These results were observed in seedlings grown from dry pretreated seeds by 10 mT for 4 h and wet pretreated seeds by 10 for 2 h in canola, and seedlings grown from wet pretreated seeds by 10 and 3 mT both for 4 h exposure in maize. Triplicate of ten 7 days of these seedlings were chosen randomly. Then root and shoot were separated and were used for protein extraction in order to assay activity of stress enzymes.

### 2.1. Activity Enzyme Assay

The assays of SOD and APX were performed on 7 days seedlings. Ten 7 days seedlings from each replicate were randomly taken exhibiting no visible injury symptoms. 0.5 gr fresh weight of root and shoot were homogenized separately in 5 ml of 100 mM potassium phosphate buffer (pH 7) containing 0.1 mM $Na_2EDTA$ and 1% (w/v) soluble PVP-10.5 mM. 0.2 mM ASA were added for measurement of APX activity. The slurry was centrifuged at $15000 \times g$ for 20 min. The supernatant was filtered (Millipore, Mitex 0.5 µm) and used for enzyme analysis. CAT activity was determined by monitoring the decrease in A240 for 1 min in 3 ml of reaction mixture containing 100 mM potassium phosphate buffer (pH 7.5) and 25 mM $H_2O_2$ and shoot or root extract. APX activity was determined by monitoring the decrease in A290 for 4 min in 3 ml of reaction mixture containing 100 mM potassium phosphate buffer (pH 7), 0.5 mM ASA, 0.1 mM $Na_2EDTA$, 0.1 mM $H_2O_2$ and shoot or root extract. Correction were made for the oxidation of AsA in the absence of $H_2O_2$. The SOD activity was measured by spectrophotometer as described by Beyer and Fridovich (1989). The reaction mixture contained 50 mM potassium phosphate buffer (pH 7.8), 9.9 mM methionine, 57 µm nitro blue tetrazolium (NBT) and 0.9 µm riboflavin and 0.025% (w/v) triton X-100 and appropriate amount of shoot or root extract. The A560 was recorded after a 10 min illumination period. In this assay 1 unite of SOD is defined as the amount required to inhibit the photo reduction of nitro blue tetrazolium by 50%. The specific activity of SOD was expressed as units $mg^{-1}$ fresh weight of plant (Calatayud, A. et al., 2003)

### 2.2. Statistical Analyses

Statistical analyses were conducted using SPSS for windows. After testing the normality of the data distribution, the variance analyses (ANOVA) was used to test the main effects of electromagnetic field pretreatment. A student test t-test was done to find the significant differences between each treatment and corresponding control. Means were compared to detect differences between parameters of pretreated seeds and plants grown from them with control ($P < 0.05$).

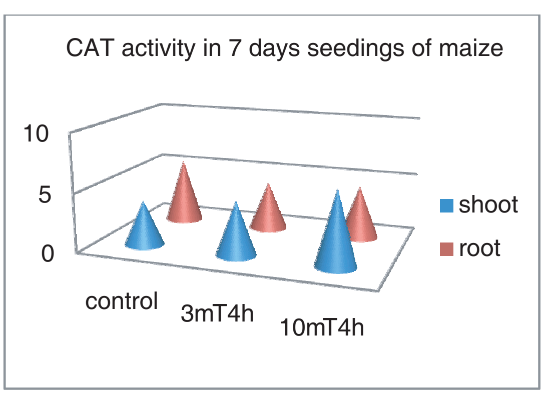



Figure 1.

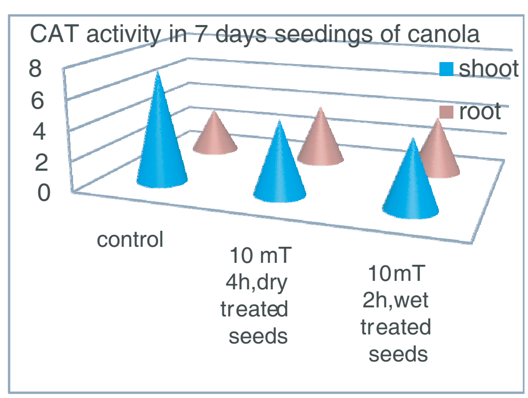



Figure 2.

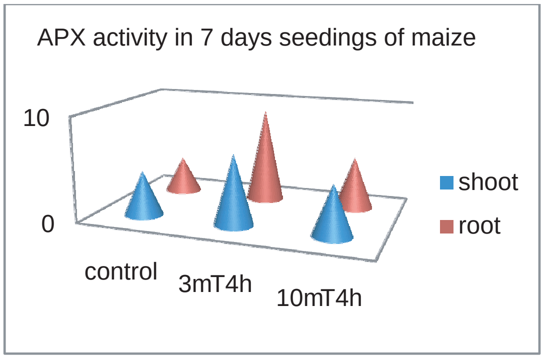



Figure 3.

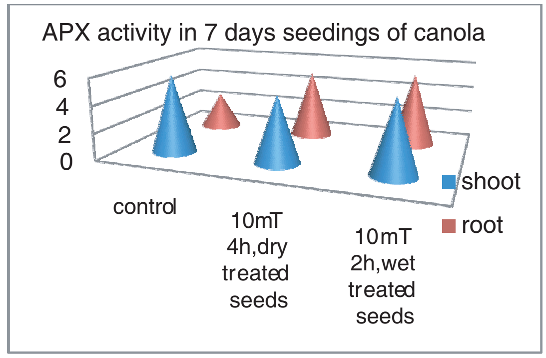



Figure 4.

Table 1. Data are the means SE of experiment performed in triplicate and expressed as percentage of control in maize. Asterisk (*) denotes significant differences between plants exposed to EMFs and their respective control using the students t-test ($P < 0.05$).

| Exposure case | CAT activity in shoot | CAT activity in root (μg) | APX activity in shoot | APX activity in root | SOD activity in shoot | SOD activity in root |
|---|---|---|---|---|---|---|
| Wet treated seeds (3 mT, 4 h) | $6.27 \pm 0.32$* | $4.4 \pm 0.36$* | $4.8 \pm 0.31$* | $5 \pm 0.12$* | $17.77 \pm 3.5$* | $14.9 \pm 3.24$* |
| Wet treated seeds (10 mT, 4 h) | $4.5 \pm 1.5$ | $4.1 \pm 0.94$* | $6.67 \pm 0.35$* | $9.1 \pm 0.35$* | $48.5 \pm 8.3$* | $13.13 \pm 5.48$* |
| Control | $3.77 \pm 0.53$* | $5.5 \pm 0.35$* | $4.27 \pm 0.18$* | $3.47 \pm 0.52$* | $20.3 \pm 7.6$ | $2.13 \pm 5.9$* |

*Significant from control at 0.05 level (t-test)

## 3. RESULTS

Based on statistical analysis, the most and the least growth were observed in seedlings grown from dry pretreated seeds by 10 mT for 4 h and wet pretreated seeds by 10 for 2 h in canola, and seedlings grown from wet pretreated seeds by 3 and 10 mT both for 4 h exposure in maize. Activity enzymes were expressed based on absorbance against mg fresh weight of root and shoot tissues separately. Both treatments caused significantly decreased in Catalase (CAT) activity in root tissue of maize. But we observed significant increase activity of CAT in shoot (Fig. 1). Ascorbate Peroxidaes (APX) activity increased significantly in both treatment and both root and shoot tissues (Fig. 3). We observed increase of Superoxide dismutase (SOD) activity in root tissue of maize in both treatment, but 3 mT intensity did not show significant difference. In shoot tissue of maize, we observed significantly increase of SOD activity by 3 mT treatment and significantly decrease by 10 mT treatment (Fig. 5 & Table 1). On the other hand, We observed significantly increase in CAT activity in root tissue of canola and significantly decrease of this enzyme activity in shoot tissue of canola (Fig. 2). APX activity results were similar to CAT activity in both tissues of this species (Fig. 4). Activity of SOD showed increase in both treatment of root tissue of canola. In contrast, SOD activity decreased in shoot tissue of canola in both treatments (Fig. 6). But treatment by 10 mT intensity did not cause significant difference of this enzyme activity in both tissues of canola (Table 2) ($P < 0.05$).These figures from number 1 to 6 shows alteration in different stress enzymes activity of 7 days seedlings grown from pretreated seeds by EMFs comparing to their corresponding control in two species.

## 4. DISCUSSION

Different growth parameters of two species (C3 and C4 plants) under effects of EMFs have changed, which was reported in the previous study [13]. Another parameter is activity enzyme changes. Activity of all stress enzymes in root tissue of canola increased and in shoot tissue of this species decreased. In contrary, activity of the same stress enzymes in the same tissue of maize showed opposite results, except APX activity in the root tissue, which showed similar result in both species ($P < 0.05$).

Table 2. Data are the means SE of experiment performed in triplicate and expressed as percentage of control in canola. Asterisk (*) denotes significant differences between plants exposed to EMFs and their respective control using the students t-test ($P < 0.05$).

| Exposure case | CAT activity in shoot | CAT activity in root (μg) | APX activity in shoot | APX activity in root | SOD activity in shoot | SOD activity in root |
|---|---|---|---|---|---|---|
| Wet treated seeds (3 mT, 4 h) | $4.8 \pm 0.8$* | $3.7 \pm 0.83$* | $4.9 \pm 0.18$* | $5.1 \pm 0.18$* | $23.8 \pm 13.8$ | $2.1 \pm 3.24$ |
| Wet treated seeds (1 0mT, 4 h) | $4.5 \pm 0.7$* | $3.9 \pm 0.76$* | $5.4 \pm 0.35$* | $5.4 \pm 0.51$* | $19.2 \pm 4.4$ | $13.13 \pm 5.48$ |
| Control | $7.5 \pm 1.3$* | $3 \pm 0.15$* | $5.8 \pm 0.31$* | $2.7 \pm 0.58$* | $42.7 \pm 25.4$* | $14.9 \pm 3.2$* |

*Significant from control at 0.05 level (t-test)

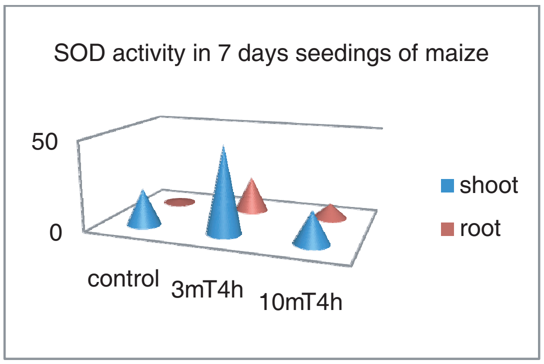



Figure 5.

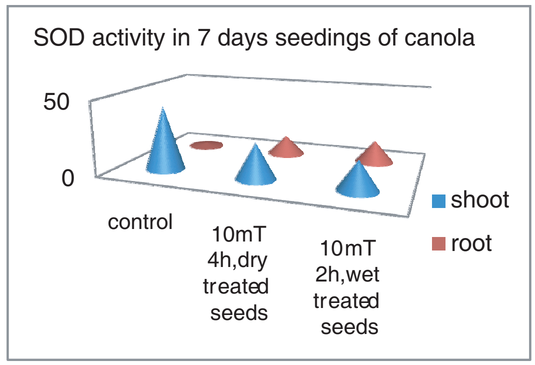



Figure 6.

Since plants are sessile organisms they display a wide spectrum of developmental and biochemical responses contributing to stress adaptation. A common responses is increasing in reactive oxygen species (ROS) production [2]. ROS can damage plants by oxidizing proteins, nucleic acids, photosynthetic pigments and membrane lipids. However, ROS can also act as signaling molecules and trigger a range of cellular responses important for stress tolerance [4]. The tight control of ROS level in cells is achieved by antioxidative system defense including different stress enzymes [17].

These stress enzymes scavenge free radicals. They have antioxidant mechanism. APX Scavenges $OH^{\circ-}$ and SOD also scavenges $O_2^{\circ-}$. In fact stress enzymes decrease oxidative stress. Some studies have suggested that magnetic field exposure could be due to both the increase in the concentration [9] and oscillating of free radicals [16]. Electromagnetic fields are known to effect radical pair recombination and they may increase the concentration of oxygen free radicals in living cells. Increasing the concentration of free radicals creates oxidative stress, enhances stress response and some biological reactions, such as DNA damage occurs under this concentration.

Among ROS scavenger enzymes, CAT is the key enzyme that effectively eliminates $H_2O_2$, as regulates the activity of APX [7]. It was also reported that more reduction in CAT activity (the more accumulation of $H_2O_2$) was accompanied by the less decrease in APX activity.

Catalysis, which is main $H_2O_2$ scavenging enzymes in plants, effectively eliminates $H_2O_2$, thereby regulates the activity of APX. It also converts $H_2O_2$ to water using different cellular substrates. Ascorbate Peroxidase is the key enzyme of the ascorbate cycle, eliminating peroxidase by converting ascorbic acid to dehydroascorbate [12].

In conclusion, antioxidant system of canola as a C3 plant and maize as a C4 plant react differently against EMFs.

## REFERENCES

1. Antonow, G., N. Armjanov, and T. Todorov, "Untersuchungen zum einfluß des magnetfeldes auf die keimenergie von samen und den ertrag (bulg.)," *Selskostopanska Techn.*, Vol. 19, 5–11, Sofija, 1982.
2. Arora, A., R. K. Sairam, and G. C. Srivastava, "Oxidative stress and antioxidative system in plants," *Current Science*, Vol. 82, No. 10, 2002.

3. Arrigoni, O. and M. de Tulho, *J. Bioenerg. Biomembr.*, 1994.
4. Dat, J., et al., "Dual action of the active oxygen species during plant stress responses," *Cellular and Molecular Life Sciences*, Vol. 57, No. 5, 779–795, 2000.
5. Drobig, J., "Saatgut im elektrmagnetischen feld — Zu einigen internationalen untersuchungen," *Arch Acker-Pflanzenbau Bodenkd*, Vol. 9, 619–626, 1988.
6. Foyer, C. H., M. Lelandais, E. A. Edwards, and P. M. Mullineaux, "Cum tonics," *Plant Physiol*, Vol. 6, 131–144, 1991.
7. Ghanati, F., A. Morita, and H. Yokota, "Effects of aluminum on the growth of tea plants and activation of antioxidant system," *Plant Soil*, Vol. 276, 133–141, 2005.
8. Hossain, M. A. and K. Asada, *Plant Cell Physiol.*, Vol. 25, 85592, 1984.
9. Jajte, J. M., "Programmed cell death as a biological function of electromagnetic fields at a frequency of (50/60 Hz)," *Medycyna Pracy*, Vol. 51, 383–389, 2000.
10. Kato, R., "Effects of a magnetic field on the growth of primary roots of zea mays," *Plant Cell Physiol.*, Vol. 29, 1215–1219, 1988.
11. Nakano, Y. and K. Asada, "Hydrogen peroxide is scavenged by ascorbate — Specific peroxidase in spinach chloroplasts," *Plant Cell Physiol.*, Vol. 22, 867–880, 1981.
12. O'Farrell, P. H., *J. Biol. Chem.*, Vol. 250, 40074021, 1975.
13. Shabrangi, A. and A. Majd, "Comparing effects of electromagnetic fields (60 Hz) on seed germination and seedling development in monocotyledons and dicotyledons," *PIERS Proceedings*, 704–709, Moscow, Russia, Aug. 18–21, 2009.
14. Shabrangi. A., et al., "Effect of magnetic fields on growth and antioxidant systems in agricultural plants," *PIERS Proceedings*, Beijing, China, Mar. 23–27, 2009.
15. Salin, M. L., *Physiol Plant*, Vol. 72, 681689, 1987.
16. Scaiano, J. C., F. L. Cozens, and N. Mohtat, "Development of a model and application of the radical pair mechanism to radicals in micelles," *Photochem. Photobiol.*, Vol. 62, 818–829, 1995.
17. Tkalec, M., K. Malaric, and B. Pevalek-Kozlina, "Exposure to radiofrequency radiation induces oxidative stress in duckweed Lemna minor L," *Science of the Total Environment*, Vol. 388, No. 1–3, 78–89, 2007.

# Zero Reflection from a PEC Plate Coated by Double Zero (DZR) Metamaterials

**Homayoon Oraizi**[1], **Ali Abdolali**[1], **and Noushin Vaseghi**[2]

[1]Department of Electrical Engineering, Iran University of Science and Technology
Tehran 1684613114, Iran
[2]Department of Electrical and Computer Engineering, K. N. Toosi University of Technology
Tehran 16315-1355, Iran

**Abstract**— We consider a perfect electric conductor (PEC) plate covered by a layer of DZR metamaterial coatings under an oblique plane wave incidence of TM polarization. Several analytical formulas are derived for the realization zero reflection from such structures. Exact formulas will be derived for zero reflection from a DZR coated PEC plate. Then several examples of the applications of DZR metamaterials as zero reflection coatings are provided.

## 1. INTRODUCTION

Common materials called double positive (DPS) have both the real parts of their permittivities ($\varepsilon$) and permeabilities ($\mu$) positive. On the other hand, metamaterials are classified as double negative (DNG) where both the real parts of their ($\varepsilon$) and ($\mu$) are negative and epsilon negative (ENG) and mu negative (MNG), where only their $\varepsilon$ and only their $\mu$ are negative, respectively [1–6].

We intend to introduce a class of materials called double zero (DZR) metamaterials by parameters $\mathrm{Re}(\varepsilon) = 0$ & $\mathrm{Re}(\mu) = 0$. In an earlier paper, we investigated the propagation of radio waves in some DZR metamaterial structures [7]. However, in this paper we intend to investigate the properties of wave propagation incident onto a perfect electric conductor coated by a layer of DZR metamaterials. Also we develop an exact mathematical formulation for zero reflection from such structures. Some uncommon phenomena will appear for DZR metamaterials, which have not been observed for common materials and metamaterials. Exact formulas [8, 9] will be derived for zero reflection from a DZR coated PEC plate for TM polarizations at a particular angle of incidence. The DZR metamaterials are assumed dispersionless media. We use the full-wave matrix method for the analysis of this structure [10–13].

## 2. PROBLEM DESCRIPTION AND EXACT FORMULAS

Consider a perfect electric conductor (PEC) plate coated by a layer of double zero (DZR) metamaterials of thickness d as drawn in Fig. 1. A plane wave of TM polarization is obliquely incident at an angle of incidence $\theta_0$ onto the structure. The structure and fields are independent of the variable "y" of Cartesian coordinate axes shown in Fig. l.

The forward and backward transverse travelling waves in layer $l$ are [14, 15]:

$$\begin{cases} H_{ly} = A_l e^{-jk_l \cos\theta_l z} + B_l e^{+jk_l \cos\theta_l z} \\ E_{lx} = \eta_l \cos\theta_l \left( A_l e^{-jk_l \cos\theta_l z} - B_l e^{+jk_l \cos\theta_l z} \right) \end{cases} \tag{1}$$

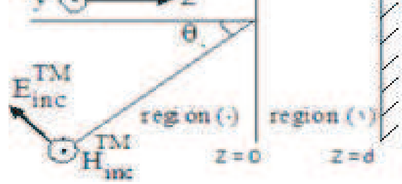



Figure 1: A perfect electric conducting (PEC) plate covered by DZR metamaterial layers.

where in the left half space $l = 0$ and the permittivity and permeability are

$$\begin{cases} \varepsilon_l = \varepsilon_l' - j\varepsilon_l'' \\ \mu_l = \mu_l' - j\mu_l'' \end{cases} \Rightarrow \begin{cases} k_l = \omega\sqrt{\mu_0\varepsilon_0\mu_l\varepsilon_l} = \pm k' \pm jk'' \\ \eta_l = \sqrt{\dfrac{\mu_0\mu_l}{\varepsilon_0\varepsilon_l}} = \pm\eta' \pm j\eta'' \end{cases} \tag{2}$$

The incident wave amplitude is assumed equal to unity ($A_0 = 1$), the reflection coefficient is $R$. The permittivity, permeability, wave number and intrinsic impedance of the layer $l$ are $\varepsilon_l$, $\mu_l$, $k_l$ & $\eta_l$, respectively. In DZR metamaterials, we have [7]

$$\begin{cases} \varepsilon = 0 - j\varepsilon'' \\ \mu = 0 - j\mu'' \end{cases} \Rightarrow \begin{cases} k = \omega\sqrt{-\mu''\varepsilon''} = 0 - jk'' \quad (k' = 0) \\ \eta = \sqrt{\dfrac{\mu''}{\varepsilon''}} = +\eta' + j0 \quad (\eta'' = 0) \end{cases} \tag{3}$$

where the parameters $\varepsilon''$, $\mu''$, $k''$ & $\eta'$ are assumed positive. The boundary conditions lead to the following matrix equation for wave amplitudes

$$\begin{bmatrix} +1 & -1 & -1 \\ \dfrac{\eta_0\cos\theta_0}{\eta_1\cos\theta_1} & +1 & -1 \\ 0 & e^{-jk_1\cos\theta_1 d} & -e^{+jk_1\cos\theta_1 d} \end{bmatrix} \begin{bmatrix} R \\ A_1 \\ B_1 \end{bmatrix} = \begin{bmatrix} -1 \\ \dfrac{\eta_0\cos\theta_0}{\eta_1\cos\theta_1} \\ 0 \end{bmatrix} \tag{4}$$

The vanishing of the reflection coefficient $R = 0$ leads to the following equation

$$\eta_1\cos\theta_1\left(1 - e^{-j2k_1\cos(\theta_1)d}\right) = \eta_0\cos\theta_0\left(1 + e^{-j2k_1\cos(\theta_1)d}\right) \tag{5}$$

For TM polarization, Eq. (5) may be separated into two relations by equating the real and imaginary parts

$$\begin{cases} \eta' - \eta' r\cos\alpha - \eta'' r\sin\alpha \equiv s + sr\cos\alpha \\ \eta'' - \eta'' r\cos\alpha + \eta' r\sin\alpha \equiv -sr\sin\alpha \end{cases} \tag{6}$$

where

$$\begin{aligned} r &= e^{-2k''\cos(\theta_1)d} \\ \alpha &= 2k'\cos(\theta_1)d \\ s &= \frac{\eta_0\cos\theta_0}{\cos\theta_1} \end{aligned} \tag{7}$$

These equations may be solved for $\eta''$ & $\eta'$

$$\begin{cases} \eta' = \dfrac{s\left(1 - r^2\right)}{1 + r^2 - 2r\cos\alpha} \\ \eta'' = \dfrac{-2\,s\,r\sin\alpha}{1 + r^2 - 2r\cos\alpha} \end{cases} \tag{8}$$

Now, considering the conditions for DZR metamaterials, namely $k' = 0$ & $\eta'' = 0$ in Eq. (3), we have from Eq. (8)

$$\alpha = 0 \quad \& \quad r = \frac{\eta' - s}{\eta' + s} \tag{9}$$

Since $s \neq 0$ and $r \neq 0$ according to Eq. (7) and also $\alpha = 0$ leads to $k' = 0$. Consequently, the angle of zero reflection for a PEC plane coated by a DZR materials for the TM polarization may be obtained by Eq. (9) and Eq. (7)and the Snell's law

$$k_0\sin\theta_0 = -jk''\sin\theta_1 \tag{10}$$

leading to

$$\cos\theta_0 = \sqrt{1+\frac{k''^2}{k_0^2}-\frac{k''^2}{k_0^2}\cos^2\theta_1}$$
$$\cos\theta_1 = \sqrt{1+\frac{k_0^2}{k''^2}-\frac{k_0^2}{k''^2}\cos^2\theta_0} \tag{11}$$

## 3. NUMERICAL EXAMPLE

We consider several examples of zero reflection from a PEC plane coated by DZR metamaterial. The computations are based on Eqs. (9) and (4). For example consider a TM plane wave of frequency 12 GHz incident on a PEC plane coated with a layer of DZR metamaterial with thickness $d = 1\,\text{cm}$ and nondispersive characteristics $\varepsilon = -j0.1$ and $\mu = -j0.3$. The incident angle of no reflection is computed equal to $\theta_0 = 6.60°$, by Eqs. (9) and (11). The reflected power is computed by Eq. (4) for various values of incident angles and drawn in Fig. 2, which shows zero reflection at the same angle $\theta_0 = 6.60°$.

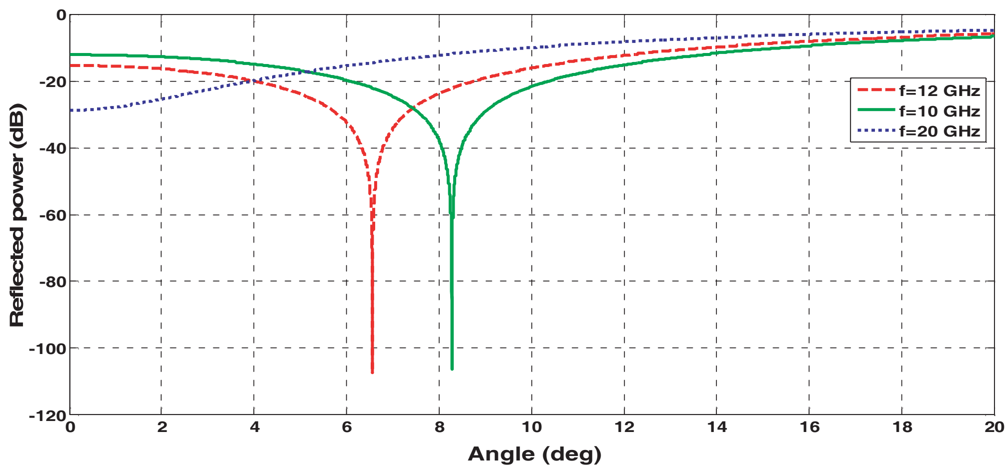



Figure 2: Reflected power (in dB) from the structure in Fig. 1 with characteristic given in example versus the angle of incidence at frequencies $f = 12$, 10, 20 GHz for TM polarization.

The same calculations are repeated for frequency $f = 10\,\text{GHz}$, which give the angles of zero reflection equal to $\theta_0 = 8.25°$. Note that Eqs. (9) and (11) may do not have response, in other word at some frequencies there is not zero reflection condition, for example in this example for frequency $f = 20\,\text{GHz}$ the zero reflection does not appear as shown in Fig. 2. Again they are verified by the computation of reflected power. However, the frequency response of reflection coefficient is narrow band.

## 4. CONCLUSION

In this paper we introduced double zero (DZR) metamaterials, which are essentially lossy media. A PEC plate coated by DZR metamaterials are studied analytically by Full wave methods for their zero reflection properties, under oblique incidence of plane waves with TM polarization. Exact formulas were derived for no reflection from a PEC plate coated by DZR metamaterials. It is shown that the application of common materials and metamaterials in such structures is not effective for the realization of zero reflection from them, but DZR metamaterials are required for the design of such no reflection surfaces. As a result from this paper we can investigate their applications for radar absorbing materials (RAMs) for the reduction of reflected power.

## ACKNOWLEDGMENT

This research was in part supported by the Iran Telecommunication Research Center, Tehran, Iran.

## REFERENCES

1. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of $\varepsilon$ and $\mu$," *Soviet Physics Uspekhi*, Vol. 10, No. 4, 509–514, Jan.–Feb. 1968.

2. Shelby, R. A., D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science*, Vol. 292, No. 5514, 77–79, 2001.
3. Marques, R., F. Martin, and M. Sorolla, *Metamaterials with Negative Parameters Theory, Design, and Microwave Applications*, 1st Edition, Wiley, 2008.
4. Smith, D. R., W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, "Composite medium with simultaneously negative permeability and permittivity," *Phys. Rev. Lett.*, Vol. 84, No. 18, 4184–4187, May 2000.
5. Caloz, C. and T. Itoh, *Electromagnetic Metamaterials, Transmission Line Theory and Microwave Applications*, Wiley, IEEE Press, Hoboken, NJ, 2005.
6. Lai, A., C. Caloz, and T. Itoh, "Composite right/left-handed transmission line metamaterials," *IEEE Microwave Magazine*, Vol. 5, No. 3, 34–50, Sep. 2004.
7. Oraizi, H. and A. Abdolali, "Some aspects of radio wave propagation in double zero metamaterials having the real parts of epsilon and mu equal to zero," *Journal of Electromagnetic Waves and Applications*, Vol. 23, Nos. 14–15, 1957–1968, 2009.
8. Oraizi, H. and A. Abdolali, "Analytical determination of zero reflection conditions for oblique incidence on multilayer planar structures," *Mediterranean Microwave Symposium*, 196–199, Damascus, Syria, Oct. 2008.
9. Oraizi, H. and A. Abdolali, "Mathematical formulation for zero reflection from multilayer metamaterial structures and their notable applications," *IET Microwaves, Antennas* & *Propagation Journal*, Vol. 3, No. 6, 987–996, Sep. 2009.
10. Cory, H. and C. Zach, "Wave propagation in metamaterial multilayered structures," *Microwave and Optical Technology Letters*, Vol. 40, No. 6, 460–465, Mar. 2004.
11. Oraizi, H. and A. Abdolali, "Combination of MLS, GA & CG for the reduction of Rcs of multilayered cylindrical structures composed of dispersive metamaterials," *Progress In Electromagnetics Research B*, Vol. 3, 227–253, 2008.
12. Ziolkowski, R. W. and E. Heyman, "Wave propagation in media having negative permittivity and permeability," *Phys. Rev. E, Tat. Phys. Plasmas Fluids Relat.*, Vol. 64, No. 5, 056–625, 2001.
13. Alù, A. and N. Engheta, "Pairing an epsilon-negative slab with a Mu-negative slab: Resonance, tunneling and transparency," *IEEE Trans. on Antennas and Propagation*, Vol. 51, No. 10, 2558–2571, Oct. 2003.
14. Kong, J. A., "Electromagnetic wave interaction with stratified negative isotropic media," *Progress In Electromagnetics Research*, PIER 35, 1–52, 2002.
15. Oraizi, H. and A. Abdolali, "Design and optimization of planar multilayer antireflection metamaterial coatings at Ku band under circularly polarized oblique plane wave incidence," *Progress In Electromagnetics Research C*, Vol. 3, 1–18, 2008.

# Analysis of Electromagnetic Guided Waves on Curved Conducting Biological Surface by Conformal Mapping Method

**Yasumitsu Miyazaki**
Department of Media Informatics, Aichi University of Technology
50-2 Manori, Nishihasama-cho, Gamagori 443-0047, Japan

**Abstract**— High frequency signal transmissions using human body sensors and transmission antennas have been studied, as biomedical communication systems of high frequency electromagnetic guided waves on biological surfaces of lossy conducting media. The first section of transmission is consisting of human body sensor and transmission antenna on the body surface, and receive terminal antenna, through transmission waveguide with body surface boundary of body media. The second section of transmission is consisting of transmission source output and usual information transmission line of pair cables.

In this paper, electromagnetic eigen characteristics of high frequency fields on conducting biomedical body surface are discussed. By using conformal mapping method and integral equation with Green's function, electromagnetic guided characteristics of electromagnetic fields on curved conducting biological body surfaces are investigated. Fundamental properties of medical application of high frequency signal transmission on human body surfaces are shown.

## 1. INTRODUCTION

Recently, signal and information processing technologies have been developed for bio-medical systems with electronic and optical diagnosis, based on high speed computers and data processing softwares. Particularly, sensing technologies and processing technologies have been investigated with high functional hardwares and softwares. Interface devices and transmission systems between human bodies and machines are required for high effective and functional bio-medical systems.

High frequency signal transmissions using human body sensors and transmission antennas have been studied, as biomedical communication systems of high frequency electromagnetic guided waves on biological surfaces of lossy conducting media for medical diagnosis and information communication [1, 2]. In these biomedical systems, the first section of transmission is consisting of human body medical sensor and transmission antenna on the body surface that is signal transmission source, and receive terminal antenna that is receiver, through transmission waveguide with body surface boundary of body media. The second section of transmission is consisting of transmission source output that is receiver of the first section, and usual information transmission line of pair cables or coaxial and optical cables and further, transmission antenna [3–5].

In this paper, electromagnetic eigen characteristics of high frequency fields in lossy waveguides consisting of lossy human body with conducting bio-medical body surface are discussed.

Electromagnetic characteristics such as permittivities and conductivities of human tissues of bloods, muscles and bones are discussed for several frequencies. Electromagnetic eigen fields in lossy waveguides consisting of human bodies, such as arms and legs, with bio-medical body surfaces are studied by approximate sub-stationary analysis and exact hybrid field analysis for high frequencies. Propagation constants with phase and attenuation characteristics of hybrid modes are derived by eigen-equations for lossy waveguides consisting of conducting human body with skin surfaces. Green's dyadic functions for lossy hybrid fields in the human body are studied by vector wave equations and boundary conditions of electromagnetic fields on skin surfaces [6–8].

By using conformal mapping method with Green's dyadic functions for straight lossy waveguides of human body with skin surfaces, electromagnetic fields in curved lossy waveguides of curved arms and legs corresponding to human elbows and knees are discussed. Curved lossy waveguide in physical space is transformed to straight lossy inhomogeneous waveguide in mapped space by conformal mapping with analytic function. By using conformal mapping method and integral equation with Green's function that can be applied to analysis of complicated electromagnetic field boundary problems developed by the author [2–7], electromagnetic guided characteristics of electromagnetic fields on curved conducting biological body surfaces are investigated. Fundamental properties of medical application of high frequency signal transmission on human body surfaces are shown.

## 2. ELECTRO-MAGNETIC CHARACTERISTICS OF BIO-MEDICAL MEDIA

Electromagnetic material properties of human bodies are given by bio-medical measurements for low and microwave frequency bands. Dielectric constants, permittivities and conductivities of bio-medical materials and organizations in the frequency band of 10 MHz and 10 GHz are shown as $\varepsilon_r = 5$–$150$, and $\sigma = 0.2$–$5\,\mathrm{S/m}$.

Skin depths and penetration depths of electromagnetic fields $\delta = \sqrt{\frac{2}{\omega\sigma\mu}}$, on human body surface for bio-medical media are about 1–10 cm around 1–10 MHz, when the conductivity of human body is $\sigma = 0.2\,\mathrm{S/m}$.

## 3. EIGEN ELECTROMAGNETIC FIELD CHARACTERISTICS OF CONDUCTING BIO-MEDICAL MEDIA

Models of conducting waveguides for high frequency transmission in bio-medical media are shown in Fig. 1, and waveguide boundary is conducting circular of radius $a$ for conducting bio-medical skin surface. Inner region of conducting cylindrical media is $\Omega_1$ with electromagnetic material constants $\varepsilon_1$, $\mu$, $\sigma$ and exterior region of cylindrical media is $\Omega_2$ with $\varepsilon_2$, $\mu$. Wave equations for electric fields $\mathbf{E}_1$ and $\mathbf{E}_2$ in region $\Omega_1$ and $\Omega_2$ are shown in Eq. (1), where, electric currents sources are $\mathbf{J}_1$ and $\mathbf{J}_2$.

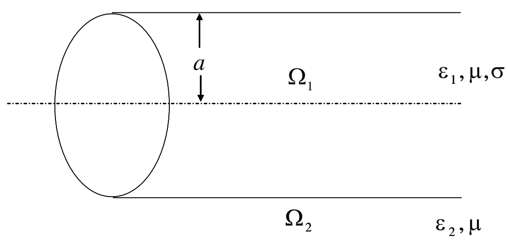



Figure 1: Conducting waveguide with bio-medical boundary surface.

$$\begin{aligned} \nabla\times\nabla\times\mathbf{E}_1(\mathbf{r},\mathbf{r}') - (k_1^2)\,\mathbf{E}_1(\mathbf{r},\mathbf{r}') = \mathbf{J}_1(\mathbf{r},\mathbf{r}') \quad \mathbf{r}\in\Omega_1 \quad \mathbf{r}'\in\Omega_1 \\ \nabla\times\nabla\times\mathbf{E}_2(\mathbf{r},\mathbf{r}') - (k_2^2)\,\mathbf{E}_2(\mathbf{r},\mathbf{r}') = \mathbf{J}_2(\mathbf{r},\mathbf{r}') \quad \mathbf{r}\in\Omega_2 \quad \mathbf{r}'\in\Omega_2 \end{aligned} \tag{1}$$

where, $k_1^2 = \omega^2\mu\varepsilon_1^*$, $\varepsilon_1^* = \varepsilon_1 + \sigma/j\omega$, $k_2^2 = \omega^2\mu\varepsilon_2$.

Boundary conditions for electromagnetic fields on $\partial\Omega_1 = \partial\Omega_2 = S$ are

$$\begin{aligned} \mathbf{n}\times\mathbf{E}_1 &= \mathbf{n}\times\mathbf{E}_2 \\ \frac{1}{-j\omega\mu}\mathbf{n}\times(\nabla\times\mathbf{E}_1) &= \frac{1}{-j\omega\mu}\mathbf{n}\times(\nabla\times\mathbf{E}_2) \end{aligned} \tag{2}$$

In case of current sources $\mathbf{J}_1 \neq 0$, $\mathbf{J}_2 = 0$, if $\varepsilon_1^* = \varepsilon_{10}$, $k_1 = k_{10}$, using operator expression, we have

$$L_{10} = \nabla\times\nabla\times - (k_{10}^2)\,,\ L_2 = \nabla\times\nabla\times - (k_2^2)\,,\ \begin{pmatrix} L_{10} & 0 \\ 0 & L_2 \end{pmatrix}\begin{pmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{J}_1 \\ 0 \end{pmatrix},\ L_0 = \begin{pmatrix} L_{10} & 0 \\ 0 & L_2 \end{pmatrix} \tag{3}$$

Using inverse operator $L_0^{-1}$, electric field $E$ is

$$E = L^{-1}J = GJ = \int \begin{pmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{pmatrix}\begin{pmatrix} \mathbf{J}_1 \\ \mathbf{J}_2 \end{pmatrix} dv \tag{4}$$

Here, if $\mathbf{J}_2 = 0$, $\mathbf{G}_{11} = \mathbf{G}_1$, $\mathbf{G}_{21} = \mathbf{G}_2$, $\mathbf{G}_{12} = \mathbf{G}_{22} = 0$, we have

$$GJ = \int \begin{pmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{pmatrix}\begin{pmatrix} \mathbf{J}_1 \\ 0 \end{pmatrix} dv$$

when $\varepsilon_1^* = \varepsilon_{10}$, in $\Omega_1$ Green functions $\mathbf{G}_1$, $\mathbf{G}_2$ satisfy

$$L_0\begin{pmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{pmatrix} = \begin{pmatrix} I\delta(\mathbf{r}-\mathbf{r}') \\ 0 \end{pmatrix} \quad \mathbf{r}\in\Omega, \quad B\begin{pmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{pmatrix} = 0 \quad \text{on} \quad S \tag{5}$$

For $\mathbf{r} \in \Omega$,

$$\mathbf{G}_1\left(\mathbf{r},\mathbf{r}';k_{10}\right) = \mathbf{G}_1^{(0)}\left(\mathbf{r},\mathbf{r}';k_{10}\right) + \mathbf{G}_1^{(1)}\left(\mathbf{r},\mathbf{r}';k_{10}\right) \quad \mathbf{r} \in \Omega \tag{6}$$

Here, transverse spectrum $\lambda_1 = \sqrt{k_{10}^2 - \beta^2}$. $\mathbf{G}_1^{(0)}\left(\mathbf{r},\mathbf{r}';k_{10}\right)$ is Green function in free space as

$$\begin{aligned}\mathbf{G}_1^{(0)} &= \left(I - \frac{1}{k_{10}^2}\nabla\nabla'\right) G_0,\\ \mathbf{G}_1^{(0)}(\mathbf{r},\mathbf{r}';k_{10}) &= \left(I - \frac{1}{k_{10}^2}\nabla\nabla'\right)\frac{(-1)}{4\pi j}\frac{1}{2}\sum_{m=-\infty}^{\infty} e^{-jm(\theta-\theta')}\int_0^{\infty} J_m\left(\lambda_1 r'\right) H_m\left(\lambda_1 r\right)\cos\beta\left|z-z'\right| d\beta\end{aligned} \tag{7}$$

where $\left(\nabla^2 + k_{10}^2\right) G_0 = -\delta\left(\mathbf{r}-\mathbf{r}'\right)$. Green function $\mathbf{G}_1^{(1)}\left(\mathbf{r},\mathbf{r}';k_{10}\right)$ presents reflection and transmission effects on bio-medical boundary surface, as

$$\begin{aligned}\mathbf{G}_1^{(1)}\left(\mathbf{r},\mathbf{r}';k_{10}\right) \;=\; \sum_{m=-\infty}^{\infty}\int_0^{\infty}\Bigg[&\left\{-j\beta\nabla_t + \mathbf{i}_z\nabla_t^2\right\} J_m\left(\lambda_1 r\right) e^{-jm(\theta-\theta')}\frac{\mathbf{V}_{(m)}^{(1)}}{\beta}\\ &+\left\{\mathbf{i}_z \times \nabla_t J_m\left(\lambda_1 r\right) e^{-jm(\theta-\theta')}\right\}\mathbf{V}_{(m)}^{(1)}\Bigg]\cos\beta\left|z-z'\right| d\beta\end{aligned} \tag{8}$$

Further, $\mathbf{G}_2^{(1)}\left(\mathbf{r},\mathbf{r}';k_{10}\right)$, $\mathbf{r} \in \Omega_2$ is given by, if $\lambda_2 = \sqrt{k_2^2 - \beta^2}$,

$$\mathbf{G}_2^{(1)}\left(\mathbf{r},\mathbf{r}';k_{10}\right) = \mathbf{G}_1^{(1)}\Big]_{\substack{\lambda_1\to\lambda_2\\ J_m\to H_m^{(2)}\\ \mathbf{V}^{(1)}\to\mathbf{V}^{(2)}}} \tag{9}$$

Vector functions $\mathbf{V}^{(1)}$, $\mathbf{V}^{(2)}$ are determined by the boundary condition at boundary surface $r = a$.

$$\mathbf{v} = \begin{pmatrix} \mathbf{V}_{(m)}^{(1)}/\beta \\ \mathbf{V}_{(m)}^{(1)} \\ \mathbf{V}_{(m)}^{(2)} \\ \mathbf{V}_{(m)}^{(1)}/\beta \end{pmatrix} \tag{10}$$

Here, we define

$$F = \begin{bmatrix} -\frac{m\beta}{a}J_m\left(\lambda_1 a\right) & -\lambda_1 J_m'\left(\lambda_1 a\right) & \frac{m\beta}{a}H_m^{(2)}\left(\lambda_2 a\right) & \lambda_2 H_m^{(2)'}\left(\lambda_2 a\right) \\ -\lambda_1 J_m\left(\lambda_1 a\right) & 0\lambda_2 H_m^{(2)}\left(\lambda_2 a\right) & 0 & \\ -j\omega\varepsilon_{10}\lambda_1 J_m'\left(\lambda_1 a\right) & -\frac{m\beta}{j\omega\mu a}J_m\left(\lambda_1 a\right) & j\omega\lambda_2 H_m^{(2)'}\left(\lambda_2 a\right) & \frac{m\beta}{j\omega\mu a}H_m^{(2)}\left(\lambda_1 a\right) \\ 0 & \frac{\lambda_1}{j\omega\mu}J_m\left(\lambda_1 a\right) & 0 & -\frac{\lambda_2^2}{j\omega\mu}H_m^{(2)}\left(\lambda_2 a\right) \end{bmatrix} \tag{11}$$

$$J = \frac{1}{4\pi j}\frac{1}{2}e^{jm(\theta-\theta')+j\beta|z-z'|}\begin{bmatrix} \left(\mathbf{i}_\theta - \frac{1}{k_{10}^2}\frac{\partial}{a\partial\theta}\nabla'\right) \\ \left(\mathbf{i}_z - \frac{1}{k_{10}^2}\frac{\partial}{\partial\theta}\nabla'\right) \\ \frac{1}{-j\omega\mu}\left(\frac{\partial}{\partial z}\mathbf{i}_r - \frac{\partial}{\partial r}\mathbf{i}_z\right) \\ \frac{1}{-j\omega\mu}\left(\frac{\partial}{\partial r}\mathbf{i}_\theta + \frac{1}{r}\frac{\partial}{\partial\theta}\mathbf{i}_r\right) \end{bmatrix} J_m\left(\lambda_1 r'\right) H_m^{(2)}\left(\lambda_1 r\right) e^{-jm(\theta-\theta')} e^{-j\beta|z-z'|} \tag{12}$$

When $\hat{J} = J\Big]_{r=a}$, $\mathbf{v}$ is given by $\mathbf{v} = F^{-1}J(v_i) = \left(\sum_{k=1}^{4} F_{ik}J_k\right)$.

Here, $F_{k_i}$ is co-factor of $F_{k_i}$ and $F_{k_i} = \frac{F_{k_i}}{|F|}$. Eigen electromagnetic fields are given by the eigen equation $|F| = 0$,

$$\begin{aligned}|F| \;=\; & J_m^2\left(\lambda_1 a\right) H_m^{(2)^2}\left(\lambda_2 a\right)\frac{\left(\lambda_1 a\right)^4}{\mu_1\mu_2 a^2}\left[-m^2\beta^2\left(\frac{1}{\left(\lambda_1 a\right)^2} - \frac{1}{\left(\lambda_2 a\right)^2}\right) - \frac{\omega^2\varepsilon_2\mu}{\left(\lambda_2 a\right)^2}\left\{\frac{H_m^{(2)'}\left(\lambda_2 a\right)}{H_m^{(2)}\left(\lambda_2 a\right)}\right\}^2\right.\\ &\left. -\frac{\omega^2\varepsilon_{10}\mu}{\left(\lambda_1 a\right)}\left\{\frac{J_m'\left(\lambda_1 a\right)}{J_m\left(\lambda_1 a\right)}\right\}^2 + \frac{J_m'\left(\lambda_1 a\right) H_m^{(2)'}\left(\lambda_2 a\right)}{J_m\left(\lambda_1 a\right) H_m^{(2)}\left(\lambda_2 a\right)}\frac{\omega^2\varepsilon_2\mu + \omega^2\varepsilon_1\mu}{\left(\lambda_1 a\right)\left(\lambda_2 a\right)}\right]\end{aligned} \tag{13}$$

When current source $\mathbf{J}_2$ exists in region $\Omega_2$, $\mathbf{G}_{12}$, $\mathbf{G}_{22}$ are similarly shown.

Fundamental mode of lossy conducting waveguide $\mathrm{TM}_{00}$ is, in bio-medical media $\Omega_1$

$$E_{1z} = AJ_0(\lambda_1 r)e^{-j\beta z}, \quad E_{1r} = jA\frac{\beta}{\lambda_1}J_1(\lambda_1 r)e^{-j\beta z}, \quad H_{1\theta} = jA\frac{k_{10}^2}{\omega\mu\lambda_1}J_1(\lambda_1 r)e^{-j\beta z} \tag{14}$$

and in the exterior region $\Omega_2$

$$E_{2z} = BH_0^{(2)}(\lambda_2 r)e^{-j\beta z}, \quad E_{2r} = jB\frac{\beta}{\lambda_2}H_1^{(2)}(\lambda_2 r)e^{-j\beta z}, \quad H_{2\theta} = jB\frac{k_2^2}{\omega\mu\lambda_2}H_1^{(2)}(\lambda_2 r)e^{-j\beta z} \tag{15}$$

From boundary condition, $\frac{\lambda_1}{k_{10}^2}\frac{J_0(\lambda_1 a)}{J_1(\lambda_1 a)} = \frac{\lambda_2}{k_2^2}\frac{H_0^{(2)}(\lambda_2 a)}{H_1^{(2)}(\lambda_2 a)}$ when $\sigma$ is very large,

$$\lambda_1 \cong \sqrt{-j\omega\mu\sigma} \tag{16}$$

Electromagnetic field is surface wave with skin depth on the bio-medical boundary surface. Eigen propagation constant $\beta_{00}$ of $\mathrm{TM}_{00}$ mode is

$$\beta_{00} \cong k_2\left[1 + \frac{(1+j)}{\sqrt{\omega\mu\sigma}2\pi a}\right] \tag{17}$$

## 4. ELECTROMAGNETIC FIELDS IN CURVED CONDUCTING WAVEGUIDE

Propagations of high frequency signal in bio-medical media are mainly concerned with field characteristics in curved parts of human bodies corresponding curved lossy waveguide can be studied by integral equation with Green's functions for normal waveguide and conformal mapping.

We consider fields of curved waveguide in physical coordinates $(X, Y, Z)$ and $Y$ coordinate plane in curved waveguide of radius $a$ with curvature of radius $R$ in Fig. 2. On the $Y$ coordinate plane and $y$ coordinate plane, we consider complex plane $\dot{Z} = X + jZ$ and mapping from complex plane $\dot{Z} = X + jZ$ into complex plane $\dot{z} = x + jz$ by analytic function $X + jZ = R\left(1 + \frac{x}{R}\right)e^{j\frac{\dot{z}}{R}}$. Curved waveguide with radium $a$ in $(X, Y, Z)$ coordinates is mapped in straight waveguide with radius $a$ in new $(x, y, z)$ coordinates.

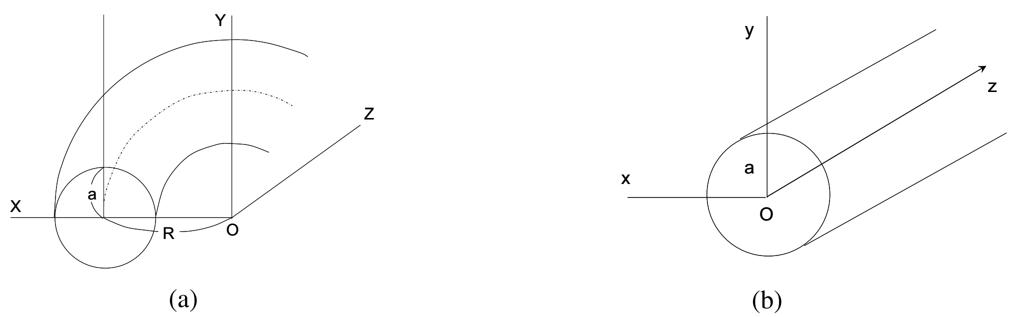



Figure 2: Curved lossy waveguide and conformal mapping. (a) Curved lossy waveguide in $(X, Y, Z)$ space. (b) Straight lossy waveguide in mapped $(x, y, z)$ space.

Here, line element $ds^2$ is $ds^2 = dX^2 + dY^2 + dZ^2 = dx^2 + dy^2 + h^2dz^2$ and metric coefficient $h$ is $h^2 = \left(1 + \frac{x}{R}\right)^2$.

Electric fields $\mathbf{E}$ of Eq. (1) are, in new coordinate space $(x, y, z)$, using unit vector $\mathbf{i}_x$, $\mathbf{i}_y$, $\mathbf{i}_z$, if $\mathbf{E}_t = E_x\mathbf{i}_x + E_y\mathbf{i}_y$,

$$\begin{aligned}
&\nabla_{x,y,z} \times \nabla_{x,y,z} \times \mathbf{E} - \beta^2\mathbf{E} = \\
&-\left(1 - \frac{1}{h^2}\right)\frac{\partial^2\mathbf{E}_t}{\partial z^2} + \left(\frac{1}{h}\frac{\partial h}{\partial x}\right)\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right)\mathbf{i}_y - \left\{\left(1 - \frac{1}{h}\right)\cdot\frac{1}{h}\frac{\partial h}{\partial x}\frac{\partial E_x}{\partial z} - \frac{1}{h^2}\frac{\partial h}{\partial x}\frac{\partial E_x}{\partial z}\right\}\mathbf{i}_z \\
&-\left[\left\{\frac{1}{h^2}\cdot\frac{\partial}{\partial x}\left(h\frac{\partial E_z}{\partial z}\right) - \frac{\partial^2 E_z}{\partial x\partial z}\right\}\mathbf{i}_x + \left\{\left(\frac{1}{h} - 1\right)\frac{\partial^2 E_z}{\partial z\partial y} - \left(\frac{1}{h} - 1\right)\frac{1}{h}\frac{\partial^2 E_z}{\partial z^2}\right\}\mathbf{i}_y - \frac{\partial}{\partial x}\left(\frac{1}{h}\frac{\partial h}{\partial x}E_z\right)\mathbf{i}_z\right] \\
&= F(\mathbf{E})
\end{aligned} \tag{18}$$

The right hand term of Eq. (18) shows inhomogeneous term due to the curve effect and is defined as $F(\mathbf{E})$. Inhomogeneous term of $F(\mathbf{E})$ shows inhomogeneous media in straight waveguide or equivalent sources.

Considering inhomogeneous media effects or source current effects in straight waveguide, we apply the Green's formula using Green's function in straight waveguide of radius $a$. The inhomogeneous term $F(\mathbf{E})$ in Eq. (18) corresponds to $\mathbf{J}$ in Eq. (1), we have operator expression $L_0E = J$. For scattered field $\mathbf{E}_{scatt} = \mathbf{E}_{totl} - \mathbf{E}_{inc}$, where total field $\mathbf{E}_{totl}$ and incident field $\mathbf{E}_{inc}$, when $\mathbf{E}_{inc}$ satisfying $\nabla \times \nabla \times \mathbf{E}_{inc} - \beta^2 \mathbf{E}_{inc} = 0$ is incident wave from input terminal $S_1$, we have

$$\begin{aligned}\mathbf{E}_{totl} =& \mathbf{E}_{inc} - \int \mathbf{G} \cdot \left[ \left(1 - \frac{1}{h^2}\right) \frac{\partial^2 \mathbf{E}_t}{\partial z^2} - \left(\frac{1}{h}\frac{\partial h}{\partial x}\right) + \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right) \mathbf{i}_y + \left\{ \left( \left(1 - \frac{2}{h}\right) \frac{1}{h}\frac{\partial h}{\partial x}\frac{\partial E_x}{\partial z} \right) \mathbf{i}_x \right.\right. \\ &+ \left\{ \left( \frac{1}{h^2}\frac{\partial}{\partial x}\left(h\frac{\partial E_z}{\partial z}\right) - \frac{\partial^2 E_z}{\partial x \partial z} \right) \mathbf{i}_x + \left( \left(\frac{1}{h} - 1\right) \frac{\partial^2 E_z}{\partial z \partial y} - \left(\frac{1}{h} - 1\right) \frac{1}{h}\frac{\partial^2 E_z}{\partial z^2} \right) \mathbf{i}_y - \frac{\partial}{\partial x}\left(\frac{1}{h}\frac{\partial h}{\partial x} E_z\right) \mathbf{i}_z \right] dv \\ &+ \int_{S_1, S_2} \mathbf{n} \cdot [\mathbf{G} \times (\nabla \times \mathbf{E}_{scatt}) - \mathbf{E}_{scatt} \times (\nabla \times \mathbf{G})] dS \end{aligned} \quad (19)$$

where $\mathbf{E}_t = E_x \mathbf{i}_x + E_y \mathbf{i}_y$, and $E_x$ and $E_y$ are components of $\mathbf{E}_{totl}$. Third term in Eq. (19) can be neglected when field matching conditions are satisfied at both terminals $S_1$ and $S_2$.

When the fundamental mode $\mathrm{TM}_{00}$ of surface wave type shown in Eqs. (14)–(16) is incident at curved lossy waveguide, and not so high frequency, lower than 10 MHz, and further few field variation along $z$ direction, scattered field and converted modes are $\mathrm{TM}_{00}$ and $\mathrm{TM}_{11}$ modes, as evaluated in Eq. (20).

$$\mathbf{E}_{scatt} = \int \mathbf{G} \cdot \frac{\partial}{\partial x}\left(\frac{1}{h}\frac{\partial h}{\partial x} E_z\right) \mathbf{i}_z dv \quad (20)$$

## 5. CONCLUSION

For signal communication systems on human body surface of conducting bio-medical media for high frequency transmission, fundamental characteristics of electromagnetic fields along lossy conducting waveguides of bio-medical human skin are shown as boundary value problems of high frequency electromagnetic fields. Particularly, propagation properties of curved lossy conducting waveguides consisting of bio-medical media, corresponding to general parts of human bodies are studied by conformal mapping method and integral equation with Green's functions. Based on theoretical results derived by this theory, the optimum system designs of high frequency communication system including bio-medical sensor for diagnosis may be accomplished to practical clinical information system.

## REFERENCES


1. Gandhi, O. P., *Biological Effects and Medical Applications of Electromagnetic Energy*, Prentice-Hall, 1990.
2. Udagawa, K. and Y. Miyazaki, "Reflection and mode conversion in tilted millimeter circular waveguides — A study by conformal mapping technique," *Jour. IECE, Japan*, Vol. 47, No. 8, 1186–1195, 1964.
3. Akao, Y. and Y. Miyazaki, "Reflection and mode conversion in curved circular waveguide for millimeter wave — A study by using conformal mapping," *Trans. IECE Japan*, Vol. 51-B, No. 11, 553–558, 1968.
4. Miyazaki, Y., "Propagation properties of optical signal waves in perturbed dielectric waveguides by conformal mapping technique," *Topical Meeting on Integrated Optics*, OSA, IEEE, Las Vegas, WA-2, 1972.
5. Miyazaki, Y., "Propagation and radiation characteristics of optical fields in plastic fibers with anisotropic inhomogeneities," *Electromagnetic Waves & Electronic Systems*, Vol. 5, No. 1, 32–45, 2000.
6. Miyazaki, Y., "Electromagnetic compatibility characteristics of information transmission in power lines," *Proc. Int. Sym. on EMC'89 Nagoya*, 9a2-A4, Vol. 2, 475–478, 1989.
7. Miyazaki, Y., "Electromagnetic analysis of circular waveguides for Tera Hertz transmission using conformal mapping," *Proc. of PIERS 2006*, 290–294, Tokyo, 2006.
8. Kikuchi, H., "Wave propagation on the ground retrun circuit in high frequency regions," *Jour. IEE, Japan*, Vol. 75, 1176–1187, 1955.

# Signal Analysis of Electromagnetic Wave Propagation for RFID Systems in In-door and Out-door

**Yasumitsu Miyazaki**[1], **Tadahiro Hashimoto**[2], **and Koichi Takahashi**[1]

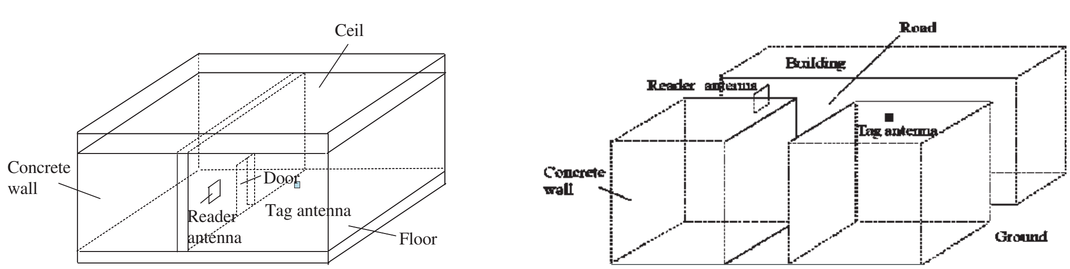
[1]Department of Media Informatics, Aichi University of Technology
50-2 Manori, Nishihasama-cho, Gamagori 443-0047, Japan
[2]Broadband Technology Department, Synclayer, Inc., 1-20 Himegaoka, Kani 509-0249, Gifu, Japan

**Abstract**— In recent years, RFID systems have received much attention in information management and security. For evaluation of RFID systems, study of electromagnetic wave propagation and scattering of UHF wave and microwave transmitted from tag antenna to reader antenna in in-door and out-door is indispensable. In this paper, we describe the characteristics of electromagnetic wave scattering, diffraction and interference by obstacles in propagation channel and show the distribution of received level using FDTD method. In transmitting and receiving points of the weak electromagnetic wave, receiving characteristics are influenced by the propagation environment. We have considered high-speed data communication by microwave in urban area using parallel FDTD computation. In RFID systems, receiving characteristics are strongly influenced by the structure of roads and buildings for out-door environment and rooms, doors and windows for in-door environment. Three-dimensional analysis is studied by numerical simulation for the optimum design of RFID system. Comparing the results of two dimensional analysis and three dimensional analysis, the effects of the earth of roads and building with finite heights for out-door application and effects of ceil and floor of buildings for in-door application can be evaluated.

## 1. INTRODUCTION

In recent years, RFID systems have received much attention in information management and security. The reader of active RFID system receives periodic weak electromagnetic wave pulses of UHF carrier band from multiple tags and recognizes multiple tags. The active RFID is able to transmit the electromagnetic wave tens of meters far from the transmitter. Therefore, the area where the reader can receive the signal from tags is wider than the area of passive RFID. In transmitting and receiving points of the weak electromagnetic wave, receiving characteristics are influenced by the propagation environment. In RFID systems, receiving characteristics are strongly influenced by the structure of roads and buildings for out-door environment and rooms, doors and windows for in-door environment. Therefore, understanding of in-door and out-door propagation characteristics of RFID is necessary to develop high performance antenna of reader. In this paper, FDTD method is applied to show the scattering, diffraction and interference by obstacles and receiving intensity of electric field is discussed. The radiation field generated at tag locations where the reader can receive weak intensity of electric fields from active RFID are investigated.

As an in-door propagation model, a building including several rooms with doors is considered. For out-door propagation model, a road model with buildings as shown in Fig. 1 is considered. Evaluation of out-door propagation characteristics of RFID is necessary to develop high performance antenna system of readers [1–7]. Figs. 1(a) and (b) show several rooms and the T-type road.




(a) RFID for in-door application (b) RFID for out-door application

Figure 1: Propagation model for RFID system.

## 2. FDTD ANALYSIS OF RF PROPAGATION IN BUILDINGS AND ROOMS

Propagation of electromagnetic wave is strongly influenced by the structure and obstacles in building and shows complex characteristics. For optimum design of radio communication system such as RFID, precise study of electromagnetic wave propagation is indispensable. Two-dimensional FDTD method is applied to show the intensity of received electric field due to the scattering, diffraction and interference by obstacles.

In FDTD simulation for in-door propagation, total-field/scattered-field formulation is applied for generation of incident plane wave which propagates into the building. The space coordinates and time are discretized by $x = (i - i_0)\Delta s$, $y = (j - (j_0 + j_1)/2)\Delta s$ and $t = n\Delta t$. The incident plane wave is represented by

$$\begin{aligned} \mathbf{r} - \mathbf{r}' &= \left(x - x'\right)\mathbf{i}_x + \left(y - y'\right)\mathbf{i}_y = (i - i_0)\,\Delta s + (j - j_1)\,\Delta s \\ \mathbf{i}_\tau &= \cos\alpha \mathbf{i}_x - \sin\alpha \mathbf{i}_y \\ \mathbf{E} &= E_z^n\,(i,j)\,\mathbf{i}_z = \mathbf{i}_z E_0 \sin\left(\omega n \Delta t - k\mathbf{i}_\tau \cdot \left(\mathbf{r} - \mathbf{r}'\right)\right) \\ \mathbf{H} &= H_x^n\,(i,j)\,\mathbf{i}_x + H_y^n\,(i,j)\,\mathbf{i}_y = \frac{k}{\omega\mu}\mathbf{i}_\tau \times \mathbf{E} \end{aligned} \tag{1}$$

where, $k = \omega/c$, $\omega = 2\pi f$, $c$ is the velocity of light in free space and $f$ is the frequency. $\alpha$ is the incident angle. Analysis space of two-dimensional FDTD method is shown in Fig. 2. In this figure, $i = 0, 1, 2, \ldots, N_i$, $j = 0, 1, 2, \ldots, N_j$ and the point $(i, j)$ is in Region 1 when $i_0 \leq i \leq i_1$

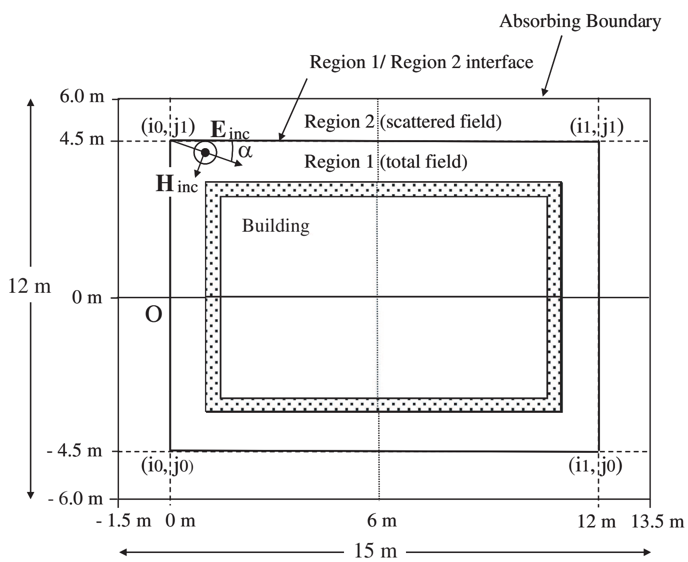
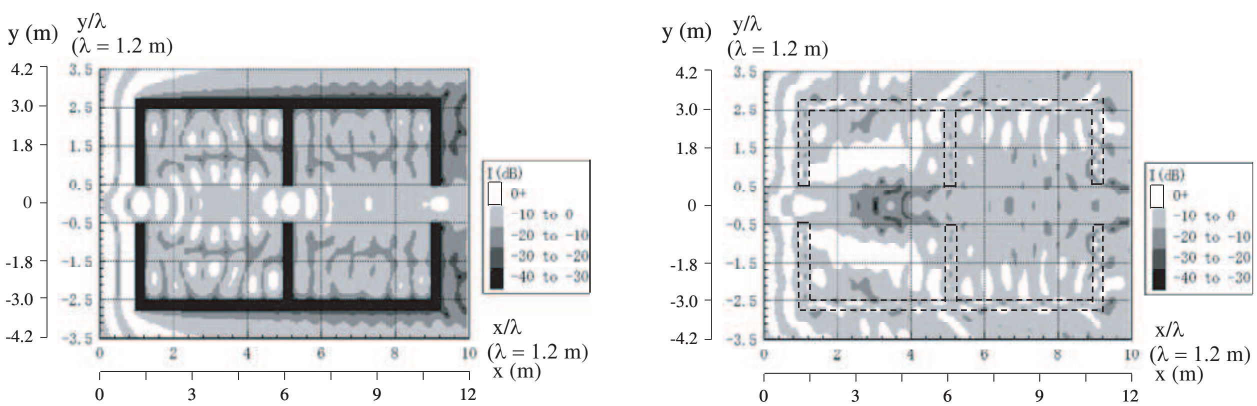



Figure 2: Analysis space of two-dimensional FDTD method for in-door propagation.




(a) Perfectly electric conductor

(b) Concrete ($\varepsilon_r^* = 5.0 - j0.1$)

Figure 3: Maximum received intensity in steady state ($f = 250$ MHz, $\alpha = 0°$) in in-door room.

and $j_0 \leq j \leq j_1$. Fig. 3 shows the maximum electric field intensity in steady state in the building with windows. In case of concrete building, complex relative dielectric constant of building wall is $\varepsilon_r^* = 5.0 - j0.1$. This figure shows that the electric field in the right side room is lower than the left side room. The very weak intensity is observed behind the window due to the interference.

## 3. FDTD ANALYSIS OF RF PROPAGATION IN OUT-DOOR

Electromagnetic fields of RFID system for complicated environment as urban area containing many houses, trees and cars on the roads, can not be easily studied by analytical methods. In three-dimensional FDTD method, electric and magnetic fields generated by current source of tag $J_{zs}$ and $J_{ys}$ are formulated by discretization of Maxwell's equations, if we define space coordinates and time parameters as $x = i\Delta s$, $y = j\Delta s$, $z = k\Delta s$ and $t = n\Delta t$. For electric fields, difference equations are given by

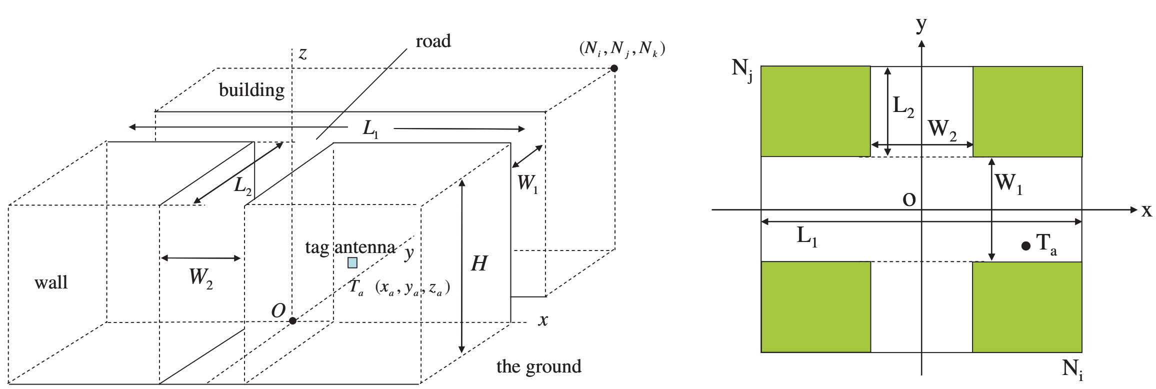
$$\begin{aligned} E_z^n\left(i,j,k+\frac{1}{2}\right) &= C_1 E_z^{n-1}\left(i,j,k+\frac{1}{2}\right) - C_2 J_{zs}^{n-\frac{1}{2}}\left(i,j,k+\frac{1}{2}\right) \\ &\quad + C_3\left\{H_y^{n-1/2}\left(i+\frac{1}{2},j,k+\frac{1}{2}\right) - H_y^{n-1/2}\left(i-\frac{1}{2},j,k+\frac{1}{2}\right)\right. \\ &\quad \left. - H_x^{n-1/2}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right) + H_x^{n-1/2}\left(i,j-\frac{1}{2},k+\frac{1}{2}\right)\right\} \end{aligned} \tag{2}$$

$$\begin{aligned} E_y^n\left(i,j+\frac{1}{2},k\right) &= C_1 E_y^{n-1}\left(i,j+\frac{1}{2},k\right) - C_2 J_{ys}^{n-\frac{1}{2}}\left(i,j+\frac{1}{2},k\right) \\ &\quad + C_3\left\{H_x^{n-1/2}\left(i,j+\frac{1}{2},k+\frac{1}{2}\right) - H_x^{n-1/2}\left(i,j+\frac{1}{2},k-\frac{1}{2}\right)\right. \\ &\quad \left. - H_z^{n-1/2}\left(i+\frac{1}{2},j+\frac{1}{2},k\right) + H_z^{n-1/2}\left(i-\frac{1}{2},j+\frac{1}{2},k\right)\right\} \end{aligned} \tag{3}$$

where, $C_1 = \frac{1-\sigma(i,j,k)\Delta t\,/\,2\varepsilon(i,j,k)}{1+\sigma(i,j,k)\Delta t\,/\,2\varepsilon(i,j,k)}$ $\quad C_2 = \frac{\Delta t\,/\,\varepsilon(i,j,k)}{1+\sigma(i,j,k)\Delta t\,/\,2\varepsilon(i,j,k)}$ $\quad C_3 = \frac{\Delta t\,/\,\varepsilon(i,j,k)}{1+\sigma(i,j,k)\Delta t\,/\,2\varepsilon(i,j,k)}\frac{1}{\Delta s}$.

$i$, $j$, $k$ indicate the position in $(x, y, z)$ space, and $n$ is the number of time steps. $\Delta s = \Delta x = \Delta y = \Delta z$ is space increment, and $\Delta t$ is time increment. $\sigma(i, j, k)$ and $\varepsilon(i, j, k)$ are conductivity and dielectric constant of air space and road wall at coordinate $(i, j, k)$, respectively.

$J_{zs}$ and $J_{ys}$ are current sources of transmitting antenna of tags. $x_a = i_a\Delta s$, $y_a = j_a\Delta s$, $z_a = k_a\Delta s$ where, $(i_a, j_a, k_a)$ is the position of tag source current that is represented by $T_a$ in Fig. 4. The analysis model for out-door propagation is shown in Fig. 4. Road models for out-door analysis consist of the straight road, the T-type road and the cross road. Parameters of T-type and cross roads are horizontal length $L_1$ and width $W_1$ and vertical length $L_2$ and width $W_2$ and the height of building $H$. The walls of buildings are considered to be concrete with relative dielectric constant of 6.0.



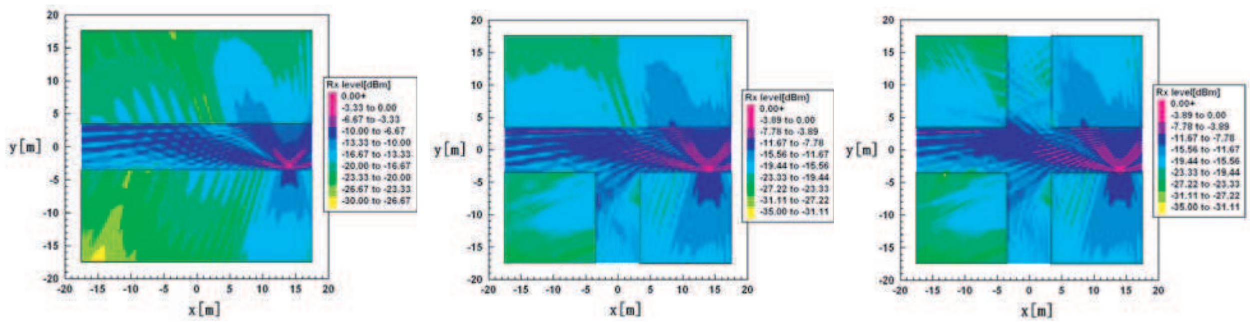

(a) Three-dimensional analysis model for T-type road

(b) Two-dimensional analysis model for cross road

Figure 4: FDTD analysis model for out-door propagation in roads.

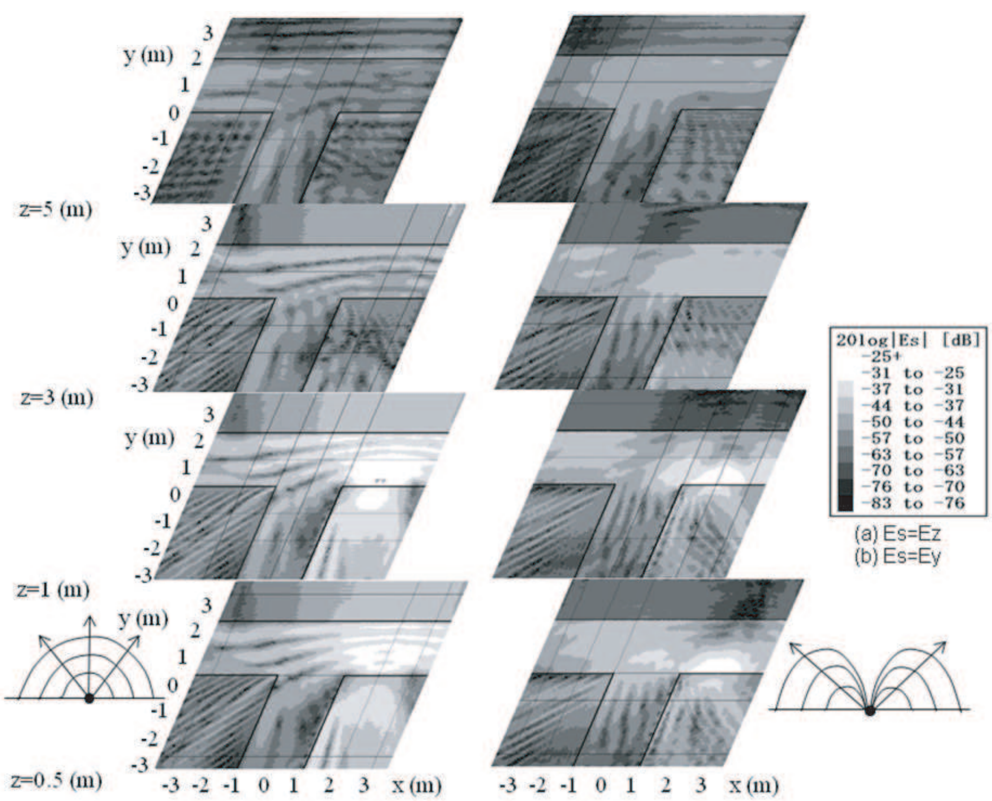
### 3.1. Two-dimensional Analysis

For examples of Active RFID systems, Low Power Radio with frequency 426 MHz is considered and recognized level of the reader is $-100\,\mathrm{dB/m}$. Numerical results of the maximum received intensity in the straight road, the T-type road and the cross road are shown in Fig. 5. In case of straight road, line-of-site communication is provided and strong received intensity is observed at almost all area as shown in Fig. 5(a). However, the intensity becomes weak periodically along $y$ direction in the road due to the interference of direct wave and reflected wave. In case of T-type road, when the tag is horizontal road the area where the received intensity becomes weak is observed behind walls in the vertical road. In case of cross road model, weak received intensity is observed in vertical road when the tag is placed at horizontal road. These results show the fundamental characteristics for the design of RFID systems.




(a) Straight road (b) T-type road (c) Cross road

Figure 5: Maximum received intensity in steady state (Exchanged Power level 1 (mW) $\to$ 0 (dBm)) Position of Tag: $x_a = +14$ (m), $y_a = -2.8$ (m), Concrete Wall ($\varepsilon_r = 6.0$), Antenna: $z$-polarization.




(a) Antenna: $z$-polarization (b) Antenna: $y$-polarization

Figure 6: Maximum received intensity of steady state in T-type road.

### 3.2. Three-dimensional Analysis

The effects of floor, ceil and the ground are evaluated by three-dimensional analysis. Analysis model of T-type road for out-door propagation is shown in Fig. 4(a). In this figure, $W_1$ and $W_2$ are the width of roads in $x$ direction and $y$ direction, respectively. $H$ is the height of the building. $W_1 = W_2 = 2\,(\mathrm{m})$ and $H = 5\,(\mathrm{m})$ are used. The building is a concrete with relative dielectric constant $\varepsilon_r = 6.0$. The relative dielectric constant of the ground is $\varepsilon_{rg} = 6.0$. Tag antenna is a dipole placed at $(x_a, y_a, z_a) = (2\,\mathrm{m},\ 0.14\,\mathrm{m},\ 1\,\mathrm{m})$ and transmits radio wave of 426 MHz frequency. Current source of tag antenna is given by

$$J_s^n\left(i_a, j_a, k_a\right) = I_0 \sin\left(2\pi f n \Delta t\right), \quad s = z, y \tag{4}$$

where $f = 426\,\mathrm{MHz}$ and $I_0 = 1\,\mathrm{A/m^2}$. Space increment $\Delta s = \lambda/15 = 0.047\,\mathrm{m}$ and time step $\Delta t = T/26 = 1/(26f) = 0.09\,\mathrm{ns}$ are used. For large distance from dipole antenna, radiation fields are given by simple field interference effects. At the receiving point, electric field is given by $E = E_a + E_r^{(1)} + E_r^{(2)} + \ldots$ where $E_a$ is a radiation wave from actual antenna and $E_r^{(i)}$ $(i = 1, 2, \ldots)$ is the reflected waves from walls considered as radiation waves from images depending on antenna directivity characteristics due to polarizations and interferences between original radiation fields and reflected fields corresponding to image radiation fields.

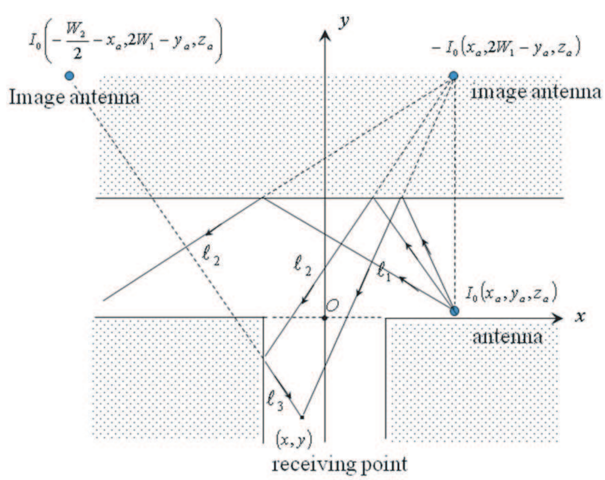



Figure 7: Interference of direct wave and reflected wave at receiving point in road.

## 4. CONCLUSION

For the optimum design of RFID system, we analyzed the characteristics of electromagnetic wave propagation for in-door and out-door active RFID system consisting of multiple tags and reader using FDTD method. In this paper, we formulated FDTD method for electromagnetic fields radiated from source current of RFID Tag. We have studied computer simulation using FDTD formulation, and make clear the characteristics of propagation. Particularly, the interference which depends on the polarization of tag antenna is shown by three-dimensional analysis. These simulations are very useful to obtain the fundamental data for optimum position system design of readers. As a next step, we perform field experiments using actual tags and compact reader in out-door and compare numerical results with the experimental results. These results yield basic foundation for the design of RFID system, and system design of resolving collision problems for multiple pulses. We are also studying parallel FDTD computation for three dimensional wide area analysis and results by three dimensional analysis and two dimensional analysis are compared precisely.

## REFERENCES

1. Kaiser, U. and W. Steinhagen, "A low-power transponder IC for high-performance identification systems," *IEEE Journal of Solid State Circuits*, Vol. 30, No. 3, 306–310, 1995.
2. Masuda, T. and Y. Miyazaki, "Scattering characteristics of electromagnetic environment by construction on the ground," *Trans. IEE Japan*, Vol. 118-C, No. 1, 112–117, 1998.

3. Miyazaki, Y. and K. Takahashi, "Electromagnetic compatibility characteristics of wireless LAN for indoor environment of buildings by FDTD method," *Proc. Asia-Pacific Conference on Environmental Electromagnetics*, 251–256, Shanghai, China, 2000.
4. Rodriguez, G., Y. Miyazaki, and N. Goto, "Matrix-based FDTD parallel algorithm for big areas and its applications to high-speed wireless communications," *IEEE Trans. Antennas and Propagation*, Vol. 54, No. 3, 785–796, 2006.
5. Taki, K. and Y. Miyazaki, "Input impedances and current distributions for meander line antennas with planar coupled parasitic meander element," *PIERS Proceedings*, 1700–1703, Beijing, China, March 26–30, 2007.
6. Masuda, T., Y. Miyazaki, and Y. Kashiwagi, "Analysis of electromagnetic wave propagation in out-door active RFID system using FD-TD method," *PIERS Proceedings*, 1713–1715, Beijing, China, March 26–30, 2007.
7. Miyazaki, Y., T. Hashimoto, and K. Takahashi, "FDTD analysis of electromagnetic wave propagation for out-door active RFID system," *PIERS Proceedings*, 1429–1433, Beijing, China, March 23–27, 2009.

# Dimensional Effects on Electric Potentials and Fields in High-permittivity Thin Films and Interfaces

**Rainer Dick**
Physics & Engineering Physics, University of Saskatchewan
116 Science Place, Saskatoon SK S7N 5E2, Canada

**Abstract**— An effective Lagrange density for electromagnetic fields in the presence of a high-permittivity thin film or interface yields an analytic expression for electromagnetic potentials which interpolates between the two-dimensional logarithmic Coulomb potential at short distances and the three-dimensional Coulomb potential at large distances. It is shown that the analytic expression for the dimensionally hybrid electromagnetic potential is a very good approximation to the infinite series solution from image charges.

## 1. INTRODUCTION

We report an investigation on the impact of quasi two-dimensional and three-dimensional behavior on electromagnetic interactions in interfaces and thin films of high dielectric permittivity.

We recall that the Green's function in $d$ spatial dimensions satisfies

$$\Delta G_d(r) = -\delta(\vec{r}), \tag{1}$$

and therefore determines electrostatic and exchange interactions through the electrostatic potential

$$\Phi_d(r) = \frac{q}{\epsilon} G_d(r). \tag{2}$$

Here $q$ is an electric charge in a dielectric material of permittivity $\epsilon$.

In a more general context,

$$G_d(r) \equiv \langle \mathbf{x}|G_d(E=0)|\mathbf{x}'\rangle\big|_{|\mathbf{x}-\mathbf{x}'|=r}$$

is also known as the zero energy Green's function, see [1, 2]. In vacuum or a homogeneous material it is given by

$$G_d(r) = \begin{cases} -\,r/2, & d=1, \\ -\,(2\pi)^{-1}\ln(r/a), & d=2, \\ \Gamma\left(\dfrac{d-2}{2}\right)\left(4\sqrt{\pi}^d r^{d-2}\right)^{-1}, & d\geq 3. \end{cases} \tag{3}$$

Therefore, the number of spatial dimensions has a profound impact on distance laws of interaction potentials (2).

The motivation for investigating the behavior of interaction potentials in the presence of high-permittivity thin films or interfaces resulted on the one hand from our previous results on Schrödinger theory Green's functions in the presence of thin interfaces with different effective electron mass. In these cases the existence of low-dimensional structures in bulk materials has a strong impact on the density of states of particles [1] and on impurity scattering [2]. It was a natural question to ask whether similar results would obtain in the presence of quasi two-dimensional modifications of permittivity instead of effective electron mass. On the other hand, high-permittivity thin films are important for the manufacturing of high-capacitance domains in highly integrated nanotechnology and microelectronics devices. Applications involve highly integrated low frequency or low damping oscillators, memory devices, low voltage thin film transistors, and efficient energy storage in mesoscopic or nanoscale devices.

To elucidate possible implications of dimensionality on forces in quasi two-dimensional systems, a Lagrangian model is introduced and motivated in Section 2 to describe the behavior of electromagnetic fields in high-permittivity thin films. This model yields an analytic potential $\Phi(r)$ which turns out to interpolate between two-dimensional logarithmic behavior at mesoscopic distance scales and three-dimensional $r^{-1}$ behavior at large distance between charges. The model can be motivated from the standard three-dimensional quantum electronics Hamiltonian if the presence of a thin high-permittivity region is taken into account. Given the geometry of a thin dielectric slab, the emergence of quasi two-dimensional behavior at least for certain distance scales may not be completely unexpected. However, the best justification for the model arises from comparison with the electrostatic potential from image charges. The potential in a dielectric slab can be calculated as a converging infinite series using image charges (although it cannot be written in closed form e.g., in terms of known special functions). We find very good numerical agreement between the analytic model and the potential from image charges both at mesoscopic and large distances. The analytic model therefore yields a good approximation which shows the deviations from the $r^{-1}$ distance law at mesoscopic scales.

## 2. EMERGENCE OF A DIMENSIONALLY HYBRID POTENTIAL IN HIGH-PERMITTIVITY THIN FILMS

The system of interest for us consists of non-relativistic charged (quasi-)particles of mass $m$ and charge $q$ and photons in a high-permittivity thin film or interface. We assume planar geometry, and splitting of three-dimensional vectors into two-dimensional vectors parallel to the interface, and normal components perpendicular to the interface, will be indicated through boldface notation for the parallel vectors, e.g., $\vec{x} = (\boldsymbol{x}, z)$, $\vec{B} = (\boldsymbol{B}, B_z)$, $\vec{\nabla} = (\boldsymbol{\nabla}, \partial_z)$. Vector components $A_z$, $E_z$, $B_z$ orthogonal to the interface or thin film will be denoted as *normal* or *perpendicular* components. This must not be confused with *transverse* three-dimensional vector fields which have vanishing divergence, e.g., $\vec{A}$ will be transverse in Coulomb gauge, $\vec{\nabla}\cdot\vec{A} = 0$, and the corresponding contribution $\vec{E}_\perp = -\partial\vec{A}/\partial t$ to the electric field will be transverse, but generically these vector fields will *not* be orthogonal to the thin film or interface.

The charged particles are described by field operators $\psi$ and $\psi^+$, and the $\vec{x}$-space field operator for the photons is the transverse vector potential $\vec{A}$, $\vec{\nabla}\cdot\vec{A} = 0$. The Lagrange density

$$\begin{aligned}\mathcal{L} \;=\; & \frac{\mathrm{i}\hbar}{2}\left(\psi^+\cdot\frac{\partial}{\partial t}\psi - \frac{\partial}{\partial t}\psi^+\cdot\psi\right) - q\psi^+\Phi\psi + \mu_q(z)\psi^+\vec{\sigma}\cdot\vec{B}\psi \\ & + \frac{1}{2m(z)}\left(\mathrm{i}\hbar\vec{\nabla}\psi^+ - q\psi^+\vec{A}\right)\cdot\left(\mathrm{i}\hbar\vec{\nabla}\psi + q\vec{A}\psi\right) + \frac{\epsilon(z)}{2}\vec{E}^2 - \frac{1}{2\mu(z)}\vec{B}^2 \end{aligned} \tag{4}$$

yields a corresponding Hamiltonian density

$$\mathcal{H} = \frac{1}{2m(z)}\left(\hbar\vec{\nabla}\psi^+ + \mathrm{i}q\psi^+\vec{A}\right)\cdot\left(\hbar\vec{\nabla}\psi - \mathrm{i}q\vec{A}\psi\right) - \mu_q(z)\psi^+\vec{\sigma}\cdot\vec{B}\psi + \frac{\epsilon(z)}{2}\vec{E}^2 + \frac{1}{2\mu(z)}\vec{B}^2, \tag{5}$$

where $m(z)$, $\mu_q(z)$, $\epsilon(z)$ and $\mu(z)$ are piecewise constant functions of $z$ due to the presence of an interface layer or thin film at $-a < z < a$. The values in the layer will be denoted by an asterisk, e.g.,

$$\epsilon(z) = \begin{cases} \epsilon, & z < -a \text{ or } z > a, \\ \epsilon_*, & -a < z < a. \end{cases}$$

Effects of dimensionally hybrid Hamiltonians for non-relativistic particles with different effective masses $m_*$ inside a thin film or interface and $m$ in the surrounding matrix were discussed in references [1–3]. Here we will focus on the implication for the effective Coulomb and exchange interaction integrals of charged particles in the thin film.

The following equations of motion and junction conditions follow directly from $\delta S = \delta\int d^4x\,\mathcal{L} = 0$. $\delta S/\delta\Phi = 0$ yields both the bulk equation

$$\vec{\nabla}\cdot\epsilon(z)\vec{E} = q\psi^+\psi \tag{6}$$

and the junction conditions

$$\epsilon E_z\Big|_{z\to -a-0} = \epsilon_* E_z\Big|_{z\to -a+0}, \quad \epsilon_* E_z\Big|_{z\to a-0} = \epsilon E_z\Big|_{z\to a+0}. \tag{7}$$

$\delta S/\delta\vec{A} = 0$ yields both the bulk equation

$$\begin{aligned}\vec{\nabla}\times\frac{\vec{B}}{\mu(z)} \;=\; & \frac{q}{2\mathrm{i}m(z)}\left(\psi^+\cdot\hbar\vec{\nabla}\psi - \hbar\vec{\nabla}\psi^+\cdot\psi - 2\mathrm{i}q\psi^+\vec{A}\psi\right) \\ & + \vec{\nabla}\times\left(\mu_q(z)\psi^+\vec{\sigma}\psi\right) + \epsilon(z)\frac{\partial\vec{E}}{\partial t}\end{aligned} \tag{8}$$

and the junction conditions

$$\begin{aligned}
\left[\frac{1}{\mu}\boldsymbol{B} - \mu_q \psi^+ \boldsymbol{\sigma}\psi\right]_{z\to -a-0} &= \left[\frac{1}{\mu_*}\boldsymbol{B} - \mu_{q*} \psi^+ \boldsymbol{\sigma}\psi\right]_{z\to -a+0}, \\
\left[\frac{1}{\mu_*}\boldsymbol{B} - \mu_{q*} \psi^+ \boldsymbol{\sigma}\psi\right]_{z\to a-0} &= \left[\frac{1}{\mu}\boldsymbol{B} - \mu_q \psi^+ \boldsymbol{\sigma}\psi\right]_{z\to a+0},
\end{aligned} \tag{9}$$

and $\delta S/\delta\psi^+ = 0$ yields both the bulk equation

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = \left(\mathrm{i}\hbar\vec{\nabla} + q\vec{A}\right)\cdot\frac{1}{2m(z)}\left(\mathrm{i}\hbar\vec{\nabla} + q\vec{A}\right)\psi + q\Phi\psi - \mu_q(z)\vec{\sigma}\cdot\vec{B}\psi \tag{10}$$

and the junction conditions

$$\begin{aligned}
\frac{1}{m}\left[(\mathrm{i}\hbar\partial_z + qA_z)\,\psi\right]_{z\to -a-0} &= \frac{1}{m_*}\left[(\mathrm{i}\hbar\partial_z + qA_z)\,\psi\right]_{z\to -a+0}, \\
\frac{1}{m_*}\left[(\mathrm{i}\hbar\partial_z + qA_z)\,\psi\right]_{z\to a-0} &= \frac{1}{m}\left[(\mathrm{i}\hbar\partial_z + qA_z)\,\psi\right]_{z\to a+0}.
\end{aligned} \tag{11}$$

Additional junction conditions usually follow from the analysis of Equations (6), (8), (10), but we emphasize the junction conditions (7), (9), (11) which follow directly from $\delta S = \delta\int d^4x\,\mathcal{L} = 0$, because these junction conditions imply that the $\delta$ function terms in Equations (6), (8) and (10) vanish, such that we can pull the derivatives through the piecewise constant parameters. This yields

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = \frac{1}{2m(z)}(\mathrm{i}\hbar\vec{\nabla} + q\vec{A})^2\psi + q\Phi\psi - \mu_q(z)\vec{\sigma}\cdot\vec{B}\psi,$$

$$\frac{1}{\mu(z)}\vec{\nabla}\times\vec{B} = \frac{q}{2\mathrm{i}m(z)}\left(\psi^+\cdot\hbar\vec{\nabla}\psi - \hbar\vec{\nabla}\psi^+\cdot\psi - 2\mathrm{i}q\psi^+\vec{A}\psi\right) + \mu_q(z)\vec{\nabla}\times\left(\psi^+\boldsymbol{\sigma}\psi\right) + \epsilon(z)\frac{\partial\vec{E}}{\partial t},$$

and

$$\epsilon(z)\vec{\nabla}\cdot\vec{E} = q\psi^+\psi. \tag{12}$$

We solve Equation (12) in Coulomb gauge,

$$\Phi(\vec{x},t) = \frac{q}{\epsilon}\int d^3\vec{x}'\,G(\vec{x},\vec{x}')\psi^+(\vec{x}',t)\psi(\vec{x}',t) \tag{13}$$

where the Green's function has to satisfy

$$\begin{aligned}
\frac{\epsilon(z)}{\epsilon}\Delta G(\vec{x},\vec{x}') &= \left[\Theta(-a-z) + \Theta(z-a) + \Theta(a+z)\Theta(a-z)\frac{\epsilon_*}{\epsilon}\right]\Delta G(\vec{x},\vec{x}') \\
&= -\delta(\vec{x}-\vec{x}').
\end{aligned} \tag{14}$$

We consider a thin film in the sense that transverse excitations can be neglected inside the thin film, $|k_\perp|a \ll 1$. In this limit we can approximate Equation (14) with

$$\Delta G(\vec{x},\vec{x}') + \Theta(a+z)\Theta(a-z)\frac{\epsilon_* - \epsilon}{\epsilon}\boldsymbol{\nabla}^2 G(\vec{x},\vec{x}') = -\delta(\vec{x}-\vec{x}'). \tag{15}$$

Substitution of the Fourier transform

$$\langle\boldsymbol{x},z|G|\boldsymbol{x}',z'\rangle = \frac{1}{4\pi^2}\int d^2\boldsymbol{k}\int d^2\boldsymbol{k}'\,\langle\boldsymbol{k},z|G|\boldsymbol{k}',z'\rangle\exp\left[\mathrm{i}\left(\boldsymbol{k}\cdot\boldsymbol{x} - \boldsymbol{k}'\cdot\boldsymbol{x}'\right)\right]$$

yields

$$\left[\partial_z^2 - \boldsymbol{k}^2 - \Theta(a+z)\Theta(a-z)\frac{\epsilon_* - \epsilon}{\epsilon}\boldsymbol{k}^2\right]\langle\boldsymbol{k},z|G|\boldsymbol{k}',z'\rangle = -\delta(\boldsymbol{k}-\boldsymbol{k}')\delta(z-z').$$

This implies with

$$\langle\boldsymbol{k},z|G|\boldsymbol{k}',z'\rangle = \langle z|G(\boldsymbol{k})|z'\rangle\delta(\boldsymbol{k}-\boldsymbol{k}')$$

the condition

$$\left[\partial_z^2 - \boldsymbol{k}^2 - \Theta(a+z)\Theta(a-z)\frac{\epsilon_* - \epsilon}{\epsilon}\boldsymbol{k}^2\right]\langle z|G(\boldsymbol{k})|z'\rangle = -\delta(z-z').$$

Fourier transformation with respect to $z$ yields

$$\left(k_\perp^2 + \boldsymbol{k}^2\right)\langle k_\perp|G(\boldsymbol{k})|z'\rangle + \frac{\epsilon_* - \epsilon}{\epsilon}\boldsymbol{k}^2 \int d\kappa_\perp \frac{\sin[(\kappa_\perp - k_\perp)a]}{\pi(\kappa_\perp - k_\perp)}\langle \kappa_\perp|G(\boldsymbol{k})|z'\rangle = \frac{1}{\sqrt{2\pi}}\exp(-\mathrm{i}k_\perp z').$$

The Dirichlet kernel $\sin[(\kappa_\perp - k_\perp)a]/[\pi(\kappa_\perp - k_\perp)]$ is localized around $k_\perp$. Both for an interface which is thin compared to any transverse excitations of interest or for small wave numbers we have $|k_\perp|a \ll 1$, and the Dirchlet kernel is approximately constant between $k_\perp - d^{-1}$ and $k_\perp + d^{-1}$. On the other hand, the Green's function will certainly vanish $\sim k_\perp^{-2}$ for large $|k_\perp|$. Therefore we can approximate the previous equation with

$$\left(k_\perp^2 + \boldsymbol{k}^2\right)\langle k_\perp|G(\boldsymbol{k})|z'\rangle + \frac{\ell}{\pi}\int d\kappa_\perp \boldsymbol{k}^2 \langle \kappa_\perp|G(\boldsymbol{k})|z'\rangle = \frac{1}{\sqrt{2\pi}}\exp(-\mathrm{i}k_\perp z'), \tag{16}$$

where

$$\ell \simeq \frac{\epsilon_* - \epsilon}{\epsilon}a.$$

Equation (16) also appears in the calculation of 2-point correlation functions of particles in the presence of a thin interface and has been solved in references [1–3]. The solution yields

$$\langle k_\perp|G(\boldsymbol{k})|z'\rangle = \frac{1}{\sqrt{2\pi}(k_\perp^2 + \boldsymbol{k}^2)}\left(\exp(-\mathrm{i}k_\perp z') - \frac{\mathrm{k}\ell\exp(-\mathrm{k}|z'|)}{1+\mathrm{k}\ell}\right),$$

where $\mathrm{k} \equiv |\boldsymbol{k}|$. The Fourier transformed solutions are

$$\langle z|G(\boldsymbol{k})|z'\rangle = \frac{1}{2\mathrm{k}}\left(\exp(-\mathrm{k}|z-z'|) - \frac{\mathrm{k}\ell\exp(-\mathrm{k}|z| - \mathrm{k}|z'|)}{1+\mathrm{k}\ell}\right)$$

and

$$\begin{aligned}\langle z|G(\boldsymbol{x})|z'\rangle &= \int_0^\infty d\mathrm{k}\int_0^{2\pi} d\varphi\,\frac{\exp(\mathrm{i}\mathrm{k}|\boldsymbol{x}|\cos\varphi)}{8\pi^2} \\ &\quad\times\left(\exp(-\mathrm{k}|z-z'|) - \frac{\mathrm{k}\ell\exp(-\mathrm{k}|z| - \mathrm{k}|z'|)}{1+\mathrm{k}\ell}\right) \\ &= \int_0^\infty \frac{d\mathrm{k}}{4\pi}\left(\exp(-\mathrm{k}|z-z'|) - \frac{\mathrm{k}\ell\exp(-\mathrm{k}|z| - \mathrm{k}|z'|)}{1+\mathrm{k}\ell}\right)J_0(\mathrm{k}|\boldsymbol{x}|).\end{aligned} \tag{17}$$

The Green's function in the thin film is given in terms of a Struve function and a Neumann function[1],

$$G(r) = \langle 0|G(r = |\boldsymbol{x} - \boldsymbol{x}'|)|0\rangle = \int_0^\infty \frac{d\mathrm{k}}{4\pi}\frac{J_0(\mathrm{k}r)}{1+\mathrm{k}\ell} = \frac{1}{8\ell}\left[\mathbf{H}_0\!\left(\frac{r}{\ell}\right) - Y_0\!\left(\frac{r}{\ell}\right)\right]. \tag{18}$$

This yields logarithmic behavior of interaction potentials at small distances $r \ll \ell$ and $1/r$ behavior for large separation $r \gg \ell$ of charges in high-permittivity thin films,

$$r \ll \ell: \quad G(r) = \frac{1}{4\pi\ell}\left[-\gamma - \ln\!\left(\frac{r}{2\ell}\right) + \frac{r}{\ell} + \mathcal{O}\!\left(\frac{r^2}{\ell^2}\right)\right], \tag{19}$$

$$r \gg \ell: \quad G(r) = \frac{1}{4\pi r}\left[1 - \frac{\ell^2}{r^2} + \mathcal{O}\!\left(\frac{\ell^4}{r^4}\right)\right], \tag{20}$$

see also Figure 1.

[1]Our notations for special functions follow the conventions of Abramowitz and Stegun [4].

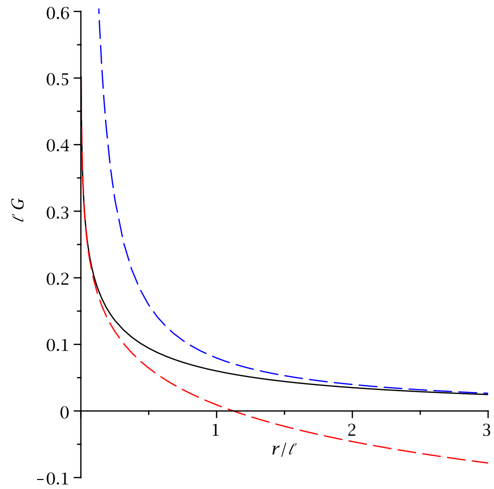



Figure 1: The upper dotted (blue) line is the three-dimensional Green's function $(4\pi r)^{-1}$ in units of $\ell^{-1}$, the continuous line is the Green's function (18) in units of $\ell^{-1}$, and the lower dotted (red) line is the two-dimensional logarithmic Green's function $\ell \cdot G = -(\gamma + \ln(r/2\ell))/(4\pi)$.

## 3. THE APPROACH THROUGH IMAGE CHARGES

We again consider a dielectric slab $-a \le z \le a$ of permittivity $\epsilon_*$ in a bulk of permittivity $\epsilon$. The image charge solution for the potential of a charge $q$ at $\boldsymbol{x} = 0$, $z = 0$ proceeds through the *ansatz*

$$
\begin{aligned}
|z| \le a: \quad \Phi &= \frac{1}{4\pi\epsilon_*}\left[\frac{q}{\sqrt{r^2+z^2}} + \sum_{n=1}^{\infty} q_n \left(\frac{1}{\sqrt{r^2+(z-2na)^2}} + \frac{1}{\sqrt{r^2+(z+2na)^2}}\right)\right] \\
&= \sum_{n=-\infty}^{\infty} \frac{q_{|n|}}{4\pi\epsilon_*\sqrt{r^2+(z-2na)^2}}, \\
z > a: \quad \Phi &= \frac{1}{4\pi\epsilon}\left(\frac{Q}{\sqrt{r^2+z^2}} + \sum_{n=1}^{\infty} \frac{Q_n}{\sqrt{r^2+(z+2na)^2}}\right) \\
&= \sum_{n=0}^{\infty} \frac{Q_n}{4\pi\epsilon\sqrt{r^2+(z+2na)^2}},
\end{aligned}
$$

and symmetric continuation to $z < -a$.

This yields electric fields

$$
\begin{aligned}
|z| \le a: \quad E_r &= \sum_{n=-\infty}^{\infty} \frac{q_{|n|} r}{4\pi\epsilon_*\sqrt{r^2+(z-2na)^2}^3}, \\
E_z &= \sum_{n=-\infty}^{\infty} \frac{q_{|n|}(z-2na)}{4\pi\epsilon_*\sqrt{r^2+(z-2na)^2}^3}, \\
z > a: \quad E_r &= \sum_{n=0}^{\infty} \frac{Q_n r}{4\pi\epsilon\sqrt{r^2+(z+2na)^2}^3}, \\
E_z &= \sum_{n=0}^{\infty} \frac{Q_n(z+2na)}{4\pi\epsilon\sqrt{r^2+(z+2na)^2}^3},
\end{aligned}
$$

and the junction conditions at $z = a$ yield for $n \ge 0$ from the continuity of $E_r$,

$$\frac{q_n + q_{n+1}}{\epsilon_*} = \frac{Q_n}{\epsilon},$$

and from the continuity of $D_z$,

$$q_n - q_{n+1} = Q_n.$$

These conditions can be solved through

$$q_n = \left(\frac{\epsilon_* - \epsilon}{\epsilon_* + \epsilon}\right)^n q, \quad Q_n = \frac{2\epsilon}{\epsilon_* + \epsilon}\left(\frac{\epsilon_* - \epsilon}{\epsilon_* + \epsilon}\right)^n q,$$

$$|z| \le a: \quad \Phi = \frac{q}{4\pi\epsilon_*}\sum_{n=-\infty}^{\infty}\left(\frac{\epsilon_* - \epsilon}{\epsilon_* + \epsilon}\right)^{|n|}\frac{1}{\sqrt{r^2 + (z - 2na)^2}},$$

$$z > a: \quad \Phi = \frac{q}{2\pi(\epsilon_* + \epsilon)}\sum_{n=0}^{\infty}\left(\frac{\epsilon_* - \epsilon}{\epsilon_* + \epsilon}\right)^{n}\frac{1}{\sqrt{r^2 + (z + 2na)^2}}.$$

In particular, the potential at $z = 0$ is

$$\Phi(r) = \frac{q}{4\pi\epsilon_* r} + \frac{q}{2\pi\epsilon_*}\sum_{n=1}^{\infty}\left(\frac{\epsilon_* - \epsilon}{\epsilon_* + \epsilon}\right)^{n}\frac{1}{\sqrt{r^2 + 4n^2a^2}}. \tag{21}$$

We have

$$\sum_{n=1}^{\infty}\left(\frac{\epsilon_* - \epsilon}{\epsilon_* + \epsilon}\right)^{n} = \frac{1}{1 - \dfrac{\epsilon_* - \epsilon}{\epsilon_* + \epsilon}} - 1 = \frac{\epsilon_* + \epsilon}{2\epsilon} - 1 = \frac{\epsilon_* - \epsilon}{2\epsilon}$$

and therefore for $\epsilon_* > \epsilon$

$$\frac{q}{4\pi\epsilon_* r} < \Phi(r) \le \Phi(r)\Big|_{a=0} = \frac{q}{4\pi\epsilon r}.$$

The solution from image charges is in very good agreement with the analytic model from Section 2 for distances $r \gtrsim a/2$, where both the image charge solution and the analytic model show strong deviations from the bulk $r^{-1}$ behavior. This is illustrated in Figure 2 by plotting the reduced electrostatic potential for a charge $q$ in the plane $z = 0$,

$$\frac{a\epsilon}{q}\Phi(r) = aG(r).$$

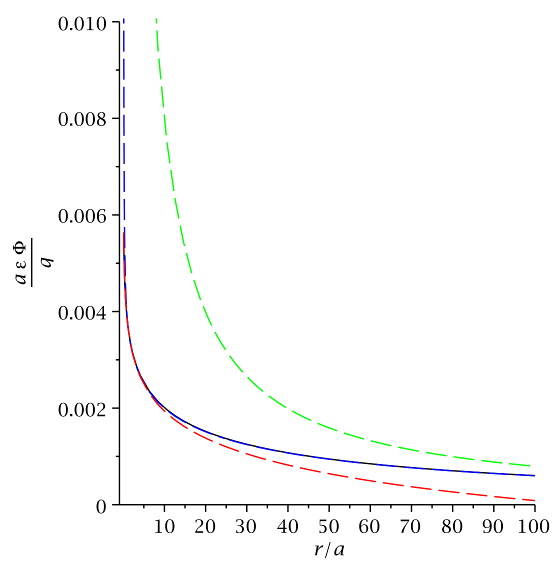



Figure 2: The different reduced electrostatic potentials are plotted for $\epsilon/\epsilon_* = 0.01$. The upper dotted (green) line is the three-dimensional reduced potential $a/(4\pi r)$. The central dotted (blue) line is the reduced potential following from the image charge solution (21). The solid (black) line is the potential from the analytic model (18). The lower dotted (red) line is the reduced logarithmic potential. The reduced potentials from our analytic model and from image charges are indistinguishable for $r \gtrsim a/2$.

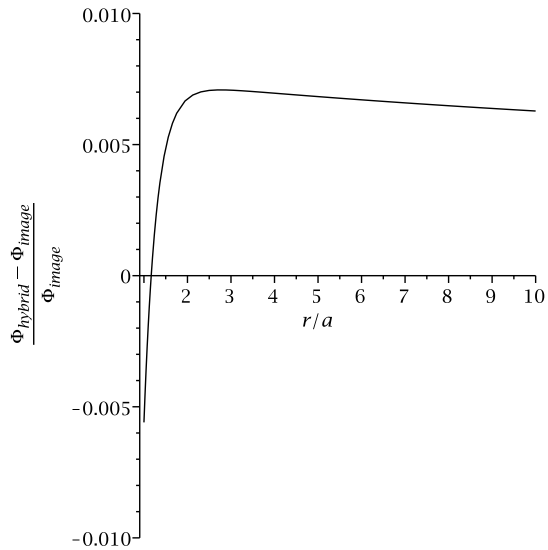



Figure 3: The relative deviation $(\Phi_{hybrid} - \Phi_{image})/\Phi_{image}$ between the dimensionally hybrid potential from (18) and the potential (21) from image charges for $\epsilon/\epsilon_* = 0.01$.

It is also instructive to plot the relative deviation $(\Phi_{hybrid} - \Phi_{image})/\Phi_{image}$ between the dimensionally hybrid potential $\Phi_{hybrid}(r) = qG(r)/\epsilon$ which follows from (18) and the potential $\Phi_{image}$ (21) from image charges.

Figure 3 shows that for $r \gtrsim a/2$, the dimensionally hybrid model is a very good approximation to the potential from image charges with accuracy better than $7 \times 10^{-3}$ if $\epsilon_*/\epsilon = 100$. For $\epsilon_*/\epsilon = 10$ the accuracy is still better than $5 \times 10^{-2}$.

## 4. SUMMARY

An analysis of the model (5, 4) yields dimensionally hybrid electrostatic potentials (13, 18) which interpolate between two-dimensional distance laws at short distances and three-dimensional distance laws at large distances. The analytic model is in very good agreement with the infinite series solution already for small distance scales $r \gtrsim a/2$, where the potential strongly deviates from the standard bulk $r^{-1}$ potential. The transition length between two-dimensional and three-dimensional behavior is predicted to be of order $\ell \simeq a \times (\epsilon_* - \epsilon)/\epsilon$, where $2a$ is the thickness of the film. At distance scales smaller than $a/2$, $r^{-1}$ behavior seems to dominate again, in agreement with expectations that for distances which are small compared to the lateral extension of a dielectric slab, bulk behavior should be restored. However, note that neither the analytic model from Section 2 nor the solution from image charges is trustworthy for very small distances, because both models rely on a continuum approximation through the use of effective permittivities, but the continuum approximation should break down at sub-nanometer scales.

## ACKNOWLEDGMENT

This research was supported by NSERC Canada.

## REFERENCES

1. Dick, R., "Dimensionally hybrid Green's functions and density of states for interfaces," *Physica E*, Vol. 40, 524–530, 2008.
2. Dick, R., "Dimensionally hybrid Green's functions for impurity scattering in the presence of interfaces," *Physica E*, Vol. 40, 2973–2976, 2008.
3. Dick, R., "Hamiltonians and Green's functions which interpolate between two and three dimensions," *Int. J. Theor. Phys.*, Vol. 42, 569–581, 2003.
4. Abramowitz, M. and I. A. Stegun (editors), *Handbook of Mathematical Functions*, 9th printing, Dover Publications, New York, 1970.

# Solution of Axisymmetric Potential Problem in Spherical Coordinates Using Exodus Method

**O. D. Momoh, M. N. O. Sadiku, and C. M. Akujuobi**
College of Engineering, Prairie View A&M University, Prairie View, TX 77446, USA

**Abstract**— In this paper, we implement the use of Exodus method to compute potential distribution in a conducting spherical shell. The results obtained perfectly matched with those obtained using exact solution and explicit finite difference solution method. Also, the approach adopted to overcome singularities encountered in the co-latitude coordinates of the spherical system was highlighted.

## 1. INTRODUCTION

The Exodus method which is a different version of Monte Carlo method was first introduced by Emery and Carson [1, 2]. Unlike the fixed random-walk Monte Carlo method, the Exodus method is not dependent on a random number generator. The fixed random walk and Exodus methods are the most frequently used Monte Carlo solutions of the heat conduction and Poisson equations. Both are based on random walkers moving on the mesh of finite difference approximations of these partial differential equations in different directions of the space with probabilities derived from the corresponding finite difference operators. Unlike finite difference approximation method, the two Monte Carlo techniques compute the solution of a problem at a point in time thereby greatly saving computational resources. However, the Exodus method is preferred to fixed random walk method because of its computational efficiency. It has been found to yield more accurate results with less computing time as compared to the original Monte Carlo method [1, 3]. The Monte Carlo Methods have been applied with great success to the solution of electromagnetic and heat conduction problems in both Cartesian and cylindrical coordinates systems, but have not been widely used in the solution of same problems in spherical coordinates system [4–8]. Specifically, the authors have not come across any literature in which the Exodus method has been used to solve problems in spherical coordinates system.

This paper introduces Exodus method for numerically computing the potential distribution in a conducting spherical shell. In so doing, an existing numerical procedure for dealing with singularity was modified to solve the presence of singularities in the co-latitude ($\phi = 0, \pi$) coordinates.

The rest of the article is organized as follows. In Section 2, we treat the numerical computation of potential distribution. This involves the determination of transition probabilities for the Exodus method. Section 3 discusses the results obtained followed by some conclusions in Section 4.

## 2. COMPUTATION OF POTENTIAL DISTRIBUTION

### 2.1. Finite Difference Scheme for Laplace's Equation in Spherical Coordinates

The Laplace's equation in spherical coordinate systems is

$$0 = \nabla^2 V = \frac{\partial^2 V}{\partial r^2} + \frac{2}{r}\frac{\partial V}{\partial r} + \frac{1}{r^2}\frac{\partial^2 V}{\partial \theta^2} + \frac{\cot\theta}{r^2}\frac{\partial V}{\partial \theta} + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 V}{\partial \phi^2} \tag{1}$$

In this research work, the last term on the right side of (1) may be ignored due to the rotational symmetry about the vertical axis [9, 10]. Therefore (1) reduces to

$$0 = \nabla^2 V = \frac{\partial^2 V}{\partial r^2} + \frac{2}{r}\frac{\partial V}{\partial r} + \frac{1}{r^2}\frac{\partial^2 V}{\partial \theta^2} + \frac{\cot\theta}{r^2}\frac{\partial V}{\partial \theta} \tag{2}$$

Equation (2) was discretized by finite difference method and the respective mesh size in each of the two coordinates are $\Delta r$ and $\Delta\theta$ [11]. The formulation of finite difference scheme in (2) is necessary to obtain the transition probabilities required for the Exodus method. The sampling points are defined by the following coordinate values;

$$\begin{aligned} r_i &= i\Delta r, \quad i = 0, 1, 2, 3, \dots \\ \theta_j &= j\Delta\theta, \quad j = 0, 1, 2, 3, \dots \end{aligned} \tag{3}$$

Therefore the finite difference equation in the spherical coordinates becomes

$$\begin{aligned} V(i,j) &= \frac{0.5}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)} * \left[\left(1+\frac{1}{i}\right)V(i+1,j)+\left(1-\frac{1}{i}\right)V(i-1,j)\right. \\ &\left. +\left(\frac{1}{(i\Delta\theta)^2}+\frac{\cot(j\Delta\theta)}{2i^2\Delta\theta}\right)V(i,j+1)+\left(\frac{1}{(i\Delta\theta)^2}-\frac{\cot(j\Delta\theta)}{2i^2\Delta\theta}\right)V(i,j-1)\right] \end{aligned} \quad (4)$$

### 2.2. Treatment of Singularity Problems

Equation (4) shows the presence of singularities at the center ($i = 0$) and at the poles ($\theta = 0, \pi$) of the sphere. The singularity due to the radial coordinate ($r = 0$) was non-existence in this study since we are dealing with a spherical shell — the radial coordinate does not assume value at the origin, it assumes values in the range

$$r_a \leq r \leq r_b \quad (5)$$

where $r_a$ and $r_b$ are inner and outer radius of the spherical conducting shell respectively [9]. The values of $r_a$ and $r_b$ chosen for the conducting spherical shell used for this study are 1 and 2 m respectively. Also, the potential values (Dirichlet condition) at $r = r_a$ and $r = r_b$ are 0 and 100 V respectively.

The singularity at the poles is treated by taking a quarter of a spherical conducting shell as shown in Figure 1. From the figure, there are two lines of symmetries along $\theta = 0$ and $\theta = \frac{\pi}{2}$ respectively. On these lines of symmetries, the condition $\frac{\partial V}{\partial \theta} = 0$ is imposed [9, 12].

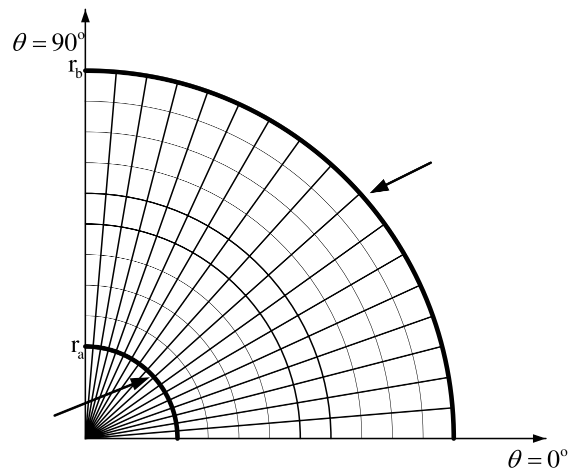



Figure 1: Potential distribution inside a quarter sphere.

The finite difference equations in the two lines of symmetries are as follows:
Along the line $\theta = 0$, $j = 0$

$$V(i,0) = \frac{0.5}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)}\left[\left(1+\frac{1}{i}\right)V(i+1,0)+\left(1-\frac{1}{i}\right)V(i-1,0)+2\left(\frac{1}{(i\Delta\theta)^2}\right)V(i,0)\right] \quad (6)$$

Along the line $\theta = \frac{\pi}{2}$, $j = j_{\max}$

$$V(i,j_{\max}) = \frac{0.5}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)}\left[\left(1+\frac{1}{i}\right)V(i+1,j_{\max})+\left(1-\frac{1}{i}\right)V(i-1,j_{\max})+2\left(\frac{1}{(i\Delta\theta)^2}\right)V(i,j_{\max}-1)\right] \quad (7)$$

This strategy led to the elimination of the singularity causing term $\left(\frac{\cot(j\Delta\theta)}{2i^2\Delta\theta}\right)$, at the spherical poles as evident in (4).

Equations (4), (6), and (7) are used for the finite difference simulation for generating the potential distribution along equipotential surfaces. The potential distribution as illustrated in Figure 1 is typical of each of the remaining three-quarter cross-section of the sphere.

### 2.3. Transition Probabilities Determination

Equation (4) can be re-written as

$$V(i,j) = P_{r+}V(i+1,j) + P_{r-}V(i-1,j) + P_{\theta+}V(i,j+1) + P_{\theta-}V(i,j-1) \tag{8}$$

where

$$P_{r+} = \frac{0.5\left(1+\frac{1}{i}\right)}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)} \tag{9a}$$

$$P_{r-} = \frac{0.5\left(1-\frac{1}{i}\right)}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)} \tag{9b}$$

$$P_{\theta+} = \frac{0.5\left(\frac{1}{(i\Delta\theta)^2}+\frac{\cot(j\Delta\theta)}{2i^2\Delta\theta}\right)}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)} \tag{9c}$$

$$P_{\theta-} = \frac{0.5\left(\frac{1}{(i\Delta\theta)^2}-\frac{\cot(j\Delta\theta)}{2i^2\Delta\theta}\right)}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)} \tag{9d}$$

Note that $P_{r+} + P_{r-} + P_{\theta+} + P_{\theta-} = 1$. Equation (8) can be given a probabilistic interpretation. If a random-walking particle is instantaneously at the point $(r_0,\, \theta_0)$, it has probabilities $P_{r+}$, $P_{r-}$, $P_{\theta+}$, and $P_{\theta-}$ of moving to points $(r+\Delta r,\, \theta)$, $(r-\Delta r, \theta)$, $(r, \theta+\Delta\theta)$, and $(r, \theta-\Delta\theta)$ respectively.

For the random walk computation along the line of symmetry ($\theta = 0$), (6) becomes

$$V(i,0) = P_{r+}V(i+1,0) + P_{r-}V(i-1,0) + P_{\theta+}V(i,1) \tag{10}$$

where

$$P_{r+} = \frac{0.5\left(1+\frac{1}{i}\right)}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)} \tag{11a}$$

$$P_{r-} = \frac{0.5\left(1-\frac{1}{i}\right)}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)} \tag{11b}$$

$$P_{\theta+} = \frac{\left(\frac{1}{(i\Delta\theta)^2}\right)}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)} \tag{11c}$$

Note that $P_{r+} + P_{r-} + P_{\theta+} = 1$ along this line of symmetry. If a random-walking particle is instantaneously at the point $(r_0, 0)$, it has probabilities $P_{r+}$, $P_{r-}$, and $P_{\theta+}$ of moving to points $(r+\Delta r, 0)$, $(r-\Delta r, 0)$, and $(r, \Delta\theta)$ respectively.

For the random walk computation along the line of symmetry ($\theta = 90^\circ$), (7) becomes

$$V(i, j_{\max}) = P_{r+}V(i+1, j_{\max}) + P_{r-}V(i-1, j_{\max}) + P_{\theta-}V(i, j_{\max}-1) \tag{12}$$

The radial parts of the transition probabilities ($P_{r+}$ and $P_{r-}$ ) in (12) are the same as in (11a) and (11b). The transition probability for the co-latitude component of (12) is

$$P_{\theta-} = \frac{\left(\frac{1}{(i\Delta\theta)^2}\right)}{\left(1+\frac{1}{(i\Delta\theta)^2}\right)} \tag{13}$$

Note also that $P_{r+} + P_{r-} + P_{\theta-} = 1$. If a random-walking particle is instantaneously at the point $(r_0,\, j_{\max})$, it has probabilities $P_{r+}$, $P_{r-}$, and $P_{\theta-}$ of moving to points $(r+\Delta r, 0)$, $(r-\Delta r, 0)$, and $(r, 90^\circ - \Delta\theta)$ respectively.

Suppose the potential at a specific point $(r_0,\, \theta_0)$ is to be determined. By defining the transition probabilities $P_n$ as the probability that a random walk starting at the point of interest $(r_0,\, \theta_0)$ ends at the boundary node $(r_n,\, \theta_n)$ with prescribed potential $V_b(n)$.

If there are $M$ boundary or fixed nodes (excluding the corner points since a random walk never terminates at these points), the potential at the starting internal node $(r_0,\, \theta_0)$ of the random walks is

$$V(r_0, \theta_0) = \sum_{n=1}^{M} P_n V_b(n) \tag{14}$$

Since $V_b(n)$ is specified at the fixed nodes (boundary), the remaining task is finding $P_n$. The value of $V(r_0,\, \theta_0)$ in (14) would be "exact" if only the transition probability $P_n$ are calculated exactly. However, Exodus method provides a numerical procedure of finding $P_n$.

To apply the Exodus method, let $P(i,j)$ be the large particles dispatched at point $(i,j)$ in the solution region R. We begin the application of the Exodus method by setting $P(i,j) = 0$ at all nodes (both fixed and free) except at free (internal) node $(r_0,\, \theta_0)$, where $P(i,j)$ assumes a large value $N$. By scanning the mesh as is usually done in finite difference analysis, the particles are dispatched at each free node to its neighboring nodes according to the random walk (transition) probabilities determined in (9), (11) and (13). New $P(i,j)$ becomes zero at the $(r_0,\, \theta_0)$ node, while old $P(i,j)$ is shared among the neighboring nodes. Following each scanning operation, the number of particles that reached the two boundaries with Dirichlet conditions are absorbed and recorded. It is imperative to mention that at the two lines of symmetry, the particles are not absorbed but reflected back into the conduction region to continue the walk [5]. This is due to the Nuemann boundary condition ($\frac{\partial V}{\partial \theta} = 0$) imposed. This scenario portrays the two lines of symmetry as "insulators" borrowing analogy from heat conduction. The scanning process continues in a manner similar to the iterative process applied in finite difference solution until a set number of particles (say 99.99% of $N$) have reached the two boundaries with the Dirichlet conditions. If $N_n$ is the number particles that have reached the boundary n, the probability that a random walk terminates on the boundary is

$$P_n = \frac{N_n}{N} \tag{15}$$

Therefore (15) becomes

$$V(r_0, \theta_0) = \frac{\sum\limits_{n=1}^{M} N_n V_b(n)}{N} \tag{16}$$

## 3. RESULTS

Table 1 compares the results obtained using Exodus method with those obtained from finite difference solution and exact solution. For this study, the value of $N$ is chosen to be equal to $10^6$. The mesh scanning operation (simulation) was performed 200 times which is quite less than 500 times performed in the finite difference solution to obtain the same results. $\Delta r$ and $\Delta\theta$ are chosen as 0.1

Table 1: Comparing Exodus solution with finite difference and exact solutions.

| Point | | Exact | Finite Difference | Exodus |
|---|---|---|---|---|
| $r$ (meter) | $\theta$ (Degree) | | | |
| 1.1 | 40 | 18.1818 | 18.1818 | 18.1818 |
| 1.2 | 30 | 33.3333 | 33.3333 | 33.3333 |
| 1.3 | 45 | 46.1538 | 46.1538 | 46.1538 |
| 1.4 | 50 | 57.1429 | 57.1429 | 57.1429 |
| 1.5 | 60 | 66.6667 | 66.6667 | 66.6667 |
| 1.6 | 65 | 75.0000 | 75.0000 | 75.0000 |
| 1.7 | 70 | 82.3529 | 82.3529 | 82.3529 |
| 1.8 | 80 | 88.8889 | 88.8889 | 88.8889 |

and $9^0$ respectively. However, the value chosen for $\Delta\theta$ does not matter because the accuracy of the computed results does not depend on it. Consequently, it should be noted in Table 1 that the computed potentials at any instance of $r$ (constant spherical surface) remains the same (equipotential) irrespective of the value of $\theta$ used for the computation.

The equation for computing the exact solution is

$$V = V_o \frac{\left[\frac{1}{r_a} - \frac{1}{r}\right]}{\left[\frac{1}{r_a} - \frac{1}{r_b}\right]} \tag{17}$$

The values of $V_0$, $r_a$, and $r_b$ as reflected in Fig. 1 are 100 V, 1 m, and 2 m respectively.

## 4. CONCLUSION

The use Exodus method to compute potential distribution in spherical coordinates has been implemented in this paper. The results obtained agreed totally with those obtained using finite difference (FD) solution and the exact solution. This method can be said to be almost exact when compared to the fixed random walk Monte Carlo method because it does not depend on random number generation.

## REFERENCES

1. Naraghi, M. H. N. and S.-C. Tsai, "A boundary-dispatch Monte Carlo (Exodus) method for analysis of conductive heat transfer problems," *J. Numerical Heat Transfer*, Part B, Vol. 24, 1993, 475-487.
2. Sadiku, M. N. O., *Monte Carlo Methods for Electromagnetics*, 127–159, CRC Press, Boca Raton, 2009.
3. Grigoriu, M., "A Monte Carlo solution of heat conduction and Poisson equations," *J. Heat Transfer Trans. ASME*, Vol. 122, 40–45, Feb. 2000.
4. Thompson, J. J. and P. Y. P. Chen, "Heat conduction with internal sources by modified Monte Carlo methods," *J. Nuclear Engineering and Design*, Vol. 12, 207–214, 1970.
5. Zinsmeister, G. E. and S. S. Pan, "A modification of the Monte Carlo method for steady heat conduction problems," *Int. Journal for Numerical Methods in Engineering*, Vol. 10, 1057–1064, 1976.
6. Sadiku, M. N. O. and T. Hunt, "Solution of Dirichlet problems by the Exodus method," *IEEE Trans. Microwave Theory and Techniques*, Vol. 40, No. 1, 89–95, Jan. 1992.
7. Davey, K. R. and M. B. Nair, "A Monte Carlo Technique for solving eddy current problems," *IEEE Trans. Magnetics*, Vol. 29, No. 2, Mar. 1993.
8. Sadiku, M. N. O., C. M. Akujuobi, S. M. Musa, and S. R. Nelatury, "Analysis of time-dependent cylindrical problems using the Exodus method," *Proc. 2009 IEEE Southeastcon*, 93–97, Mar. 2009.
9. Momoh, O. D. and M. N. O. Sadiku, "An easy approach to numerical computation of spherical shell capacitance using finite difference method," *Int. J. of Modelling and Simulations (IREMOS)*, Vol. 2, No. 2, 196–200, 2009.
10. Myland, J. C. and K. B. Oldham, "Modeling diffusion to a electrode by fully explicit simulation," *J. Electroanalytical Chemistry*, Vol. 576, 353–362, 2005.
11. De Lima, A. G. B. and S. A. Nebra, "Theoretical analysis of the diffusion process inside prolate spheroidal solids," *J. Drying Technology*, Vol. 18, No. 1&2, 21–48, 2000.
12. Sadiku, M. N. O., *Elements of Electromagnetics*, 4th Edition, 29–86, 740–748, Oxford University Press, New York, 2007.

# Behavioral Modeling of Asymmetric Intermodulation Distortion of Nonlinear Amplifier

**A. H. Yuzer[1,2] and S. Demir[1]**

[1]Electrical and Electronics Engineering, Middle East Technical University, Ankara, Turkey
[2]Electrical and Electronics Engineering, Zonguldak Karaelmas University, Zonguldak, Turkey

**Abstract**— In this study, a behavioural modelling is proposed to model asymmetric IMD components produced in the two tone excitation of nonlinear amplifiers. Model parameters are extracted based on single tone and two-tone measurement results. Then measurement results and model predictions are compared for a four-tone signal.

## 1. INTRODUCTION

Amplifier characterization is an important issue especially in the design of amplifiers and systems involving amplifiers such as linearizers and wireless communication systems. Behavioral modeling is one of the characterization techniques where a mathematical relation between input and output is constructed. Conventional power series expansion is one of the methods of behavioral modeling of an amplifier. A sample power series expansion (GPSA) to model the fifth order nonlinearity is as follows:

$$V_o(t) = \sum_{k=1,3,5} \dot{a}_k * V_i^k(t) \tag{1}$$

where $V_i$ and $V_o$ represents the input and the output signals, respectively, and $\dot{a}_k$ are the model coefficients to be found. These coefficients can be taken as real numbers for AM/AM distortion modeling and as complex numbers to include AM/PM distortion [1–4]. When an amplifier is excited with a two-tone signal, inter-modulation distortion (IMD) appears as in-band distortion. Representation of the two-tone input signal and the response of the amplifier to this input signal are shown in Figure 1. Two-tone input signal and lower band IMD (IMDL) and upper band IMD (IMDU) equations are given in (2) and (3).

$$V_i(t) = V_1 \cos\omega_1 t + V_2 \cos\omega_2 t \tag{2}$$

$$\begin{aligned} V_{IMDL} &= \frac{3}{4}\dot{a}_3 V_1^2 V_2 + \frac{5}{4}\dot{a}_5\left[V_1^4 V_2 + \frac{3}{2}V_1^2 V_2^3\right] \\ V_{IMDU} &= \frac{3}{4}\dot{a}_3 V_1 V_2^2 + \frac{5}{4}\dot{a}_5\left[V_1 V_2^4 + \frac{3}{2}V_1^3 V_2^2\right] \end{aligned} \tag{3}$$

If equal amplitude excitation is assumed, IMD components will be of equal amplitude; $V_{IMDL} = V_{IMDUs}$. This means that there is no asymmetry in the IMD expression given in (3). However, there is an inherent asymmetrical distribution of magnitude and phase of IMD components in measurement, which means that (1) require correction.

In the current study, asymmetry in magnitude and phase is modelled by introducing time delay to the power series terms in (1). Namely, a power series with time delay (PSwTD) is introduced. According to our proposed model input-output relation and resultant IMD components for equal excitation tone are found as given in (4) and (5). $a_k$ are real numbers and different than the coefficients given in (3). 9th order model polynomial is used to widen validity range instead of fifth order.

$$V_o(t) = \sum_{k=1,3,5,7,9} a_k * V_i^k(t - \tau_k) \tag{4}$$

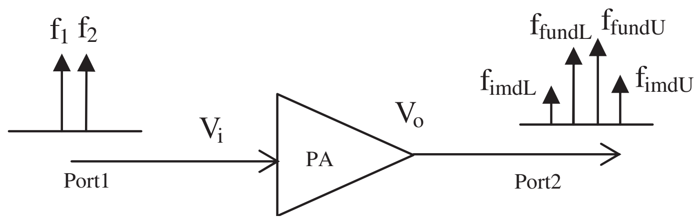



Figure 1: Two tone excitation and response of amplifier to that excitation.

$$\begin{aligned}
V_{IMDL} =& \frac{3}{4}a_3V_1^3\angle(-(2\omega_1-\omega_2)\tau_3)+\frac{25}{8}a_5V_1^5\angle(-(2\omega_1-\omega_2)\tau_5)\\
&+\frac{735}{64}a_7V_1^7\angle(-(2\omega_1-\omega_2)\tau_7)+\frac{1323}{32}a_9V_1^9\angle(-(2\omega_1-\omega_2)\tau_9)\\
V_{IMDU} =& \frac{3}{4}a_3V_1^3\angle(-(2\omega_2-\omega_1)\tau_3)+\frac{25}{8}a_5V_1^5\angle(-(2\omega_2-\omega_1)\tau_5)\\
&+\frac{735}{64}a_7V_1^7\angle(-(2\omega_2-\omega_1)\tau_7)+\frac{1323}{32}a_9V_1^9\angle(-(2\omega_2-\omega_1)\tau_9)
\end{aligned} \tag{5}$$

## 2. MEASUREMENT SETUP AND PHASE MEASUREMENT

Measurement setup is prepared to measure both magnitude and phase of the IMD components. There are three signal generators in the measurement setup. Two of them, SG1 and SG2, are used to create fundamental tones since each signal generator creates only one tone signal. 10 MHz reference signal created by SG3 is transferred to the SG2 and SG1 with shortest possible cable connection for phase locking. Both magnitudes and phases of IMD & FUND components are measured under two tone 30 MHz & 31 MHz excitation by using this setup. Measurement setup will be presented in detail [5].

## 3. FINDING THE COEFFICIENTS AND COMPARISON

The coefficients $a_k$ and $\tau_k$ used for PSwTD modeling are found by equating the measurement results and (5) for both magnitude and phase. The equations are not linear, and the coefficients cannot be found directly. Therefore, the coefficients are found by using optimization tool of MATLAB program. The final coefficients are listed in Table 1. Comparison between two tone excitation measurement and model estimation for magnitudes & phases are given in Figure 2 and Figure 1, respectively.

Error is always lower than 2.5 dB for magnitude of IMDL and IMDU except error in the magnitude of IMDU is getting bigger than 2.5 dB for lower than −5 dBm input power level. Likely error is always lower than 0.5 dB for magnitude of FL and FU as seen is Figure 2.

Although there is error in the phase of IMDU, behaviour of model ouput for phase of IMDU is similar to measurement result. Error is always lower than 0.5° for phase of FL and FU for lower than 3 dBm input power level.

Table 1: Model Coefficients.

| $a_1$ | $a_3$ | $a_5$ | $a_7$ | $a_9$ |
|---|---|---|---|---|
| 123.66 | −44.36 | 137.71 | 188.06 | 73.84 |
| $\tau_1(\mu sn)$ | $\tau_3(\mu sn)$ | $\tau_5(\mu sn)$ | $\tau_7(\mu sn)$ | $\tau_9(\mu sn)$ |
| 0 | −22.28 | −29.45 | −18.57 | −6.25 |

The model coefficients given in Table 1 are found according to two-tone excitation measurement result and comparisons are performed for two tone model estimation. For verification purposes, the model coefficients which are found according to two-tone measurement results are also used to predict the four tone excitation results as given below. Four tone signal is prepared in computer by using MATLAB and then loaded to arbitrary wave shape generator used in the measurement setup. A frequency spectrum of the measured amplifier response to the four tone excitation is given in Table 2. The differences between measured values and model estimation are given in Table 3.

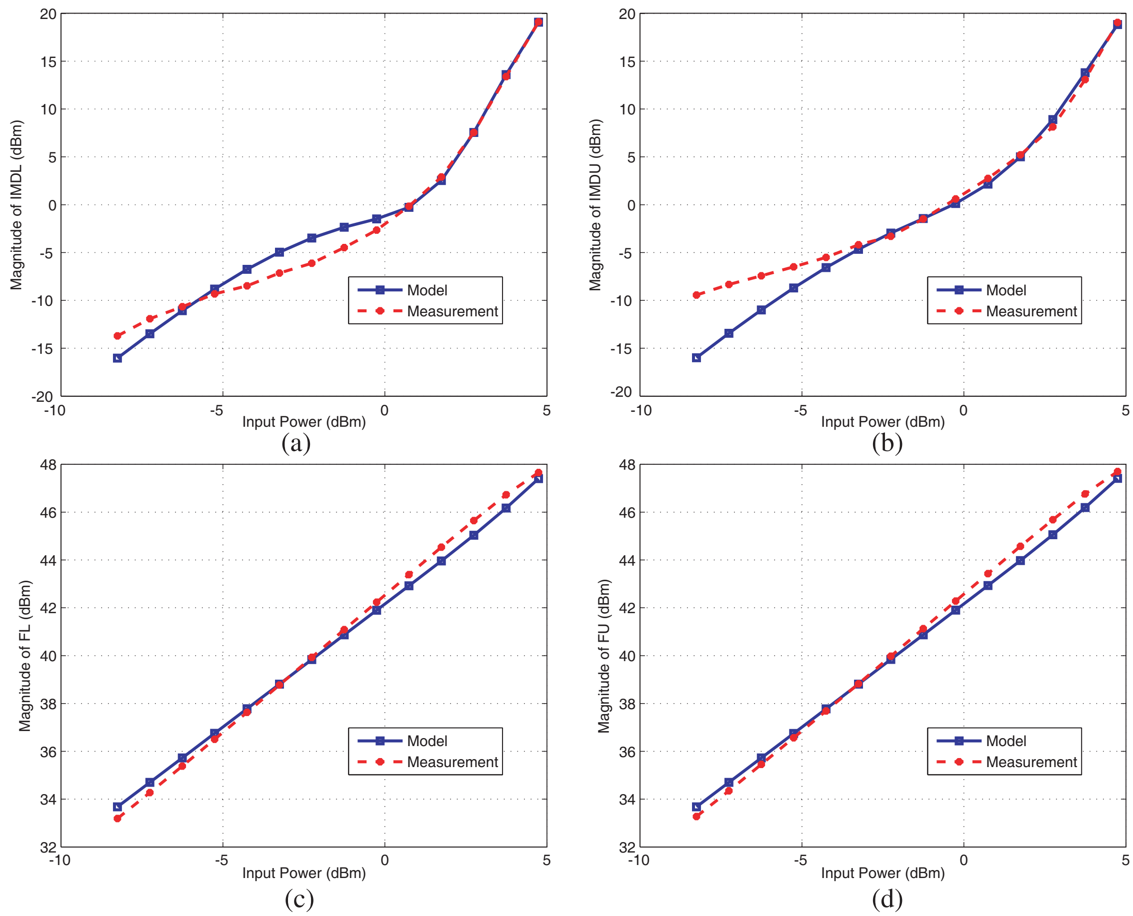



Figure 2: Comparison between model estimation and measurement results for magnitudes of (a) ImdL, (b) ImdU, (c) FundL, and (d) FundU components.

Table 2: Frequaency spectra of amplifier output.

| Index | ImdL3 | ImdL2 | ImdL1 | FundL1 | FundL2 | FundL3 | FundL4 | ImdU1 | ImdU2 | ImdU3 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | −20.26 | −10.34 | 4.12 | 32.20 | 22.89 | 22.78 | 32.17 | 4.29 | −6.18 | −14.45 |
| 2 | −15.36 | −7.01 | 6.83 | 34.34 | 25.08 | 24.93 | 34.30 | 7.11 | −3.42 | −12.14 |
| 3 | −11.20 | −4.15 | 8.99 | 36.59 | 27.30 | 27.19 | 36.55 | 9.22 | −1.15 | −10.67 |
| 4 | −7.99 | −1.10 | 11.03 | 38.90 | 29.59 | 29.49 | 38.86 | 11.25 | 1.15 | −9.10 |
| 5 | −5.90 | 2.11 | 13.62 | 41.25 | 31.96 | 31.86 | 41.20 | 13.86 | 3.73 | −7.88 |
| 6 | −5.97 | 3.14 | 15.62 | 43.60 | 34.29 | 34.20 | 43.55 | 15.82 | 6.60 | −3.55 |
| 7 | 9.18 | −2.43 | 15.95 | 45.84 | 36.35 | 36.29 | 45.77 | 16.52 | 13.08 | 11.26 |

Table 3: Difference in between measured value and model estimated.

| Index | ImdL3 | ImdL2 | ImdL1 | FundL1 | FundL2 | FundL3 | FundL4 | ImdU1 | ImdU2 | ImdU3 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 10.72 | 5.58 | 1.58 | 0.55 | 0.68 | 0.70 | 0.56 | 1.37 | 0.79 | 4.39 |
| 2 | 9.03 | 5.08 | 1.55 | 0.42 | 0.57 | 0.61 | 0.43 | 1.24 | 0.85 | 5.50 |
| 3 | 7.65 | 4.41 | 1.35 | 0.25 | 0.41 | 0.42 | 0.25 | 1.14 | 1.01 | 7.07 |
| 4 | 6.87 | 3.13 | 1.12 | 0.01 | 0.18 | 0.21 | 0.02 | 0.91 | 0.89 | 8.27 |
| 5 | 8.05 | 2.35 | 0.85 | −0.25 | −0.06 | −0.05 | −0.22 | 0.77 | 1.14 | 10.58 |
| 6 | 14.51 | 7.02 | 2.18 | −0.40 | −0.08 | −0.04 | −0.38 | 2.05 | 3.80 | 12.24 |
| 7 | 4.79 | 18.65 | 6.28 | −0.35 | 0.45 | 0.47 | −0.33 | 5.40 | 3.35 | 2.77 |

As seen in the table, error for fundamental tones is less than 0.7 dB for all excitation levels. For the IMD components, error is increasing as the frequency of the IMD component gets away from the center frequency.

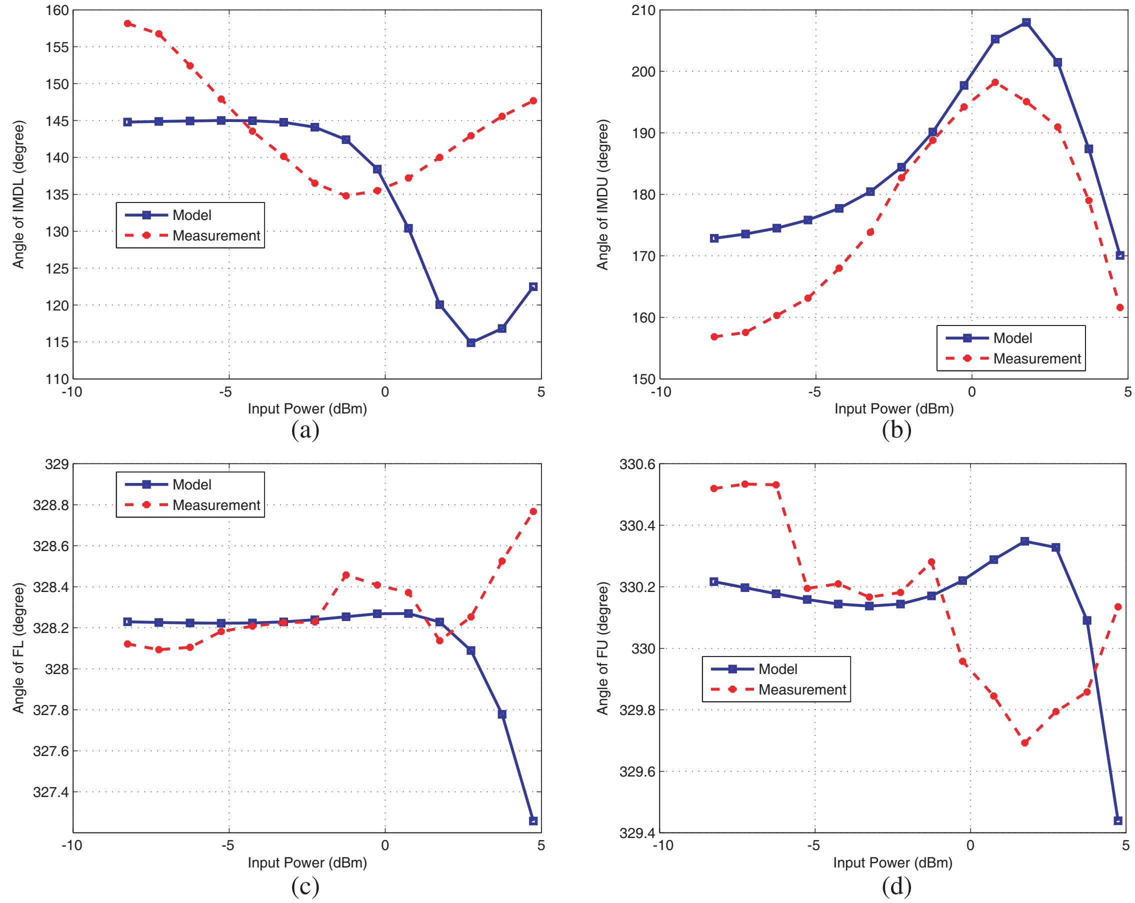



Figure 3: Comparison between model estimation and measurement results for phases of (a) ImdL, (b) ImdU, (c) FundL, and (d) FundU components.

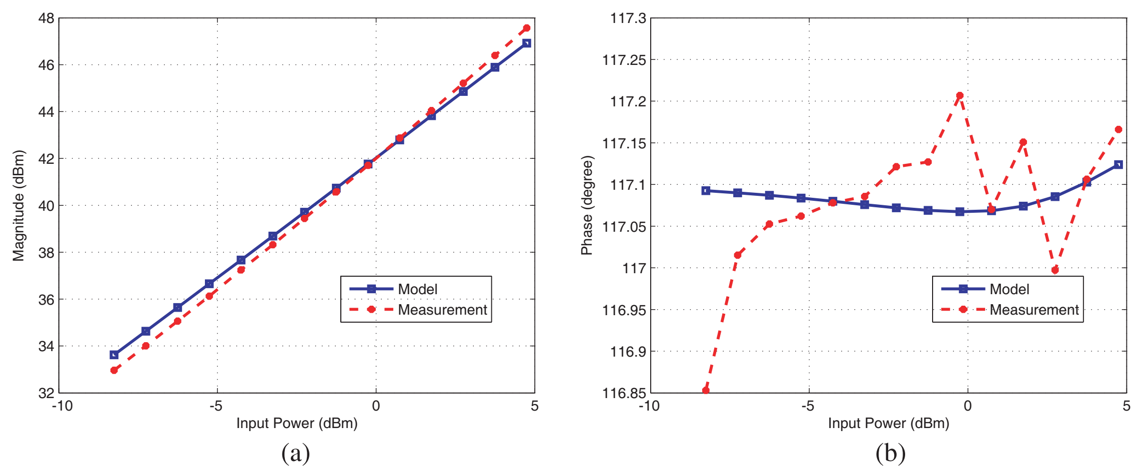



Figure 4: Comparison between model estimation and measurement results of one tone excitation for (a) magnitude and (b) phase.

## 4. CONCLUSION

In this study, a new behavioral modeling technique is introduced which can estimate the asymmetry in the IMD components. This model abbreviated as PSwTD is obtained by adding proper time delay components to the terms of the power series representation. In the presented applications only the odd order terms are considered. PSwTD model for a real sample amplifier is extracted according to single tone and two-tone measurement results. Comparison of the measurement and model results indicate that especially for higher power levels, the magnitudes are more consistent. The overall weighted error function distributes the error in different parts, as a result the phase error in lower IMD component is higher in the corresponding power levels. In order to make an independent comparison a four-tone signal is generated and the amplifier response to this signal is measured. Similarly, output of the amplifier is predicted using the model. Comparisons are given in Table 3. In four tone excitation tone comparison, maximum error for fundamental tones is 0.7 dB. Error is increasing for far-off tones, these four tone comparisons indicate that there is still a need for further optimization for model parameters. Increasing the order of the power series representation which decreases the analytical applicability of the model is not preferred. It is widely expected that the contribution of the even order terms is effective in the memory dependent asymmetry of the amplifier response. Increasing the order of odd order representation cannot accurately count for these effects. In our group, work on a model with even order terms which gives promising results is in progress.

**REFERENCES**

1. Acciari, G., et al., "A fast AM/AM and AM/PM characterization technique," *European Gallium Arsenide and Other Compound Semiconductors Application Symposium*, Munich, 2003.
2. Ciccognani, W., et al., "AM/AM and AM/PM power amplifier characterisation technique," *15th International Conference on Microwaves, Radar and Wireless Communications*, Vol. 2, 678–681, 2004.
3. Boumaiza, S. and F. Ghannouchi, "Realistic power-amplifiers characterization with application to baseband digital predistortion for 3G base stations," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, 3016–3021, 2002.
4. Jeckeln, E. G., et al., "Adaptive baseband/RF predistorter for power amplifiers through instantaneous AM-AM and AM-PM characterization using digital receivers," *2000 IEEE MTT-S International Microwave Symposium Digest*, Vol. 1, 489–492, 2000.
5. Yuzer, H. and S. Demir, "A new phase measurement technique for RF power amplifier only measuring magnitudes," Presented in *Progress In Electromagnetics Research Symposium*, Cambridge, USA, July 5–8, 2010.

# Simultaneous and Synchronous Measurement of Even and Odd Order Nonlinear Distortion Terms

**S. K. Remillard**
Physics Department, Hope College, 27 Graves Place, Holland, MI 49423, USA

**Abstract**— Nonlinear distortion of different orders is usually measured at different frequencies, such as $2f$ and $3f$ for harmonic distortion and $f_1 + f_2$ and $2f_2 - f_1$ for two-tone intermodulation distortion (IMD). This limits the physical insight available from these measurements because the distortion products are generated at different frequencies and different spatial locations in a distributed element device sample. This paper will describe a new technique to characterize nonlinear distortion in which products of even and odd order are generated at the same frequency and the same location in the driving field region. Two probing tones at frequencies $f_1$ and $f_2$, which are very similar in frequency, are introduced at a specific location with a small probe. This determines the location of the nonlinear generation. A third driving tone at $f_3$, which is much higher in frequency than $f_1$ and $f_2$, determines the frequency of the nonlinear generation. With the three tones, 2nd order intermodulation occurs at $f_3 + f_2$ and 3rd order intermodulation occurs at $f_3 + (f_2 - f_1)$, among other mixing terms. By setting $f_3$ much higher than $f_1$ and $f_2$ by at least two orders of magnitude, the 2nd and 3rd order nonlinearities are generated at virtually the same frequency.

## 1. INTRODUCTION

Recently high temperature superconductors (HTS) have had an impact in wireless telecommunications base stations, and have been seen as having a potentially disruptive impact on space based communications systems [1]. The commercial success in wireless communication has so far been in passive filtering of signals using superconductive bandpass filters, notch filters and multiplexers [2]. It has been known for some time that the HTS materials used in these high performance devices respond nonlinearly to microwave currents, generating harmonics and mixing products from the signals that they are supposed to be filtering [3]. The nonlinear response of HTS filter devices has lead to considerable interest in the scientific community due to the potential for the nonlinearity to reveal new information about the electrodynamics of HTS materials [4]. Nonlinearity models of linear devices made from nonlinear HTS materials have also drawn on the understanding gained from studies of signal distortion [5]. Nonlinearity research in HTS materials has also expanded over the years into studies of the passive nonlinear response of other materials including metals [6], dielectrics and ferrites.

Microwave nonlinearity is usually measured one of two ways. The first method is harmonic distortion whence a single tone at frequency, $f$, incident upon a device composed of the nonlinear material, generates even and odd order harmonics at $2f$, $3f$, $4f$ and so on, which are detected by a spectrum analyzer. A second technique is two-tone odd order intermodulation distortion (IMD) whence two tones at frequencies $f_1$ and $f_2$ are incident, and generate odd order IMD at numerous frequencies, for example, $2f_2 - f_1$, $3f_2 - 2f_1$, and so on.

The method reported here is expected to add to the capability of experimentalists studying the electrodynamics of nonlinear materials. The frequencies of even and odd order distortion tones are usually considerably different, being $2f$ and $3f$ for the 2nd and 3rd order harmonics respectively, and being $f_1 + f_2$ and $2f_1 - f_2$ for the 2nd and 3rd order IMD respectively. This puts the harmonic distortion terms considerably outside the frequency range of the driving signal. In the common case that either the device or the input signal probe is resonant at $f$, all odd and even order harmonics and all even order intermodulation are outside the resonance band. This paper reports on a method which results in the even and the odd order intermodulation distortion occurring at, or very near, the resonant frequency.

By using three input frequencies, the even and odd order nonlinearity can be arranged to occur at the same frequency [7]. For convenient measurement, the IMD frequencies can be offset by some small amount that exceeds the resolution of the spectrum analyzer. This approach is now being developed for probing the nonlinearity of the material in a high temperature superconductor as has been previously accomplished using harmonic distortion [8]. The ability of this method to probe locally around the superconductive structure will be demonstrated.

## 2. THREE-TONE IMD MEASUREMENT

As described in Reference [7], two "probing" tones at frequencies $f_1$ and $f_2$ impinge on the resonant sample far outside the resonance passband, which is centered at $f_r$. A third tone at $f_3 \approx f_r$ is selected to be inside the resonance passband. Second order IMD occurs at, $f_3 + f_1$ among other frequencies. Third order IMD occurs at $f_3 + (f_2 - f_1)$, among other frequencies. By choosing the frequencies, $f_1$ and $f_2$, to be very low, both the second and the third order IMD can occur within the resonance passband.

Three probes are depicted, along with the HTS resonator, in Figure 1. The probing tones at $f_1$ and $f_2$ were introduced by a magnetic dipole probe that was brought in through Port 1. The driving tone at $f_3$ was introduced by a magnetic dipole probe brought in through Port 3. A third probe which can be either of magnetic dipole or electric monopole configuration is brought in through Port 2 and detects the IMD as well as the driving tone. The power levels for $f_1$ and $f_2$ were kept fixed at +10 dBm. The power level at $f_3$ was swept between −50 dBm and +10 dBm. Depending on temperature, the HTS structure was resonant at about 912 MHz with a low power unloaded $Q$ at 102.0 Kelvin of about 500. The $Tl_2Ba_2CaCu_2O_8$ superconducting resonator exhibited a critical temperature at about 103.0 Kelvin.

## 3. RESULTS

Figure 1 highlights several locations within the HTS device from Reference [7]. Position $e$ is near the electrical center, and when the resonator is excited in its 912 MHz resonance, a current maximum occurs at position $e$. Position $f$ is near the electrical end of the device, where the resonance mode has no current. Position c is in the high current region and the driving tone is introduced there with a magnetic dipole probe. The magnetic dipole probe from Port 1, which introduces the probing tones, was used at both Positions $e$ and $f$.

In the realization of this measurement described here, the 2nd order IMD was measured with $f_1 = 500$ kHz. For the 3rd order IMD measurement $f_1 = 400$ MHz and $f_2 = 400.2$ MHz. The effect of the probing frequency remains a subject of investigation. For both 2nd and 3rd order IMD, $f_3$ was positioned just below the center of the 912 MHz resonance, but well within the 3 dB resonance passband. With 3 dB resonant peak widths in the range of 0.5 MHz to 2.0 MHz, these frequency selections permitted both $f_3$ and the IMD frequencies to be on resonance.

The peak current in the resonator at the driving frequency, $f_3$, is calculated using IE3D from Zeland Software (Fremont, CA, USA) following the procedure published in Reference [7]. The probing frequencies, $f_1$ and $f_2$, are off resonance and thus do not propagate throughout the structure. The hypotheses advanced in Reference [7] and tested in this paper are (1) that the IMD is generated locally at the position of the probe at Port 3 which brings in $f_1$ and $f_2$, and (2) that the local point of IMD generation then serves as a signal source at the IMD frequency which itself excites the resonance of the structure.

The 2nd and 3rd order IMD signals were measured at 102.0 Kelvin. The Port 3 probe was first located at position $e$. The coupling of this probe at resonance was $\Delta S_{33} = -0.37$ dB, indicating that

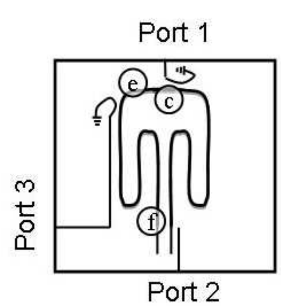



Figure 1: The folded hairpin HTS resonator. The probing signals were introduced at Points $e$ and $f$.

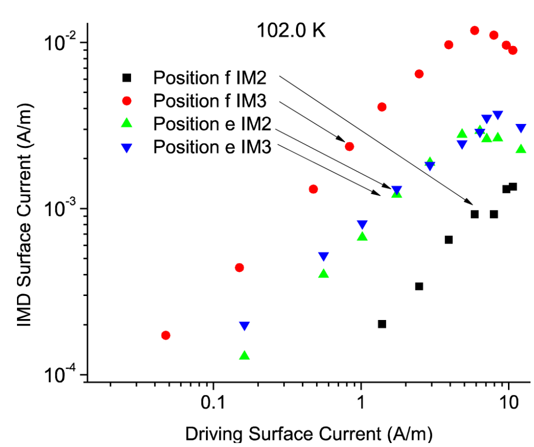



Figure 2: 2nd and 3rd order intermodulation distortion measured at Points $e$ and $f$ just below $T_C$ as a function of peak surface current.

even at resonance, the probe only induces a small amount of current on the structure. The probe was then moved to position $f$, which is very close to the electrical end of the resonator and where $\Delta S_{33} = -0.15\,\text{dB}$ on resonance. If the IMD is not generated locally, then it would be reasonable to expect less IMD when the probe is located at $f$ than at $e$, due to the probe's diminished capacity to excite current, in-band or out-of-band, at this location. However, if the IMD is generated locally, then it would be reasonable to expect higher IMD from this location, because a local IMD current originating at the electrical end corresponds to an even larger excitation of resonance by the IMD.

Figure 2 shows the measured IMD for the two probe locations. The 3rd order IMD current is about six times larger when the probe is at Point $f$ than at Point $e$. However, the 2nd order IMD current is about three times smaller. The dual implication of this observation is that (1) the IMD is generated at the point of the Port 3 probe and the resonance of the patterned HTS structure is then excited; and (2) the description of the electrodynamics is different at Point $e$, where there is higher current at the driving frequency, $f_3$, than at Point $f$, where there is very little current at $f_3$. The even order nonlinearity is much weaker than the odd order nonlinearity at Point $f$ than it is at Point $e$, indicating a lower degree of time reversal symmetry breaking at Point $f$ where there is less driving current. Whereas much of the odd order nonlinearity is expected to derive from intrinsic mechanisms, such as the nonlinear Meissner effect, even order nonlinearity is expected to derive from time reversal symmetry breaking extrinsic losses [9].

## 4. CONCLUSION

Three tone intermodulation distortion allows the synchronous measurement of even and odd order nonlinearity. When performed on a resonant structure, the currents associated with the distortion products can be computed. By using a nonresonant probe, the out-of-resonance IMD currents are generated locally at the point of the probe and then function as an excitation source of the resonance. This permits the observation of different origins of nonlinearity in different locations within the sample under test.

## ACKNOWLEDGMENT

This work was supported by a Cottrell College Science Award from the Research Corporation for Science Advancement.

## REFERENCES

1. Gayard, J.-D., "Technology disruptions in future communication payloads," *1st Symposium on Potentially Disruptive Technologies and Their Impact in Space Programs*, Marseille, France, July 4–6, 2005.
2. Remillard, S. K., A. Abdelmonem, D. S. Applegate, P. R. Radzikowski, and N. Lazzaro, "Field deployable microwave filters made from $YBa_2Cu_3O_{7-d}$ thick films," *J. Supercond. and Novel Magnetism*, Vol. 14, 47–56, 2001.
3. Remillard, S. K., L. J. Klemptner, J. D. Hodge, T. A. Freeman, P. A. Smith, and T. W. Button, "Generation of intermodulation products by granular $YBa_2Cu_3O_{7-x}$ thick films," *Proceedings of the SPIE Conf. High-temp. Microwave Superconductors and Applications*, 2559, San Diego, CA, USA, July 9–14, 1995.
4. Oates, D. E., "Overview of nonlinearity in HTS: What we have learned and prospects for improvement," *J. Supercond. and Novel Magnetism*, Vol. 20, No. 1, 3–12, 2007.
5. Mateu, J., J. C. Booth, C. Collado, and J. O'Callaghan, "Intermodulation distortion in coupled-resonator filters with nonuniformly distributed nonlinear properties — Use in HTS IMD compensation," *IEEE Trans. Microwave Theory and Tech.*, Vol. 55, No. 4, 616–624, 2007.
6. Lee, S.-C., C. P. Vlahacos, B. J. Feenstra, A. Schwartz, D. E. Steinhauer, F. C. Wellstood, and S. M. Anlage, "Magnetic permeability imaging of metals with a scanning near-field microwave microscope," *Appl. Phys. Lett.*, Vol. 77, 4404, 2000.
7. Pease, E. K., B. J. Dober, and S. K. Remillard, "Synchronous measurement of even and odd order intermodulation distortion at the resonant frequency of a superconducting resonator," *Rev. Sci. Instr.*, Vol. 81, 024701, 2010.
8. Lee, S.-C. and S. M. Anlage, "Study of local nonlinear properties using a near-field microwave microscope," *IEEE Trans. Appl. Supercond.*, Vol. 13, No. 2, 3594–3597, 2003.
9. Lee, S.-C., M. Sullivan, G. R. Ruchti, S. M. Anlage, B. S. Palmer, B. Maiorov, and E. Osquiguil, "Doping-dependent nonlinear Meissner effect and spontaneous currents in high-$T_C$ superconductors," *Phys. Rev. B*, Vol. 71, 014507, 2005.

# YIG Thin Film Used to Fabricate a Coplanar Waveguide Circulator

**Bassel Abdel Samad**[1,2]

[1]Laboratoire DIOM (Dispositifs et Instrumentation en Optoélectronique et Micro-ondes)
TELECOM Saint-Etienne, École Associée de L'Institut TELECOM
Université de Saint-Etienne, F-42000, Saint-Etienne, France
[2]Université de Lyon, F-69000, Lyon, France

**Abstract**— The requirements of mobile communication devices are today invariably associated with the miniaturization of microwave components. The development of Yttrium iron garnet (YIG) is of great interest to miniaturize these devices. The present work deals with the integration of YIG ferrite sputtered films of 10 µm thick for coplanar circulators, working at frequencies near 10 GHz. A small circulator with a coplanar structure is designed and analysed using a three dimensional finite element method. Aimed at this objective, we work on the development of a miniature planar circulator/isolator. The circulator is designed with coplanar waveguides. The structure is analysed by using a three dimensional finite-element method. The circulator is then fabricated, and its properties in the microwave range are characterised using a network analyzer and a probing system. A circulation is obtained around 10 GHz, the measured insertion loss and isolation giving interesting perspectives for the device.

## 1. INTRODUCTION

For several years, caused by growing needs of telecommunications, studies in microwave electronic materials were strongly increased [1]. In the field of wireless systems, the circulator is an important part. Generally the fabricant uses ferrite materials. It is well known that the ferrite circulator is the key element in wireless telecommunication systems, this non reciprocal components is based on the gyro-resonance properties and field displacement. This paper takes place in the field of passive microwave components. Coming from the requirements of mobile communication devices, miniaturization of microwave components is needed. The development of planar ferrite circuits is in great demand by the industry, to transferred ferrite thin film for the fabrication of monolithic microwave integrated circuit. Our aim was to design and manufactured a miniaturized circulator with magnetic thin film of YIG operating between 8 and 12 GHz. Many researchers have studied the transmission characteristics of the circulator using stripline $Y$-junctions since Bosma's work [2] but there are a few articles about coplanar circulators [3, 4]. Coplanar circulators can be designed following some approximate rules which are well known but are specific to stripline circulator. As a result, an optimization of the initial design must be performed before obtaining a functional device.

In this paper, a circulator based on the design studies in [5, 6] is fabricated with only 10 µm ferrite film and its transmission characteristics are measured by using three GSG coplanar probes connected with a vector network analyzer. The main aimed in this paper is to obtain a value of the insertion loss near the theoretical value by using YIG thin film.

## 2. EXPERIMENTAL

To fabricate a circulator, a radiofrequency magnetron sputtering was used for YIG thin film deposition. The influence of the deposition condition on the magnetic properties of the YIG film was important especially the thickness of YIG (10 µm). Under such condition, the deposition rate was $0.7\,\mu\mathrm{m}\cdot\mathrm{h}^{-1}$.

The YIG film has the following properties: A dielectric constant $\varepsilon = 15.3$, a saturation magnetization $M_s = 140\,\mathrm{mT}$, a dielectric loss tangent $\tan\delta_\varepsilon = 10^{-3}$ and a ferromagnetic resonance (FMR) line width $\Delta H = 4\,\mathrm{KA/m}$. For a conductor lines made of gold, the conductivity is $\sigma = 41.10^6$ S/m.

The real and imaginary parts of the elements of the permeability tensor are shown in Figure 1 for a 140 mT saturation magnetization and 0.01 damping factor. Figure 1 shows a frequency sweep with a fixed applied field, also shows three different zones according to the frequency. The expression below and above the frequency depend of the circulation of the circulator, and at resonance zone depend to the isolator.

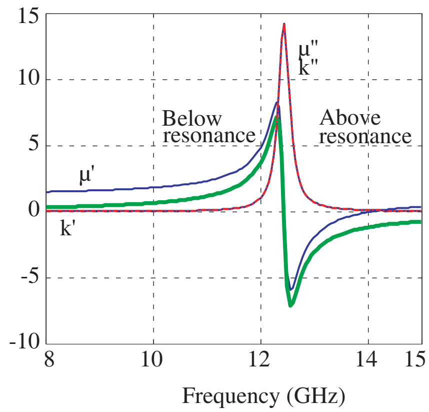



Figure 1: Ferromagnetic resonance: Real and imaginary parts of the elements of the permeability tensor, frequency sweep and 500 KA/m applied field with a small slot (G) between line and GND as shown in Figure 2.

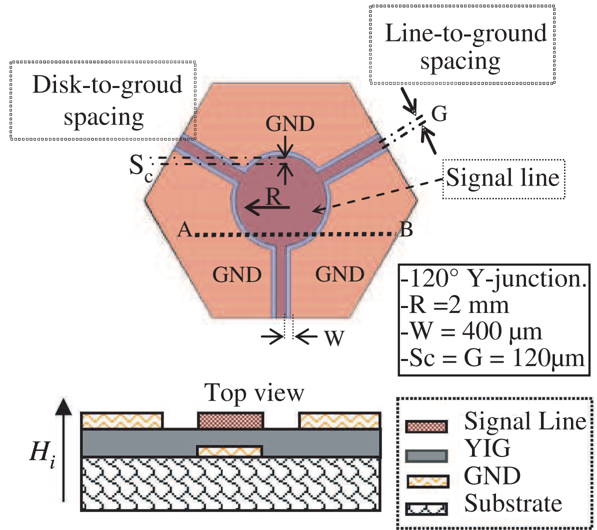



Figure 2: Structure of a circulator with CPW structure: Top view and A-B cross-sectional of circulator with YIG thin film.

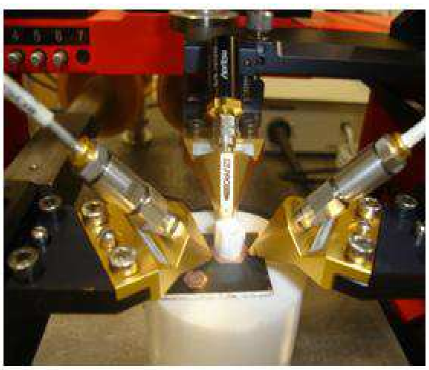


Figure 3: Circulator under test with a probe system connected two of them connected a 120° and the three ports were terminated by a register

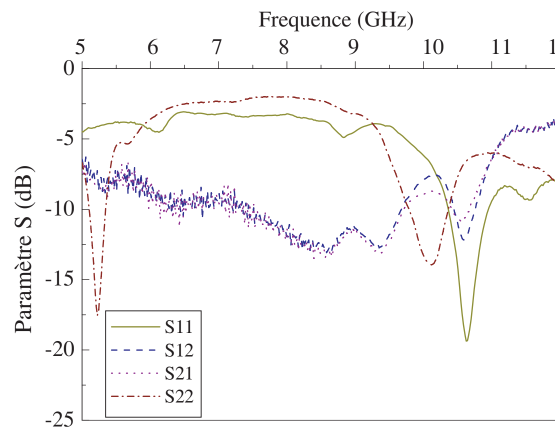



Figure 4: Experimental results of the coplanar circulator with 10 μm YIG film.

## 3. RESULTS AND DISCUSSION

The transmission characteristics of the circulator were measured by using the network vector analyser (VNA) 37397A (Anritsu). Three coplanar Picoprobe of GSG type, two probes in 120° are connected to the VNA by the HF rigid cables V-Type (Figure 3) and the third was terminated by a register of 50 Ω.

Figure 4 shows the band of frequency characteristics of the $S$-parameters for the circulator were analysed. In this analysis, the physical parameters were used. The internal magnetic field $H_{in}$ was of 335 kA/m. The Line-to-GND spacing was strictly and carefully optimised to achieve impedance matching to 50 Ω.

The experimental results when the magnetic bias field of approximately 477 kA/m was applied on the device with YIG film of 10 μm thick, the frequency characteristics of the circulator are shown in Figure 4.

A non-reciprocal effect due to the thin layer of YIG to about 2 dB, isolation −9 dB, loss of insertion of −7.5 dB in the 10 GHz band. Comparing these results with those of the previous [3, 4] show that the presence of a non-reciprocal effect is reduced due to the small amount of ferrite.

## 4. CONCLUSION

YIG thin film of 10 μm was deposited by using RF sputtering system. The properties of thecirculator were studied. The observed result of this film for the non-reciprocal effect was weak.The decrease of the non-reciprocal effect is due to the weak thickness of the YIG film.

Two proposed solution to increase the influence of this film: The use of other design or change thedeposition parameters. The nonreciprocal operation of our coplanar circulator with magnetic thinfilm was worked experimentally at 10 GHz.

## REFERENCES

1. Bayard, B., D. Vincent, C. R. Simovoski, and G. Noyel, "Electromagnetic study of a ferrite coplanar isolator suitable for integration," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, No. 7, 1809–1814, 2003.
2. Bosma, H., "On the principle of stripline circulation," *Proc. Inst. Elec. Eng.*, Vol. 109, Part B, No. 21, 137–146, 1962.
3. Zahwe, O., B. Sauviac, J. P. Chatelon, A. S. Dehlinger, M. Leberre, and S. Perrot, "Different versions of circulators with coplanar wave guide and YIG thin film," *Proceedings of OHD 2007, 19ème Colloque International Optique Hertzienne et Diélectriques*, Valence, France, Sep. 5–8, 2007.
4. Zahwe, O., B. Sauviac, J. P. Chatelon, A. S. Dehlinger, S. Perrot, and M. Leberre, "Towards a miniaturized circulator with magnetic thin film," *Proceedings of EuMC 2007*, Munich, Germany, Oct. 8–12, 2007.
5. Bosma, H., "On stripline Y-circulation at UHF," *IEEE Trans.*, Vol. 12, 61–72, 1964.
6. Polder, D., "On the theory of electromagnetic resonance," *Phil. Mag.*, Vol. 40, 99, 1949.

# Study of High Frequency Input Interference for Buck Converter

**Mao Zhang**[1], **Weiping Zhang**[2], **and Zheng Zhang**[2]

[1]School of Computing, Engineering and Physical Sciences, University of Central Lancashire, UK
[2]North China University of Technology, Shijingshang District, Beijing, China

**Abstract**— Applying the ideal analog signal sampling-recovery model, the high frequency input noise of Buck converter has been analyzed and a new type of noise, called as beat frequency noise, has been first revealed in this paper. The main contributions are as the followings: (1) Some high frequency noise can be shifted to the lower frequency — beat frequency, which could not filter out by original converter; (2) A more accurate model has been proposed to predict the output beat frequency noise; (3) A novel method to reduce the beat frequency noise has been put forward by adding a new pre-filter.

## 1. IDEAL MODEL FOR ANALYZING HIGH FREQUENCY INPUT INTERFERENCE

In the switching power DC/DC converter, there is very strong electromagnetic interferencewhich will affect the performances of DC/DC converter; Though it is thought the LPF had enough attenuation for this kind noise, this may be a mistake concept. In this paper, ideal analog signal sampling-recovery model [1] has been employed to reveal the effect of a high frequency input noise. Through analyzing the noise transformation in DC/DC converter, the beat frequency noise has been discovered.

A Buck converter, showed in Fig. 1, will be employed as example. $V_g$ is a DC input voltage source; $V_{gN}$ expresses the high frequency interference. A Buck converter can be separated into three parts, such as switch network, LPF, and the load.

An ideal analog signal sampling-recovery model has been shown in Fig. 2. In Fig. 2, a multiplier has been applied to replaces the switch network and it is sampler to convert an analog signal to a discrete signal. $v_g(t)$ is the DC input signal with a high frequency noise, $\delta_T(t)$, unity impulse train, is used to take place the drive signal. The sampling and recovery process has been illustrated in Fig. 3. The input voltage is (1)

Where, $V_g$ is DC component with a noise of $V_{gN}\cos\omega_N t$ which is a high frequency noise with frequency range $0.5\omega_s < \omega_N < \omega_s$. The Fourier transform of the input voltage:

$$v_g(t) = V_g + V_{gN}\cos\omega_N t \tag{1}$$
$$V_g(\omega) = V_{gdc}(\omega) + V_{gN}(\omega) = 2\pi V_g\delta(\omega) + \pi V_{gN}[\delta(\omega-\omega_N)+\delta(\omega+\omega_N)] \tag{2}$$

In order to simplify the analyzing process and reveal clearly the effect of a high frequency input noise, suppose that the driven signal $v_{gs}(t)$ is a sampling signal $\delta_T(t)$ of unity impulse train with frequency $\omega_s$

$$\delta_T(t) = \sum_{n=-\infty}^{\infty}\delta(t-nT_s) \quad T_s \text{ is the period of the switch} \tag{3}$$

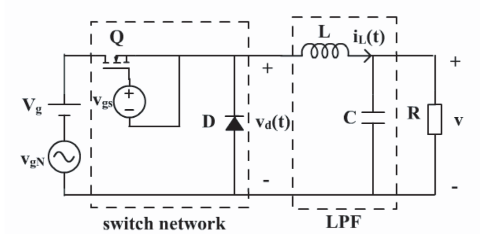



Figure 1: BUCK converter.

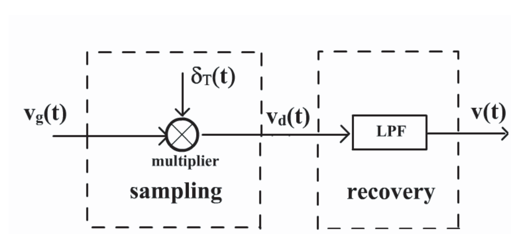



Figure 2: Ideal analog signal sampling-recovery model.

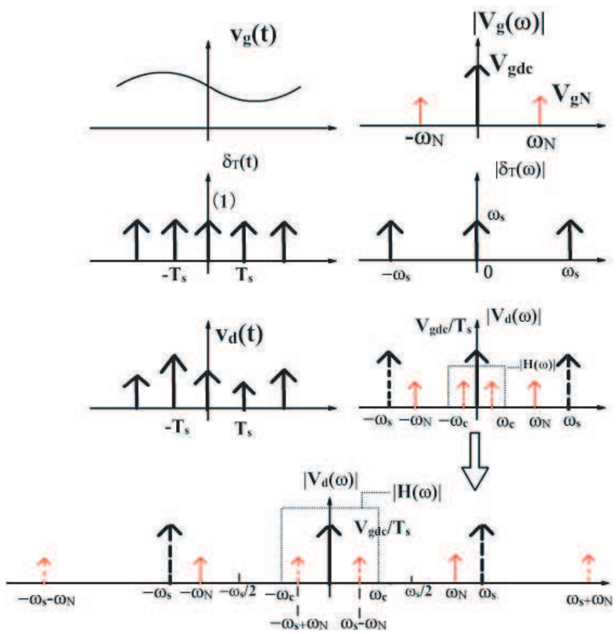



Figure 3: The process from analog signal to discrete signal.

The complex exponential form of Fourier series of unity impulse train is

$$\begin{aligned}
\delta_T(t) &= \sum_{k=-\infty}^{\infty} C_k e^{jk\omega_s t}, \\
C_k &= \frac{1}{T_s}\int_{-\infty}^{+\infty} \delta(t) e^{-jk\omega_s t} dt = \frac{1}{T_s}, \\
F[\delta_T(t)] &= \sum_{k=-\infty}^{\infty} \frac{2\pi\delta(\omega - k\omega_s)}{T_s} = \sum_{k=-\infty}^{\infty} \omega_s \delta(\omega - k\omega_s)
\end{aligned} \tag{4}$$

Applying the convolution property [1], one can obtain the frequency response of the output of the switch network,

$$V_d(\omega) = V_g(\omega) * F[\delta(t)]/2\pi = f_s \sum_{k=-\infty}^{\infty} V_g(\omega - k\omega_s) \tag{5}$$

The frequency response $V_d(\omega)$ and $V_g(\omega)$ have been shown in Fig. 3. In this Fig. 3, $V_d(\omega)$ are one sets of shifting $V_g(\omega)$ in frequency domain. The shifting rule is as the followings: (1) The shifting interval is Buck converter's switch frequency $\omega_s$, (2) The magnitude is modified by the factor $1/T_s$. (3) In the DC/DC converter, the spectrum of the high frequency noise will be moved to $|\omega_s - \omega_N|$ in lower frequency region.In this case, if $|\omega_s - \omega_N| < \omega_c$ and especially $\omega_s \approx \omega_N$, the output LPF has no way to reduce this kind noise. Therefore, in this paper the noises $V_{gN}(\omega_s - \omega_N)$ and $V_{gN}(-\omega_s + \omega_N)$ are defined as a beat frequency noise.

If the LPF is ideal filter with cutoff frequency $\omega_c$ and $\omega_c = (0.01 \sim 0.1)\omega_s$, the frequency response is specified by

$$H(\omega) = 1 \quad |\omega| < \omega_c \qquad \text{Others is equal to 0} \tag{6}$$

The output voltage and the transform function of beat frequency noise can be obtained as

$$\begin{aligned}
V(\omega) &= 2\pi f_s \left\{ V_g \delta(\omega) + V_{gN}[\delta(\omega_s - \omega_N) + \delta(-\omega_s + \omega_N)]/2 \right\} H(\omega) && |\omega| < \omega_s/2 \\
\hat{v}(\omega)/\hat{v}_{\scriptscriptstyle N}(\omega)|_{\omega=\omega_s-\omega_N} &= H(\omega_s - \omega_N)/T_s, && |\omega_s - \omega_N| < \omega_c
\end{aligned} \tag{7}$$

## 2. A MORE ACCURATE MODEL

In Fig. 2, if PAM (pulse-amplitude-modulation) signals takes place the impulse train. If the PAM-recovery model is used, a more accurate solution can be got. The Fourier transform of the input beat frequency noise is

$$\hat{v}_{gN}(\omega) = \pi V_{gN} \left\{ \delta[\omega - (\omega_s - \omega_N)] + \delta[\omega - (-\omega_s + \omega_N)] \right\} \tag{8}$$

The output voltage of the beat frequency noise:

$$\hat{v} = C_1 \pi V_{gN} H(\omega) \left\{ \delta[\omega - (\omega_s - \omega_N)] + \delta[\omega - (-\omega_s + \omega_N)] \right\}, \quad \text{where } C_1 = D \sin c(\omega_s \tau/2) \tag{9}$$

The transform function of input beat frequency noise to output is

$$\hat{v}(\omega)/\hat{v}_N(\omega)\big|_{\omega=\omega_s-\omega_N} = C_1 H(\omega_s - \omega_N), \quad \text{where } C_1 = D \sin c(\omega_s \tau/2), \quad |\omega_s - \omega_N| < \omega_c \tag{10}$$

## 3. USE OF PRE-FILTER TO REDUCE THE BEAT FREQUENCY NOISE

The high frequency noise is the high frequency ripple of the previous power stage, the frequency is $\omega_N$. The beat frequency noise can be reduced by LPF of Buck converter, if $|\omega_s - \omega_N| > \omega_c$. But this condition required the lower switching frequency of the previous stage, which would result in some troubles for designer, such as big volume inductor and capacitor as well as transformer. Therefore, a reasonable suggestion is that the switching frequency of post stage should be greater than that of the previous stage and satisfy the following condition,

$$\omega_s \gg (\omega_N + \omega_c) \tag{11}$$

If the maximum frequency of input noise is greater than half switching frequency, the beat frequency noise must occur. Therefore, the switching frequency is selected high enough so that all beat frequency noise could be moved out the pass band of LPF and doesn't act like lower-frequency input noise. However, the efficiency of DC/DC converter will go down rapidly as the switching frequency is increased. A reasonable way is to add a pre-filter between input DC source $V_g$ and the input port of Buck converter so that the beat frequency noise should be eliminated before it go through the switch network, shown in Fig. 4 and these spectra has shown in Fig. 5.

The pre-filter should be a LPF or BPF so that the high frequency noise can be eliminated and the DC component is remained while the input signal goes through it.If the frequency response of the pre-filer is $H_{pf}(\omega)$ with cutoff frequency $\omega_{cpf}$ and the Fourier transform of pre-filter output is $V_{g1}(\omega)$, then

$$V_{g1}(\omega) = H_{pf}(\omega) V_g(\omega) \tag{12}$$

Applying sampling theorem [1], the aliasing pheromone will not occur, and the original high frequency input noise will be eliminated by original LPF in Buck converter, if the cutoff frequency $\omega_{cpf}$ is smaller than $\omega_s/2$, shown in Fig. 5.

## 4. FUNDAMENTAL OF DESIGNING THE PRE-FILTER

In this case, DC voltage source and DC/DC converter are located in separate PCB board. So, a long wire is used to connect the DC source with input port of DC/DC converter. This long wire has a parasitic inductance and capacitance, shown in Fig. 6. These parasitic component will has the following functions: (1) This parasitic inductor will resonate with the parasitic capacitor to produce high frequency ring oscillation. (2) The parasitic inductor and capacitor form a inherent second order low pass filter. A long wire is used to connect the output port of PFC circuit with input port of DC/DC converter or a Rectifier output.

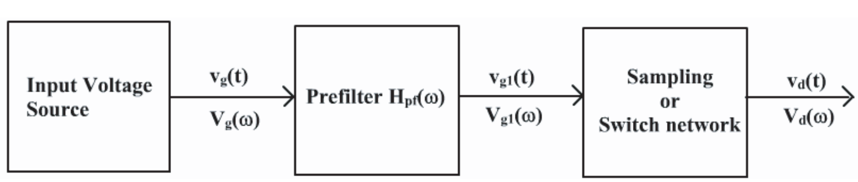



Figure 4: Use of pre-filter to reduce beat frequency noise.

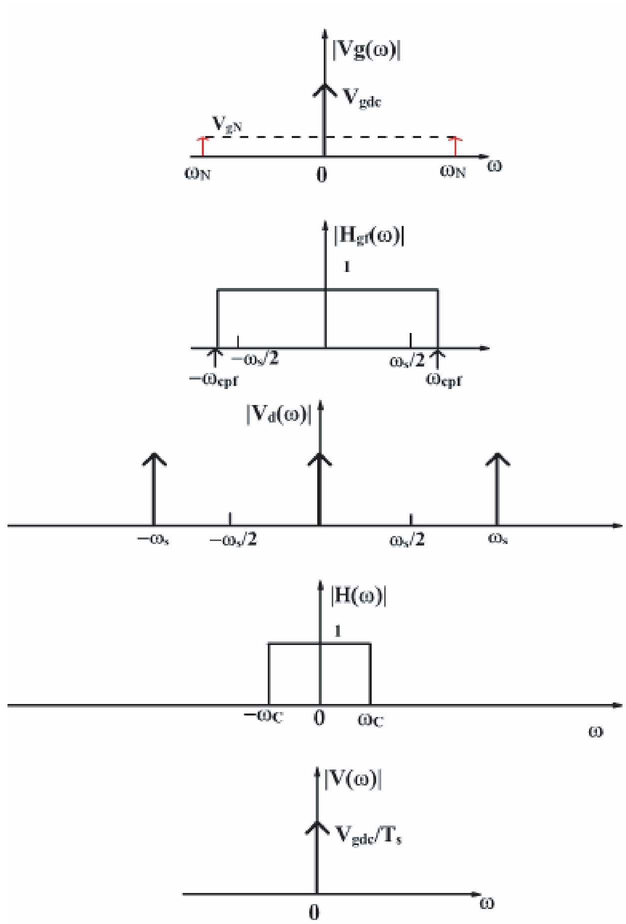



Figure 5: The frequency response by adding a pre-filter.

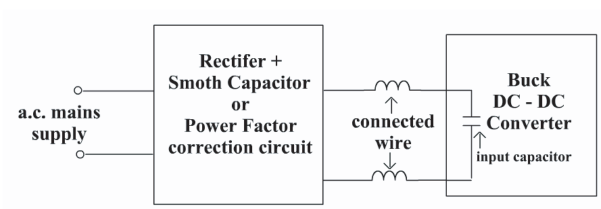



Figure 6: A wire to connect DC source with the input port of converter.

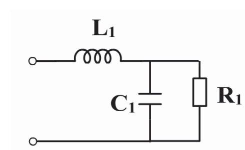



Figure 7: Second-order pre-LPF.

An input capacitor is needed in the input port of converter, which is constructed an equivalent input second-order $L.C$ filter as the pre-LPF, shown in Fig. 7. The load $R_1$ is the equivalent input resistance of Buck converter.

If $\omega_{N\,\min}$, $\omega_{N\,\max}$ are the smallest and highest frequencies of the input noises, $\omega_{cpf}$ is the cutoff frequency of the pre-filter and $\omega_{cpf} \ll \omega_{N\,\min}$, then the pre-filter can filter out the high frequency noise. Howeverbecause $\omega_{N\,\min}$ is a small frequency, for example, $\omega_{N\,\min} = 2 \times (2\pi \cdot 50) = 200\pi$, a big inductor and a capacitor are required to form the pre-filter, so it is difficult to be implemented.

In Fig. 8, the curve 1 and 2 denote the frequency response of the ideal LPF with a cut-off frequency $\omega_c$ and input noise spectrum with band $\omega_{N\,\min} \ll \omega_{N\,\max}$. The curve 3 is an upper equivalent noise spectrum which is the beat noise spectrum. The curve 4 is the upper equivalent LPF frequency response which is not really exist in practice and is sketched to explaining the bandwidth of output beat frequency noise spectrum using dash-dotted curve. The curve 5 is the output beat frequency noise frequency spectrum. In Figs. 8(a) and (b), there are three shadow parts, denoted as A and B as well as $C$. These three parts have equal frequency bandwidth,The curve 6 represses the ideal pre-filter frequency response with cut-off frequency $\omega_{cpf}$. The following

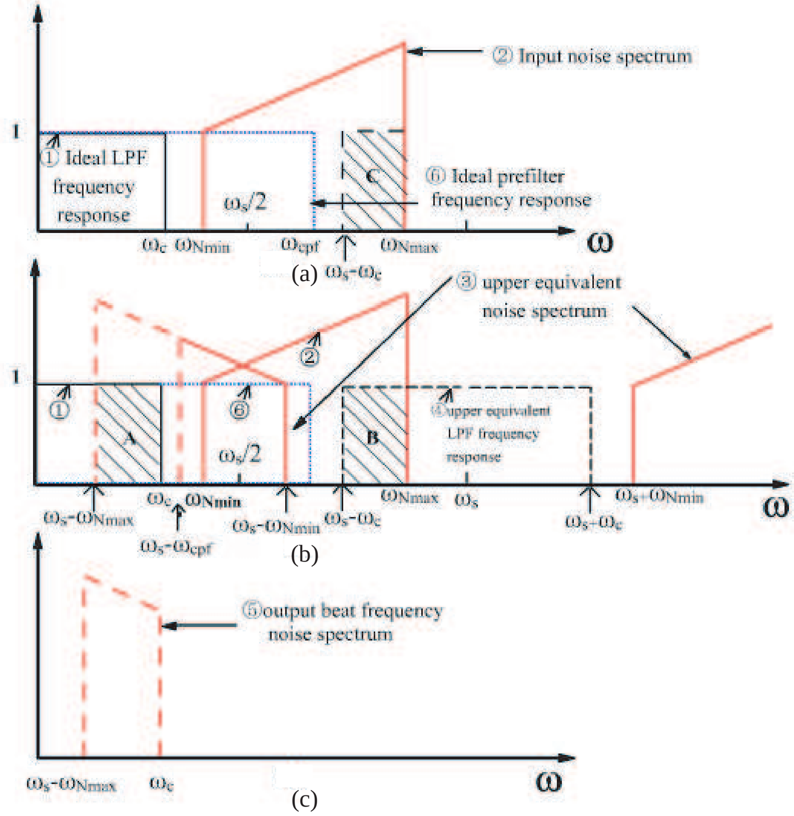



Figure 8: Selecting cutoff frequency $\omega_{cpf}$ of pre-filter.

conditions should be necessary,

$$\Delta\omega = \omega_{N\max} - (\omega_s - \omega_c) = \omega_c - (\omega_s - \omega_{N\max}) \tag{13}$$

$$0.5\omega_s < \omega_{cpf} < \omega_s - \omega_c \quad \text{and} \quad \omega_c < 0.5\omega_s \tag{14}$$

In Fig. 8, there are three shadows, which have same frequency region and plays important role on designing pre-filter according known $\omega_c$, $\omega_{N\max}$, $\omega_s$, which has given out the interrelation relationship of these frequencies.

A second-order Butterworth low-pass-filter has bee chosen to implement the pre-filter which has a magnitude frequency response described by following Equation (15). The magnitude frequency response of the pre-filer circuit shown in Fig. 7

$$|H(\omega)| = \omega_c^2/\sqrt{\omega^4 + \omega_c^4} \tag{15}$$

$$|H(\omega)| = 1/L_1C_1\sqrt{\omega^4 + (1/L_1C_1)^2 + \omega^2[(1/R_1C_1)^2 - 2/L_1C_1]} \tag{16}$$

By comparing Equations (15) and (16), the pre-filter will match the Butterworth filter if we choose

$$\omega_c = \omega_{cpf} = 1/\sqrt{L_1C_1} \tag{17}$$

$$L_1 = 2R_1^2C_1 \tag{18}$$

The design procedure for pre-filter is described as the followings: Step 1 to select the cut off frequency $\omega_{cpf}$. Step 2 calculate the input resistance $R_1$ of Buck converter [2, 3]. Step 3 solve the set of Equations (17) and (18) to get the values of $L_1$ and $C_1$. Step 4 check the magnitude frequency response by using Matlab program to calculate the Equation (15) and plot the magnitude frequency response.

## 5. CONCLUSION

In this paper, the ideal analog signal sampling-recovery model is used to analyze the input high frequency noise of Buck converter (DC/DC) and some following valuable conclusions have been arrived:

(1) If $|\omega_s - \omega_N| < \omega_c$, where $\omega_s$, $\omega_N$, $\omega_c$ repress switching frequency and noise signal frequency as well as the cutoff frequency of LPF respectively, there exist beat frequency noises which can not be eliminated by the original LPF in Buck converter. (2) The beat frequency noise can be accurately predicted by the formula derived in this paper. (3) A best way of reducing the beat frequency noise is to add a low pass or band pass pre-filter. (4) Butterworth filter has been chosen to implement the ideal low pass pre-filter and the relationship of the noise frequency, switch frequency, cut-off frequency of the original LPF as well as the cut off frequency of pre-filter.

By the way, the analyzing methods and conclusions can be easily extended to investigate the other kind of converter although we only apply Buck converter as an example.

**REFERENCES**

1. Phillips, C. L. and J. M. Parr, *Signals Systems and Transforms*, 2nd Edition, 275–277, 1998.
2. Zhang, W., *The Control and Modeling of Switch Converter*, 5–15, China Power Publishers, 2006.
3. Erickson, R. W. and D. Masksimovic, *Fundamentals of Power Electronics*, 2nd Edition, Kluwer Academic Publishers, 2001.
4. Ludeman, L. C., *Fundamentals of Digital Signal Processing*, Happer & Row Publishers, Inc., New York, 1986.

# A Novel Approach for Changing Bandwidth of FSS Filter Using Gradual Circumferential Variation of Loaded Elements

**S. M. Choudhury**, **M. A. Zaman**, **M. Gaffar**, **and M. A. Matin**
Bangladesh University of Engineering and Technology, Dhaka, Bangladesh

**Abstract**— A novel approach for varying bandwidth of Frequency Selective Surface (FSS) microwave filters has been discussed. The filter studied comprised of four-legged-loaded elements FSS filter. The unit cell of a conventional loaded element FSS filter contains only a single loaded element. The novel approach requires the unit cell to contain multiple loaded elements. The circumferences of the loaded elements of the unit cell are altered utilizing a factor named *circumferential variability*, $d$. The simulation results show that the bandwidth can be appreciably varied by changing the value of $d$.

## 1. INTRODUCTION

The concept of Frequency Selective Surface is to provide variable opacity for different values of the wavelength of incident radiation. FSS structures are commonly used in making reflector antenna more efficient [1], making dual band [2] or multiband [3] parabolic reflector and high directivity EBG resonator antenna [4]. FSS are commonly composed using periodic arrays to produce a resonant structure [5]. FSS geometries that possess greater than a two-fold rotational symmetry have reflection (transmission) coefficients independent of polarization at normal angle of incidence [6]. So elements for FSS are selected to maintain rotational symmetry. The loop elements are some common structures that are rotationally-symmetric and used in FSS structures. The loop element category include: three and four-legged loaded element which was issued patent in 1974 to Munk [7], square loops, hexagonal loops, and circular loops. A common property of loop elements are that they all have a fundamental resonance frequency when their circumference is approximately equals to $\lambda$, and a second resonance occurs when the circumference is approximately $2\lambda$ [5]. Munk suggested the alteration of bandwidth of such filter by reducing the distance of the transmission line spacing [7]. Callaghan et al. [8] observed change in bandwidth of such FSS filters with the introduction of dielectric layers.

In this paper, a different approach for altering the bandwidth is explored. As mentioned earlier, the fundamental resonance wavelength is inversely proportional to the circumference of the filter element. The bandwidth of the filter is altered by constructing a unit cell with multiple loaded elements while gradually varying the circumference of the loaded elements within the unit cell.

## 2. DESIGN METHODS

The circumference of the loaded element can be determined for a particular center frequency [5]. By changing the circumference, the center frequency of the filter is changed. The novel unit cell is defined as a combination of multiple loaded elements. For continuity of pattern, $2 \times 2$ or $3 \times 3$ elements can be chosen. The circumference of these filters are changed keeping the mean center frequency same. Guo et al. [1] proposed a method of dual bandpass FSS filter using four-legged loaded elements of two different circumferences. In this paper, a design parameter Circumferential Variability, $d$ is taken to alter circumference of the elements. For the $2 \times 2$ filter, there are four loaded elements, and their circumference are scaled with factors: $1 + d$, $1 - d$, $1 + 3d$, $1 - 3d$ respectively. For the $3 \times 3$ filter, the circumferences are scaled with factors $1 - 3d$, $1 + 4d$, $1 - 2d$, $1 - d$, 1, $1 + d$, $1 + 2d$, $1 - 4d$ and $1 + 3d$. The distance between the centers of the loaded elements are kept constant and chosen in such a manner so that the value of $d$ does not cause two adjacent elements to overlap.

## 3. SIMULATION

Multiple layer of FSS filters were proposed by Romeu [9] to have multiband operations. In this paper, two identical layer of FSS filters is used to have same passband of both the layers. This increases the order of the filter. A cross-sectional view is show in Figure 2. For dielectric layers Duroid 5880 is used, and etched copper layers can be used to implement each layer of the FSS filter. To reduce the overall simulation time, the Babinet's principle [10] has been applied to simulate the structure. At first the effect of changing circumference of a four-legged loaded element

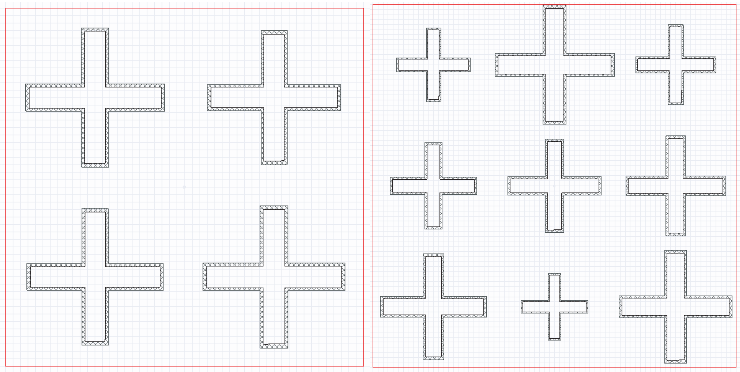


Figure 1: Unit cell of proposed $2 \times 2$ and $3 \times 3$ filter.

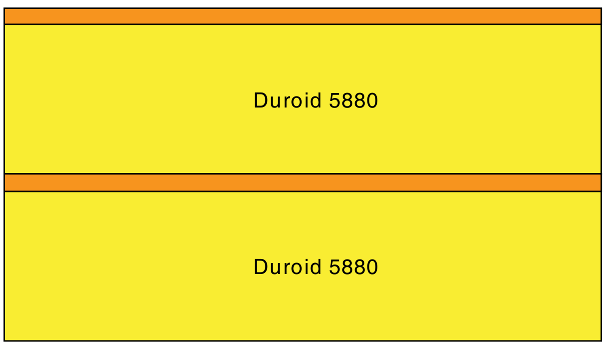



Figure 2: Layout of the filter structure.

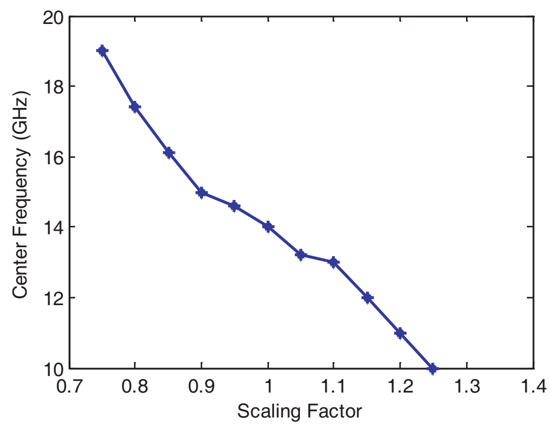



Figure 3: Center frequency vs. scaling factor of a four-legged loaded element FSS filter.

is simulated. For this simulation, the distance between the centers of two adjacent unit cells are kept constant, and all dimensions of each loaded element is scaled using a constant value. Afterwards, the circumferential variability is swept using parametric sweep and the effect of changing circumferential variability on bandwidth is studied.

## 4. RESULTS

In Figure 3, the Center Frequency vs Scaling factor curve of a single four-legged loaded element FSS filter has been shown. The center frequency gradually decreases with the scaling factor of the loaded element. This is concept which is utilized in increasing bandwidth of the filter.

For this purpose, parametric analysis of the proposed filter structures is performed.

For the $2 \times 2$ filter, the 5 dB bandwidth can be varied from 2.75 to 4.5 GHz, by changing the circumferential variance from 0.015 to 0.04. Further increasing circumferential variance causes the bandwidth to decrease. Cubic interpolating function fits the datapoints within the range. The interpolating function can be written as

$$BW = p_1 d^3 + p_2 d^2 + p_3 d + p_4$$

where the coefficients are $p_1 = -1.2136e + 005$, $p_2 = 11140$, $p_3 = -243.27$, $p_4 = 4.2228$. The Norm of residuals is 0.019317.

For the $3 \times 3$ filter, the 5 dB bandwidth vs. circumferential variance characteristics is irregular in shape. By varying the circumferential variance, the bandwidth can be varied from 1.886 to 3.705 GHz. So the $3 \times 3$ filter cannot give wider bandwidth compared to the $2 \times 2$ filter, but it can give narrower bandwidth compared to the original filter.

5 dB bandwidth vs. circumferential variance is shown on figure. The bandwidth can be appreciably varied in for the values of circumferential variance in the range 0.011 to 0.023. Two quadratic

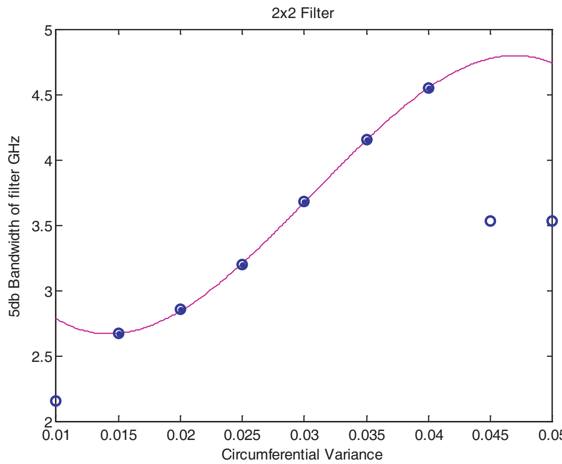



Figure 4: Bandwidth vs. circumferential variance of $2 \times 2$ filter.

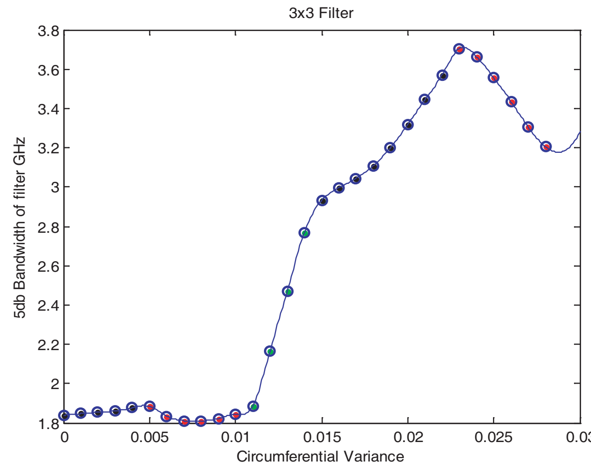



Figure 5: Bandwidth vs. circumferential variance of $3 \times 3$ filter.

functions can appreciably fit the data points, the function being

$$BW = p_1 d^2 + p_2 d + p_3$$

For the range of values of $d$ from 0.011 to 0.015, the coefficients are $p_1 = -16457$, $p_2 = 697.87$, $p_3 = -3.813$. Norm of residuals is 0.053674.

For the range of values of $d$ from 0.015 to 0.02, the coefficients are $p_1 = 7196$, $p_2 = -176.45$, $p_3 = 3.9647$. Norm of residuals is 0.025662.

## 5. CONCLUSIONS

In this paper, a design procedure for FSS filters to vary bandwidth is illustrated. This novel technique of creating filters with wider bandwidth gives designers the flexibility to choose a particular bandwidth, specific to the application of the microwave filter.

## ACKNOWLEDGMENT

The authors would like to express their sincere gratitude towards Bangladesh University of Engineering and Technology (BUET), for supporting the research work of the paper.

## REFERENCES

1. Guo, C., H. Sun, and X. Lu, "A novel dual dualband frequency selective surface with periodic cell perturbation," *Progress In Electromagnetics Research B*, Vol. 9, 137–149, 2008
2. Mittra, R., C. H. Chan, and T. Cwik, "Techniques for analyzing frequency selective surfaces-a review," *Proceedings of the IEEE*, 1988.
3. Wu, T. K., "Four-band frequency selective surface with double-square-lop patch elements," *IEEE Trans. Antennas Propagat.*, Vol. 42, No. 12, 1994.
4. Lee, Y. J., J. Yeo, R. Mittra, and W. S. Park, "Design of a high-directivity electromagnetic band gap (EBG) resonator antenna using a frequency-selective surface (FSS) superstrate," *Microwave and Optical Technology Letters*, Vol. 43, No. 6, 2004.
5. Volakis, J. L., *Antenna Engineering Handbook*, 4th Edition, McGraw-Hill, 2007.
6. Mackay, A., "Proof of polarization independence and non-existence of crosspolar terms for targets presenting $n$-fold ($n > 2$) rotational symmetry with special reference to frequency selective surfaces," *Electronics Letters*, Vol. 25, No. 24, 1624–1625, 1989.
7. Munk, B. A., *Frequency Selective Surfaces*, Wiley, New York, 2000.
8. Callaghan, P., E. A. Parker, and R. J. Langley, "Influence of supporting dielectric layers on the transmission properties of frequency selective surfaces," *IEE Proceedings*, Vol. 138, No. 5, 448–454, 1991.
9. Romeu, J., "Fractal FSS: A novel dual-band frequency selective surface," *IEEE Trans. Antennas Propagat.*, Vol. 48, No. 7, 2000.
10. Kraus D. K. and R. J. Marhefka, *Antennas for All Applications*, McGraw-Hill, 2003.

# A Compact Substrate Integrated Waveguide Band-pass Filter

**Changjun Liu and Kama Huang**
School of Electronics and Information Engineering, Sichuan University, Chengdu 610064, China

**Abstract**— A substrate integrated waveguide (SIW) filter with compact size compared to traditional SIW filters are presented in this work. A band-pass filter for general purpose is designed at 5.8 GHz with relative bandwidths of 5%. The fabricated filter is on F4B-2 substrate and about $0.5\lambda_g$ by $1\lambda_g$, which is about 50% of a conventional SIW dual-mode band-pass filter. Simulation and measurements agree well over a frequency band of 3–10 GHz. The empirical design formulae are presented as well.

## 1. INTRODUCTION

Substrate integrated waveguides (SIWs) have many advantages over microstrip lines, such as high Q-factor, small size, light weight, and easy to integrate. Since SIWs have been proposed in 1998 [1], SIWs have found many applications in various microwave circuits [2]. A lot of conventional passive microwave components and devices, such as antennas [3], directional couplers [4], diplexers [5], and phase shifters [6] have been successfully realized with SIWs.

SIW filters have also been investigated by many researchers. Some of these filters [7–12] make use of quasi-waveguide cavity resonant modes formed by metallic vias. SIWs are integrated to reduce its size, weight, and cost. Meanwhile, they greatly enhance manufacturing repeatability and reliability [12].

Most of them have very high center frequencies [9, 12] e.g., 20 GHz or higher. Even when the resonance is of high-order modes, the overall size of a SIW filter keeps compact enough to work with other components and exhibits good performances. However, a SIW filter is not acceptable, when the center frequency goes down to several GHz and, even, in low-order modes, such as $TE_{201}$ and $TE_{102}$. It limits the applications of SIW filters at low frequency part. But the SIW is a born high-pass filter due to its waveguide characteristics. If a stop band higher than its cut-off frequency is introduced, a band-pass filter is presented.

This work is focused on a compact SIW band-pass filter. We present a method to design and realize compact SIW band-pass filter, in which a hole on the top metallic surface of the SIW filter is applied for mode regulation. A SIW band-pass filter at 5.8 GHz with a fractional bandwidth of 5% is demonstrated. The measured results of the fabricated SIW filter agree well to simulations. The size of the proposed SIW filter is roughly about half of a traditional dual-mode SIW filter, e.g., with $TE_{201}$ and $TE_{102}$ modes.

## 2. PRINCIPLES

A SIW is a type of dielectric-filled waveguide which is synthesized in a planar substrate with linear arrays of metallic vias. Those vias are used to realize metal edge walls. Its basic structure is similar to a microstrip structure, which is easy to fabricate. It works in TE or TM modes, instead of the quasi-TEM mode of a microstrip line. Thus, some components are usually required at both input and output of a SIW band-pass filter. Planar structures, such as microstrip lines and coplanar waveguides (CPWs), may be integrated to SIWs as various transitions. An entire system based on SIWs is compact, and convenient to connect to other parts.

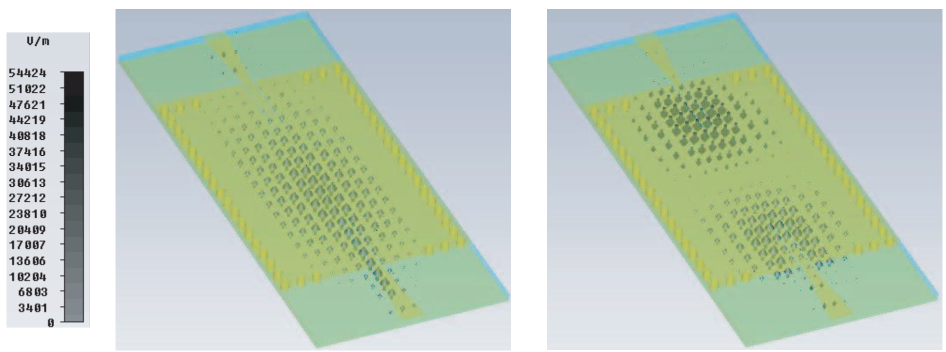
A section of SIW at given boundary conditions is equivalent to a cavity resonator. The resonant frequencies are

$$f_{m0n} = \frac{c}{2\sqrt{\varepsilon_{eff}}}\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2} \tag{1}$$

where $\varepsilon_{eff}$ is the effective permittivity, $m$ and $n$ are non-negative integers, and $a$ and $b$ are effective length along $x$ and $z$ directions, $c$ is the speed of light in free space, respectively. The resonant frequencies are corresponding to different electromagnetic field distributions and modes.

In many dual mode filter design, sizes in $x$- and $y$-direction are almost the same, so the orders of resonance modes are the same [11]. As for SIW, resonance modes of $f_{101}$ and $f_{102}$ are shown in Fig. 1. It's the first time that we have introduced a circular hole etched on the top metal layer of the SIW filter to perturb the electromagnetic field distributions of these two modes and form a filter as shown in Fig. 3. With the increment of $R_c$, resonance frequencies of $f_{101}$ and $f_{102}$ get close to each other gradually. When coupling those two modes, the cavity is resonant at dual mode. Moreover, the introduced circular hole reduces the size of the resonator by enhance the path length of currents.

A valid method is to use a section of Conductor Backed Coplanar Waveguide (CBCPW) as a transition to connect a SIW to a microstrip line. The gaps of the CBCPW transition work as chokes at a given frequency, when the gap length is about quarterwavelength, as shown in Fig. 2. It is clear that there is a short circuit formed at the end of the CBCPW from the simulated current distribution and, then, leads to a strong reflection from the choke. The CBCPW transition may used as a band-stop structure in the SIW filter with the introduced extra transmission zero, of which the frequency is dependent on the length of the CBCPW transition. It is a compact design without occupy extra space. Thus, CBCPW transitions are introduced in the proposed SIW band-pass filter design to improve its performance.




(a) $f_{101}$ (4.95 GHz) (b) $f_{102}$ (6.28 GHz)

Figure 1: Simulated electric field distributions of resonance modes.

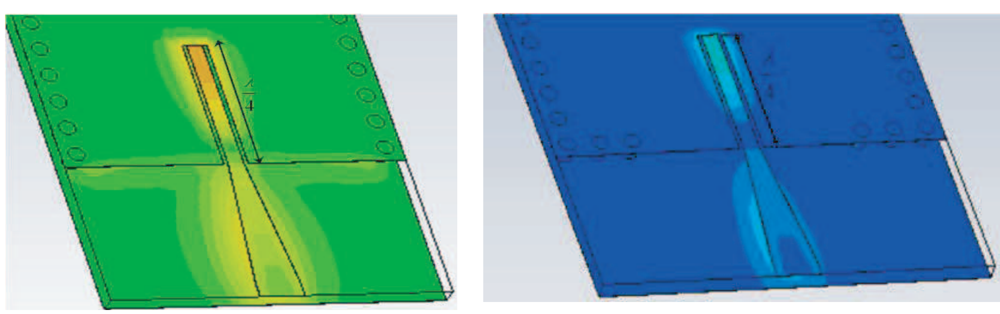


Figure 2: Current distribution of a CBCPW with a choke structure.

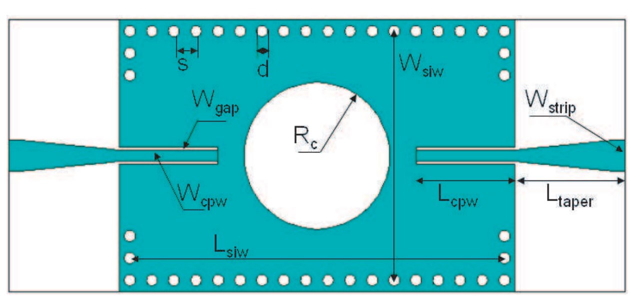



Figure 3: General filter.

## 3. SIW BAND-PASS FILTER

The SIW band-pass filter works in dual-mode with a circular hole etched on top metal layer and two CBCPW sections, as shown in Fig. 3. We first design the SIW filter prototype with roughly estimated parameters at the center frequency $f_0$. Then, we use CST to simulate and optimize the filter prototype. Finally, we give the following empirical formulae to help the design of the proposed SIW band-pass filter whose relative bandwidth of about 5% (The thickness of the substrate is 1 mm.):

$$W_{\mathrm{SIW}} = \frac{0.58c}{f_0\sqrt{\varepsilon_r}} \tag{2}$$

$$L_{\mathrm{SIW}} = 1.895 * W_{SIW} \pm 2d \tag{3}$$

$$L_{\mathrm{CPW}} = 0.46 \cdot W_{SIW} \tag{4}$$

$$R_c = 0.26 \cdot W_{SIW} \tag{5}$$

where $c$ is the light speed in free space, $\varepsilon_r$ is the relative permittivity of the substrate, $d$ is the diameter of vars. As shown in Fig. 3, metallic vias are added at corners of the SIW to reduce its insertion loss, achieve a sharper skirt and a wider stop-band. Since these vias reduce the equivalent width of the SIW, $W_{\mathrm{SIW}}$ should increase to maintain the resonance frequency $f_0$ unchanged. Parameter values of the filter are given in Table 1. Larger $R_c$ helps to get narrower pass band and steeper skirt, but leads to higher insertion loss.

Table 1: Parameters of the filter with matched port configuration, Unit: mm.

| $s$ | $d$ | $W_{\mathrm{gap}}$ | $W_{\mathrm{CPW}}$ | $L_{\mathrm{taper}}$ | $L_{\mathrm{CPW}}$ | $R_c$ | $W_{\mathrm{SIW}}$ | $L_{\mathrm{SIW}}$ | $W_{\mathrm{strip}}$ |
|---|---|---|---|---|---|---|---|---|---|
| 2.0 | 1.0 | 0.3 | 0.5 | 10.0 | 9.0 | 6.8 | 22.2 | 34.0 | 2.8 |

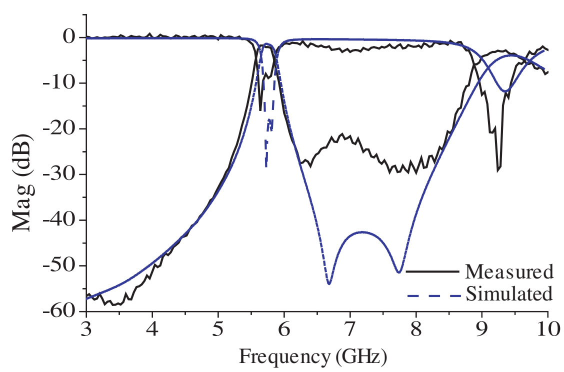



Figure 4: Simulated and measured results of the general filter.

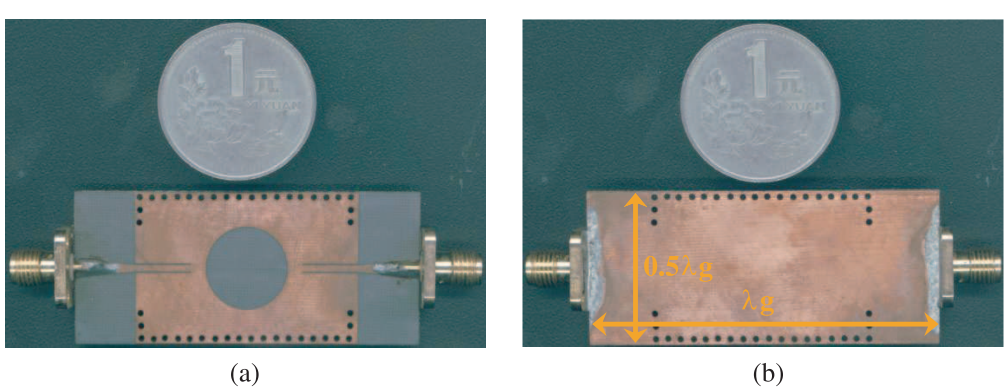



Figure 5: Front and back view of the general filter. (a) Front view, (b) back view.

Responses of full-wave simulation and measurement are shown in Fig. 4. The center frequency is 5.8 GHz and the 3-dB bandwidth is 280 MHz, which is about 5% of the center frequency. The measured return lose is better than 9 dB and the insertion is 1.8 dB in pass band. The fabricated filter is shown in Fig. 5.

There are three drawbacks of this band-pass filter, although it has a quite good performance. A) The frequency response at the cut-off frequency is not steep enough. B) The first parasitic pass band is too close, which limits the applications of the proposed band-pass filter. C) The return loss in pass band is too small.

## 4. CONCLUSION

In this research, a novel SIW filter is investigated and the empirical design formulae are presented. Filters are fabricated on an F4B-2 substrate, whose dielectric constant is 2.65 and dielectric loss tangent is 0.001. Although there is a little frequency shift by approximately 0.8%, measured results are in good agreement with simulated ones, which suggest that the proposed filters present attractive performances, especially the high selective filter. It's a difficulty for traditional SIW filters to get rid of parasite pass bands caused by lower- or higher-order modes [11], while the parasite pass band has been depressed effectively in this paper. The size of the filter is about half of a traditional dual mode SIW filter that make use of $TE_{102}$ and $TE_{201}$, as discussed in [11]. It is expected to find applications in communication systems, especially for frequencies in C-band.

## REFERENCES

1. Hirokawa, J. and M. Ando, "Single-layer feed waveguide consisting of posts for plane TEM wave excitation in parallel plates," *IEEE Transactions on Antennas and Propagation*, Vol. 46, 625–630, May 1998.
2. Deslandes, D. and K. Wu, "Integrated microstrip and rectangular waveguide in planar form," *IEEE Microwave Wireless Components Letter*, Vol. 11, 68–70, February 2001.
3. Henry, M., C. E. Free, B. S. Izqueirdo, J. Batchelor, and P. Young, "Millimeter wave substrate integrated waveguide antennas: Design and fabrication analysis," *IEEE Transactions on Advanced Packaging*, Vol. 32, No. 1, February 2009.
4. Wang, Y.-T., X.-W. Zhu, and L. Tian, "Design of crossed-SIW directional couplers with different angles," *APMC 2008*, 1–4, December 16–20, 2008.
5. Tang, H.-J., W. Hong, J.-X. Chen, G.-Q Luo, and K. Wu, "Development of millimeter-wave planar diplexers based on complementary characters of dual-mode substrate integrated waveguide filters with circular and elliptic cavities," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 55, 776–782, April 2007.
6. Chen, Y. J., W. Hong, and K. Wu, "Novel substrate integrated waveguide fixed phase shifter for 180-degree directional coupler," *Proc. of IEEE MTT-S*, 189–192, June 2007.
7. Deslandes, D. and K. Wu, "Single-substrate integration technique of planar circuits and waveguide filters," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 51, No. 2, February 2003.
8. Wang, R., L.-S. Wu, and X.-L. Zhou, "Compact folded substrate integrated waveguide cavities and bandpass filter," *Progress In Electromagnetics Research*, PIER 84, 135–147, 2008.
9. Chen, X.-P., K. Wu, and Z.-L. Li, "Dual-band and triple-band substrate integrated waveguide filters with chebyshev and quasi-elliptic responses," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 55, 2569–2577, December 2007.
10. Ismail, A., M. S. Razalli, M. A. Mahdi, R. S. A. R. Abdullah, N. K. Noordin, and M. F. A. Rasid, "X-band trisection substrate-integrated waveguide quasi-elliptic filter," *Progress In Electromagnetics Research*, PIER 85, 133–145, 2008.
11. Shen, W., X.-W. Sun, W.-Y. Yin, J.-F. Mao, and Q.-F. Wei, "A novel single-cavity dual mode substrate integrated waveguide filter with non-resonating node," *IEEE Microwave and Wireless Components Letters*, Vol. 19, No. 6, 368–370, 2009.
12. Chen, X.-P., K. Wu, and D. Drolet, "Substrate integrated waveguide filter with improved stopband performance for satellite ground terminal," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 57, No. 3, March 2009.

# CAD of Resonant Circular Iris Waveguide Filter with Dielectric Filled Cavities

**Uma Balaji**
Electrical Engineering Technology, Farmingdale State College, Farmingdale, NY 11735, USA

**Abstract**— Resonant Iris bandpass filters form compact structures. Filters composed of circular irises formed as a junction between two circular waveguide cavities has been designed. Mode matching method has been used to analyze discontinuity present in this filter. By using generalized scattering matrix method all the discontinuities that compose the filter are cascaded to obtain its performance. The filter dimensions have been further optimized to meet specifications more precisely. Filters with both empty and dielectric filled circular waveguide cavities between irises have been designed and optimized.

## 1. INTRODUCTION

Closed form solutions for the susceptance of rectangular and circular apertures in transverse plane of waveguides are available in literature [1]. Based on these solutions, filter in rectangular and circular waveguides are designed using the equivalent network theory approach [2]. This method may however yield filters whose performance does not meet specifications very well. Computer aided design based on mode matching method (MMM) is presented here to design filters in circular waveguide. Circular irises are used as coupling elements between cavities. Two filter structures have been designed for the same bandwidth. One filter uses empty sections of circular waveguide for the cavities, while the other filter uses dielectric filled circular waveguide cavities.

## 2. THEORY

The analysis of any discontinuity using MMM involves the following steps [3]. The fields on both sides of the discontinuity are expanded in terms of a series of modes of incident and reflected waves. The magnitude of power carried by each of the modes is set to unity. The continuity conditions for the tangential components of electric and magnetic fields are imposed. Using the principle of orthogonality of modes, the equations of continuity conditions are transformed into matrices relating the expansion coefficients of incident and reflected waves at the discontinuity. The matrices are rearranged and inverted suitably to obtain the generalized scattering matrix which describes the discontinuity in terms of the dominant and higher order modes. Theoretically the generalized scattering matrix is of infinite dimension corresponding to the infinite number of modes. The matrix is truncated to a finite size for numerical computations after testing the convergence of the $S$-parameters.

The filters discussed here are composed of discontinuity from a larger circular waveguide to a smaller circular waveguide as shown in Figure 1. While analyzing a discontinuity from larger to smaller circular waveguide (placed along $z$-axis) for dominant $\mathrm{TE}_{11}$ mode excitation it is sufficient to consider only $\mathrm{TE}_{1m}$ and $\mathrm{TM}_{1m}$ modes, where $m$ is an integer alone for the analysis of the discontinuity. This is due to the fact that the circular waveguides have rotational symmetry and this is maintained when the two sides of the discontinuity are placed with their longitudinal axis coinciding. The electric and magnetic fields in the two regions of the discontinuity can be obtained from the electric and magnetic potential functions $\psi^h$ and $\psi^e$ in them. For example, for the region I (empty circular waveguide of radius $a_1$) of the discontinuity as shown in Figure 1 the potential functions are written as,

$$\begin{aligned}\psi(\rho,\phi)^{Ih} &= \sum_{m=1}^{M^h} P^{Ih}_{1,m} J_1(k^{Ih}_{c1,m}\rho)\cos\phi \\ \psi(\rho,\phi)^{Ie} &= \sum_{m=1}^{M^e} P^{Ie}_{1,m} J_1(k^{Ie}_{c1,m}\rho)\sin\phi\end{aligned}$$

The coefficients $P$ in the above equations are the power normalization constants that set the power carried in each of the modes to a watt. The cutoff wavenumbers of the $\mathrm{TE}_{1m}$ and $\mathrm{TM}_{1m}$

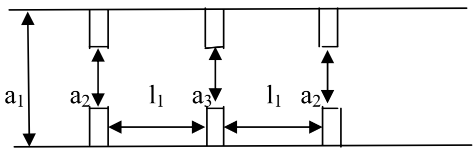



Figure 1: Longitudinal cross-section of the filter with empty circular waveguide for cavities and thickness of iris = 0.1 mm.

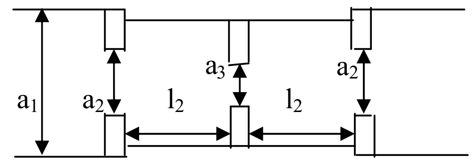



Figure 2: Longitudinal cross-section of the filter with dielectric filled and cavities, iris thickness = 0.1 mm.

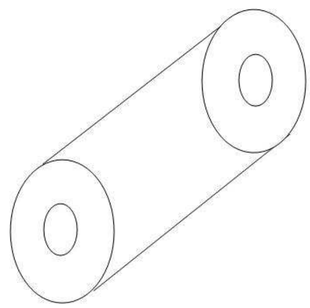


Figure 3: Circular waveguide cavities in the filter between circular irises (0.1 mm thick)

modes are $k^{Ih}_{c1,m}$ and $k^{Ie}_{c1,m}$. The number of TE and TM modes used in the analysis of the discontinuity is decided by $M^h$ and $M^e$. Similarly the potential functions for region 2 (circular iris of radius $a_2$) are also written.

The electric and magnetic fields obtained from the potential functions are expressed as a sum of incident and reflected waves of unknown co-efficient on both the sides of the discontinuity. The continuity condition is applied as follows with $a_2$being the radius of the smaller waveguide or the iris.

$$\begin{aligned} E^I(\rho,\phi) =& E^{II}(\rho,\phi); \quad \text{where } \rho \in [0,a_2];\ \phi \in [0,2\pi] \\ E^I(\rho,\phi) =& 0; \quad \text{otherwise} \\ H^I(\rho,\phi) =& H^{II}(\rho,\phi); \quad \text{where } \rho \in [0,a_2];\ \phi \in [0,2\pi] \end{aligned}$$

Applying these continuity conditions results in scattering matrix of fundamental and higher order modes is given by the following matrix equation where $U$ is a identity matrix, 0 is a zero matrix and $V$ is the sub matrix of the coupling between TE and TM modes on both sides of the discontinuity. The dimension of sub matrices depends on the number of modes used for matching fields on both the sides of the discontinuity.

$$S = \begin{bmatrix} U & 0 & V'_{hh} & V'_{eh} \\ 0 & U & 0 & V'_{ee} \\ -V_{hh} & 0 & U & 0 \\ -V_{eh} & -V_{ee} & 0 & U \end{bmatrix}^{-1} \begin{bmatrix} U & 0 & V'_{hh} & V'_{eh} \\ 0 & U & 0 & V'_{ee} \\ V_{hh} & 0 & -U & 0 \\ V_{eh} & V_{ee} & 0 & -U \end{bmatrix}$$

$$V_{hh}(m,q) = \frac{1}{2} P^{II}_{1,m} P^{I}_{1,q} \left( \int_0^{a_2} \frac{1}{\rho} J_1(k^{II}_{c1,m}\rho) J_1(k^{I}_{c1,q}\rho) d\rho + \int_0^{a_2} J'_1(k^{II}_{c1,m}\rho) J'_1(k^{I}_{c1,q}\rho) \rho d\rho \right)$$

The coupling matrix $V_{hh}$ given includes integrals of Bessel function but can be solved analytically in terms of Bessel functions. The other coupling sub matrices are similar equations with the integrals that can also be solved analytically in terms of Bessel functions. It has to be noted that if the waveguide is filled with dielectric the propagation constants will depend on the dielectric constant of the material used to fill the waveguide.

The scattering matrix of each discontinuity is cascaded using the generalized scattering matrix technique to analyze the filter [3]. The filter is further optimized to achieve the specifications desired using a practical quasi Newton algorithm [4].

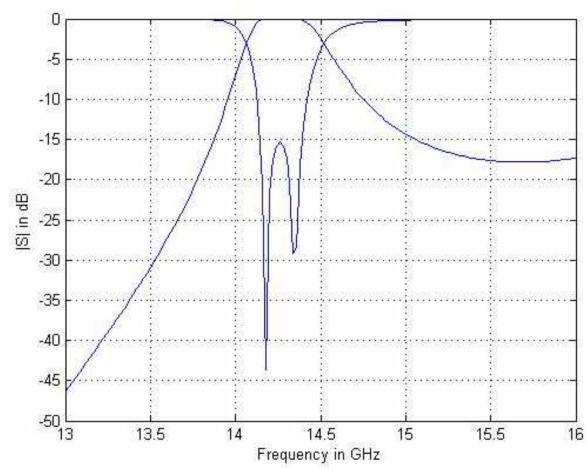



Figure 4: Performance of filter in Figure 1 $a_1 = 13.97\,\mathrm{mm}$, $a_2 = 8.68\,\mathrm{mm}$, $a_3 = 7.02\,\mathrm{mm}$, dia. of cavities $= a_1$, $l_1 = 18.7\,\mathrm{mm}$.

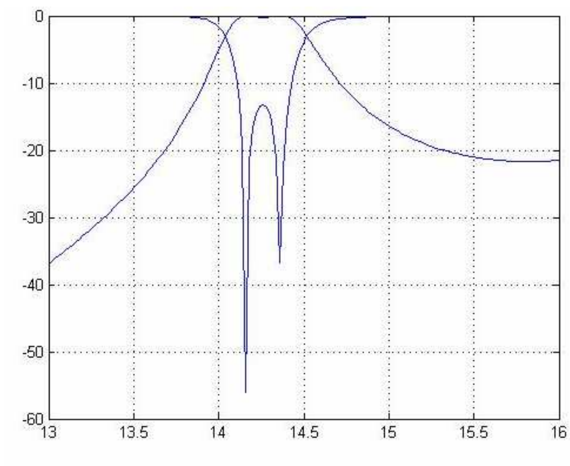


Figure 5: Performance of filter in Figure 2 $a_1 = 13.97\,\mathrm{mm}$, $a_2 = 7.5\,\mathrm{mm}$, $a_3 = 5\,\mathrm{mm}$, dia. of cavities $= 11.6\,\mathrm{mm}$, $l_2 = 17.6\,\mathrm{mm}$.

## 3. RESULTS

The discontinuity from a larger empty circular waveguide to smaller circular waveguide was analyzed with 40 TE and TM and convergence of the $S$-parameters has been observed. The filter using empty circular waveguide with two cavities has been analyzed and optimized further for the desired specifications. A return loss of better that 15 dB is observed in the passband as shown in Figure 4. A filter with cavities filled with dielectric was also analyzed and optimized for the same specifications. The cavities of this filter that were filled with dielectric were chosen to be of smaller radius in comparison to that of the filter with empty circular waveguide cavities. The length of the cavities is also reduced. The performance of the filter is shown in Figure 5. It has yielded a better performance in stopband, but the return loss in the mid passband is reduced slightly in comparison to that of the other filer. Further optimization may improve the passband performance of this filter. The performance and the dimension of both the filters are shown in Figures 4 and 5.

## 4. CONCLUSION

The Mode Matching Method used for the analysis and design of circular waveguide filters with cavities that are empty are dielectric filled is found to yield optimum solutions for the desired specifications. The filters with dielectric filled cavities have reduced dimensions. Cavities filled partially with dielectric can further be designed using this method.

## REFERENCES

1. Eastham, G. B. and K. Chang, "Analysis and closed form solutions of circular and rectangular apertures in a circular waveguide," *IEEE Trans. Microwave Theory and Techniques*, Vol. 39, No. 4, 718–723, April 1991.
2. Matthaei, G. L., L. Young, and E. M. T. Jones, *Microwave Filters, Impedance Matching-networks and Coupling Structures*, McGraw-Hill, 1964.
3. Uher, J., J. Bornemann, and U. Rosenberg, *Waveguide Components for Antenna Feed Systems: Theory and CAD*, Artech House Inc., 1193.
4. Antoniou, A., *Digital Filters, Design and Applications*, 2nd Edition, Mc-Graw Hill, 1993.

# On-chip Impedance-optimized Microstrip Transmission Line for Multi-band and Ultra-wide-band Microwave Applications

**Wayne Woods**[1], **Guoan Wang**[1], **Hanyi Ding**[1], **and Shurong Dong**[2]

[1]IBM Microelectronics, USA
[2]Zhejiang University, China

**Abstract**— This paper presents a novel on-chip passive device that provides a more constant characteristic impedance over a wide frequency band compared to conventional microstrip transmission lines. This device is designed for applications whose frequencies span the range of the microwave frequency region from 1 GHz to 30 GHz up to the millimeter-wave (MMW) frequency bands greater than 30 GHz. The novel, impedance-optimized, microstrip transmission line presented is well suited to reduce the effects of characteristic impedance mismatch in high-performance on-chip analog circuits where it is desirable for the transmission line characteristic impedance to remain relatively constant in all operating frequency bands. This new device utilizes the frequency-dependent nature of the silicon substrate capacitance to reduce variation in transmission line characteristic impedance with frequency. Using specially designed metal-to-silicon substrate capacitance structures designed in "windows" that perforate a bottom grounded microstrip return line to provide additional capacitance to the signal path at lower frequencies which compensates for the higher DC inductance of the thick metal signal lines at low frequencies caused by the resistance/inductance skin-effect. The proposed structure is ideal for multi-band and ultra-wide band applications, for example, analog circuits that operate in both the WCDMA range (2.11–2.17 GHz) and also in the MMW range ($f > 30$ GHz). An impedance-optimized transmission line is designed and studied in a 130 nm BiCMOS technology. The proposed device exhibits significantly improved characteristic impedance behavior versus frequency compared to conventional microstrip lines and can be implemented in any silicon-based analog technology.

## 1. INTRODUCTION

Highly accurate modelling and design of critical interconnect paths in state-of-the-art wireless designs is crucial for first-pass design success [1]. Microwave and MMW circuits are sensitive to the electrical parasitics of even small interconnect lines, and the optimization of the electrical performance of critical interconnect paths is important. Many signal paths in microwave and MMW circuits are shielded microstrip transmission lines [2]. In high-frequency digital applications, critical interconnect paths often take the form of coplanar waveguides. In microstrip lines, resistance increases due to skin-effect in the signal conductor and the ground plane below and is generally proportional to the square of the frequency. The line inductance of on-chip microstrips decreases at the onset of skin-effect in the signal conductor. Full-bandwidth microwave compact passive models of microstrips include the effect of skin-effect on resistance and inductance [3]. Also, the line capacitance of on-chip microstrip transmission lines is generally constant with frequency. However, in coplanar waveguides, the effect of the silicon substrate on signal capacitance must be considered when there are no crossing lines below the signal line to shield it from the silicon substrate and the spacing to the ground return lines is large. Full-bandwidth microwave compact coplanar waveguide models include accurate modelling of the capacitance over silicon as well as the metal conductor skin-effect on inductance and resistance [4]. In conductors over silicon with no shielding, the capacitance of the signal line is much more sensitive to frequency than in coplanar waveguides. Predictive equations can be found in the literature for line inductance, line capacitance, and line conductance through the silicon substrate for an isolated un-shielded line over a silicon substrate [5, 6]. In practical applications, these equations can deviate somewhat from values measured from hardware and predicted by commercial EM solvers such as HFSS [7] for a given technology. It is clear from these preceding interconnect and modeling citations that the frequency-dependent behavior of critical on-chip interconnect is understood and well documented in the literature. This paper proposes a novel on-chip microstrip transmission line that makes use of the frequency-dependent nature of the capacitance of conductors over a silicon substrate to counter the effect of the frequency-dependent inductance decrease due to skin-effect in the microstrip signal lines in order to optimize the characteristic impedance, $Z_o$, versus frequency. By designing the capacitance of the novel microstrip to decrease with frequency by roughly the same proportion as the inductance, the $Z_o$ of the proposed

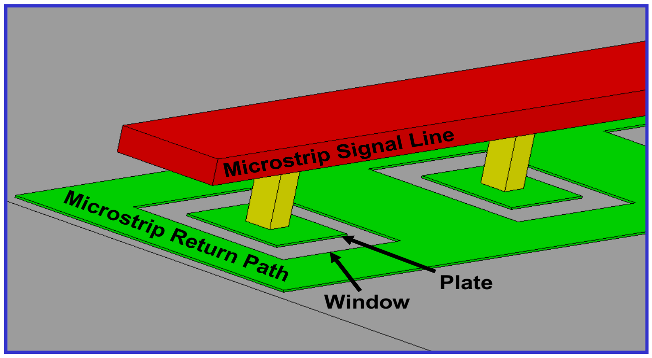



Figure 1: On-chip impedance-optimized microstrip transmission line.

microstrip is made more constant versus frequency. This makes the proposed microstrip design well suited for multi-band and ultra-wide-band circuit designs that operate from a few GHz up to MMW frequencies.

## 2. IMPEDANCE-OPTIMIZED MICROSTRIP STRUCTURE

Figure 1 shows the proposed impedance-optimized microstrip transmission line in a 130 nm BiCMOS technology. In Figure 1, there is a signal line above a ground return line, like in a conventional microstrip design. However, the ground return of the microstrip transmission line in Figure 1 is different form a conventional microstrip because there are "windows" created in the middle of the return path. In the "windows", on the same metal layer, are capacitance "plates" that are connected with metal-via posts to the signal line. These capacitance "plates" are located directly above the silicon substrate. In the example shown, the signal line is 16 μm wide and is routed in a 4 μm thick aluminum signal line that is the top metal layer of a 130 nm BiCMOS technology and the ground return line and capacitance "plates" are in the lowest metal layer which is a 0.32 μm thick copper layer that is 0.57 μm above the silicon substrate. The presence of the "plates" above the silicon substrate causes the signal line capacitance to be frequency dependent unlike the general behavior of a conventional on-chip microstrip transmission line. The ground return is 40 μm wide and the "plates" are 10 μm by 10 μm squares with a repeating pitch of 50 μm along the length of the signal line, and the windows are 20 μm by 20 μm.

Figure 2 shows the general inductance and capacitance behavior versus frequency of both a conventional and the proposed impedance-optimized microstrip transmission line. The inductance of the signal line in both cases is generally identical and is depicted as the single inductance plot versus frequency shown in Figure 2. However, the capacitance versus frequency behavior shown in Figure 2 of the two cases differs significantly. In the conventional microstrip design, the capacitance is roughly constant versus frequency. However, the proposed impedance-optimized microstrip capacitance shows the same frequency-dependent behavior as the signal inductance: a decrease with increased frequency. Ignoring resistance, the characteristic impedance of a transmission line is equal to the square root of the ratio of line inductance to line capacitance. So, in the conventional microstrip design, the characteristic impedance will decrease with frequency because the inductance decreases with frequency while the capacitance remains constant. The characteristic impedance of the proposed impedance-optimized microstrip design will remain more constant with frequency because as it's inductance decreases with frequency, the capacitance decreases as well, keeping the ratio of inductance to capacitance roughly constant. The frequency-dependent capacitance of the capacitance "plates" in Figure 1 can be determined using a field solver such as HFSS [7], however, a good intuition for what to expect can be obtained from [5]. From [5], it is clear that at very low frequencies, the silicon substrate acts as a grounded electrical conductor. So at low frequencies, the capacitance of the "plates" is primarily the parallel-plate capacitance between the 10 μm by 10 μm "plate" and the top of the silicon substrate which is acting as a grounded conductor. At high-frequencies, the silicon substrate acts more like a dielectric which eliminates the parallel-plate capacitance to the top of the silicon and the capacitance of the "plates" is primarily from the enclosing metal of the ground return path. The frequency that marks the transition of silicon from behaving like a conductor to behaving like a dielectric is the relaxation frequency and is dependent on the bulk conductivity of the silicon substrate. The relaxation frequency of the silicon substrate in

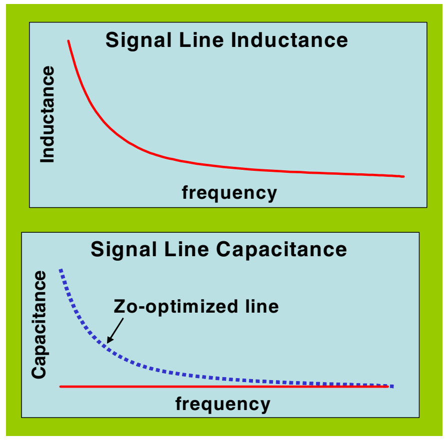



Figure 2: General inductance and capacitance behavior of conventional and impedance-optimized on-chip microstrip transmission lines.

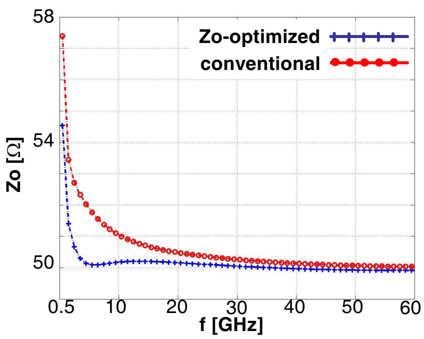



Figure 3: Simulated HFSS $Z_o$ results of both conventional and impedance-optimized microstrip transmission lines from 0.5 GHz to 60 GHz.

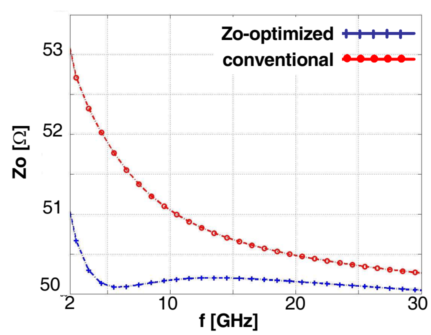



Figure 4: Simulated HFSS $Z_o$ results of both conventional and impedance-optimized microstrip transmission lines from 2 GHz to 30 GHz.

this 130 nm BiCMOS example is around 12 GHz. At frequencies well below 12 GHz, the capacitance of the "plates" is maximized and therefore the signal line capacitance of the proposed microstrip design is also maximized. Likewise, above 12 GHz, the line capacitance is minimized and remains roughly constant with increasing frequency. HFSS was used to model the impedance-optimized design shown in Figure 1. The size of the "windows": 20 μm by 20 μm and "plates": 10 μm by 10 μm was also determined using HFSS by trying several different dimensions to optimize the effect of the frequency-dependent capacitance on $Z_o$ in order to achieve a more constant $Z_o$.

The result of HFSS [7] simulations of the characteristic impedance of a conventional 50 Ω microstrip and the impedance-optimized 50 Ω microstrip in the 130 nm BiCMOS technology are shown in Figure 3 and Figure 4. Figure 3 shows the frequency range from 0.5 GHz to 60 GHz while Figure 4 shows the frequency range from 2 GHz to 30 GHz. Both of the microstrip designs were designed to be 50 Ω in the MMW frequency ranges ($f > 30$ GHz). At low frequencies in Figure 3, $Z_o$ of both the conventional and $Z_o$-optimized microstrips deviates significantly from 50 Ω : 57.5 Ω and 54.5 Ω at 0.5 GHz respectively. Around 60 GHz, the $Z_o$ of both transmission line designs approach 50 Ω. Figure 4 shows the range from 2 GHz to 30 GHz where it can be seen that the $Z_o$ of the conventional and $Z_o$-optimized microstrips is 53 Ω and 51 Ω at 2GHz and 50.5 Ω and 50.1 Ω at 30 GHz respectively. Also, at 5 GHz the impedances of the two designs are 51.9 Ω and 50.2 Ω respectively. So, at 5 GHz the percentage difference from the 50 Ω target is 3.8% for the conventional microstrip and 0.4% for the $Z_o$-optimized microstrip. The percentage differences do not seem too large in either case. However, in many high-performance designs, microstrip lines are used extensively with many different microstrips employed. So, the effects of impedance mismatches at the inputs and outputs of all the different microstrips become significant. It is clear from Figure 4 that it is desirable to use the impedance-optimized design. Also, because the $Z_o$-optimized design is design rule checklist (DRC) clean in the BiCMOS 130 nm technology, no additional expense is required in choosing the $Z_o$-optimized microstrip transmission line. The impedance change with frequency of a conventional microstrip is the result of the skin-effect inductance decrease in the signal line. In general, the thicker the signal line, the greater the skin-effect inductance decrease. So, in technologies with thick top-metals, it is advantageous to use the $Z_o$-optimized microstrip.

## 3. CONCLUSIONS

An on-chip impedance-optimized microstrip transmission line was presented in this paper that makes use of capacitance "plates" in a microstrip ground return line with "windows" opened to expose the silicon substrate below. The proposed impedance-optimized microstrip is DRC clean and requires no additional mask levels or process changes. HFSS simulations of a $50\,\Omega$ $Z_o$-optimized microstrip and a $50\,\Omega$ conventional microstrip show that the $Z_o$-optimized microstrip $Z_o$ is more constant from 2 GHz to 30 GHz giving it an advantage in multi-band and ultra-wide-band circuit designs with many microstrips where $Z_o$ matching is critical at multiple frequencies.

## REFERENCES

1. Singh, R., D. L. Harame, and M. M. Oprysko, *Silicon Germanium Technology, Modelling, and Design*, IEEE Press and Wiley, New York, 2004.
2. Valdes-Garcia, A., S. Reynolds, and U. R. Pfeiffer, "A 60 GHz class-E power amplifier in SiGe," *IEEE Asian Solid-State Circuits Conference*, 199–202, Hangzhou, China, Nov. 2006.
3. Zwick, T., Y. Tretiakov, and D. Goren, "On-Chip SiGe transmission line measurements and model verification up to 110 GHz," *IEEE Microwave and Wireless Components Letters*, Vol. 15, No. 2, 65–67, Feb. 2005.
4. Goren, D., S. Shlafman, B. Sheinman, W. Woods, and J. Rascoe, "Silicon-chip single and coupled coplanar transmission line measurements and model verification up to 50 GHz," *IEEE Workshop on Signal Propagation in Interconnects*, 21–24, Genova, Italy, May 2007.
5. Weisshaar, A., H. Lan, and A. Luoh, "Accurate closed-form expressions for the frequency-dependent line parameters of on-chip interconnects on lossy Silicon substrate," *IEEE Transactions on Advanced Packaging*, Vol. 25, No. 2, 288–296, May 2002.
6. Eo, Y. and W. R. Eisenstadt, "High-speed VLSI modelling based on s-parameter measurements," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol. 16, No. 5, 555–562, Aug. 1993.
7. Ansoft Corporation, *High-Frequency Structure Simulator User's Guide*, 2008.

# Microstrip Resonator as a Measuring Device for a Single Molecule Magnet

**Thomas Fan**[1], **Vladimir I. Tsifrinovich**[1], **and Andrew D. Kent**[2]

[1]Department of Physics, Polytechnic Institute of NYU, 6 Metrotech Center, Brooklyn, NY 11201, USA
[2]Department of Physics, New York University, 4 Washington Square Place, New York, NY 10003, USA

**Abstract**— A single molecule magnet (SMM) could be used as a qubit in quantum information processing. In this work, we suggested a method for the measurement of the SMM state with a microstrip resonator. In our method the frequency of the electron spin resonance adiabatically approaches to the fundamental frequency of the microstrip resonator. The shift of the resonator frequency is maximal near the crossing point. This shift depends on the initial state of the SMM and can be used for the measurement of the initial state. We have used the classical approach for description of the SMM dynamics and the interaction between the microstrip and SMM. We have obtained an analytical expression for the resonator maximum frequency shift and performed numerical analysis in the vicinity of the crossing point.

## 1. INTRODUCTION

In quantum information processing, single molecule magnets (SMM) can be used as qubits [2]. Measurement of the qubit state is fundamental to a successful quantum computer. A microstrip resonator recently became an important tool in quantum information processing [3]. In this paper, we suggest using a microstrip resonator as a method to measure the SMM spin state . We consider the adiabatic spin dynamics assuming that, the SMM magnetic moment remains parallel or antiparallel to the effective magnetic field. In this work, we use the classical equations of motion for the SMM magnetic moment. The microstrip was described in an equivalent circuit model. Only the rotating component of the transversal field on the SMM was taken into consideration and the anisotropy field was ignored.

## 2. SMM-MICROSTRIP SYSTEM

A diagram of a microstrip resonator is shown in Fig. 1. The microstrip resonator can be represented by an LC circuit. In terms of the physical shape of the microstrip resonator, the parameters, $C$ and $L$, can be approximated as:

$$C = \frac{\epsilon w l}{d}, \quad L = \frac{\mu_o l d}{w} \tag{1}$$

where $\epsilon$ is the permittivity of the dielectric substrate, $\mu_o$ is the permeability of free space, $w$ and $l$ are the width and the length of the microstrip, and $d$ is the distance between the microstrip and the metallic plate as shown in Fig. 1 [1].

We consider a SMM that is experiencing a permanent external magnetic field together with a transversal magnetic field created by a microstrip. The dynamics of the SMM is governed by the following equation

$$\dot{\vec{m}} = -\gamma\left(\vec{m} \times (B_1\cos(\omega t)\hat{x} + B_1\sin(\omega t)\hat{y} + B_{ext}\hat{z}\right) \tag{2}$$

where $\vec{m} = m_x\hat{x} + m_y\hat{y} + m_z\hat{z}$ is the magnetic moment of SMM, $\gamma$ is the absolute value of the electronic gyromagnetic ratio, $B_1$ is the magnitude of the transversal rotating magnetic field of frequency $\omega$ caused by the microstrip resonator and $B_{ext}$ is the magnitude of the external permanent magnetic field, which points in the positive $z$-direction. (Only the resonant rotating component of the transversal field is taken into consideration.) Letting $m_+ = m_x + im_y = \pm m_o e^{i\omega t}$, where $m_o$ is the magnitude of the transversal magnetic moment, we describe the two stationary solutions in the rotating system of coordinates, where

$$m_o = \frac{\gamma B_1 m}{\sqrt{(\gamma B_1)^2 + (\gamma B_{ext} - \omega)^2}}, \tag{3}$$

The magnitude of the magnetic moment is $m = \sqrt{m_o^2 + m_z^2}$. The two stationary solutions describe the magnetic moment, parallel or antiparallel to the effective field in the rotating system.

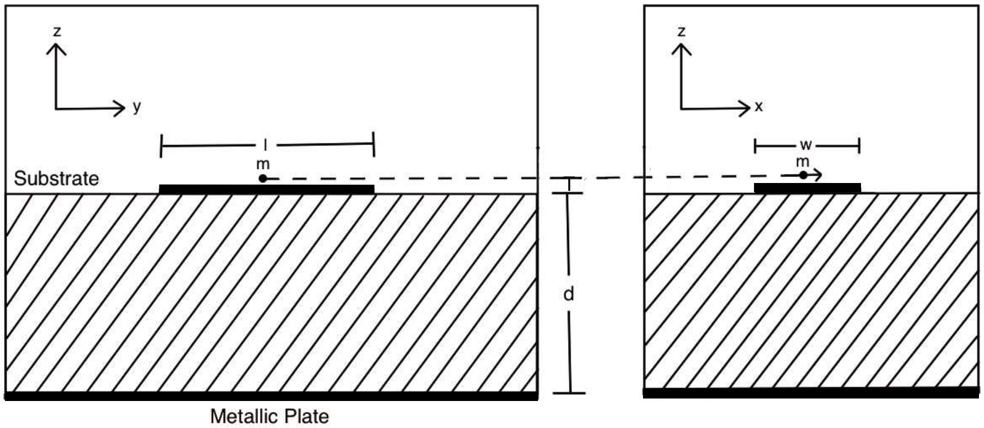



Figure 1: Microstrip resonator and magnetic moment, $\vec{m}$, of the SMM at an instant when the magnetic moment of the SMM points in the positive $x$-direction. $l$ and $w$ are the length and the width of the microstrip, $s$ is the distance between the substrate and the SMM.

## 3. EQUATION FOR RESONATOR FREQUENCY

We assume that the oscillating component of the magnetic moment, $m_x$, generates an $emf$ in the equivalent circuit of the microstrip resonator

$$emf = -\frac{d\Phi}{dt} = -k'\frac{dm_x}{dt} \tag{4}$$

where $k'$ is a parameter determined by the geometry of the microstrip. Applying Kirchoff's rules to an LC circuit with the $emf$ from Eq. (4) and taking its time derivative, we get

$$-\frac{I}{C} - L\ddot{I} - k'\ddot{m}_x = 0 \tag{5}$$

In order to find the eigenfrequency of the resonator affected by the SMM, we put

$$I = I_o e^{i\omega t}, \quad m_x = \kappa m_o e^{i\omega t} \tag{6}$$

The parameter, $\kappa = \pm 1$, respectively describes the magnetic moment parallel or antiparallel to the effective field in the rotating system. In the zeroth approximation putting $k' = 0$ we obtain the resonator frequency, $\omega = \omega_o = 1/\sqrt{LC}$.

Next, we obtain the frequency of the resonator-SMM system in the first approximation. The transversal oscillating magnetic field at the SMM location is assumed to to be $B_x = kI$, where $k$ depends on the geometry of the microstrip. We take into consideration, only the resonant rotating component of the transversal field with the magnitude $B_1 = \frac{1}{2}kI_o$. From Eq. (5) and taking into consideration Eq. (3), one obtains

$$I_o\left(-\frac{1}{C} + L\omega^2 + \frac{\kappa\gamma kk'm\omega^2}{\sqrt{\gamma^2k^2I_o^2 + 4(\gamma B_{ext} - \omega)^2}}\right) = 0 \tag{7}$$

Letting $A = \frac{\gamma kk'm}{L}$, $D = \gamma kI_o$, $r = \frac{\gamma B_{ext}}{\omega_o}$, and $x = \frac{\omega}{\omega_o}$, one gets

$$-1 + x^2 + \frac{\kappa Ax^2}{\sqrt{D^2 + 4\omega_o^2(r - x)^2}} = 0 \tag{8}$$

We will use the following parameters for estimation: $\gamma = 1.761 \times 10^{11}\ s^{-1}T^{-1}$, $s = 252\,\mathrm{nm}$ (we assume the distance between the SMM and the microstrip is $2\,\mathrm{nm}$, and the thickness of the microstrip is $250\,\mathrm{nm}$), $w = 0.4\,\mathrm{mm}$, $m = 1.85 \times 10^{-22}\,\mathrm{JT^{-1}}$ (this corresponds to SMM, $Mn_{12}$-Acetate, with $m = 20\mu_b$, i.e., $S = 10$), $l = 1.5\,\mathrm{mm}$, and $d = 5\,\mathrm{mm}$, $\epsilon = 13\epsilon_o$. With these parameters, we have $k = 1.57 \times 10^{-3}\,\mathrm{TA^{-1}}$, $k' = 3.02 \times 10^{3}\,\mathrm{Hm^{-2}}$, $L = 23.6\,\mathrm{nH}$, $C = 13.8\,\mathrm{fF}$, $\omega_o/2\pi = 8.82\,\mathrm{GHz}$, $\frac{D}{I_o} = 2.77 \times 10^{8}\,\mathrm{s^{-1}A^{-1}}$ , and $A = 6.57 \times 10^{-3}\,\mathrm{s^{-1}}$.

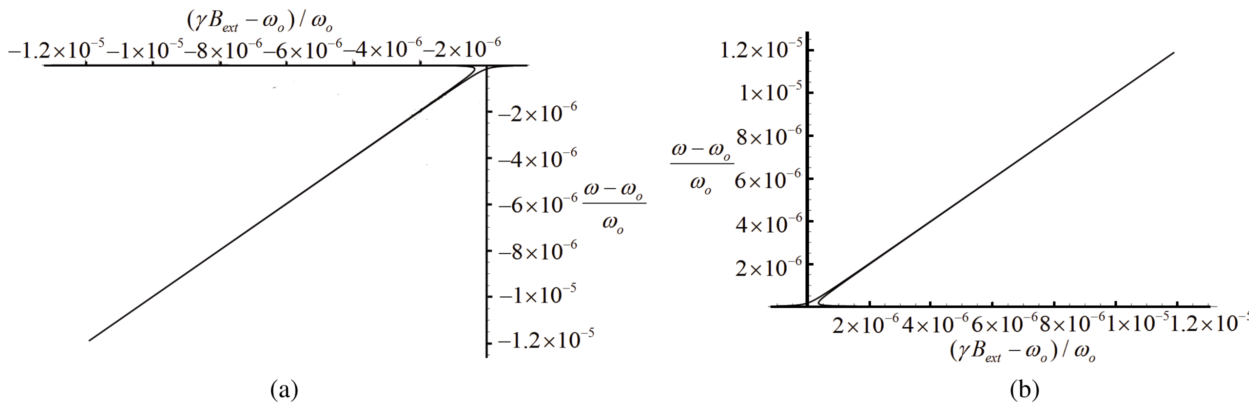



Figure 2: The dependence of the relative frequency shift on $\frac{\gamma B_{ext}}{\omega_o}$ for $I_o = 1\,\mu\text{A}$. (a) $\kappa = 1$ and (b) $\kappa = -1$.

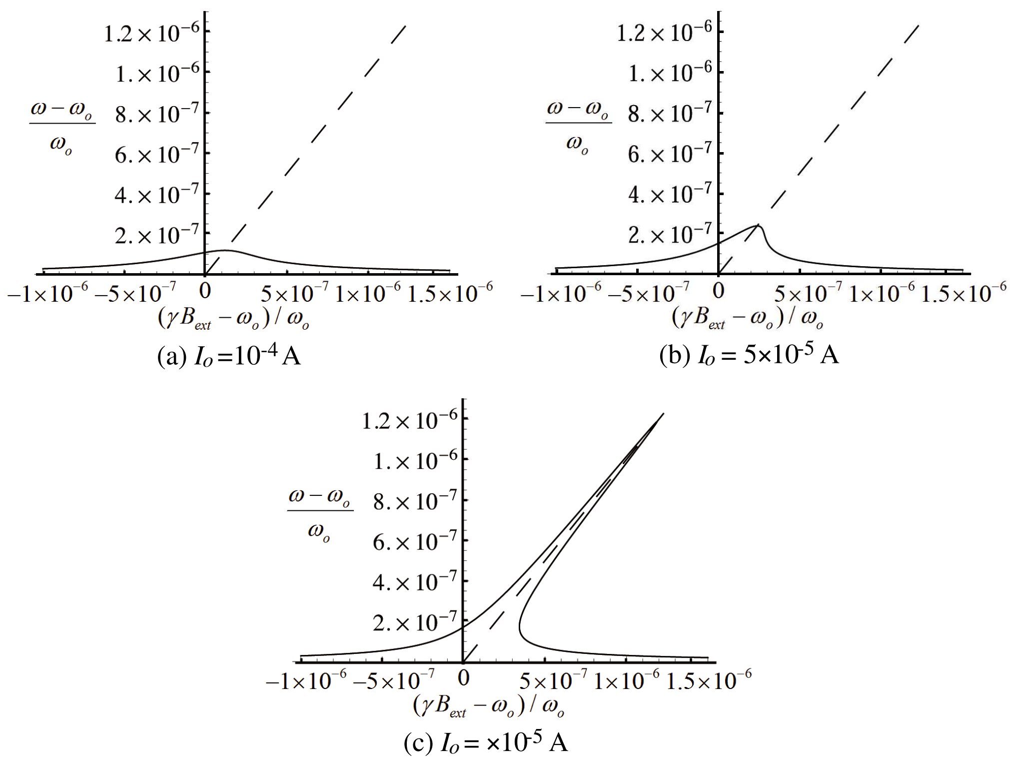



Figure 3: Relative frequency shift for different current amplitudes and $\kappa = -1$. All other parameters are the same as in Fig. 2. The function, $\omega = \gamma B_{ext}$, is represented by the dashed line.

At the crossing point, ($x = 1$), the relative frequency shift, $|\epsilon|$, is maximum when $D \to 0$ i.e., when the current approaches zero

$$|\epsilon(I_o \to 0)| = \frac{1}{2}\sqrt{\frac{A}{\omega_o}} = \frac{1}{2}\sqrt{\frac{\gamma k k' m}{L\omega_o}} \tag{9}$$

With our parameters, the maximum relative frequency shift at the crossing point is approximately $1.72 \times 10^{-7}$.

The classical description of the microstrip is valid if the microstrip energy is much greater than its energy quantum: $\frac{1}{2}LI_o^2 \gg \hbar\omega_o$. Taking $\frac{1}{2}LI_o^2/2 = 10\hbar\omega_o$, we obtain the lower bound for the current amplitude: $I_o \geq I_o' = \sqrt{20\hbar\omega_o/L}$. For our parameters, we have $I_o' = 7.04 \times 10^{-8}$ A. For current amplitude, $I_o'$, the relative frequency shift is approximately the same as for $I_o \to 0$.

## 4. HYSTERESIS BEHAVIOR OF THE MICROSTRIP FREQUENCY

With our parameters and current amplitude $I_o = 1\,\mu\mathrm{A}$, we numerically solved Eq. (8). The relative frequency shift of the resonator as a function of the ratio $\frac{\gamma B_{ext}-\omega_o}{\omega_o}$ is shown in Fig. 2. When $\kappa = 1$, Eq. (8) has three real positive solutions in the region, $-1.18\times10^{-5} \leq r-1 \leq -3.45\times10^{-7}$, and one real positive solution outside this region. When $\kappa = -1$, Eq. (8) has three positive real solutions in the region, $3.45\times10^{-7} \leq r-1 \leq 1.18\times10^{-5}$, and one real positive solution outside the region. One can see that the maximum resonator frequency shift is not at the crossing point. The maximum relative frequency shift, $|\epsilon|_{\max}$, for our parameters is approximately $1.18\times10^{-6}$ for $\kappa = \pm1$. For the minimum value of the "classical current amplitude", $I_o = I'_o$, we have approximately $1.69 \times 10^{-4}$. This value is the maximum frequency shift one can expect to observe.

We numerically determined that the hysteresis behavior disappears with current, $I_o$, approximately greater than $10^{-4}$ A. Fig. 3 shows the relative frequency shift for different current amplitudes.

Our numerical computations show that when the current amplitude decreases, the relative frequency shift $|\epsilon|$ monotonically increases. We numerically infer that the extremum (maximum for $\kappa = 1$ and minimum for $\kappa = -1$) value of $x$ in Eq. (8), occurs when $x = r$ i.e., $\omega = \gamma B_{ext}$. With this substitution, we obtain the exact expression for the resonator frequency:

$$x_{ext} = \sqrt{\frac{D}{D+\kappa A}} = \sqrt{\frac{LI_o}{LI_o+\kappa k'm}} \tag{10}$$

Thus, the maximum frequency shift, $|\epsilon|_{\max}$, is given by the expression

$$|\epsilon|_{\max} = \left|\sqrt{\frac{LI_0}{LI_0+\kappa k'm}} - 1\right| \tag{11}$$

For $I_o = 1\,\mu\mathrm{A}$, the maximum frequency shift for a $20\mu_B$ SMM is estimated as $|\epsilon|_{\max} \approx 1.18 \times 10^{-6}$ for $\kappa = \pm1$.

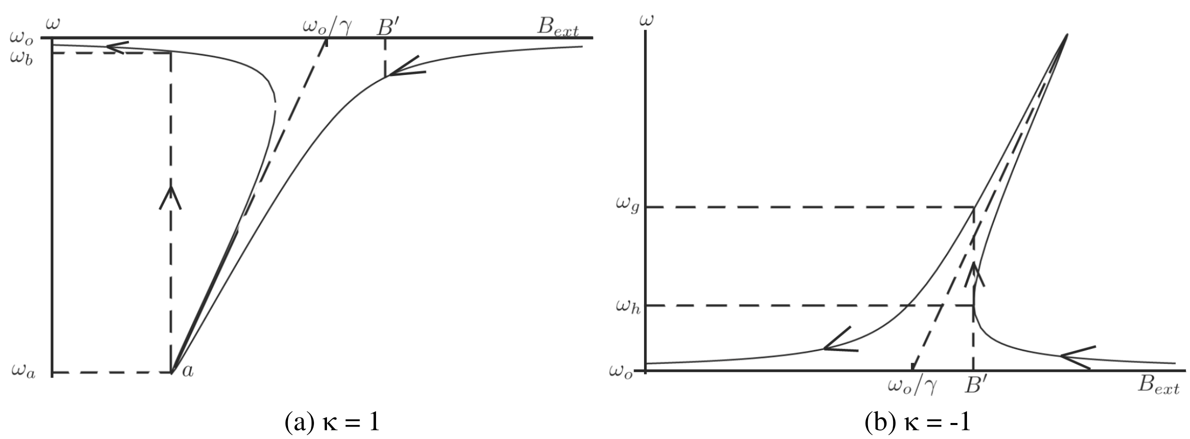



Figure 4: Expected behavior of the resonator frequency for a decreasing magnetic field $\kappa = \pm1$. The dashed line corresponds to $\omega = \gamma B_{ext}$.

## 5. CONCLUSION

We have developed the theory of the SMM measurement with the microstrip resonator using a classical description of the microstrip resonator and the SMM spin dynamics. We derived an analytical expression for the maximum frequency shift of the resonator (Eq. (11)) and numerically determined the behavior of the relative frequency shift, $\frac{\omega-\omega_o}{\omega_o}$, as a function of the ratio, $\frac{\gamma B_{ext}}{\omega_o}$. The hysteresis behavior of the resonator frequency was demonstrated. The maximum possible relative frequency shift for a $20\mu_B$ SMM is estimated as $|\epsilon|_{\max} \approx 1.69 \times 10^{-4}$.

Next, we propose two methods to use the resonator frequency shift to find the spin state of the SMM. We consider only two initial states of the SMM, ones where the magnetic moment is parallel or anti-parallel to the external magnetic field. The two states can be distinguished using the frequency shift of the microstrip resonator. Let the initial Larmor frequency, $\gamma B_{ext}$, be much greater than the frequency of the microstrip resontator, $\omega_o$. Then the Larmor frequency is slowly (adiabatically) lowered. The frequency shift, $\omega - \omega_o$, is negative if the magnetic moment was initially parallel to the external magnetic field. Otherwise, the frequency shift will be positive. Thus, using the sign of the resonator frequency shift one can distinguish these two states of the SMM: the ground state, $m_z = m$, and the upper state, $m_z = -m$. Note that the spin state can be measured non-destructively if we do not cross the point of frequency jump and adiabatically return to the initial value of the magnetic field $B_{ext}$.

Suppose that an experimental apparatus does not allow one to measure the sign of the frequency shift, $\omega - \omega_o$, but only its absolute value, $|\omega - \omega_o|$. In this case, the spin state of the SMM can still be deteremined. Assume that the external magnetic field adiabatically decreases from the values $\gamma B_{ext} \gg \omega_o$ to the value $B'$ in Fig. 4(a) and Fig. 4(b). The frequency jump is avoided so that the spin state can continue to change adiabatically. For current amplitude, $I_o = 1\,\mu\text{A}$, the relative frequency shift, $|(\omega - \omega_o)/\omega_o|$, is $7.13 \times 10^{-8}$ when $m_z = m$ and $4.15 \times 10^{-7}$ when $m_z = -m$. Thus, the relative frequency shift, $|\epsilon|$, at $B_{ext} = B'$ differs more than 5 times for the two initial states of the magnetic moment. Afterwards, the external magnetic field is increased back to its initial value. Therefore, a nondestructive measurement of the spin state of the SMM can be performed regardless of the sign of the frequency shift.

If the spin state of the SMM need not to be preserved, the magnetic field can be further decreased. If $m_z = m$, then the frequency shift, $|\omega - \omega_o|$ reaches its maximum value, $|\omega_a - \omega_o|$ in Fig. 4(a). This maximum value will be observed at $\gamma B_{ext} < \omega_o$. In the opposite case, $m_z = -m$, the maximum frequency shift for the decreasing field $B_{ext}$ is $|\omega_g - \omega_o|$ in Fig. 4(b). It will be observed at $\gamma B_{ext} > \omega_o$. Thus, decreasing the external magnetic field, $B_{ext}$, one can observe the frequency shift whose sign, absolute value and position are different for the two SMM spin states. Despite the fact that this method destructively measures the spin state of the SMM, the frequency shift, $|\epsilon|$, is permitted to reach its maximum possible value.

## ACKNOWLEDGMENT

This work was supported by the NYU-Poly seed funds.

## REFERENCES

1. Simon Ramo, John R. Whinnery, Theodore Van Duzer, *Fields and Waves in Communication Electronics.*, John Wiley and Sons, Inc., 1994.
2. G. de Loubens, G. D. Chaves-O'Flynn, A. D. Kent, C. Ramsey, E. del Barco, C. Beedle, and D. N. Hendrickson, "Magnetization and EPR studies of the single molecule magnet Ni4 with integrated sensors.," *Journal of Applied Physics*, Vol. 101, 2007.
3. H. M. Quddusi, C. M. Ramsey, J. C. Gonzalez-Pons, J. J. Henderson, E. del Barco, G. de Loubens and A. D. Kent, "On-chip Integration of High-Frequency Electron Paramagnetic Resonance Spectroscopy and Hall-Effect," *Review of Scientific Instruments*, Vol. 79, 2008.

# Electromagnetic Sources and Observers in Motion III — Derivation and Solution of the Electromagnetic Motional Wave Equation

**S. E. Wright**
ECASS Technologies Ltd., Moor Lane Laboratory, Kirkburton, Huddersfield, HD8 0QS, UK

**Abstract**— The evidence for the existence of the propagation medium (ether) for the transmission of electromagnetic (EM) waves (light), is considered in Paper I PIERS Xi'an. Einstein's relativity based on no propagation medium, and therefore not supported by the medium based Lorentz transform (LT), is considered in Paper II PIERS Xi'an. These two papers have led to the requirement of an EM motional wave equation, based entirely on the medium, in this Paper III. It appears that those aspects of Einstein's special relativity (SR), based on no medium, have not been proven theoretically or confirmed experimentally. Researchers claiming to have verified SR have usually verified the medium based Lorentzian (time and space compression) aspect of SR, not the absence of the medium. Without a propagation medium, and having no ability to distinguish between source and observer motion, it appears that SR cannot be a solution of the medium based wave equation. It appears to be inconsistent; Einstein claims no medium, with all the ambiguities and paradoxes that are implied. But uses a medium in his solution of the wave equation, in his motional electrodynamics, which cannot be derived without a medium. Therefore, SR appears to be an incomplete, ambiguous description of the EM propagation process. A general solution for source and observer motion moving relative to the propagation medium, is developed for the first time, removing ambiguities and predicting the complete causal radiation process. The derivation uses the medium based Maxwell's equations, and the kinematic effect is included through the Lorentz contraction of space and time with motion relative to the medium. The new visual theory, an extension of the LT, distinguishes between source and observer motion and between stationary and constantly moving systems, which, without a medium, SR cannot do. That part of SR based on the propagation medium seems correct, but those aspects based on no medium, (concept of relativity and relative motion) appear false. Reinstating the medium forbids material time travel, allowing causally only visual time travel to the past. This prevents the absurdity that the future can be observed before it occurs.[1]

## 1. INTRODUCTION

Maxwell [1], Lorentz [2], and Poincaré [3] supported the propagation medium, and Einstein [4] opposed it. For clarity, the medium is defined here as the means that transmits or propagates the wave disturbances, described by its causal wave equation. By having an EM propagation medium satisfying the wave equation, observations will have a well defined reference, be observable and predict a consistent sequence of events (cause always precedes the effect). To effectively remove the propagation medium, SR uses oblique time and space axes simulating simultaneity (symmetrical propagation times upstream and down), which has not been observed. Thus the theories of Maxwell, Lorentz, and Poincaré are medium based, satisfy the wave equation, have rectangular axes, are causal and observable. Whereas, Einstein's SR, assuming no medium, can not satisfy the wave equation, have oblique axes, and are non causal and non observable. The medium based LT predicts asymmetrical propagation times, (not symmetrical ones predicted by the ether-less SR), which have been confirmed through Sagnac [5]. SR therefore does not appear to represent complete reality.

The existence of the propagation medium [6] and the lack of credibility of Einstein's relativity [7] have been considered in Xi'an. *To predict events, according to cause and effect, the classical wave equation should be derived and solved with respect to its propagation medium.* This is true for vibrations in structures, sound propagation in air and EM waves in its propagation medium. The only difference between the classical and EM wave equation is that time and space shrink through motion according to the LT. But of course this does not affect its causality, providing the medium is retained. Through the medium's assumed absence, this derivation has not been attempted previously. The medium based theory established here is in accord with the LT, it is extended for both source and observer motion. It supports all the motional properties of the LT, plus important new motional optics resulting from individual source and observer motion.

[1]Author's email: selwyn.wright@ntlworld.com.

## 2. EM WAVE EQUATION

The electrical medium properties are: permeability (equivalent electrical density) $\mu$, dielectric permittivity (electrical rigidity) ę, EM impedance to the passage of light $z = \mu c$ and propagation speed $c = (\mu ę)^{-1/2}$. To describe EM disturbances propagating relative to its medium, Maxwell's [1] four electromagnetic field equations are used, vis:

$$\mathrm{div}D = \rho \quad \text{Gauss electrical} \tag{1a}$$
$$\mathrm{div}B = 0 \quad \text{Gauss magnetic} \tag{1b}$$
$$\mathrm{curl}E = -\partial B/\partial t \quad \text{Faraday inductive} \tag{1c}$$
$$\mathrm{curl}H = \partial D/\partial t + J \quad \text{Ampere-Maxwell} \tag{1d}$$

In the above equations, $D = ęE$, where $D$ is the electrical displacement and $E$ is the electric field. $B = \mu H$ where $B$ is the magnetic flux density and $H$ is the magnetic field, $\rho$ is the electric charge density and $J$ is the current density. In a vacuum, where there is no charge or current, $\rho = 0$, $J = 0$, $\blacktriangledown \cdot D = 0$, $\blacktriangledown \cdot B = 0$, then the last two equations above reduce to:

$$\blacktriangledown \times E = -\partial B/\partial t, \quad \blacktriangledown \times B = \mu ę \partial E/\partial t, \tag{2}$$

Here we have used the shorthand symbols div $\equiv (\blacktriangledown\cdot)$ and curl $\equiv (\blacktriangledown\times)$. If an electromagnetic source singularity of charge strength $Q(t)$ is introduced at the origin of the coordinate system, then one has:

$$\blacktriangledown \times E = -\partial B/\partial t, \quad \blacktriangledown \times B = \mu ę \partial E/\partial t + Q(t) \tag{3}$$

Again $1/\mu ę = c^2$ corresponds to the propagating speed squared, relative to the propagation medium and its electric field. Taking the curl of both sides of the first equation in (3) one has:

$$\blacktriangledown\times(\blacktriangledown\times E) = \blacktriangledown\times(-\partial B/\partial t) \quad \text{where} \quad \blacktriangledown\times(\blacktriangledown\times E) = \blacktriangledown(\blacktriangledown\cdot E) - \blacktriangledown^2 E = -\blacktriangledown^2 E \quad \text{as} \quad \blacktriangledown(\blacktriangledown\cdot E) = 0 \tag{4}$$

Inserting the right hand side (RHS) of Equation (3), into Equation (4), gives:

$$-\blacktriangledown^2 E = \blacktriangledown \times (-\partial B/\partial t) = -\partial/\partial t(\blacktriangledown \times B) = -c^{-2}\partial^2 E/\partial t^2 - \partial Q(t)/\partial t, \tag{5}$$

Thus the 'stationary' electromagnetic wave equation becomes:

$$\blacktriangledown^2 E - c^{-2}\partial^2 E/\partial t^2 = \partial Q(t)/\partial t, \quad \text{or} \quad \Box^2 E[t] = \partial Q[t]/\partial t, \quad \text{where} \quad \Box^2 = \blacktriangledown^2 - c^{-2}\partial^2/\partial t^2 \tag{6}$$

Generally, the outward solution for this second order differential equation is then:

$$E[\tau] = (4\pi R)^{-1}\partial Q[\tau]/\partial t, \quad \text{where} \quad \tau = t - R/c \tag{7}$$

Thus, we now have to distinguish between $t$ the familiar universal (observer) time and the observed source event time $\tau$. $R$ is the disturbance propagation distance between the source emission and the observer reception position in the medium. $\tau$ is sometimes called the retarded time because it can be expressed in terms of time $t$ and the propagation delay time $R/c$ between the source and observer, where $c$ is the propagation speed of light. $Q$ is the charge strength and the rate of change of charge $dQ/dt$ is the source strength.

## 3. EFFECT OF MOTION (CLASSICAL, GALILEAN)

A solution in general for classical sources and observers in motion relative to a propagation medium is available, Wright [8]. *According to this solution, one can assume the disturbance to be invariant and time to change through motion.* Equation (7), is then Galilean transformed as follows. Let the source and observer move at velocity '$s$' and '$o$', respectively, relative to the propagating medium. Let the source and observer time be represented by $t_s$ and $t_o$, respectively, and $t_p$ be the time in the stationary propagating medium space $x_p$. Following the format of Equation (6), the motional wave equation can then be written as:

$$\Box_p^2(\Delta_o E[t_p]) = \partial(\Delta_s Q[t_p])/\partial t_p, \quad \Box_p^2 = \blacktriangledown_p^2 - c^{-2}\partial^2/\partial t_p^2 \tag{8}$$

where the local motion operator $\Delta_o$ transforms, through observer motion, its local time as:

$$\begin{gathered}\Delta_o E[t_p] = \int^{d\to o} \delta_o(x - (ot_p - nd))H_o(x)H_o(d-x)E[t_p]dx = E[t_o],\\ \delta_o(\beta) = \infty \text{ if } \beta = 0, \quad \delta_o(\beta) = 0 \text{ if } \beta \neq 0, \quad H_o(\gamma) = 1 \text{ if } \gamma > 0, \quad H_o(\gamma) = 0 \text{ if } \gamma < 0\end{gathered} \tag{9}$$

For those who are not mathematically inclined, the $\delta$ function defines motion of an impulse representing a point source moving across a window $d$, and the $H$ functions confine motion within the window. Performing the integration in Equation (9), the relationship between $t_o$ and $t_p$ becomes:

$$t_o = \varepsilon_o^{-1} t_p, \quad \varepsilon_o = (1 - M_o), \quad M_o = o/c \tag{10}$$

Similarly for source motion:

$$\begin{aligned} &\Delta_s Q[t_p] = \int^{d \to o} \delta_s(x - (st_p - nd)) H_s(x) H_s(d - x) Q[t_p] dx = Q[t_s], \\ &t_s = \varepsilon_s^{-1} t_p, \quad \varepsilon_s = (1 - M_s), \quad M_s = s/c \end{aligned} \tag{11}$$

Therefore, the motional transforms $\Delta_o$ and $\Delta_s$ describe individually point observers and sources moving repeatedly across separate windows, in the propagating medium. On reaching the edge of the window they reappear at the start. They have, therefore, associated with their motion a frequency band or time spread. As the flight path $d$ goes to zero for a constant velocity, $n$ goes to infinity. The frequency band becomes a single Doppler shifted frequency or single time shift. *Thus time is transformed from that of a stationary to a moving frame.* These point moving (local motion) transforms have the property of describing motion at a specific point in the medium exactly. The dynamic motional operators $\varepsilon_s$ and $\varepsilon_o$ describe the effect of motion on source and observer time, respectively. Thus the stationary EM wave Equation (6), can be extended for both source and observer motion relative to the propagation medium, based on the local motion transforms, as:

$$\Box_p^2 E[t_o] = \partial(Q[t_s])/\partial t_p, \quad \Box_p^2 = \blacktriangledown_p^2 - c^{-2}\partial^2/\partial t_p^2 \tag{12}$$

Its solution, in the format of Equation (7), is then:

$$E[\tau_o] = (4\pi R)^{-1} \partial Q[\tau_s]/\partial t_p, \quad \tau_o = t_o - R_p/c \tag{13}$$

$\tau_o$ is the observed source event time, which is a different scale than the source event time $\tau_s$. The relationships between these temporary instantaneous event times, from Equations (10) and (11), now become:

$$\varepsilon_o \tau_o = \tau_p = \varepsilon_s \tau_s, \quad \text{or} \quad \tau_o = \varepsilon_s \varepsilon_o^{-1} \tau_s, \tag{14}$$

For more general arbitrary motion the situation is illustrated in Figure 1 where:

$$\varepsilon_s = 1 - M_s \cos\sigma_s, \quad \varepsilon_o = 1 - M_o \cos\sigma_o \tag{15}$$

Here $\sigma_s$ and $\sigma_o$ are the angles made between the individual source and observer motion directions and a line (propagation path) joining the emission and reception positions. $R_i$ is the initial separation distance, $R_p$ is the initial propagation distance and $R^*$ is the instantaneous propagation distance, between the source and observer. $t_o$ is the observer time which is equal to the propagation time $R_p/c$, plus the observer event time $\tau_o$. $M_s = s/c$ and $M_o = o/c$, where $M$ is used in honour of Mach [9], it is preferred to $\beta$ that is sometimes used.

## 4. KINEMATIC (LORENTZIAN) MOTION

We now include the permanent accumulative effect of Lorentz's time and space contraction through motion, relative to the medium. *It should be emphasized that it's not the medium that changes through motion. It's the material (the electronic structure of matter), in the moving system, that compresses in space and time*, i.e., any kind of system, made of atoms or molecules, including life itself. The fundamental effect of Lorentz's system transform, is to maintain the invariance of the propagating speed in the moving observer frame $c_o = x_o/\tau_o$ and the stationary propagation medium $c_p = x_p/\tau_p$. The theory is extended to include the moving source frame $c_s = x_s/\tau_s$. To accomplish this invariance, systems have to compress (time and space contract) in the moving frame, relative to the stationary propagation medium, by the same ratio $\alpha$, exactly, i.e.,

$$c_s = x_s/\tau_s = \alpha_s x_p/\alpha_s \tau_p = x_p/\tau_p = c_p = \alpha_o x_p/\alpha_o \tau_p = x_o/\tau_o = c_o \tag{16}$$

where

$$\tau_s = \alpha_s \tau_p, \quad \tau_o = \alpha_o \tau_p, \quad x_s = \alpha_s x_p, \quad x_o = \alpha_o x_p, \quad \alpha_s = (1 - M_s^2)^{1/2}, \quad \alpha_o = (1 - M_o^2)^{1/2} \tag{17}$$

Giving the kinematic time and space relationships between the source and observer as:

$$\tau_s/\alpha_s = \tau_p = \tau_o/\alpha_o \quad x_s/\alpha_s = x_p = x_o/\alpha_o \tag{18}$$

The conventional Mach numbers for source ($M_s$) and observer ($M_o$) motion relative to the medium are:

$$s = (x_p/\tau_p)_s, \quad M_s = s/c = (x_p/\tau_p)_s/c, \quad o = (x_p/\tau_p)_o, \quad M_o = o/c = (x_p/\tau_p)_o/c \tag{19}$$

And the hybrid (across frame) Mach numbers ($M_s^*$) and ($M_o^*$) between the stationary medium space $x_p$, and times $\tau_s$ and $\tau_o$ in the moving source and observer frames respectively, also can be defined as:

$$M_s^* = (x_p/\tau_s)_s/c = (x_p/\tau_p)_s/c\alpha_s = M_s/\alpha_s, \quad M_o^* = (x_p/\tau_o)_o/c = (x_p/\tau_p)_o/c\alpha_o = M_o/\alpha_o \tag{20}$$

*Although it is presumed that the speed of light cannot be exceeded, i.e., $M < 1$, within frames, it appears that it can be exceeded across frames, $M^* > 1$.* The speed and distance capability is much greater in the hybrid frame by $\alpha^{-1}$, where distance is measured in the stationary frame (real world space) and the slower time in the moving astronaut frame. As $M \to 1$, $\alpha \to 0$, and $M^* \to \infty$, allowing huge distances to be achieved in space travel, in a shorter time in the moving frame, without exceeding the speed of light in the stationary medium. Super-'lightic' speed across frames $M^* = 1$, occurs when $M \approx 0.7$. Combining the dynamic (Galilean) and kinematic (Lorentzian) motional effects, assuming the effects are multiplicative, in accordance with the Lorentz transform (LT), one has from Equations (14) and (18), the final form of the motional event time transform $K_t$ (rate of change of observer time compared to the source time) as:

$$(\varepsilon_s\tau_s)/\alpha_s = \tau_p = (\varepsilon_o\tau_o)/\alpha_o, \quad \text{or} \quad \tau_o/\tau_s = K_t = K_d \cdot K_k = \left(\varepsilon_o^{-1}\varepsilon_s\right)_d \left(\alpha_o\alpha_s^{-1}\right)_k \tag{21}$$

From Equations (13), (17) and (21) one has:

$$t_o = R_o/c + \tau_o, \quad R_o = \alpha_o R_p, \quad \tau_o = \int \left[\varepsilon_s\alpha_s^{-1}\varepsilon_o^{-1}\alpha_o\right] d\tau_s \tag{22}$$

## 5. ARBITRARY MOTION

Equation (22) gives the accumulative observer event time for given continuous arbitrary source and observer flight paths through the universe. The dynamic directional time change, affected by $\varepsilon_s\varepsilon_o^{-1}$, is caused through the wave dynamics (instantaneous Doppler) including the generation, transmission and reception of the EM waves. The kinematic omni-directional time change, determined by $\alpha_s^{-1}\alpha_o$, is caused through the instantaneous Lorentzian time contraction at the moving systems. *The dynamic and kinematic time changes are indistinguishable during motion.* However, the former ceases (is temporary), the latter remains (is permanent) after motion stops. The relationships between event time scales $\tau_s$, and $\tau_o$, for arbitrary source and observer motion, are illustrated in Figure 2. The vertical and horizontal axes represent time $t_p$ and space $x_p$, in the propagation

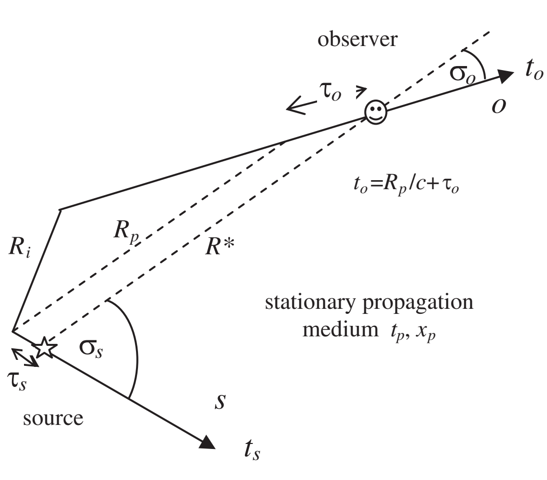



Figure 1: Source and observer flight paths, not in the same direction, moving with respect to a stationary propagation medium.

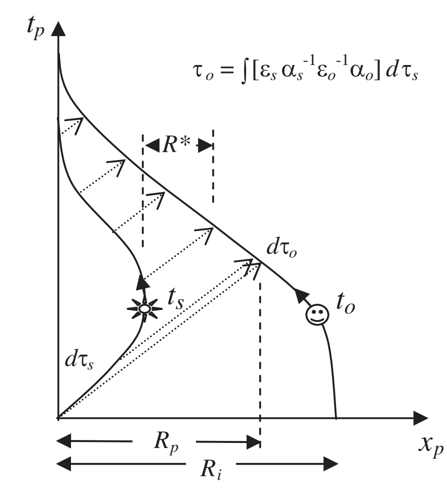



Figure 2: Relationship between source time $\tau_s$ and observer time $\tau_o$ for motion relative to a stationary propagation medium of time $t_p$.

medium respectively. Corresponding event times $\tau_s$ and $\tau_o$ are connected by propagation light paths. These paths represent light moving forward in time, from source to observer, going from left to right in the direction of the arrows. The gradient of the light paths is $1/c$ ($45^\circ$ if $x_p$ is in light years). *If the source and observer were interchanged, the light paths would go from right to left and have a negative gradient, giving an entirely different observed time history for the same flight paths.* This clearly distinguishes between source and observer motion, which Einstein's relativity cannot do. In these particular flight paths, the $t_s$ and $t_o$ curves start at the same time but at a different place. They end at a different time, but at the same place. *Solving the motional wave equation with respect to the propagation medium establishes causality, giving a definite direction of time, future and past.* Time can be quickened, slowed and reversed, according to the speed and direction of the source and observer relative to the medium. Einstein's SR without a medium, cannot satisfy causality or distinguish between source and observer motion.

*To actually participate in or change the source events as they happen, the observer must, of course, come in physical contact (interact) with the source, at the source event time. By definition this rules out any kind material or active participatory time travel.* According to this new causal theory, only visual time travel to the past is valid, defined to happen when the event time transform $K_t$ in Equation (21) is negative. This allows one to observe past events, but not to interfere with them, because they have already happened. If the actual speed of light can be avoided or bypassed, then perhaps one could observe an historic battle, or see one's ancestors rising from their graves, but not to interact or communicate with them. On the other hand, it is not possible (non causal) to participate presently in future source events, materially or visually, because they have not yet occurred. *Although generally it is not considered possible to travel faster than the speed of light in the medium, effective speeds greater than the speed of light appear possible across hybrid reference frames.* Here space is measured in the familiar universal propagation space, and time in the slower astronaut's time. In this way, vast distances can be achieved across the universe in a relatively short time. This can be expressed by $M^* = M/\alpha$, If $\alpha$ is say 0.01, then the motion can potentially be 100 times that of the speed of light. This indicates that distant galaxies, thought too far to reach in a lifetime, can now be colonized within a reasonable practical time, without exceeding the actual speed of light in the propagation medium. *However, again this is not time travel, it is just getting there quicker and ageing less.* Applications of the new theory, demonstrating the use of Figure 2, are given in [10].

## 6. CONCLUSION

A causal motional wave equation, for electromagnetic sources and observers in motion, with respect to its propagation medium has been derived and solved. Whereas, those aspects of Einstein's special relativity based on no propagation medium are considered irrational (non causal).

## REFERENCES


1. Maxwell, C. M., "A dynamical theory of the electromagnetic field," *Philosophical Transactions of the Royal Society*, Vol. 155, 459–512, London, 1865.
2. Lorentz, H. A., "Simplified theory of electrical and optical phenomena in moving systems," *Proc. Acad. Science Amsterdam*, Vol. 1, 427–442, 1899.
3. Poincaré, H., "La théorie de Lorentz et le principe de réaction," *Archives Néerlandaises des Sciences Exactes et Naturelles*, Vol. 5, 253–278, 1900.
4. Einstein, A., "On the electrodynamics of moving bodies," *Annalen der Physik*, Vol. 17, 891–921, 1905.
5. Sagnac, G., "L'éther lumineux démontré par l'effet du vent relatif d'éther dans un interféromètre en rotation uniforme," *Comptes Rendus*, Vol. 157, 708–710, 1913.
6. Wright, S. E., "Electromagnetic sources and observers in motion I — Evidence supporting the EM propagation medium for the transmission of light," *PIERS Proceedings*, 71–76, Xi'an, China, March 22–26, 2010.
7. Wright, S. E., "Electromagnetic sources and observers in motion II — Einstein's ether-less relativity versus Lorentz's medium based theory," *PIERS Proceedings*, 77–81, Xi'an, China, March 22–26, 2010.
8. Wright, S. E., "Sources and observers in motion I: Time variant analysis and implications to aerodynamic sound," *Journal of Sound and Vibration*, Vol. 108, No. 3, 361–378, 1986.
9. Mach, E., Mach's Principle, Mach's archive, Deutches Museum, Munich, 1887.
10. Wright, S. E., *Problems with Einstein's Relativity*, Trafford Publishing, 2010.

# Electromagnetic Sources and Observers in Motion IV — The Nature of Gravity and Its Effect on the Propagation Medium

**S. E. Wright**
ECASS Technologies Ltd., Moor Lane Laboratory, Kirkburton, Huddersfield, HD8 0QS, UK

**Abstract**— Considerable data indicates that the propagation medium (ether) exists (Paper I PIERS Xi'an). Also it appears that Einstein's ether-less relativity is not supported by the medium based Lorentz transform (Paper II PIERS Xi'an). Whereas, a third companion Paper III, at this conference, establishes a motional wave equation for both source and observer motion relative to the medium. This Paper IV investigates the effect of gravity, based on the medium's existence. Here it appears that source motion compresses source space and time, observer motion effectively expands them, and gravity compresses the medium and systems within. The medium therefore provides a link between Lorentz, accelerating frames and gravity, giving an optical explanation of the Equivalence Principle. The nature of gravity is considered, where it recognisees that finite structures of atoms and molecules can avoid singularities in their field equations. This allows gravity to be considered as a residual difference electrical field from dissimilar charges within these finite distributed atomic structures, linking large scales of gravity with the small atomic scale. Distinction is made also between binding mass, (energy required in constructing particle structures), and gravitational matter, (mass of complete atoms and molecules capable of generating the residual difference electrical fields producing gravity). It appears the medium is attracted locally by these gravitational fields, surrounding and moving with gravitational bodies. The gravitational field pervading throughout the universe, from all matter, will then create a net zero vector field, but a finite universal scalar field. This potential energy field appears to provide the absolute reference for the medium at rest in space and create the inertia for mass in motion.

## 1. INTRODUCTION

The propagation medium is shown to exist [1], its presence justified theoretically in [2], and the motional wave equation derived with respect to the medium in [3]. In this paper it is shown that the medium provides the common basis between the Lorentz transform [4], accelerating systems and gravity [5], providing an optical explanation of the Equivalence Principle. It also provides evidence that gravity is the residual difference electrical field between dissimilar charges within finite distributed atomic structures, providing the link between the electric field, large scales of gravity and the small atomic scale.

## 2. LORENTZ TRANSFORM, ACCELERATING SYSTEMS AND GRAVITY

According to the Lorentz transform (LT), systems compress (systems contract and time slows) moving relative to the medium, the medium remaining unchanged. This leads to the source compressing (time slowing) through motion, but effectively expanding (source time quickening) through observer motion (observer time actually slows, making the source time appear to quicken). This is an important property of observer motion, which Einstein's special relativity (SR) [6], cannot recognise. In the absence of gravity, *the system compression is based purely on motion with respect to the medium*, as summarized in Figure 1(a), considered in the companion Paper III and in further detail in [7]. There is nothing distinctive regarding motion, whether it be constant or accelerative motion. *Time slowing and system compression is concerned with the system's actual velocity at any instant with respect to the medium, compared to the speed of light. Not relative motion between systems, unable to detect constant motion, as considered by SR.*

Whereas, gravity actually compresses the medium and everything within it, according to Schwarzschild (8), considered in some detail in Section 3. As a simple example, Figure 1(b) illustrates the presence of a gravitational field '$g$', compressing the medium in the direction of the field (gravitational body). Here space contracts and time slows, bending a light ray perpendicular to the field, to the left, in the direction of the field, as depicted in the figure. If now an observer frame is accelerated by '$a$' to the left, in the direction of the field, time slows and space compresses at the observer. This makes the stationary medium appear to expand relatively, and its time quicken, making the light ray appear to bend to the right (away from the direction of motion), in the moving observer frame. When $a = g$ (freely falling) the ray will be vertical. It is taken for granted in a

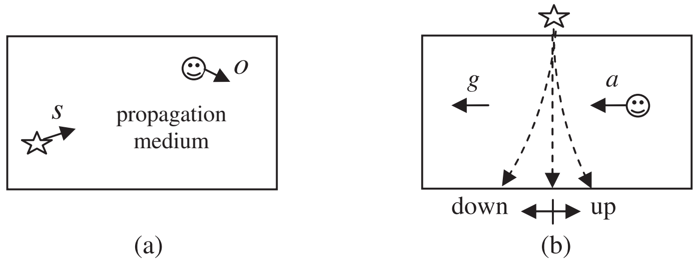



Figure 1: (a) Individual source and observer motion. Individual source speed '$s$' and observer speed '$o$', relative to the propagation medium locally compress time and space. Observed event times compress with source motion and expand with observer motion. (b) The effect of gravity '$g$' and observer acceleration '$a$'. Gravity compresses medium, observer acceleration effectively expands medium. Light ray bends left towards gravity '$g$', and to the right in the moving frame, through acceleration '$a$'. Ray is vertical when $a = g$.

falling lift (elevator), that there will be no visual effect of motion, everything will appear to be accelerating together.

However, to actually see individual items 'floating' in the moving frame there appears to be a further condition. The ray bending by the Earth's medium compression must be offset exactly by the ray bending through observer expansion in the accelerating lift. This is the physical explanation of the Equivalence Principle [9], based on an observer and a propagation medium. Einstein, not recognising the existence of the propagation medium, or the special expansion properties of a moving observer, cannot explain this phenomenon optically. The observed effect clearly depends on distinguishing between source and observer motion relative to the medium, which special relativity cannot do. *It is not that gravity attracts light; it is the medium through which the light travels that is compressed. If there was no medium there could be no gravitational medium compression or accelerating observer frame expansion of the medium.*

## 3. GRAVITATIONAL AND ACCELERATIVE TIME

Consider the source event time $\tau_{sg}$, in a gravitational field, in the direction of the field, compared to the general free field propagation time $\tau_p$. Because of the equivalent compression effect through motion, gravitational time can be expressed in a similar format used to calculate Lorentz's system time and space compression through motion, described in the companion paper III. For gravity, the medium and imbedded systems compress in the direction of the field (medium contracts, time slows), according to a gravitational compression operator $\delta_{sg}$, thus for $M \ll 1$:

$$\tau_{sg}=\delta_{sg}\tau_p,\ x_{sg}=\delta_{sg}x_p,\ \delta_{sg}=(1-M'_{sg})^{1/2}\approx(1-M'_{sg}/2),\ \Delta\tau_{sg}=\Delta\delta_{sg}\tau_p=-(M'_{sg}/2)\tau_p=-M_{sg}\tau_p \quad (1)$$

$$M'_{sg}=2Gm/c^2R=2gR/c^2=2(gR/c)/c=2v/c=2M_{sg} \text{ where } g=Gm/R^2,\ v=gt,\ t=R/c,\ M_{sg}=v/c \quad (2)$$

$M'_{sg}$ is equivalent to twice the effective Mach number (velocity over the speed of light, $M_{sg} = v/c$), its value was established by Schwarzschild, $g$ is the acceleration due to gravity (at the surface of the Earth $g = 9.8\,\mathrm{m/s^2}$). $G$ is the gravitational constant $6.67 \times 10^{-11}\,\mathrm{Nm^2/kg^2}$ ($G/c^2 = 7.4 \times 10^{-28}$), $m$ is the gravitational mass of the body and $R$ is the radial distance from the body centre. The event horizon $R'$ radius from the centre of a gravitational body or singularity (black hole), inside of which light cannot escape, is then defined when $M'_{sg} = 1$. From Equation (2):

$$R' = 2Gm/c^2 = M'_{sg}R \quad (3)$$

From Equations (1) and (2), the reduced time ageing (time slowing) through gravitational medium compression, is given approximately by $\Delta\tau_{sg} = \Delta\delta_{sg}\tau_p$ where $\Delta\delta_{sg} \approx -M_{sg} = -Gm/c^2R$. For example in Figure 2 the gravity on the Sun's surface of mass $m = 2\times10^{30}$ kg and radius $R = 7\times10^8$ m gives $\Delta\delta_{sg} = -(G/c^2)m/R = 7.4 \times 10^{-28} \times 2 \times 10^{30}/7 \times 10^8 = 2.1 \times 10^{-6}$ and for a 24 hour period $\tau_p = 8.6 \times 10^4$ s, $\Delta\tau_{sg} = -2.1 \times 10^{-6} \times 8.6 \times 10^4 = -180$ ms/day slowing compared to free space. On the Earth's surface, from its own gravity, $m = 5.98 \times 10^{24}$ kg, $R = 6.37 \times 10^6$ m, $\Delta\tau_{sg} = -0.06$ ms/day slowing. Also on the Earth's surface a distance of $R = 1.5 \times 10^{11}$ m from the Sun, $\Delta\tau_{sg} = -0.74$ ms/day.

A gravitational potential change $-M_{sg}$ across a small distance $\Delta R = R_1 - R_2$ in the direction of field ($R_1 > R_2$), from Equations (1) and (2) is given below. Light propagating a distance $\Delta R$

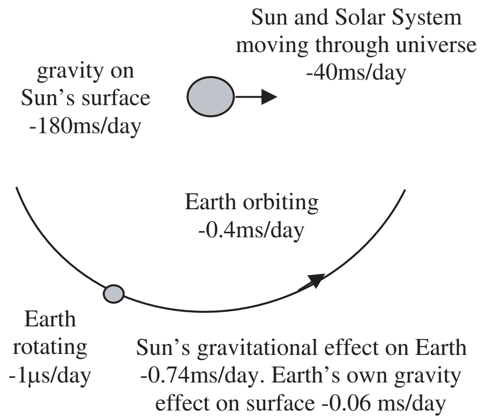



Figure 2: Earth orbiting around the Sun illustrating time slowing through gravity and motion.

in the medium, against (+) (wavelength increase) or with the direction of the gravitational field (-) (wavelength decrease), is stretched or compressed, its wavelength changing by $\Delta\lambda$, giving red shifted or blue shifted light respectively, thus:

$$\Delta\tau_{sg}/\tau_p = -\Delta M_{sg} = -\left(Gm/c^2\right)(1/R_2 - 1/R_1) \approx -\left(Gm/c^2\right)(R_1 - R_2)/R^2 = -g\Delta R/c^2 \quad (4)$$

$$\text{giving} \quad \Delta\lambda/\lambda = \Delta\tau_{sg}/\tau_p \approx \pm\Delta M_{sg} \approx \pm g\Delta R/c^2 \quad (5)$$

For example, a vertical height rise of $\Delta R = 10\,\text{km}$ above the Earth's surface, against the direction of the Earth's field, $\Delta M_{sg} = g\Delta R/c^2 = 9.8 \times 10^4/(3 \times 10^8)^2 = 1.1 \times 10^{-12}$ and for a 24 hour period, $\Delta\tau_{sg} = 1.1 \times 10^{-12} \times 8.6 \times 10^4 = 95 \times 10^{-9}\,\text{s}$, i.e., a time difference (quickening) of 95 ns/day. *The effect of gravitational acceleration (compression of the medium) is dependent on how close an effective velocity (acceleration through a given distance) is compared to the velocity of light, similar to systems in motion in Paper III.* For a system of acceleration '$a$' moving a distance $d$, the time $t$ for a light wave to reach $d$ is $d/c$. The velocity attained by the system in this time is $v = at = ad/c$. The Mach number is then $M = v/c = ad/c^2$. Thus gravity and accelerating systems are equivalent, their time and space compression depend on an equivalent velocity relative to the medium. For $M \ll 1$, the time change for example an accelerating observer, through acceleration, in the format of Equation (1), becomes:

$$\Delta\tau_{oa} = \Delta\delta_{oa}\tau_p, \quad \Delta\delta_{oa} = -M_{oa}, \quad \text{where} \quad M_{oa} = ad/c^2 \quad (6)$$

Equation (6) is equivalent to Equation (4), where $M_{oa} = M_{sg}$, $a = g$, and $d = \Delta R$. Thus the effect of medium compression through gravity and expansion through observer acceleration, are equivalent, but opposite effects. One can now obtain, in a simple potential addition, the total time change $\Delta\tau_T$, though time slowing caused by constant source motion (Lorentz compression) $\Delta\tau_{s\alpha}$, through time quickening caused by observer acceleration, (expansion) $\Delta\tau_{o\alpha}$, and time slowing through local gravity (compression) $\Delta\tau_{sg}$, thus:

$$\Delta\tau_T \approx \Delta\tau_{s\alpha} - \Delta\tau_{oa} + \Delta\tau_{sg} \quad (7)$$

where for $M \ll 1$

$$\Delta\tau_{s\alpha} = \Delta\alpha_{s\alpha}\tau_p, \quad \Delta\alpha_{s\alpha} \approx -M_{s\alpha}^2/2, \quad M_{s\alpha} = s/c, \quad \text{(constant source motion)} \quad (8)$$

$$\Delta\tau_{oa} = \Delta\delta_{oa}\tau_p, \quad \Delta\delta_{oa} \approx -M_{oa}, \quad M_{oa} = ad/c^2 \quad \text{(observer acceleration)} \quad (9)$$

$$\Delta\tau_{sg} = \Delta\delta_{sg}\tau_p, \quad \Delta\delta_{sg} \approx -M_{sg}, \quad M_{sg} = Gm/c^2R = g\Delta R/c^2 \quad \text{(gravity)} \quad (10)$$

In the absence of gravity $g = o$, time and space compression are given according to the first and second terms respectively, on the right hand side in Equation (7). Equations (8) and (9) can be interchanged depending on whether the source or observer is accelerating or moving at constant velocity. It can be seen that acceleration is dependant on $M$, whereas constant motion (LT) is dependent on $M^2$, which is much weaker for $M \ll 1$. The Equivalence Principle is given by the last two terms in Equation (7). For $\Delta R = d$ and $g = a$ (free fall), gravity is completely neutralized. *Thus for objects to be observed as freely falling (floating), as mentioned earlier, there is a further requirement. The light deflection on Earth through compression of the medium by gravity, must be equalled through the kinematic expansion by observer motion. All three effects, gravity, accelerating frames and the Lorentz transform (constant motion) are based on motion or equivalent motion with respect to the propagation medium.* In Figure 2, using Equation (8), illustrates the time slowing for the Sun's motion ($M \approx 1 \times 10^{-3}$) through the universe. Here $\Delta\tau_{s\alpha} = \Delta\alpha_{s\alpha}\tau_p$, $\Delta\alpha_{s\alpha} \approx -M_{s\alpha}^2/2 = -(1 \times 10^{-3})^2/2 = -0.5 \times 10^{-6}$, and for a 24 hour period $\tau_p = 8.6 \times 10^4$ s giving $\Delta\tau_{s\alpha} = -0.5 \times 10^{-6} \times 8.6 \times 10^4\,\mathrm{s} \approx -40\,\mathrm{ms/day}$. The Earth's orbital speed ($M \approx 1 \times 10^{-4}$) around the Sun gives $\Delta\tau_{s\alpha} \approx -0.4\,\mathrm{ms/day}$, and the Earth's surface rotational velocity ($M \approx 1.5 \times 10^{-6}$) becomes $\Delta\tau_{s\alpha} \approx -0.1\,\mu\mathrm{s/day}$. From crude simple addition the total time slowing on the Sun's surface compared to free space is $\approx -180 - 40 = -220\,\mathrm{ms/day}$ and on the Earth's surface time slowing $\approx -40.0 - 0.40 - 0.74 - 0.06 \approx -41.20\,\mathrm{ms/day}$. Without a medium none of these effects could occur. Further examples of source and observer motion are given in [7].

## 4. THE NATURE OF GRAVITY

The fact that gravity has similarities to the electric field (intensity reduces according the inverse square law with distance, and its effect is not instantaneous at a distance i.e., it is delayed according to the light propagation retarded time — distance divided by the speed of light, one must suspect that the two are connected. Trying to unite the large scales of gravity with the small scales of atoms and molecules, using 'point' atomic models, causes singularities (infinities) in their field equations. *However, it appears that the singularities can be avoided, quite naturally, by replacing the compact models with non-compact finite atomic structures, creating a link between the electric field, gravity and the atomic structure.* Rather than representing matter by hypothetical strings, which has been one approach, perhaps a more physical and realistic description is to consider the actual finite dimensions of these distributed atomic structures.

According to the Standard Model these structures are considered to contain 16 elementary subatomic particles, 12 fermions which have mass, and 4 bosons which are force carriers, plus a second set of antiparticle equivalents. In simple terms, the fermions have 6 varieties of quarks (up-down, strong-charm and bottom-top), and 6 flavours of leptons (electron, muon, tau, and neutrino versions of the electron, muon, tau). Fortunately, most of the matter in the universe is thought to be comprised of just negatively charged electrons and two types of charged quarks, having fractional values of charge $q$, an up quark '$u$' ($q = 2/3$) and a down quark '$d$' ($q = -1/3$). Regular atoms have nuclei of positively charged protons ($uud$), $q = 2/3 + 2/3 - 1/3 = 1$ and neutrons ($ddu$), $q = -1/3 - 1/3 + 2/3 = 0$. Equalled by negative electrons ($-q$) orbiting around the nucleus as there are protons ($+q$) in the nucleus.

Generally, it appears that a distribution of positive and negative charges, will always produce internally an attractive force between dissimilar charges, and externally a micro comb like (alternate polarity), residual difference electric field. This field will tend to zero as the distribution size goes to zero. *But for a finite size atomic distribution, this field, although small, remains finite and is capable of attracting similar, dissimilar charge distributions (gravity), explaining for the first time how gravitational matter always attracts itself.* Recognizing the finite distribution of real subatomic particles, atoms and molecules; their opposite charges and dissimilar fields not quite coincident at a point, produce a weak, steady, but finite atomic residual difference field (ARDF). Although there could be weak residual fields from individual quarks, the main contributor to this field appears to be the proton $q = 1$ and the electron $q = -1$.

*This internal ARDF appears to cause the close attraction between atoms and molecules (mercury beads and soap films pulling together), and a very weak external ARDF creating gravity at a distance.* The larger the atomic number (number of orbiting electrons), the larger the interaction between atoms, the stronger the ARDF and the heavier the element tends to become. Equating the electric field $E$, given by $E = kqr'/r^2$ where $k = 8.99 \times 10^9\,\mathrm{Nm^2/C^2}$ and the gravity field given by $g = Gmr'/r^2$ where $G = 6.67 \times 10^{-11}\,\mathrm{Nm^2/kg^2}$, and $r'$ is the unit vector along the distance radius $r$, the ratio of gravitational matter $m$ to charge $q$ becomes $m/q = k/G = 8.99 \times 10^9/6.67 \times 10^{-11} = 1.35 \times 10^{20}\,\mathrm{kg/C}$. It can be seen that gravitational mass $m$ is exceedingly less attractive than charge $q$, i.e., vastly more $m$ is required than $q$ to create the same strength $g = E$, illustrating the weakness of gravitational (residual electric) fields.

## 5. MASS, MATTER AND REFERENCE FIELD

As Einstein kept reminding us, all forms of energy (potential and kinetic) are equivalent to mass. It appears that binding mass of subatomic particles, complete atoms and molecules results through the energy required to construct these systems. For example, in the proton only 2% of its mass is from its individual components, the other 98% is accountable through the strong binding force (energy) of the gluons holding the proton together. Inertia appears to be the propagation medium's resistance to the change in motion (acceleration) of mass. Examples are accelerating electrons in orbit around an atom, J. J. Thompson's e/m experiment and ions moving in mass spectrometers. Here gravity has no part, just the acceleration of mass (no $g$, or $G$ involved). *Therefore mass does not have to have anything to do with gravity. Gravitational attraction seems to arise mainly through finite distributions of dissimilar charges within gravitational matter (atoms and molecules).* However, even neutrons, which are nominally neutral for a 'point' structure, their arithmetic charge being zero, should result in a very weak but finite residual field from individual quarks. Although neutron stars are nominally neutral, to create their intense magnetic and gravitational fields, it appears, under their huge compacting pressures, they assume some kind of charge reconstitution of their quarks, resulting in charges, currents and fields.

Whatever the actual details of the atomic residual difference field (ARDF), it can be considered to be finite from the almost limitless gravitational matter throughout the universe. It appears same phase potential fields reduce according to the inverse of distance and dissimilar fields reduce with the inverse square of their distance from gravitational sources. Whereas, the effect of known and inferred gravitational matter from any point in the universe tends to increase at least as the square of the distance, appearing to maintain the field strength. Only through symmetry, from all directions in space, at any point, does the ARDF intensity tend to zero, whereas its potential energy remains finite. *Thus from the total gravitational matter in the universe, the ARDF should provide a well defined universal gravitational reference field (UGRF), whose intensity becomes zero but retains a uniform finite potential energy throughout space. This scalar field, at rest in space, appears to contribute to the dark energy, provide the mass inertia and the reference for the propagation medium.* Unlike EM and gravity waves which are time varying, the UGRF, from steady charges throughout the universe appears to be a steady continuous (non-quantised) diffuse field, making it difficult to detect. This field in some respects, resembles the elusive Higgs field [10].

Newton [11] claimed that accelerative motion was with respect to absolute space. Mach [12] claimed that it was with respect to the distant matter of the universe. They both appear to be right. There will be local variations, but for most reference purposes, absolute space, the propagation medium at rest in space and the universe are considered to be the same thing. *In this sense, it appears that Newton's claim of absolute acceleration, can be confirmed, and extended to all forms of motion, including constant motion with respect to the medium and its UGRF. This enables absolute motion to be measured relative to free space and the medium.* In summary, i) Light is a varying $E$ field from varying charges. ii) The ARDF is the residual steady difference $E$ field from basically proton and electron charges. iii) Gravity is generated from the ARDF from large quantities of dissimilar charges in atoms and molecules. iv) The UGRF is the ARDF from the total gravitational matter in the universe. v) The UGRF provides the reference (energized field) for inertia and the propagation medium at rest in space. *All these effects have a common basis i.e., they are all electrical, they all seem to depend on the electric charge.* Without charge there could be no electric field, light, ARDF (gravity) or UGRF (universal reference).

## 6. CONCLUSION

It appears that the propagation medium provides a common link between gravity, accelerating frames and the LT, through an equivalent velocity. In the LT, time and space compress in systems moving relative to the medium. Observer acceleration, with respect to the medium, effectively expands the medium's time and space. Whereas, acceleration through gravity compresses time and space, in both the propagation medium and the systems within, offsetting the observer acceleration effect. This explains the Equivalence Principle through observations relative to the propagation medium. Taking into account the finite structure of atoms and molecules, rather than using point models, removes the singularities in the field equations. Gravity then appears to be the residual difference electric field from the dissimilar charges within these finite atomic structures, linking the electric field, the large scales of gravity and the small atomic scale. Distinction is made between binding mass (energy required to build atomic particles and structures) and gravitational matter (mass of complete atoms and molecules providing the residual difference field from dissimilar charges within). The gravitational field, from the total matter in the universe, appears to generate a field throughout the universe with zero net intensity but finite energy. This field which could be involved in the dark energy problem, appears to provide the reference for the medium at rest in space and the inertia for mass in accelerative motion.

**REFERENCES**


1. Wright, S. E., "Electromagnetic sources and observers in motion I — Evidence supporting the EM propagation medium for the transmission of light," *PIERS Proceedings*, 71–76, Xi'an, China, March 22–26, 2010.
2. Wright, S. E., "Electromagnetic sources and observers in motion II — Einstein's ether-less relativity versus Lorentz's medium based theory," *PIERS Proceedings*, 77–81, Xi'an, China, March 22–26, 2010.
3. Wright, S. E., "Electromagnetic sources and observers in motion III — Derivation and solution of the electromagnetic motional wave equation," *Progress in Electromagnetics Research Symposium*, Cambridge, USA, July 5–8, 2010, submitted.
4. Lorentz, H. A., "Simplified theory of electrical and optical phenomena in moving systems," *Proc. Acad. Science Amsterdam*, Vol. 1, 427–442, 1899.
5. Einstein, A., *Die Feldgleichungen der Gravitation*, 844–847, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, Berlin, 1915.
6. Einstein, A., "On the electrodynamics of moving bodies," *Annalen der Physik*, Vol. 17, 891–921, 1905.
7. Wright, S. E., *Problems with Einstein's Relativity*, Trafford Publishing, 2010.
8. Schwarzschild, K., "Über das gravitationsfeld eines massenpunktes nach der einstein'schen theorie," *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, Vol. 1, 189–196, 1916.
9. Einstein, A., "Über das relativitätsprinzip und die aus demselben gezogene folgerungen (on the relativity principle and the conclusions drawn from it)," *Jahrbuch der Radioaktivitaet und Elektronik*, Vol. 4, 1907.
10. Higgs, P. W., "Brocken symmetries, mass-less particles and gauge fields," *Phys. Lett.*, Vol. 12, 132, 1964.
11. Newton, I., *Philosophiæ Naturalis Principia Mathematica*, 1689.
12. Mach, E., Mach's Principle, Mach's archive, Deutches Museum, Munich, 1887.

# Matrix Converter Induction Motor Drive Employing Direct Torque Control Method

**J. Lettl and D. Kuzmanovic**
Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

**Abstract**— The presented paper deals with the implementation of the Switching Table Based Direct Torque Control (ST-DTC) in the matrix converter-fed induction motor drive. In order to improve the control of the matrix converter input power factor, the use of the simplified Indirect Space Vector Modulation (ISVM) was proposed. To verify the feasibility of the proposed solution, experimental results were presented.

## 1. INTRODUCTION

The matrix converter belongs to direct AC/AC converters, i.e. converters that transform energy from one AC power system into another one without an intermediate DC link with energy storage elements. It enables achieving sinusoidal output current and voltage while maintaining sinusoidal input current with adjustable input power factor. The absence of the energy storage elements reduces the size of the converter, which makes it suitable for applications with limited installation space, like in electric vehicles.

The power part of the matrix converter consists of a $3 \times 3$ matrix of bi-directional switches (Fig. 1) which enables switching of 27 switching states (vectors): 3 synchronous rotating vectors, 3 inverse rotating vectors, 18 active (pulsating) vectors, and 3 zero vectors. The synchronous vectors correspond to direct connection of the input and output terminals. Active vectors are stationary with amplitude depending on the selected input line-to-line voltage. They can have one of the six possible directions, mutually shifted by 60 (Fig. 2(a)). If permanently switched, one such combination would produce a vector which pulsates between its maximum positive and its maximum negative values. The input current vectors corresponding to these switching states are shown in Fig. 2(b).

The ST-DTC or Takahashi DTC was proposed by Takahashi and Noguchi in 1986 [3]. The block diagram of the ST-DTC controller is presented in Fig. 3. The estimated torque and flux are compared with their reference values, and if they are out of their tolerance bands, an adequate voltage vector is applied to the motor stator windings. The flux is controlled by a two-level hysteresis controller and the torque is controlled by a three-level hysteresis controller. The outputs of these controllers ($c_\Psi$ and $c_T$), together with the current flux vector sector ($k_{S\text{-}\Psi}$), are used in the *Voltage Vector Selection Table* block to select the appropriate voltage vector from Table 1.

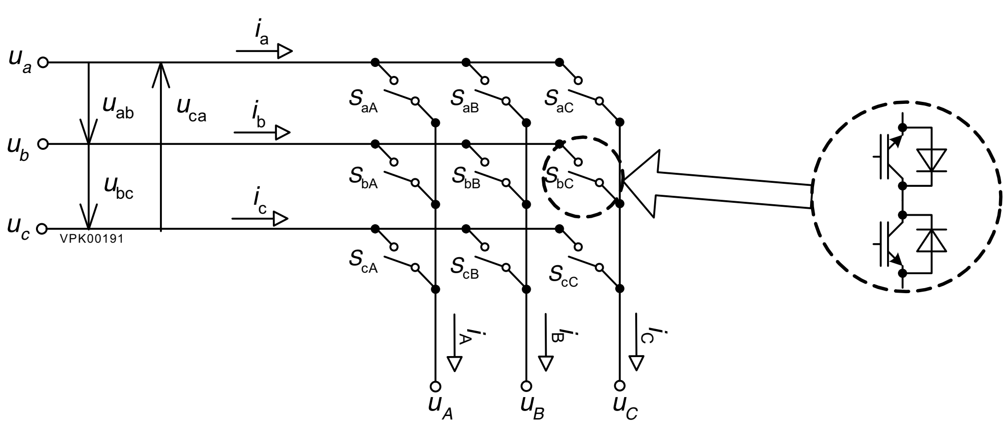
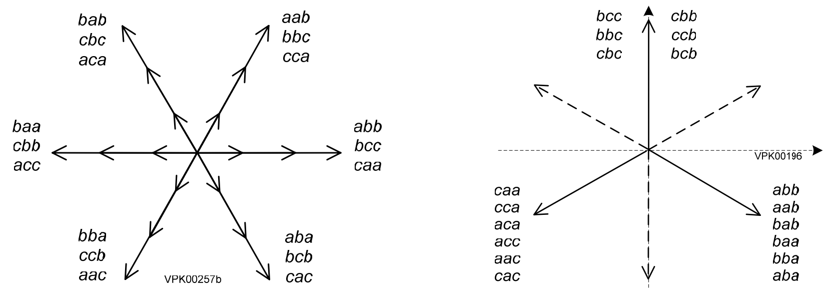



Figure 1: Matrix converter.

## 2. DTC USING MATRIX CONVERTER

While the implementation of the Field Oriented Control (FOC) in the matrix converter-fed drives is quite straightforward, in terms that it doesn't differ much from the implementation with the VSI [7], the implementation of the DTC requires modifications of the modulation algorithm.




(a) (b)

Figure 2: (a) Overview of output voltage vectors and (b) input current vectors.

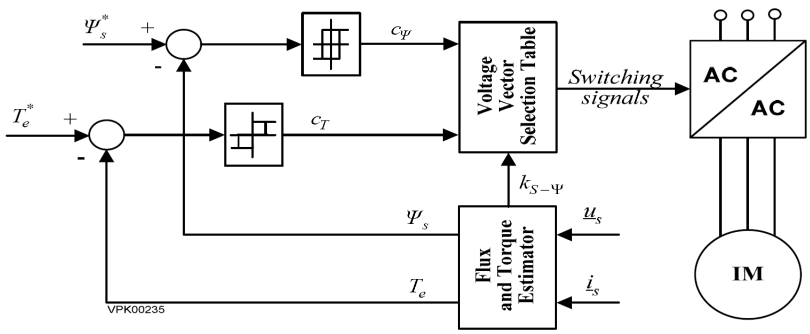



Figure 3: ST-DTC block diagram.

Table 1: Voltage vectors selection table.

| $k_{S\text{-}\Psi}$ | | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $c_\Psi = 0$ | $c_T = 1$ | $\underline{u}_2$ | $\underline{u}_3$ | $\underline{u}_4$ | $\underline{u}_5$ | $\underline{u}_6$ | $\underline{u}_1$ |
| | $c_T = 0$ | $\underline{u}_7$ | $\underline{u}_0$ | $\underline{u}_7$ | $\underline{u}_0$ | $\underline{u}_7$ | $\underline{u}_0$ |
| | $c_T = -1$ | $\underline{u}_6$ | $\underline{u}_1$ | $\underline{u}_2$ | $\underline{u}_3$ | $\underline{u}_4$ | $\underline{u}_5$ |
| $c_\Psi = 1$ | $c_T = 1$ | $\underline{u}_3$ | $\underline{u}_4$ | $\underline{u}_5$ | $\underline{u}_6$ | $\underline{u}_1$ | $\underline{u}_2$ |
| | $c_T = 0$ | $\underline{u}_0$ | $\underline{u}_7$ | $\underline{u}_0$ | $\underline{u}_7$ | $\underline{u}_0$ | $\underline{u}_7$ |
| | $c_T = -1$ | $\underline{u}_5$ | $\underline{u}_6$ | $\underline{u}_1$ | $\underline{u}_2$ | $\underline{u}_3$ | $\underline{u}_4$ |

A DTC scheme for the matrix converter-fed induction motor drive was first proposed in [1] and [2]. It uses a capability of the matrix converter to generate 18 active voltage vectors (unlike 6 in the case of VSI) to control the average value of the input power factor in addition to the torque and flux. The hysteresis torque and flux control works in the same way as in the ST-DTC for VSI-fed induction motors (Fig. 3). Based on the torque and flux errors and the flux vector sector, a suitable voltage vector is selected from Table 1. In the case of the VSI, this vector is generated by switching one of the six active and two zero switching combinations. As seen in Fig. 2(a), for every VSI vector, there are always three matrix converter switching combinations that can generate a parallel vector with the same polarity. From these three switching combinations, the two that will produce vectors with maximum modules are used by the DTC. From this analysis it can be seen that for every vector selected by the basic DTC algorithm, there are two corresponding matrix converter switching combinations. This redundancy is used to control the average value of $\sin\varphi_{in}$, where $\varphi_{in}$ is the displacement angle of the input current vector towards the input voltage vector.

The average value of $\sin\varphi_{in}$ is controlled by an additional two-level hysteresis controller. The average value of $\sin\varphi_{in}$ is obtained by applying a low-pass filter to its instantaneous value. If the average value of $\sin\varphi_{in}$ is positive ($c_{\sin\varphi_{in}} = +1$), which means that the input current vector is lagging the input voltage vector, the switching combination that will decrease $\sin\varphi_{in}$ will be applied. If the average value of $\sin\varphi_{in}$ is negative ($c_{\sin\varphi_{in}} = -1$), which means that the input current vector is leading the input voltage vector, the switching combination that will increase $\sin\varphi_{in}$ will be applied.

## 3. IMPROVED DTC USING MATRIX CONVERTER

A problem with the DTC scheme for the matrix converter described in the previous section is a relatively slow and indeterministic control of the displacement angle $\varphi_{in}$ of the input current vector towards the input voltage vector, especially at low speed, where torque increases sharply under the influence of an active vector, but declines slowly when a zero vector is switched. Namely, under the influence of an active vector, the torque often increases from its lower tolerance band limit to the upper one within only one sampling period, but its decrease can take a few sampling periods, which means that zero vectors are switched most of the time. Under these conditions, it is hard to achieve a good control of $\varphi_{in}$, which is influenced only by the active vectors. The result is a high harmonic content in the input current.

In this paper, an ST-DTC scheme using simplified ISVM was proposed. The ISVM is based on the indirect representation of the matrix converter proposed in [4]. It treats the matrix converter as a combination of an input virtual rectifier and an output virtual inverter stages, connected by a virtual DC link (Fig. 4). The indirect matrix converter representation enables applying conventional SVM techniques to the rectifying and inverting stages (Fig. 5) [5, 6]. The virtual rectifier stage SVM generates the input current vector with selected phase displacement with respect to the input voltage vector and at the same time generates the virtual DC link voltage. The virtual inverter stage then uses this virtual DC link voltage to generate the output voltage vector. The operation of the virtual rectifier and inverter stages is synchronized by applying a suitable switching pattern.

In the proposed scheme, the selection of the output voltage vector from Table 1 is done in the same way as in the conventional ST-DTC for VSI-fed induction motors. This vector will then be generated by the simplified ISVM. In the simplified ISVM, the virtual rectifier stage will insure proper orientation of the input current vector and therefore will keep the reference power factor. The inverter stage will switch the selected voltage vector over the entire sampling period. The virtual rectifier stage duty cycles are calculated from Equation (1), where $\theta_{S\text{-}I}$ is the input current vector angle within a sector and its value can be between 0 and $\pi/3$. The structure of the switching

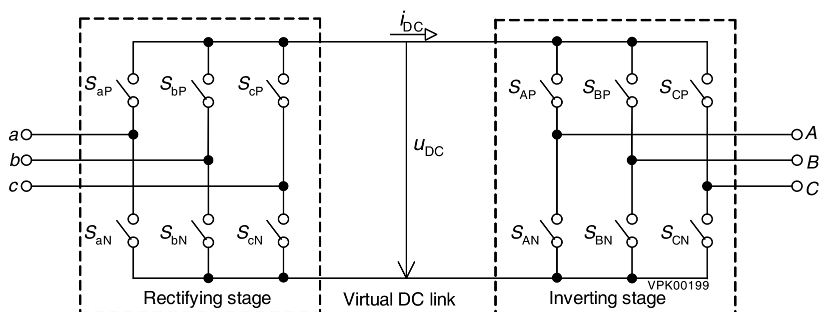



Figure 4: Matrix converter equivalent circuit for indirect modulation.

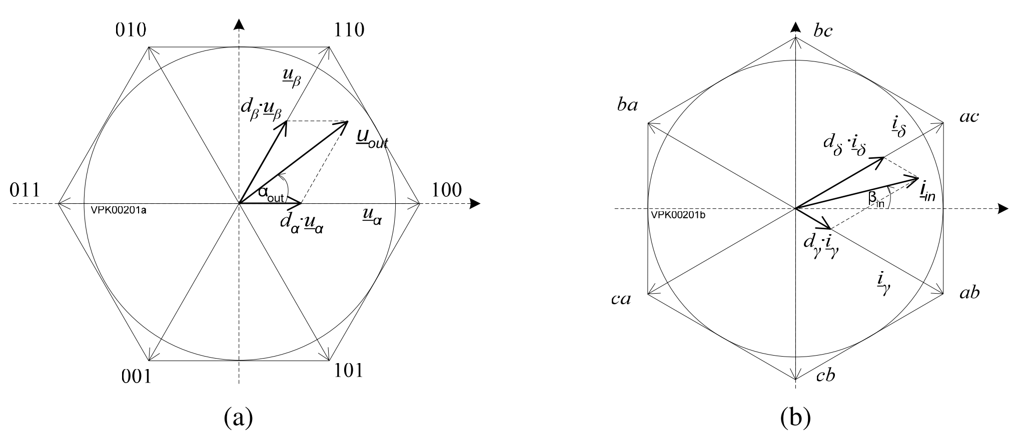



Figure 5: (a) Inverter stage SVM and (b) rectifier stage SVM.

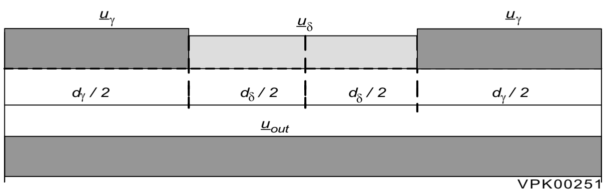



Figure 6: Simplified ISVM switching pattern structure.

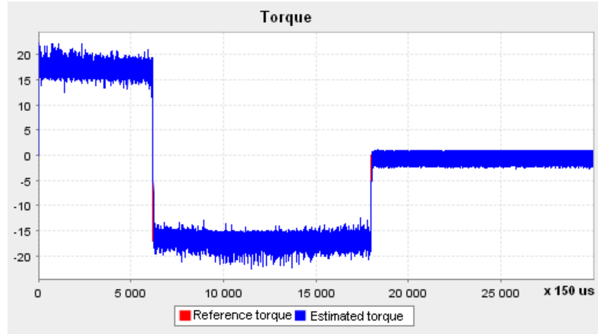
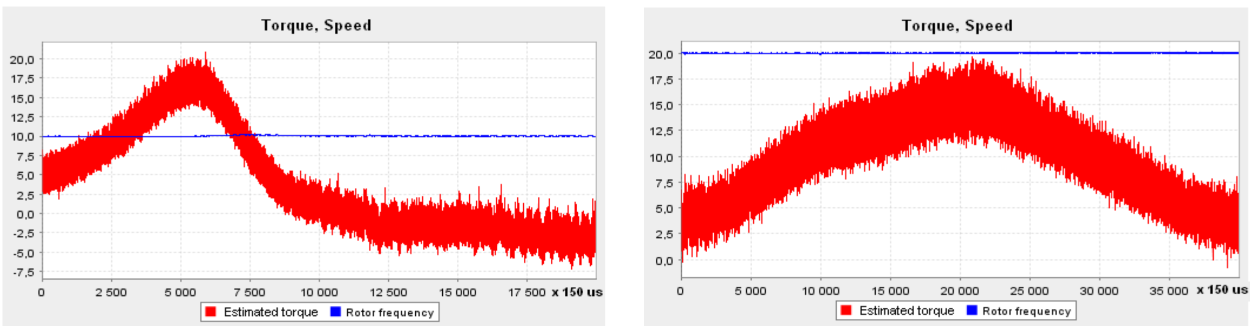



Figure 7: Torque control [Nm] (reference torque changes in the following sequence $0 \rightarrow 17 \rightarrow -17 \rightarrow 0$ Nm).




(a) (b)

Figure 8: (a) Speed control [rps]: load increase at 10 rps and (b) load increase at 20 rps.

pattern is illustrated in Fig. 6

$$d_\gamma = \frac{\sin\left(\frac{\pi}{3} - \theta_{S\text{-}I}\right)}{\sin\left(\frac{\pi}{3} - \theta_{S\text{-}I}\right) + \sin(\theta_{S\text{-}I})} \quad d_\delta = 1 - d_\gamma \tag{1}$$

## 4. EXPERIMENTAL RESULTS

The proposed scheme was first verified by numerical simulations [8] and then implemented on a test bed consisting of a 12 kW matrix converter and a standard 3 kW, 4-pole, 50 Hz cage induction motor. The results of the practical implementation are presented in Fig. 7 and in Fig. 8.

## 5. CONCLUSIONS

An improved ST-DTC scheme for the matrix converter using the simplified ISVM was proposed in this paper. The virtual rectifier stage SVM inherently insures proper input current vector orientation and maintaining the reference power factor. The virtual inverter stage switches the selected voltage vector in the same way as in the inverter-fed drives. The experimental results show good torque and speed control.

## REFERENCES


1. Casadei, D., G. Grandi, and G. Serra, "The use of matrix converters in direct torque control of induction machines," *Proceedings of the 15th Annual Conference of the IEEE Industrial Electronics Society, IECON 98*, 744–749, Philadelphia, USA, November 1998.
2. Casadei, D., G. Serra, and A. Tani, "The use of matrix converters in direct torque control of induction machines," *IEEE Transactions on Industrial Electronics*, 1057–1064, December 2001.
3. Takahashi, I. and T. Noguchi, "A new quick-response and high-efficiency control strategy of an induction motor," *IEEE Transaction on Industry Applications*, 1820–1827, September 1986.
4. Ziogas, P. D., S. I. Khan, and M. H. Rashid, "Analysis and design of forced commutated cycloconverter structures with imporved transfer characteristics," *IEEE Transactions on Industrial Electronics*, 271–280, August 1986.

5. Huber, L. and D. Borojevic, "Space vector modulator for forced commutated cyclo-converters," *IEEE Industry Applications Society Annual Meeting*, 871–876, October 1989.
6. Huber, L. and D. Borojevic, "Space vector modulation with unity input power factor for forced commutated cycloconverters," *IEEE Industry Applications Society Annual Meeting*, 1032–1041, September 1991.
7. Kuzmanovic, D. and J. Lettl, "Matrix converter-fed induction motor drive controlled by field oriented control," *Proceedings of the 16th Electronic Devices and Systems — IMAPS CS International Conference, EDS 09*, 473–478, Brno, Czech Republic, September 2009.
8. Kuzmanovic, D. and J. Lettl, "Matrix converter-fed induction motor drive controlled by takahashi direct torque control," *Proceedings of the 8th Electronic Conference of Young Research and Scientific Workers, TRANSCOM 09*, 55–58, Zilina, Slovak Republic, June 2009.

# Some Consequences of the Non-constancy of the Speed of Light in Vacuum for Different Galilean Reference Systems

**Namik Yener**
Technology Faculty, Umuttepe Campus, Kocaeli University, Izmit, Kocaeli 41380, Turkey

**Abstract**— Having established in previous articles that the principle of the constancy of speed of light of Special Relativity Theory is false in general, we examine some consequences of the non-constancy of speed of light in vacuum for different Galilean reference systems. We consider the modified Lorentz transformation incorporating different speeds of light for different inertial frames and which is a result of the non-constancy of speed of light in vacuum for different Galilean reference systems, i.e., reference systems in uniform rectilinear motion with respect to each other. Some consequences of this transform including the limit speed concept, transformation of lengths, summation and transformation of velocities are considered. Results of the transformation of the electromagnetic field are given. Also discussed are some consequences of the relativity of the electromagnetic field, all under the modified Lorentz transformation. These include variance of the electric charge, transformation of the potentials, transformation of the force.

Some of the highlights of these consequences are as follows. The transformation of lengths abide by the same rule as in Special Relativity Theory, at least in form. Transformation of time intervals and velocities now involve the speeds of light in vacuum for both inertial frames in distinction from Special Relativity Theory. The invariance of charge no longer applies and charge measured depends on the reference frame the measurement is made from. Force acting on a charge $q$ in an external electromagnetic field has the same form as in Special Relativity Theory, but now because charge is no longer invariant, the force transformation also has to be modified accordingly.

## 1. INTRODUCTION

It has been established for various media with dissipation that the speed of light in vacuum for that medium becomes dependent on the speed of the medium with respect to another when the two media are in uniform rectilinear motion with respect to each other [1–5]. This entails requirement of a modified Lorentz transformation which incorporates different speeds of light in vacuum for different Galilean reference systems. This transformation is given in [1]. This paper presents some consequences of this transformation. We start by the statement of the transformation.

$$x = \alpha\left(x' - c'rt'\right) \tag{1a}$$
$$y = y' \tag{1b}$$
$$z = z' \tag{1c}$$
$$t = \alpha\left(-\frac{r}{c}x' + \frac{c'}{c}t'\right) \tag{1d}$$

$$x' = \alpha\left(x + crt\right) \tag{2a}$$
$$y' = y \tag{2b}$$
$$z' = z \tag{2c}$$
$$t' = \alpha\left(\frac{r}{c'}x + \frac{c}{c'}t\right) \tag{2d}$$

It should also be pointed out that this paper is exclusively based on [6]. The analysis given therein for Special Relativity Theory [7] has been adapted here for the fact of the non-constancy of speed of light in vacuum. Therefore the ideas of nearly all the proofs have been taken directly from this reference.

## 2. SOME CONSEQUENCES OF THE MODIFIED LORENTZ TRANSFORMATION

### 2.1. The Limit Speed Concept

In order for formulas (1) and (2) to be meaningful $v_1 < c$ and $-v_2 < c'$ must hold. This implies that if system $K'$ is receding from $K$ with a speed $v_1$ greater than $c$, observers in $K$ can not observe and evaluate the physical phenomena occurring in $K'$. Similarly, observers in $K'$ can not evaluate the phenomena in $K$ if $-v_2 > c'$. This means message transmission with a speed greater than the speed of light in vacuum for the respective inertial frame is not possible. Since transmitting a message is in one sense transmitting matter or energy, we can state that transmission of matter or energy with speed greater than the speed of light in vacuum for the respective inertial frame is not possible. On the other hand because $r = -v_1/c = v_2/c'$ holds, one has, $\frac{v_1}{v_2} = -\frac{c}{c'}$, which implies that $v_1 > c'$ and $-v_2 > c$ can be satisfied individually, but not simultaneously in general if $v_1 < c$ and $-v_2 < c'$ hold, which must under all conditions as discussed above.

### 2.2. Transformation of Lengths

In frame $K'$, let the coordinates of the end points of a rod placed on the $O'x'$ axis be $A'(x'_1, 0, 0)$ and $B'(x'_2, 0, 0)$ (we assume $x'_2 > x'_1$). Whatever $t'$ is, the $l'$ value which is defined by $l' = x'_2 - x'_1$, equals the length of the stationary rod. Let us try to define the length of this rod as observed from $K$ in the following way. The $l$ value which is defined by $l = x_2 - x_1$, equals the length of the rod as observed from $K$. This is called 'the length of the rod in motion', whereas $l'$ is called the 'length at rest'. Let us observe that $x_2 - x_1 = \alpha[x'_2 - x'_1 - c'r(t'_2 - t'_1)]$. But $t'_2 - t'_1 = \alpha[c/c'(t_2 - t_1) + r/c'(x_2 - x_1)]$ where $t_1$ and $t_2$ are the times the observer measures $x_1$ and $x_2$. To measure $l = x_2 - x_1$ accurately, measure $x_1$ and $x_2$ simultaneously, so that $t_1 = t_2$. Thus $t'_2 - t'_1 = \alpha[r/c'(x_2 - x_1)]$, so that

$$l = x_2 - x_1 = \alpha\left[x'_2 - x'_1 - \alpha r^2 (x_2 - x_1)\right] = \alpha\left(l' - \alpha r^2 l\right)$$

From which we have

$$l = \frac{\alpha l'}{1 + r^2\alpha^2} = \frac{l'}{\alpha}. \tag{3a}$$

This indicates that the rod appears shorter in the $K$ frame than in $K'$ where it is stationary, i.e., the motion has caused the rod to contract. Length contraction is independent of $c$ and $c'$ as the formula (3a) indicates. It depends only on $r$ which is the same for both $K$ and $K'$. Therefore length contractions are not different for observers in frames $K$ and $K'$ for lengths in $K'$ and $K$ respectively.

We can easily verify that for every rod which is parallel to $O'x'$, a contraction as in (3a) is valid. However contraction does not occur for rods parallel to $O'y'$ and $O'z'$. This leads to the following result. Objects in motion contract in direction of motion. Because contraction is in one dimension only, volumes of moving objects must contract as follows:

$$V' = \alpha V \tag{3b}$$

Here $V'$ is the volume of object at rest, $V$ is the apparent volume in motion.

### 2.3. Transformation of Time Intervals

Let us suppose two different events occur at instants $t'_1$ and $t'_2$ at a fixed point $A'(x', 0, 0)$ in system $K'$. When the first event takes place point $A'$ corresponds to a point $A_1(x_1, 0, 0)$ of $K$, and in this case the clock at $A_1$ indicates a value of time such as $t_1$. When the second event occurs, point $A'$ will correspond to a point $A_2(x_2, 0, 0)$ and at this point the clock will indicate a value of time as $t_2$. The difference $\tau' = t'_2 - t'_1$ is the time difference between two events as measured in $K$ between the same events. If (1d) is written for $t'_1$ and $t'_2$ at point $A'$, and subtracted side by side,

$$\tau = \alpha\left(\frac{c'}{c}\right)\tau' \tag{4a}$$

will be found. This shows that the time interval between two events as measured in $K$ is proportional to the corresponding time interval in $K'$, with proportionality constant of $\alpha(\frac{c'}{c})$.

Based on (4a) we can state that an event that occurs with frequency $\upsilon'$ in system $K'$ will be observed as if occurring with a frequency $\upsilon$ in $K$ such that

$$\upsilon' = \alpha\left(\frac{c'}{c}\right)\upsilon. \tag{4b}$$

Since $c$, $c' > 0$, if $\upsilon > 0$, then $\upsilon' > 0$, because $\alpha > 0$.

### 2.4. Summation and Transformation of Velocities

We assumed the system $K'$ moves in the $Ox$ direction with a constant speed $v_1$. If an object moving in the $O'x'$ direction in $K'$ has speed $u'$ as measured from $K'$, what will the speed of this object be according to measurements of observers in $K$? It must be stressed that we have to find the total (or resultant) of two speeds measured from different Galilean systems. We can immediately state that the result will be different from $u' + v_1$.

Let the position of a point $A'$ in motion in $K'$ be given by the functions

$$x' = f_1\left(t'\right), \quad y' = f_2\left(t'\right), \quad z' = f_3\left(t'\right) \tag{5a}$$

that are dependent on $t'$. Components of the $\vec{u}'$ velocity vector of this point are as follows:

$$u_1' = \frac{dx'}{dt'}, \quad u_2' = \frac{dy'}{dt'}, \quad u_3' = \frac{dz'}{dt'}. \tag{5b}$$

During the motion, the position of the point $A \in K$ corresponding to $A'$, will be a function of $t$. The functions $x(t)$, $y(t)$ and $z(t)$ that determine this position, can be found by eliminating $t'$ between (5a) and (1). The velocity of the point $A$ in system $K$ is called the velocity of $A'$ observed from $K$, or its 'relative velocity' with respect to $K$. The components $u_1 = \frac{dx}{dt}$, $u_2 = \frac{dy}{dt}$, $u_3 = \frac{dz}{dt}$ of this velocity can be compared as follows using (1) and (5b):

$$u_1 = \frac{dx}{dt} = \frac{dx' - c'rdt'}{\frac{c'}{c}dt' - \frac{r}{c}dx'} = \frac{u_1' - v_2}{\frac{c'}{c} - \frac{r}{c}u_1'} = \frac{u_1' - v_2}{\frac{c'}{c} + \frac{v_1}{c^2}u_1'} \tag{6a}$$

$$u_2 = \frac{dy}{dt} = \frac{dy'}{\left(\frac{c'}{c}dt' - \frac{r}{c}dx'\right)\alpha} = \frac{u_2'}{\left(\frac{c'}{c} - \frac{r}{c}u_1'\right)\alpha} \tag{6b}$$

$$u_3 = \frac{dz}{dt} = \frac{dz'}{\left(\frac{c'}{c}dt' - \frac{r}{c}dx'\right)\alpha} = \frac{u_3'}{\left(\frac{c'}{c} - \frac{r}{c}u_1'\right)\alpha} \tag{6c}$$

## 3. TRANSFORMATION OF THE ELECTROMAGNETIC FIELD

Transformation of the electromagnetic field as per the modified Lorentz transformation of [1] was given in [1]. We repeat these results here again for completeness.

$$E_x' = E_x \tag{7a}$$
$$H_x' = H_x \tag{7b}$$

$$E_y' = \alpha(E_y + rcB_z) \tag{8a}$$
$$H_y' = \alpha\left(H_y - rcD_z\right) \tag{8b}$$

$$E_z' = \alpha(E_z - rcB_y) \tag{9a}$$
$$H_z' = \alpha\left(H_z + rcD_y\right) \tag{9b}$$

$$D_x' = \frac{c}{c'}D_x \tag{10a}$$
$$B_x' = \frac{c}{c'}B_x \tag{10b}$$

$$D_y' = \alpha\frac{c}{c'}\left(D_y + \frac{r}{c}H_z\right) \tag{11a}$$
$$B_y' = \alpha\frac{c}{c'}\left(B_y - \frac{r}{c}E_z\right) \tag{11b}$$

$$D_z' = \alpha\frac{c}{c'}\left(D_z - \frac{r}{c}H_y\right) \tag{12a}$$
$$B_z' = \alpha\frac{c}{c'}\left(B_z + \frac{r}{c}E_y\right) \tag{12b}$$

$$\rho' = \alpha \frac{c}{c'} \left( \rho + \frac{r}{c} J_x \right) \tag{13}$$

$$J'_x = \alpha (J_x + r c \rho) \tag{14a}$$

$$J'_y = J_y \tag{14b}$$

$$J'_z = J_z \tag{14c}$$

In the above $r = \frac{\sqrt{\alpha^2 - 1}}{\alpha}$, $E$ and $H$ stand for the electric and magnetic fields, $D$ and $B$ stand for the displacement and magnetic flux densities and $\rho$ and $J$ are the volume charge and surface current densities, with the subscripts on them representing the respective components in Cartesian coordinates. These results were obtained using the standard approach as under the classical Lorentz transformation and adhering to the principle of relativity.

## 4. SOME CONSEQUENCES OF THE RELATIVITY OF THE ELECTROMAGNETIC FIELD

Knowledge of the relations between the expressions of a certain electromagnetic field in systems $K$ and $K'$, will help reveal the relations between other quantities expressed in terms of these quantities. Here below we shall touch upon some of these.

### 4.1. Variance of the Electric Charge

Consider an electrified object $\vartheta'$ moving with velocity $\vec{v}_1$ with respect to frame $K$. Here $\vartheta'$ takes the place of the reference frame $K'$ mentioned in above paragraphs. Assume the density of charge in $\vartheta'$ is independent of $t'$. As per the inverse of relation (13) the expression of this charge as observed from $K$ is $\rho = \alpha \frac{c'}{c} \rho'$. In this case because of Section 2.2

$$dx\, dy\, dz = dx'\, dy'\, dz' / \alpha$$

and one has $\rho\, dx\, dy\, dz = \alpha \frac{c'}{c} \rho'\, dx'\, dy'\, dz'/\alpha$ or $\rho\, dx\, dy\, dz = \frac{c'}{c} \rho'\, dx'\, dy'\, dz'$. Integration of the left side of this equality in $\vartheta$ gives $Q$, and of the right side in $\vartheta'$ gives $\frac{c'}{c} Q'$. Here $Q'$ is the total charge the object carries. Because the region $\vartheta$ has been obtained by means of the modified Lorentz transformation, $Q$ is the expression of the total charge carried by the object as observed from $K$ and because of the last equality, $Q = \frac{c'}{c} Q'$. This equation shows that electric charge does not have the same value in all Galilean systems.

### 4.2. Transformation of the Potentials

Let the expressions of the potentials that an electromagnetic field derives from, be $V(x, y, z, t)$, $\vec{A}(x, y, z, t)$ and $V'(x', y', z', t')$, $\vec{A}'(x', y', z', t')$ in $K$ and $K'$ respectively. Because of (7a),

$$\begin{aligned} \frac{\partial V}{\partial x} + \frac{\partial A_1}{\partial t} = \frac{\partial V'}{\partial x'} + \frac{\partial A'_1}{\partial t'} &= \left\{ \frac{\partial V'}{\partial x} \frac{\partial x}{\partial x'} + \frac{\partial A'_1}{\partial x} \frac{\partial x}{\partial t'} \right\} + \left\{ \frac{\partial V'}{\partial t} \frac{\partial t}{\partial x'} + \frac{\partial A'_1}{\partial t} \frac{\partial t}{\partial t'} \right\} \\ &= \frac{\partial}{\partial x} \left\{ \frac{V' - v_2 A'_1}{\sqrt{1 - r^2}} \right\} + \frac{\partial}{\partial t} \left\{ \frac{(c'/c) A'_1 + (v_1/c^2) V'}{\sqrt{1 - r^2}} \right\} \end{aligned}$$

can be written and

$$V = \frac{V' - v_2 A'_1}{\sqrt{1 - r^2}} \tag{15}$$

$$A_1 = \frac{(c'/c)\, A'_1 + \left( v_1/c^2 \right) V'}{\sqrt{1 - r^2}} \tag{16a}$$

can be found. Similarly from (10b) we have:

$$\frac{\partial A_3}{\partial y} - \frac{\partial A_2}{\partial z} = \frac{c'}{c} \left( \frac{\partial A'_3}{\partial y'} - \frac{\partial A'_2}{\partial z'} \right) = \frac{c'}{c} \left( \frac{\partial A'_3}{\partial y} - \frac{\partial A'_2}{\partial z} \right),$$

or

$$A_2 = \frac{c'}{c} A'_2, \tag{16b}$$

$$A_3 = \frac{c'}{c} A'_3. \tag{16c}$$

### 4.3. Transformation of the Force (Special Case)

Consider a point charge $q$ which can freely move in an external electromagnetic field (see [8] for the concept of 'external electromagnetic field'). If the velocities of this charge in systems $K$ and $K'$ are denoted by $\vec{u}$ and $\vec{u}'$, the expressions in $K$ and $K'$ of the force acting on it will be as follows, as per the relativity principle and Section 4.1:

$$\vec{F} = q\left(\vec{E} + \vec{u} \times \vec{B}\right) \tag{17a}$$

$$\vec{F}' = \beta q\left(\vec{E}' + \vec{u}' \times \vec{B}'\right) \tag{17b}$$

Here $\beta$ is a proportionality constant that is determined in Subsection 4.1. There we proved that the factor $\beta$ is equal to $c/c'$ for an arbitrary electrified object. It will equally apply for a point charge as well.

If by considering (7a), (8a), and (9a) along with (10b), (11b), (12b), we can eliminate $\vec{E}$, $\vec{E}'$, $\vec{B}$ and $\vec{B}'$ in (17), we shall obtain a relation between $\vec{F}$ and $\vec{F}'$. This relation, which shall depend on $\vec{u}$ and $\vec{u}'$, shall be a general relation that will transform expressions in $K$ and $K'$, of any force field into each other. We shall not consider this general case here. We want to consider the special relation that will transform $\vec{F}$ and $\vec{F}'$ into each other, when the speed of the charge in system $K'$ is zero.

In this special case, $\vec{u}' = 0$, and by Section 2, $\vec{u} = \vec{v}_1$ will hold. In this case, if the expression for $\vec{E}'$, given by (7a), (8a) and (9a) is carried into (17b),

$$\vec{F}' = \beta q\left[\alpha\vec{E} + (1-\alpha)\left(\vec{E}\cdot\vec{v}_1\right)\frac{\vec{v}_1}{v_1^2} + \alpha\vec{v}_1 \times \vec{B}\right] \tag{17c}$$

will result. On the other hand noting that $\vec{F}\cdot\vec{v}_1 = q\vec{E}\cdot\vec{v}_1$ and that $\vec{v}_1 \times \vec{B} = \frac{\vec{F}}{q} - \vec{E}$, we obtain

$$\vec{F}' = \beta\left[\alpha\vec{F} + (1-\alpha)\left(\vec{F}\cdot\vec{v}_1\right)\frac{\vec{v}_1}{v_1^2}\right]. \tag{17d}$$

This is the transformation formula for the force acting on a point charge with zero speed in $K'$. If we scalar multiply both sides of (17d) by $\vec{v}_1$, we shall find

$$\vec{F}'\cdot\vec{v}_1 = \beta\vec{F}\cdot\vec{v}_1. \tag{17e}$$

Taking this into consideration, if we solve for $\vec{F}$ from (17d),

$$\vec{F} = 1/\beta\left\{(1/\alpha)\,\vec{F}' + [1-(1/\alpha)]\left(\vec{F}'\cdot\vec{v}_1\right)\frac{\vec{v}_1}{v_1^2}\right\} \tag{17f}$$

shall be obtained. While (17d) and (17f) are used, it should be borne in mind that it is assumed that in the system where $\vec{F}'$ is measured, the velocity of the point is zero at the instant of measurement. On the other hand, as noted above $\beta = c/c'$ through out.

**REFERENCES**


1. Yener, N., "On the non-constancy of speed of light in vacuum for different Galilean reference systems," *Journal of Electromagnetic Waves and Applications*, Vol. 21, No. 15, 2241–2255, 2007.
2. Yener, N., "Determination of speeds of light in vacuum for different Galilean reference systems," *PIERS Proceedings*, 115–119, Beijing, China, March 23–27, 2009.
3. Yener, N., "The non-constancy of speed of light in vacuum for different Galilean reference systems (revisited)," *PIERS Proceedings*, Xi'an, China, March 22–26, 2010.
4. Yener, N., "Non-constancy of speed of light in vacuum for different Galilean reference systems and momentum and energy of a particle," *PIERS Proceedings*, Xi'an, China, March 22–26, 2010.
5. Yener, N., "Non-constancy of speed of light in vacuum for different Galilean reference systems in case of an impulsive plane wave," *PIERS Proceedings*, Xi'an, China, March 22–26, 2010.
6. Idemen, M., *Elektromagnetik Alan Teorisinin Temelleri*, Literatür Yayıncılık, Istanbul, 1996.
7. Einstein, A., "On the electrodynamics of moving bodies," *The Principle of Relativity*, Dover, New York, 1952.
8. Jackson, J. D., *Classical Electrodynamics*, John Wiley & Sons, New York, 1975.

# New Approach to Modeling of Diffuse Reflection and Scattering for Millimeter-wave Systems in Indoor Scenarios

**L. Subrt, P. Pechac, and S. Zvanovec**
Department of Electromagnetic Field, Czech Technical University in Prague, Czech Republic

**Abstract**— A new approach to modeling diffuse reflection and scattering for millimeter-wave band based on a fast and simple semi-deterministic principle is presented. The model is based on the combination of pre-computed radiation patterns and a fast semi-deterministic algorithm. The principle and function of this model are explained as well as tested in a common indoor scenario represented by a room equipped with furniture. The results obtained through simulations are compared with results received from a classical ray-based model and significant differences are discussed in detail.

## 1. INTRODUCTION

Nowadays, users of modern mobile radio communication systems demand higher and higher data rates for their applications. One possible way to achieve high network capacity is to use large bandwidth and therefore to situate systems into millimeter-wave bands. For future applications, the indoor high-speed systems working at frequencies around 60 GHz are supposed [1]. A different approach needs to be utilized for electromagnetic wave propagation predictions in millimeter bands [1, 2] when compared with lower frequencies (i.e., frequencies up to 10 GHz occupied by present personal wireless communication systems).

At the high frequencies, dimensions of common obstacles (e.g., furniture) in indoor scenarios are comparable or even smaller than the wavelength and therefore the obstacles which are normally neglected have to be taken into account [3]. High free space loss and strong attenuation caused by walls usually limit the size of the area covered by a transmitter to a single room. On the other hand, this feature helps to avoid interference among multiple transmitters deployed in a building. The last significant feature of propagation within millimeter bands is that the surface roughness has a substantial effect on the impinging wave. The roughness of common surfaces is not negligible when compared with the wavelength of the impinging wave, i.e., the wave is scattered significantly more than at lower frequencies. The main motivation for this work arose from the above mentioned propagation issues connected with diffuse reflection and scattering phenomena.

## 2. SCENARIO

To demonstrate the functionality of our model, we chose a small office room (Figure 1) inside a typical office building measuring $5 \times 3 \times 2.5$ m and featuring several types of obstacles such as concrete walls, a floor, ceiling, wooden wardrobe, table and door. Table 1 shows the parameters of the materials used during the simulations. Empirical parameters (transmission and reflection loss) were used for standard ray launching simulations. The material constants were utilized in the new model to compute scattering patterns as will be shown in the following section. People and other minor obstacles were not considered in this scenario.

The wireless system in the scenario consists of two devices. The transmitting part is formed by a transmitter and an antenna placed in the upper corner of the room (the red wheel in Figure 1). The receiving part is represented by a laptop located on the table and equipped with a 60 GHz wireless network card. The distance between the transmitter and receiver is 3.91 m. Both transmitting and receiving antennas are considered as omni-directional. The situation including spatial orientation is depicted in Figure 1.

Table 1: Material constants used in simulations.

| Part of room | Material | Width [mm] | Transmission Loss [dB] | Reflection Loss [dB] | Relative Permittivity [-] | Relative Permeability [-] | Conductivity [S/m] |
|---|---|---|---|---|---|---|---|
| walls, ceiling | concrete | 100 | 6.89 | 7.51 | 6.00 | 1.00 | 0.078 |
| door, wardrobe, table | wood | 50 | 2.48 | 15.09 | 2.05 | 1.00 | 0.039 |
| window | glass | 2 | 1.69 | 7.53 | 6.00 | 1.00 | 0.006 |

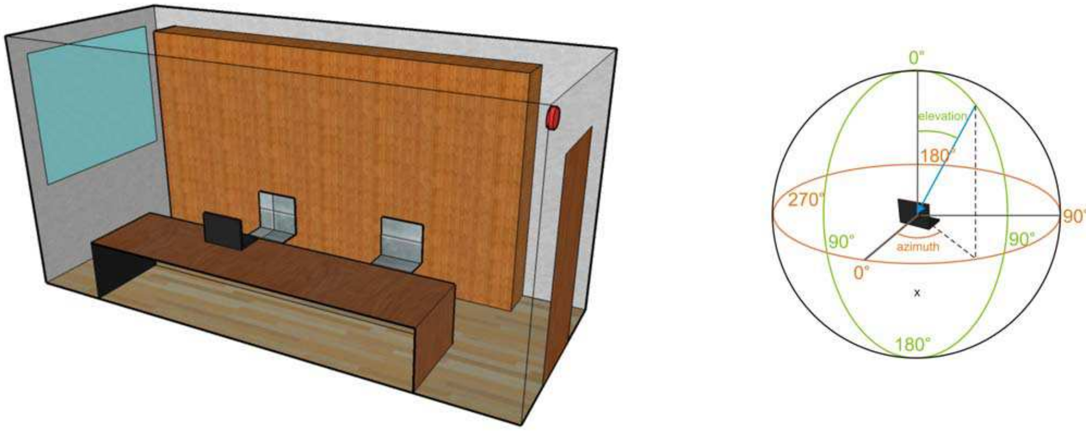
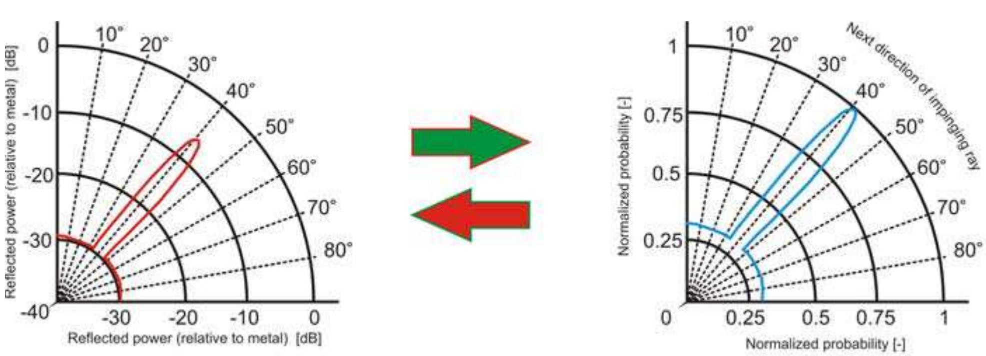



Figure 1: Scenario used for modeling and spatial orientation.

## 3. MODELING

The modeling can be divided into two main parts. Firstly, we have modeled and simulated the above mentioned scenario by means of the classical 3D ray launching model while using the empirical parameters given in Table 1. The rays are launched from the transmitter until the requested precision is reached. The model considers neither diffuse reflection nor diffuse scattering. If a ray impinges upon an obstacle, the next direction of the ray is always determined on the basis of geometrical optics. Diffraction phenomena are not considered in this case. This approach is rapid, because there is no need to solve complex and very time consuming equations. On the other hand, it can result in quite serious errors. This simulation was chosen in order to show possible inaccuracies caused by neglecting roughness of the materials and to discuss advantages of the new model.

In the second case, the same scenario was simulated using the new approach. We used a 3D ray-based model (more information can be found in [4]) similar to the previous case but with a significant difference in solving ray/obstacle interaction. The idea of the model is based on pre-computed, or measured, scattering patterns (Figure 2(a)). The scattering pattern needs to be computed or measured for each impact angle. It describes the relationship between an angle and the intensity of the electric field or power. It is also important to mention that such a characteristic (Figure 2(a)) cannot be used directly for ray-based models. The electric field intensity values (or power in this case) are transformed into probability function (Figure 2(b)) which allows us to solve ray/obstacle interaction as follows: When a ray impinges upon the obstacle surface, the corresponding scattering pattern (Figure 2(b) shows the pattern for 40°) is chosen and the next direction of the impinging ray is determined on this basis. This method provides us with precise results and extremely efficient computations.




(a) (b)

Figure 2: Scattering pattern, probability radiation pattern and corresponding distribution function. Impact angle is 40°, material is 5cm wide aerated concrete [6].

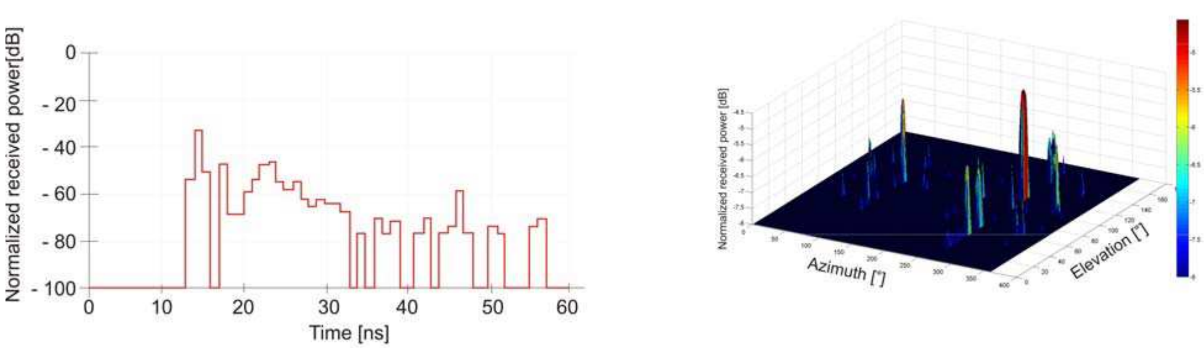
As mentioned previously, the scattering patterns for each angle of arrival can be measured or computed just once in advance (before starting the computations) using the complex equations based on Gaussian beams [5] and therefore it is not necessary to solve complex equations when a ray impacts an obstacle. Let us emphasize that Figure 2(a) depicts an example of a horizontal cut of the scattering pattern (taken from [6] and approximated), but the model always considers the general case — the three dimensional scattering patterns determined for each material from Table 1.

## 4. SIMULATION RESULTS

Figure 3(a) depicts the impulse response prediction and Angle of Arrival (AoA) prediction for the above mentioned case of a purely geometrical simulation. As can be seen, energy (represented by rays) impacts the transmitter at a rate of 13 ns at first. There is an existing line of sight (LOS) between the transmitter and the receiver and thus these rays travel the shortest distance between the transmitter and the receiver with no obstacle impacting. The time of arrival of all rays is proportional to the distance they travelled. All the other rays impacted in a time of more than 13 ns; they did not travel the shortest path, i.e., they interacted with one or more obstacles at least once and therefore the distance they travelled was longer.

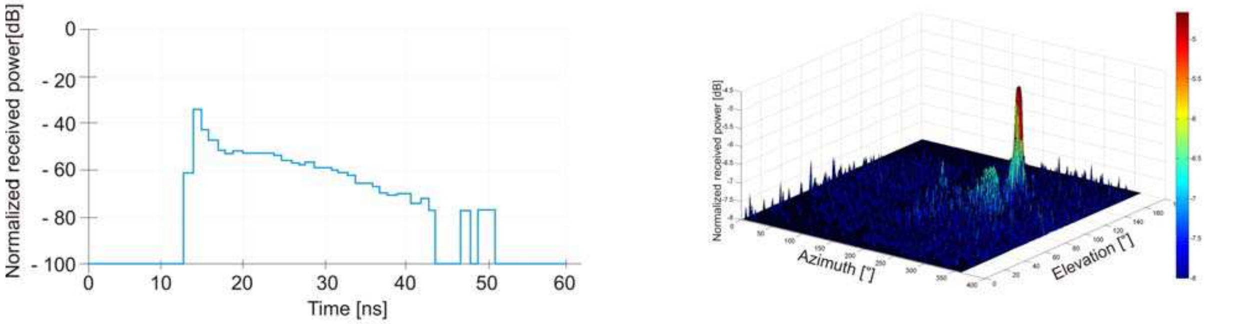
The impulse response prediction gives us invaluable information about the conditions in the scenario. But information about propagation environment provided by impulse response is not complete without an AoA prediction which gives us more precise information concerning the direction of incoming rays. The AoA prediction computed for the scenario (if we use simple geometrical optics) can be observed in Figure 3(b). As can be seen, most of energy arrives from the direction corresponding to an azimuth of 263° and an elevation of 116°. It is obvious (according to Figure 1) that these angles correspond to a direct connection line between the transmitter and the receiver. We can find several other directions with significant contributions (reflections from table, wardrobe, ceiling), but the rest of the contributions from other directions are negligible. We can conclude that energy comes to the receiver from a relatively small number of other directions. These results are given by the method of computation (simple geometrical optics).

On the other hand, if the new model is used, we obtain results which differ markedly from the results obtained by the simple geometrical approach. Firstly, the impulse response curve (Figure 4(a)) seems to be more continuous. It is due to the scattering which is taken into account in contrast to the previous case which was simply geometrical. Many new propagation paths emerge due to diffuse scattering. Each of these paths represents a different distance which the ray needs to travel on the way from the transmitter to the receiver. The path lengths in this case differ slightly and this is why the arrival times of different rays are sometimes very similar. Figure 4(b) shows the AoA characteristic. From the comparison of Figure 4(b) and Figure 3(b), it can be determined that the largest amount of energy comes directly from the transmitter (again LOS). But the rest of energy comes from many various directions in contrast to Figure 3(b). It is clear that the diffuse scattering has a direct influence on resulted AoA.




(a) (b)

Figure 3: Impulse response and AoA prediction computed by classical model.




(a) (b)

Figure 4: Impulse response and AoA prediction computed by the new model.

## 5. CONCLUSIONS

A new simulations model for diffuse scattering and reflection was presented. The model is based on a semi-deterministic approach taking advantage of precise pre-computed data and their fast usage during simulations. The model is able to provide both impulse response and angle of arrival prediction while having a short computation time. Further it is able to take into account roughness of the materials, which, as shown, needs to be taken into account. The advantage of this model is that complex and time-consuming computations are performed just once (before the simulation is started). The results obtained by means of the previously mentioned method are then used during the simulation where a very fast semi-deterministic algorithm determines the subsequent directions of impinging rays. Further it was shown that neglecting surface in the millimeter band can cause serious inaccuracies. Therefore only models which are able to take into account can be used for predicting in this frequency band.

## ACKNOWLEDGMENT

This work has been supported by the Czech Ministry of Education, Youth and Sports within the framework of the project OC09075 “Novel Emerging Wireless Systems” in the frame of COST IC0803 and by the Czech Science Foundation under the doctoral project No. 102/08/H027.

## REFERENCES

1. Smulders, P. F. M. and L. M. Correia, “Characterisation of propagation in 60 GHz radio channels,” *Electronics & Communication Engineering Journal*, Vol. 9, 73–80, 1997.
2. Suiyan, G., J. Kivinen, Z. Xiongwen, and P. Vainikainen, “Millimeter-wave propagation channel characterization for short-range wireless communications,” *IEEE Transactions on Vehicular Technology*, Vol. 58, 3–13, 2009.
3. Collonge, S., G. Zaharia, and G. El Zein, “Influence of the furniture on 60 GHz radio propagation in a residential environment,” *International Symposium on Signals, Circuits and Systems, SCS*, 413–416, Vol. 2, 2003.
4. Fluerasu, A. and C. Letrou, “Gaussian beam based simulations of indoor radio propagation channels,” *IEEE Antennas and Propagation Society International Symposium*, Vol. 2, 102–105, 2003.
5. Piesiewicz, R., T. Kleine-Ostmann, N. Krumbholz, D. Mittleman, M. Koch, J. Schoebel, and T. Kurner, “Short-range ultra-broadband terahertz communications: Concepts and perspectives,” *IEEE Antennas and Propagation Magazine*, Vol. 49, 24–39, 2007.
6. Langen, B., G. Lober, and W. Herzig, “Reflection and transmission behaviour of building materials at 60 GHz,” *5th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Wireless Networks — Catching the Mobile Future*, Vol. 2, 505–509, 1994.

# A Measurement System for Propagation Measurements at 300 GHz

**S. Priebe**[1,3], **C. Jastrow**[2,3], **M. Jacob**[1,3], **T. Kleine-Ostmann**[2,3], **T. Schrader**[2,3], **and T. Kürner**[1,3]

[1]Institut für Nachrichtentechnik, TU Braunschweig, Schleinitzstraße 22, 38106 Braunschweig, Germany
[2]Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany
[3]Terahertz Communications Lab (TCL), Braunschweig, Germany

**Abstract**— We present a 300 GHz measurement system based on commercially available subharmonic Schottky diode mixers and a vector network analyzer. Apart from a detailed description of the measurement setup and a system characterization, this paper also shows first results of ultra broadband channel measurements in typical indoor environments over short ranges.

## 1. INTRODUCTION

The ever increasing demand for higher data rates is accompanied by the quest for new spectrum. It has become obvious that wireless data rates of several 10 Gbit/s will be required in a few years from now [1]. While current standards for wireless data transfer at 60 GHz aim at data rates exceeding 5 Gbit/s [2], the available spectrum around this carrier frequency is still limited to 7–9 GHz depending on varying national regulations [3]. Therefore, the shift to carrier frequencies of 300 GHz and beyond is the most promising alternative for future wireless LANs as it will make currently unregulated bandwidths of up to 47 GHz [4] accessible. The frequencies beyond 300 GHz are called Terahertz frequencies. In order to gain insight into basic propagation phenomena at THz frequencies like reflection and absorption, THz time domain spectroscopy has been applied at the Terahertz Communications Lab (TCL) to determine material parameters of typical building materials [5, 6]. Furthermore, a transmission system has been used to show the feasibility of wireless data communication at 300 GHz by presenting an analog video transmission in [7]. Nevertheless, the derivation of complex channel models necessary for system simulations of future wireless communication systems will require more comprehensive propagation measurements. For this purpose, a 300 GHz measurement system has been set up at the TCL based on the components of the 300 GHz transmission system.

The paper is structured as follows. Section 2 will focus on the metrological characterization of the measurement system whereas Section 3 will introduce first channel measurements with regard to a possible wireless interconnection of different electronic devices operated on a desktop. Several channel transfer functions and the corresponding channel impulse responses will be shown for different distances and different antenna misalignments to account for non-perfect device positioning and orientation.

## 2. THE 300 GHZ MEASUREMENT SYSTEM

The transmission system introduced in [7] consists of autarkic transmitter (TX) and receiver (RX) units both based on double sideband subharmonic Schottky diode mixers. At the TX, a signal with a frequency $f$ between 0 and 10 GHz is upconverted from the baseband to the passband at 300 GHz using a local oscillator signal of 16.66 GHz, which is tripled twice and then fed into the mixer. To overcome the high path loss at 300 GHz [4], 26 dBi horn antennas are used both at TX and RX.

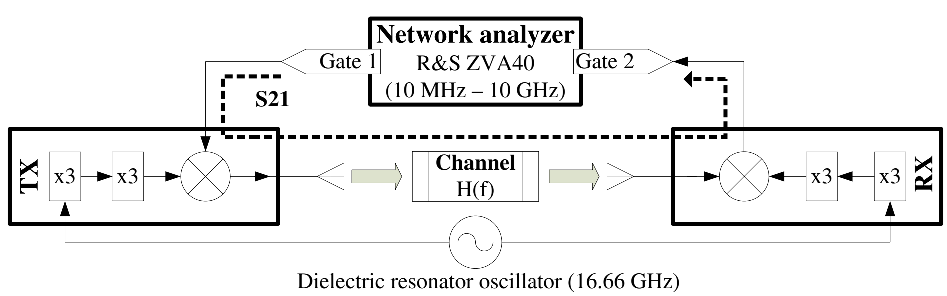



Figure 1: The 300 GHz measurement system setup.

Instead of downconverting the received signal to an intermediate frequency of 5 GHz as proposed in [7] for data transmission, the RX module is also fed with the TX local oscillator signal in the measurement setup, resulting in a homodyne signal conversion directly to the baseband. At the same time, phase coherence between TX and RX as obligatory for vectorial network analysis is achieved. On the basis of a Rohde & Schwarz ZVA 40 vector network analyzer (VNA, used from 10 MHz to 10 GHz) connected to the baseband in-/output of the TX and RX, this configuration allows channel measurements with a maximum bandwidth of 9.99 GHz in the frequency range from 300.01 GHz to 310 GHz (Fig. 1) by recording the complex scattering parameter $S_{21}$. However, a detailed system characterization is required for absolute measurements. In addition, the homodyne downconversion has to be investigated in detail due to different free space losses in the upper sideband (300.01–310 GHz) compared to the lower sideband (290–299.99 GHz) regarding the path loss.

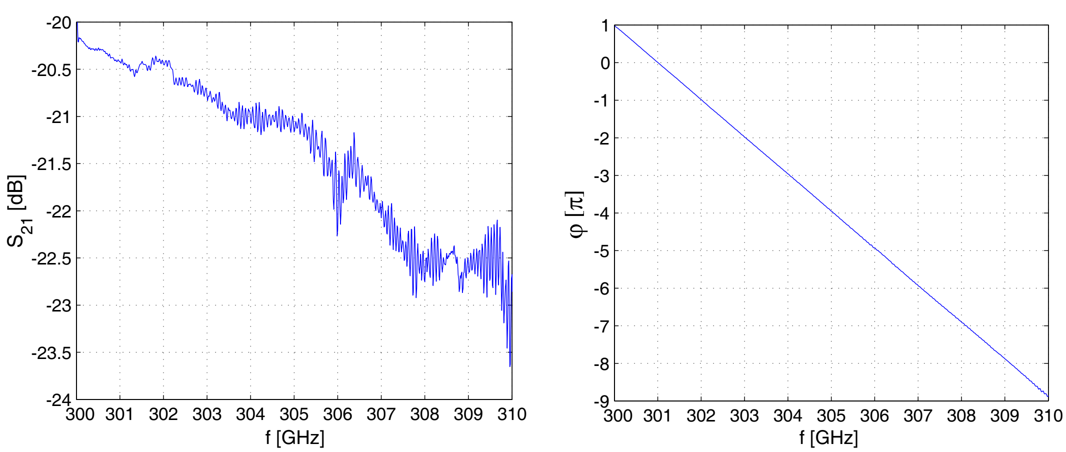
Assuming the mixing processes as ideal multiplications [8] and calculating the spectrum of a complex downconverted sinusoidal test signal at the receiver output, one can derive the spectrum

$$\begin{aligned} S(f) \;=\; & \frac{P(f)}{2} \cdot \left(\frac{c}{4\pi r(f_{LO}+f)}\right)^2 \cdot e^{j\left(\frac{2\pi f}{c}d+\varphi_0\right)} \\ & \cdot \left[\delta(f_{Test}+f) \cdot \left\{\cos\left(\frac{2\pi f_{LO}}{c}d\right) + j \cdot k(f) \cdot \sin\left(\frac{2\pi f_{LO}}{c}d\right)\right\}\right. \\ & \left. +\delta(f_{Test}-f) \cdot \left\{k(f) \cdot \cos\left(\frac{2\pi f_{LO}}{c}d\right) + j \cdot \sin\left(\frac{2\pi f_{LO}}{c}d\right)\right\}\right] \end{aligned} \tag{1}$$

where $c$ corresponds to the lightspeed, $\varphi_0$ is the phase response of the measurement system, $f_{LO}$ is the local oscillator frequency (300 GHz), $d$ is the distance between TX and RX, $f_{Test}$ is the frequency of the test signal (10 MHz to 10 GHz), $P(f)$ includes the input power, the system transfer function as well as the antenna gains and $k(f)$ means

$$k(f) = \left(\frac{f_{LO}-f}{f_{LO}+f}\right)^2. \tag{2}$$

$k(f)$ ranges from 1 for $f = 0$ GHz to 0.8751 for $f = 10$ GHz, thus resulting in a slight log linear amplitude distortion between 0 dB at 300 GHz and 0.28 dB at 310 GHz whereas the phase remains unchanged. Compared to the influence of the error in measurement due to non-perfect module alignment, which can result in several dB additional attenuation as shown in Section 3, this inherent error is almost negligible. Nevertheless, it can be easily corrected alongside with the amplitude and phase response of the measurement system itself, which are shown in Fig. 2. Here and in the following $S_{21}$ given in dB always implicitly means the amplitude of the respective complex transfer function. For the depicted reference measurements the waveguide outputs of both mixers were connected directly, eliminating any channel influence.




(a) Amplitude response (b) Phase response

Figure 2: Amplitude and phase response of the 300 GHz measurement system.

Table 1: Measurement parameters.

| Parameter | Symbol | Value |
|---|---|---|
| Measurement points | N | 801 |
| Intermediate frequency filter bandwidth | $\Delta f_{IF}$ | 10 kHz |
| Average noise floor | $P_N$ | −113.97 dBm |
| Noise standard deviation | $\sigma$ | 6.74 dB |
| Power of test signal | $P_{VNA}$ | −5 dBm |
| Bandwidth | B | 9.99 GHz |
| Start frequency | $f_{Start}$ | 10 MHz |
| Stop frequency | $f_{Stop}$ | 10 GHz |
| Time domain resolution | $\Delta t$ | 0.1 ns |
| Smallest resolvable distance | $\Delta l$ | 3 cm |
| Maximum excess delay | $\tau_m$ | 80 ns |
| Maximum detectable path length | $l_m$ | 24 m |

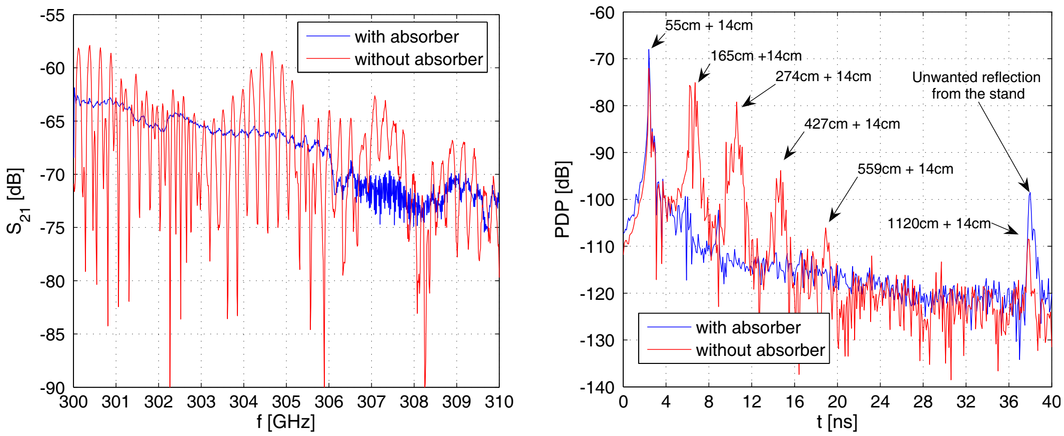
Due to the frequency dependence of the conversion losses, connection losses and impedance mismatches, a considerably increasing system loss can be observed with increasing frequency, ranging from 20.2 dB at 10 MHz to 23.7 dB at 10 GHz. Regarding the phase, the system shows a nearly perfect linear phase shift that corresponds to an electrical length of $l_{el} = 0.147$ m or to a delay of $t = 0.491$ ns in time domain, respectively. The initial offset of $0.98\pi$ at $f = 300.01$ GHz includes both $\varphi_0$ and the phase shift due to the phase term $\frac{2\pi f_{LO}}{c}d$ reduced to the interval $[-\pi; \pi]$ whereas the negative linear slope only represents the phase term $\frac{2\pi f}{c}d$ of the test frequencies.

All measurements were taken using the parameters given in Table 1. With the aim of the highest possible temporal resolution the maximum bandwidth of 9.99 GHz was chosen. As only short path lengths were expected due to high path losses the number of sweep points was limited to 801, allowing for a maximum detectable distance of 24 m. Due to the input power restriction of the TX Schottky mixer of −3 dBm [7], the output power $P_{VNA}$ of the VNA was chosen to −5 dBm, still providing received powers well above the average noise floor of −114 dBm in all investigated scenarios. In addition, each recorded complex transfer function was averaged over 10 successive measurements to avoid any unspecific time dependent influence.

Apart from considering the system influence it has proven necessary to mount electromagnetic absorber panels at the front of both the TX and RX module. Without absorbers, multiple reflections between the two modules can be observed in the power delay profiles (PDP), each following the direct path at a time corresponding to approximately twice the module distance. This behavior is shown in an exemplary situation with best possible module alignment at a distance of 55 cm with and without absorbers (Fig. 3). In addition, the electrical length of the system itself (14.7 cm, see above) has to be taken into account, which in this case was not corrected for illustration. Without absorbers the rays are reflected at different components on the modules each having different spacing from the antenna, which leads to a pulse broadening over time. The paths are attenuated increasingly with each reflection as their path length increases. Here five multipath components cannot be neglected. In the frequency domain, the unwanted multipaths lead to deep fading dips of up to 30 dB below the average signal level. This effect cannot be avoided even if absorber panels are mounted. After several reflections at the stand, a ray reaches the RX antenna under a flat incidence angle and hence is able to pass the absorber panel. As a consequence, the stand also has to be shielded in certain situations to obtain correct channel transfer functions. Those cases can be identified instantly using the time domain option of the VNA. Nevertheless, the aspect of reflection between the two devices will also have to be considered when designing the enclosure of future THz communication systems.

## 3. CHANNEL MEASUREMENTS

Possible scenarios of future 300 GHz data transmission systems mainly include the high speed wireless extension of local area networks or the interconnection of electronic devices on a desktop. An example is fast data exchange with portable flash drives [4]. The latter case requires channel




(a) Transfer functions (b) Power delay proles

Figure 3: Transfer functions and power delay profiles with and without absorbers at a module distance of 55 cm.

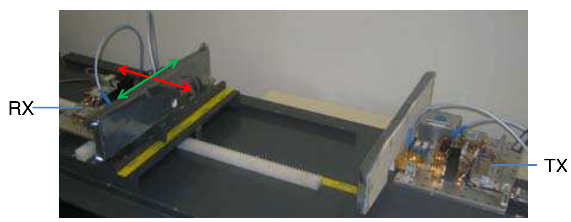
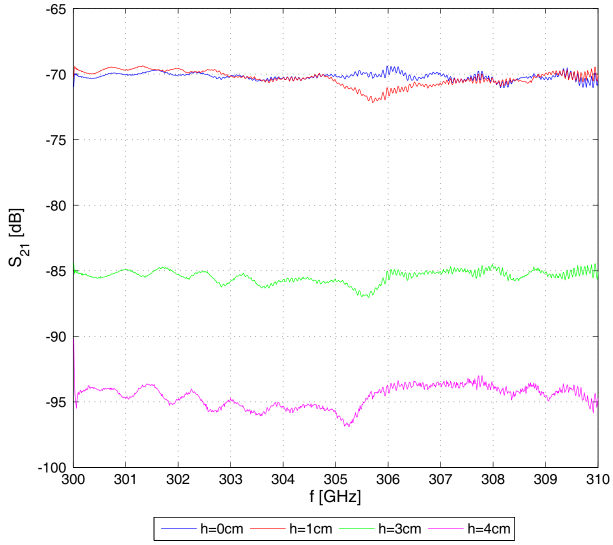



Figure 4: The short range channel measurement setup.

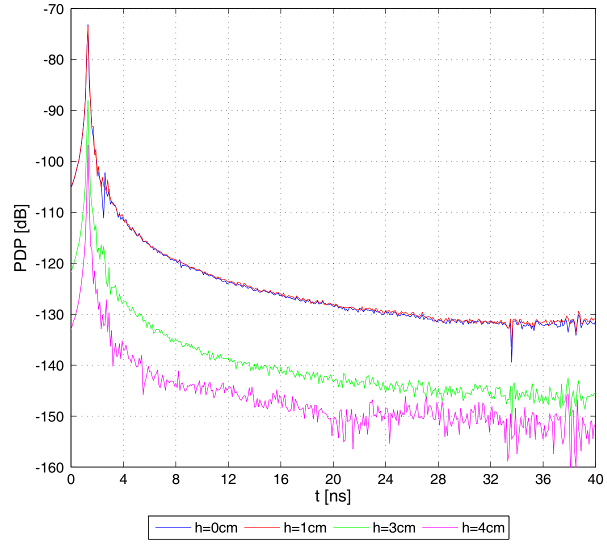
measurements with regard not only to distance but also to antenna misalignment as users cannot be expected to place their devices perfectly. Especially when using highly directive antennas as required to compensate the high path loss [9], even slight misalignments may lead to considerable additional attenuation as shown in the following. Based on the measurement setup depicted in Fig. 4, the antenna misalignment was simulated shifting the modules orthogonally to each other (green arrow) while the module distance could be varied simultaneously (red arrow).

Typical distances of $d = 20\,\mathrm{cm}$ and $d = 40\,\mathrm{cm}$ with orthogonal displacements of $l_{20\,\mathrm{cm}} = 0\,\mathrm{cm}$, 1 cm, 3 cm, 4 cm and $l_{40\,\mathrm{cm}} = 0\,\mathrm{cm}$, 2 cm, 4 cm, 8 cm were chosen, respectively, each displacement distance relating to the best possible alignment (Fig. 5). The transfer functions and power delay profiles were corrected considering the system loss, the antenna gains and the distortion due to the homodyne downconversion process. Despite the corrections, the measured $S_{21}$ parameters for $l = 0\,\mathrm{cm}$ still differ from the theoretically expected free space loss of 68 dB for $d = 20\,\mathrm{cm}$ and 74 dB for $d = 40\,\mathrm{cm}$, respectively. This behavior can be explained by an antenna mismatch not included in the calibration measurements and a non-perfect module alignment even in the reference position at $l = 0\,\mathrm{cm}$. As the absorber plates are to be considered as non-perfect, non-negligible reflections between the modules induce a slight frequency dependence of the channel behavior. A maximum amplitude fluctuation of 2.9 dB peak-to-peak over the whole bandwidth of 9.99 GHz can be observed for $d = 20\,\mathrm{cm}$ and $l = 4\,\mathrm{cm}$. Increasing the module distance from 20 cm to 40 cm does not alter the general channel characteristics apart from the additional free space loss of approximately 6 dB. Comparing the transfer functions for different displacements, one can observe a significantly increasing attenuation with an increasing misalignment which is a consequence of the high antenna directivity. A displacement of 1 cm at $d = 20\,\mathrm{cm}$ has almost no influence on the channel whereas $l = 4\,\mathrm{cm}$ leads to an additional attenuation of approximately 24 dB. At $d = 40\,\mathrm{cm}$ the same displacement only causes an increase in attenuation of about 6 dB. This aspect will have to be included in the link budgets of future short-range THz communication systems by introducing an application-specific link margin.

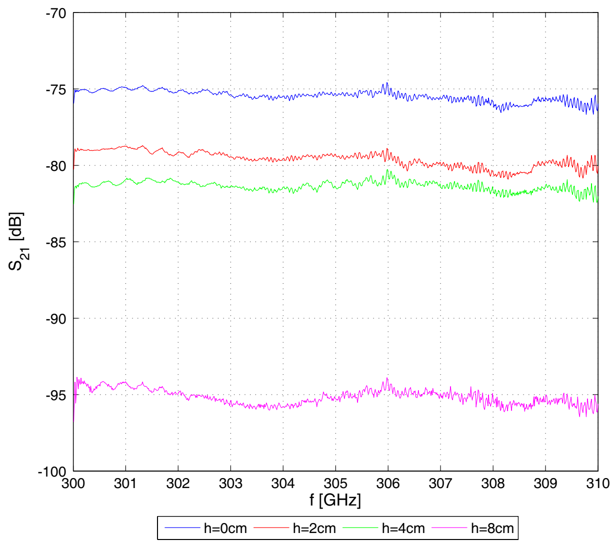
In addition, the RMS delay spread $\tau_{RMS}$ [10] has been calculated from the power delay profiles (Table 2). For the purpose of noise suppression, only signal components with a power of no more than 40 dB below the detected maximum have been considered. The computed values range from



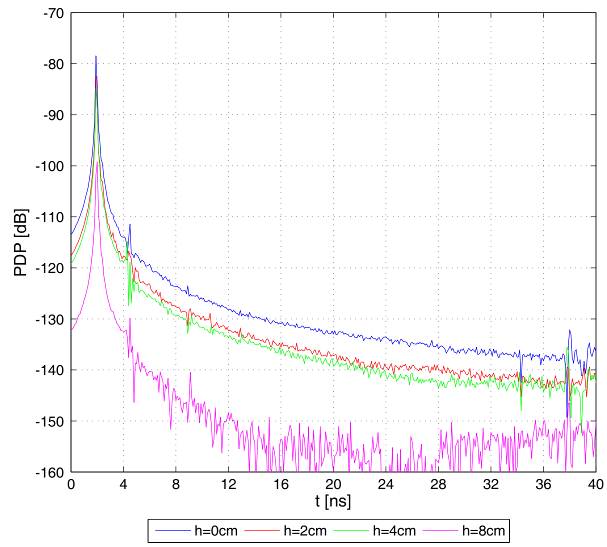

(a) Transfer functions for $d = 20$ cm




(b) Power delay profiles for $d = 20$ cm




(c) Transfer functions for $d = 40$ cm




(d) Power delay profiles for $d = 40$ cm

Figure 5: Channel transfer functions and power delay profiles for different distances and module misalignments.

Table 2: RMS delay spread for the measured power delay profiles.

| | $d = 20$ cm | | | | $d = 40$ cm | | | |
|---|---|---|---|---|---|---|---|---|
| | $l = 0$ cm | $l = 1$ cm | $l = 3$ cm | $l = 4$ cm | $l = 0$ cm | $l = 2$ cm | $l = 4$ cm | $l = 8$ cm |
| $\tau_{RMS}$ | 0.13 ns | 0.15 ns | 0.14 ns | 0.12 ns | 0.18 ns | 0.17 ns | 0.2 ns | 0.2 ns |

0.13 ns ($d = 20$ cm, $l = 4$ cm) to 0.2 ns ($d = 40$ cm, $l = 4$ cm), being similar for all channels. Basically, no correlation between the module displacement and the RMS delay spread can be observed. The slight differences are rather explained by the varying amplitude of the unwanted reflection between the modules. Regarding digital data transmission, such low delay spreads allow symbol rates of several GSymbols/s without intersymbol interference, which become necessary if aiming at data rates of more than 10 GBit/s. Besides the available bandwidth, the temporal channel characteristics hence make the frequency band from 300 GHz upwards an ideal candidate for future high speed wireless data communication systems.

## 4. CONCLUSION

Using commercially available components of a 300 GHz transmission system, we have set up a 300 GHz measurement system. The system capabilities have been investigated in detail. It has become clear that absorber panels shielding the front of both transmitter and receiver will be necessary to obtain correct measurement results. At the same time, the measurements taken without absorber panels have served as proof that valid channel impulse responses can be derived using the measurement system. Ultra broadband channel measurements at short distances and with slight module displacements have been shown, relating to the use case of wirelessly connected devices on a desktop. Nevertheless, further measurements in indoor environments will have to be performed in order to derive a first empiric 300 GHz channel model.

## REFERENCES

1. Cherry, S., "Edholm's law of bandwidth," *IEEE Spectrum*, Vol. 41, No. 7, 58–60, 2004.
2. 802.15.3c-2009, IEEE Std., 10 2009.
3. Yong, S. K. and C.-C. Chong, "An overview of multigigabit wireless through millimeter wave technology: Potentials and technical challenges," *EURASIP J. Wirel. Commun. Netw.*, Vol. 2007, No. 1, 50–50, 2007.
4. Piesiewicz, R., et. al., "Short-range ultra-broadband terahertz communications: Concepts and perspectives," *IEEE Ant. and Prop. Magazine*, Vol. 49, No. 6, 24–39, 2007.
5. Piesiewicz, R., et. al., "Properties of building and plastic materials in the THz range," *International Journal of Infrared and Millimeter Waves*, Vol. 28, No. 5, 363–371, 2007.
6. Piesiewicz, R., et. al., "Measurements and modeling of multiple reflections effect in building materials for indoor communications at THz frequencies," *Proc. German Microwave Conference GEMIC*, 3089–3092, Hamburg, March 2008.
7. Jastrow, C., et. al., "300 GHz transmission system," *Electron. Lett.*, Vol. 44, 213–214, 2008.
8. Pozar, D., *Microwave Engineering*, 3rd Edition, John Wiley & Sons, Inc., 2005.
9. Kürner, T., et al., "An integrated simulation environment for the investigation of future indoor THz communication systems," *Proc. of 2007 Interntl. Symp. on Performance Evaluation of Computer and Telecommunication Systems*, 2007.
10. Sarkar, T., et al., "A survey of various propagation models for mobile communication," *IEEE Antennas and Propagation Magazine*, Vol. 45, No. 3, 51–85, 2003.

# An Evaluation of Approaches for Modeling of Terrestrial, HAP and Satellite Systems Performance during Rain Events

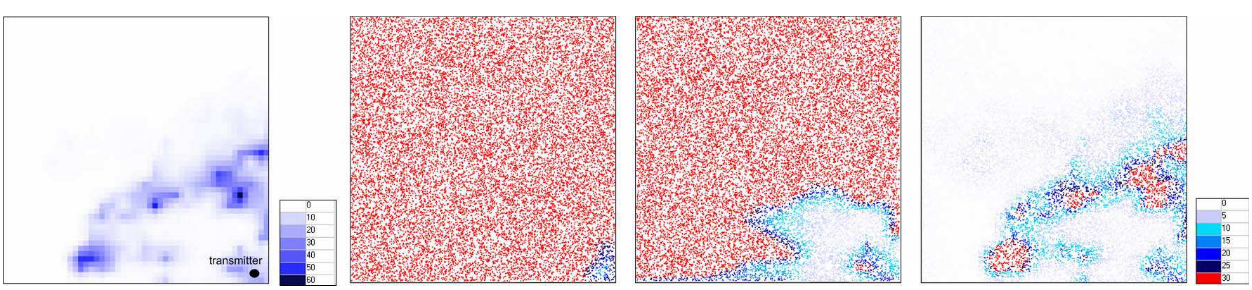
**S. Zvanovec, L. Subrt, and P. Pechac**
Department of Electromagnetic Field, Czech Technical University in Prague, Czech Republic

**Abstract**— This paper presents the results from site-specific simulations of route diversity performance in case of High Altitude Platform system, terrestrial system and satellite system during specific rain events. Based on the radar rainfall database, route diversity dependences on altitude of transmitter were studied using the simulation tool. The main goal of the paper is to adopt concept for HAP systems from investigations of both terrestrial point-to-multipoint systems and satellite systems operating in millimeter frequency bands.

## 1. INTRODUCTION

The proper choice of the modeling approaches for electromagnetic wave propagation within the millimeter wave band is crucial. This is particularly due to the high cost of millimeter wave equipments associated with transmitting power. Considerably high attenuation can be caused by interaction of transmitted waves with rain drops. There can be found several approaches to enumeration and compensation of signal fades caused by rain [1, 2]. Investigations of millimeter wave propagation aspects have been conducted out within the frame of European international projects such as CRABS [3] and COST Action 280 [4] and COST Action 297 [5]. Nevertheless, due to the variability of systems configurations, these models can be used only for investigation of alone terrestrial, satellite or High Altitude Platforms (HAPs) links. Spatial distribution of rain attenuation observed by users in the area 50 km × 50 km from transmitter placed either at terrestrial mast (10 m above ground) or at HAP (deployed on quasi-stationary unmanned vehicles/airships in the stratosphere at the altitude of 20 km) or at LEO satellite (altitude of 320 km) is depicted in Figure 1. These results clearly demonstrate, how particular systems are able to connect with users behind a rain event.

One of the rain fade mitigating approaches used in millimeter waveband introduces a route diversity [2] — i.e., two joint links to one HAP station from two different ground localities — during a storm, when a rain cell moves across the service area, any outage of a terminal can be avoided when it is able to connect to two different HAP stations. Nevertheless [2], involves only the parameters given based on the measurements of Earth-space links within particular geographical localities. The paper brings a comparative study of modeling approaches for propagation of millimeter waves within rain events and their adaption to HAP systems. The rain database from the period of 2002–2005 was utilized including 250 km × 250 km rain scans from Czech meteoradars (rain rate distributions with 1 km grid resolution and 1 minute time steps). Particular results of HAP system simulations performed at the frequency of 48 GHz will be discussed. Two evaluation methods were tested based on simulation results: the first following terrestrial approaches [6] and the second one from the other side utilizing a satellite approach [2].



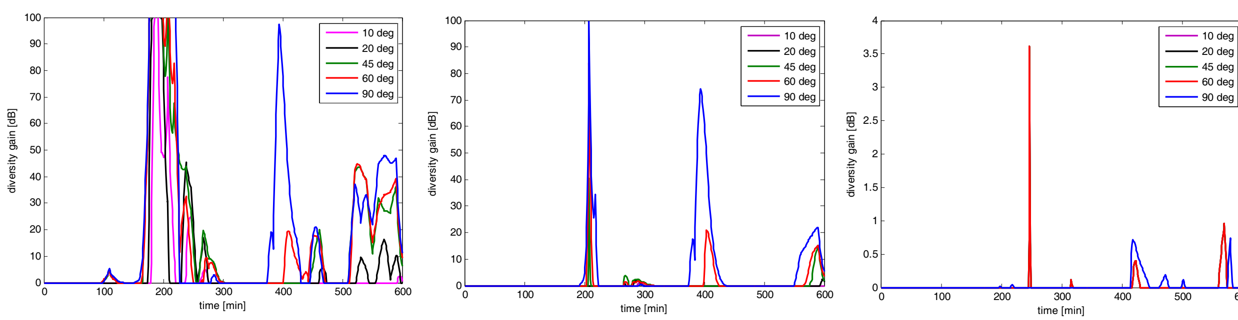

(a) (b) (c) (d)

Figure 1: (a) rain rate distribution in [mm/h] and corresponding rain attenuation in [dB] for coverage from (b) terrestrial transmitter; (c) HAP; (d) satellite.

## 2. SYSTEM PERFORMANCE COMPARISON

Analyze of dependences of rain spatial parameters on system performance for both terrestrial and HAP systems was elaborately performed in [7] with the result — the rain spatial classification for the evaluation of rain influences on terrestrial systems can be with particular corrections adapted in the case of stratospheric systems. Let us now consider a dependence of transmitter altitude. At first, performance of two joint links was analyzed for terrestrial, HAP and satellite systems. Common transmitting frequency of 48 GHz was assumed (the frequency allocated worldwide to HAP systems [8]). Significant rain events recorded by meteoradars in the Czech Republic on 13th of July 2002 was chosen from the rain database [9]. In Figure 2, an illustrative example from performed analyses — diversity gain dependences on angular separations of route diversity links (the main link length 20 km and the diversity link length 40 km) are depicted. As it can be seen, utilization of route diverzity is very efficient in case of terrestrial systems. In case of HAP systems, diverzity gain concept can be utilized for statistically shorter time a year (see example of results in Figure 2(b)). Nevertheless, it was observed that the HAP system with a particular route diversity scheme is able to efficiently combat the rain attenuation.

Contrary to terrestrial and HAP systems, links of satellite systems cross due to the high elevations rainy layer only within near proximity of user (note the average rain height determined for Europe is of 3.36 km [10]). This has impact on used route diversity scheme, where diversity gains are substantially smaller (see Figure 2(c)). In next step, a diversity gain was determined based on base station altitude. Example for approximately 5 hour rain event is shown in Figure 3. There can be clearly demonstrated iterative approaches from terrestrial and satellite models to HAP systems. From simulations of 3 year period (years 2002–2004), it was derived that terrestrial models can be applied with very slight deviances in results for transmitter up to altitudes of 1–10 kilometers. For higher altitudes route diversity expressed by diversity gain becomes differentiate (in most cases rapidly decreases) for more than 3 dB. This follows gradually with increasing altitude. Surprisingly, it was observed HAP systems propagation dependences statistical behavior




(a) (b) (c)

Figure 2: Example of time dependence of diverzity gain for (a) terrestrial system; (b) HAP system; (c) satellite system.

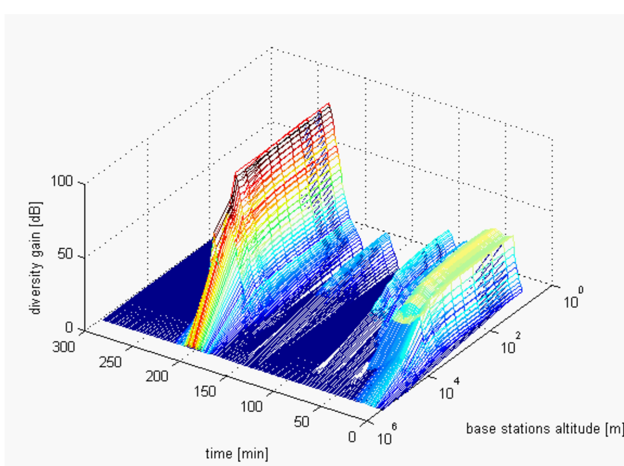



Figure 3: Dependence of diversity gain on altitude.

has higher correlation with terrestrial link statistics (correlation coefficient 0.85) than with satellite link statistics (correlation coefficient 0.14).

## 3. CONCLUSIONS

Rain represents one of the main limitations of millimeter wave band systems regardless altitude of transmitter. In the paper, a comparison of route diversity utilization for terrestrial, HAP and satellite links was accomplished based on actual rainfall radar data. HAP link statistics proved to be more related to terrestrial ones than to satellite statistics. Based on the results discussed in this paper, detailed relationships on altitudes of transmitter will be subsequently derived.

## ACKNOWLEDGMENT

The simulations and presentation of results were supported by Czech Science Foundation grant 102/08/P346. The research is a part of the activities of the Department of Electromagnetic Field of the Czech Technical University in Prague in the frame of the research project No. OC09075 of the Ministry of Education, Youth and Sports of the Czech Republic.

## REFERENCES

1. "Specific attenuation model for rain for use in prediction methods," *International Telecommunications Union, ITU-R Recommendation*, 838.
2. "Propagation data and prediction methods required for the design of Earth-space telecommunication systems," *International Telecommunications Union, ITU-R Recommendation*, 618-8, 2003.
3. "Report from ACTS project 215 — Cellular radio access for broadband services (CRABS)," *Propagation Planning Procedures for LMDS*, 1999.
4. Usman, I. S., M. J. Willis, and R. J. Watson, "Route diversity analysis and modelling for millimetre wave point to multi-point systems," *1st Int. Workshop of COST Action 280*, Jul. 2002.
5. "COST297 — HAPCOS high altitude platforms for communications and other services," [Online], Available: http://www.hapcos.org/.
6. "Propagation data and prediction methods required for the design of terrestrial line-of-sight systems," *International Telecommunications Union, ITU-R Recommendation*, 530-11, 2005.
7. Zvanovec, S. and P. Pechac, "Validation of rain spatial classification for high altitude platform systems," *IEEE Trans. on Antennas and Propagation*, submitted for publication.
8. "Preferred characteristics of systems in the fixed service using high-altitude platform stations operating in the bands 47.2–47.5 GHz an 47.9–48.2 GHz," ITU-R Recommendation F.1500, 2000.
9. Zvanovec, S. and P. Pechac, "Rain spatial classification for availability studies of point-to-multipoint systems," *IEEE Trans. on Antennas and Propagation*, Vol. 54, No. 12, 3789–3796, 2006.
10. "Rain height model for prediction methods, international telecommunications union," ITU-R Recommendation P.839-3, 2001.

# Gas Absorption Measurement of Selected Stratospheric Substances by Fabry-Perot Resonator

**P. Piksa**[1], **S. Zvanovec**[1], **P. Cerny**[1], **J. Libich**[1], **J. Varga**[2], **and J. Koubek**[2]
[1]Department of Electromagnetic Field, Czech Technical University in Prague, Czech Republic
[2]Department of Analytical Chemistry, Institute of Chemical Technology, Czech Republic

**Abstract**— A simple Fabry-Perot resonator based measurement system was utilized for measurements of attenuation characteristics of stratospheric gases. Representative of very-short lived gases — dibromomethane — as well as Carbonyl sulfide were measured in selected microwave bands occupied by Satellite to ground or High Altitude Platform Systems.

## 1. INTRODUCTION

Observations and investigations of gas processes within the ozone layer and interaction of transmitting electromagnetic waves in the stratosphere belong nowadays to main interests of human beings. A Fabry-Perot interferometer provides an interesting tool enabling an enhancement of the sensitivity of the absorption as well as emission measurements in the region of microwave spectroscopy [1]. Measurement in the evacuated resonator allows setting particular molecular concentrations corresponding to the height above ground and adjusting concentration correspondent to the atmosphere layer. For monochromatic radiation, the Fabry-Perot interferometer can be tuned to a resonance at which the constructive interference of the multiple-reflected electromagnetic waves enables to accumulate the radiative energy.

The main advantage of the Fabry-Perot resonator comparing to the other laboratory measurements is its relatively higher sensitivity to weak absorptions due to an apparent lengthening of the optical path length by means of above mentioned multiple reflections. The effective optical path length corresponding to the quality factors (according to [2]) of the bellow described evacuated resonator equals 100 m. Therefore a precise measurement of gas attenuation can be reached. The absorption spectra of particular very short-lived (VSL) gases substances from the ozone layer were tested via the Fabry-Perot resonator. Their influences on particular millimeter wave systems, whose links intersect the stratosphere (i.e., High Altitude Platforms, Mobile to Satellite Systems), are detailed discussed in the paper.

## 2. MICRAWAVE SYSTEMS PASING STRATOSPHERE

Several millimeter wave links passing atmosphere can be temporarily influences by VSL gasses or impermanent gas clouds/regions emerging from fires. It is true especially in case of links connecting terrestrial and satellite segments, i.e., Mobile-Satellite Services (MSS) and Fixed-Satellite Services (FSS). The Federal Communications Commission (FCC) has allocated [3] several millimeter bands for these systems including, e.g., 27 GHz, 38 GHz, and 71–76 GHz, 81–86 GHz bands. Satellite links connecting land stations at specified fixed points are also authorized when used in conjunction with radionavigation-satellite service or Space Tracking and Surveillance System (STSS). Within other systems whose links passes through variant ozone layer belong High Altitude Platforms (HAPs) [4]. These systems can provide very fast coverage of areas or substitute conventional terrestrial or satellite systems from altitudes around 20 km. For HAP systems the frequency band 47/48 GHz was allocated worldwide.

Many papers introducing gaseous attenuation measurements can be found in the literature [5]. The attenuation introduced by atmospheric gases can either be described using an accurate physical model, such as Liebe's model [6] for frequencies ranging from 1 GHz up to 1 THz, or it can be approximated by probabilistic models such as the ITU-R P. 676 [7]. The ITU-R P.676 includes two models for the calculation of gaseous attenuation — a complete line-by-line method which sums the contributions from 44 oxygen lines and 30 water-vapor lines below 1000 GHz, and a simplified algorithms based on a curve-fitting to the line-by-line calculation. Nevertheless, in the real HAP or satellite scenario different situations can occur from the gas attenuation point of view. Differences between the measured/observed and theoretically calculated (via, e.g., ITU-R P.676) values of atmospheric gas attenuation can be caused by additional gas molecules in the measured gas medium, which are not considered in ITU (the standard comprises oxygen and water-vapour lines only). Parameters of ozone layer are gradually changing due to the depletion by free radicals like atomic nitric oxide, hydroxyl, nitrous oxide bromine etc.

## 3. EXAMPLE OF MEASUREMENT RESULTS

Several measurements of system performance and gas attenuation were performed with the Fabry-Perot resonator in the Department of electromagnetic field CTU in Prague. The sensitivity of the Fabry-Perot resonant cavity is the result of its very high quality factor. The absorption measurement in this case comes to measurement and consequential evaluation of quality factor of the empty and gas filled resonator. The main body of the used Fabry-Perot resonator is constructed from a stainless steel. The equipment itself comprises from a tube-shaped cavity having length of 0.55 m and diameter of 0.17 m, from two positioned spherical mirrors, from a dielectric coupling foil placed inside a cavity, and from specially designed input and output windows with dielectric lenses. The mirrors have a 0.455 m radius of curvature and are made from bronze with a golden layer with thickness of 5 μm. One of the mirrors is fixed, while the other is movable with a reversible motor, which is suitably geared down in order to provide a fine, automated adjustment of the cavity length with a 0.05 μm step. The mirror spacing is tunable from 0.495 to 0.510 m. More detailed description of measurement technique can be found in previous work of authors [8, 9].

As the first measured gas Carbonyl sulfide (OCS), which emissions are observed for instance during biomass burning or that is emitted from volcanoes, was chosen. Absorbance of OCS was measured by Fabry-Perot resonator at 48 GHz — the band allocated for HAP systems — at the pressure of 15 μ bar. From Figure 1, it can be clearly distinguished that partial regions of atmosphere containing carbonyl sulfide can cause attenuation of transmitting electromagnetic waves up to 5.3 dB/km. Nevertheless, it has to be emphasized that contrary to oxygen or water vapors only fractions of links can be influenced as well as the measured gas is present in the atmosphere at extremely low concentrations.

Figure 2 afterward depicts frequency dependence of attenuation coefficient of dibromomethane ($CH_2Br_2$) measured at the pressure of 10 μ bar. The frequency band of 71 GHz has been chosen — this band is according to the FCC [3] allocated for downlink of Mobile-Satellite Services and Fixed-Satellite Services. These systems can sometimes have due to the cost of equipments very strict power balance and therefore an additional attenuation can give rise to system failures. One can note negative values of attenuation coefficient, which would be incorrectly physically represented as gain in links; these ripples are caused by limitation of the measurement sensitivity. The measuring system is nowadays able to correctly recognize gas attenuation peaks higher than 0.2 dB/km.

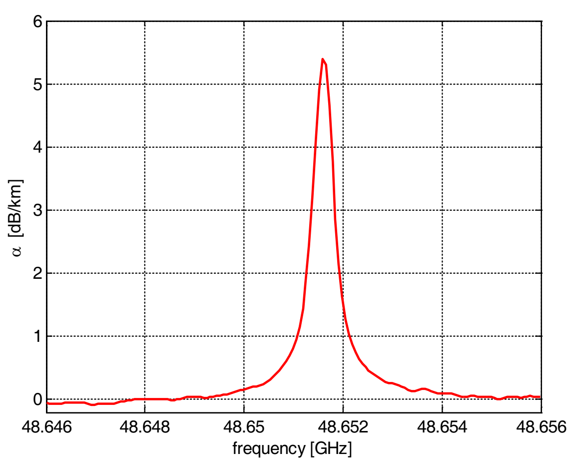



Figure 1: Measured frequency dependence of attenuation coefficient of OCS.

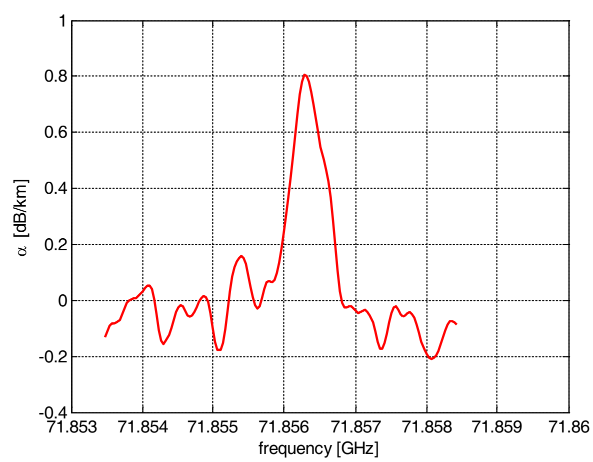



Figure 2: Measured frequency dependence of attenuation coefficient of dibromomethane.

## 4. CONCLUSIONS

Measurement system based on the Fabry-Perot resonator was presented. The resonator represents a very effective tool to ensure homogeneity of the particular gas composition, the specific gas distribution, and pressure. What more, very weak gas absorption peaks can be determined. First results from measurement campaign focused on gases which can influence stratospheric or satellite links were presented in the paper.

## ACKNOWLEDGMENT

The research is a part of the activities of the Department of Electromagnetic Field of the Czech Technical University in Prague and Department of Analytical Chemistry, Institute of Chemical Technology in the frame of the research project of the Ministry of Education, Youth and Sports of the Czech Republic No. LC06071 "Centre of quasi-optical systems and terahertz spectroscopy" and within Czech Science Foundation grant 102/08/P346 "The influence of the atmosphere on electromagnetic wave propagation for broadband stratospheric links".

## REFERENCES

1. Zvanovec, S., P. Cerny, P. Piksa, T. Korinek, P. Pechac, M. Mazanek, J. Varga, J. Koubek, and S. Urban, "The use of the Fabry-Perot interferometer for high resolution microwave spectroscopy," *Journal of Molecular Spectroscopy*, Vol. 256, No. 1, 141–145, 2009.
2. French, I. P. and T. E. Arnold, "High-Q Fabry-Perot resonator for nitric oxide absorption measurements at 150 GHz," *Review of Scientific Instruments*, Vol. 38, 1604–1607, 1967.
3. FCC Online Table of Frequency Allocations, Revised on January 25, 2010, available online http://www.fcc.gov/oet/spectrum/table/fcctable.pdf.
4. "Preferred characteristics of systems in the fixed service using high-altitude platform stations operating in the bands 47.2–47.5 GHz an 47.9-48.2 GHz," ITU-R Recommendation F.1500, 2000.
5. Townes, C. H. and A. L. Schawlow, *Microwave Spectroscopy*, McGraw-Hill, London, 1955.
6. Liebe, H. J., "Modelling attenuation and phase of radio waves in air at frequencies below 1000 GHz," *Radio Science*, Vol. 16, 1183–1199, 1981.
7. International Telecommunications Union, "Attenuation by atmospheric gases," ITU-R Recommendation P.676-6, 2005.
8. Zvanovec, S., P. Piksa, M. Mazanek, and P. Pechac, "A study of gas and rain propagation effects at 48 GHz for HAP scenarios", *EURASIP Journal on Wireless Communications and Networking*, Vol. 5, No. 1, 1–7, 2008, http://www.hindawi.com/GetArticle.aspx?doi=10.1155/2008/734216.
9. Piksa, P., S. Zvanovec, and P. Cerny, "Specific millimeter-wave features of Fabry-Perot resonator for spectroscopic measurements," *Microwave and Millimeter Wave Technologies: From Photonic Bandgap Devices to Antenna and Applications*, Edited by I. Minin, in print, In-Teh, Viena, 2010.

# Transient Electromagnetic Field of an Electric Line Source above a Plane Drude Model Plasmonic Half-space

**B. J. Kooij**
IRCTR, Delft University of Technology, The Netherlands

**Abstract**— The enhancement of the electromagnetic field at the surface of a Drude metal is classified as the so called surface plasmon effect. This surface plasmon effect in a classical Drude model has been investigated extensively in the frequency domain. However, to understand what is actually happening to the electromagnetic field at the near vicinity of a Drude metal a thorough analysis in the space-time domain is needed. Once the space-time domain counterpart of the surface plasmon effect is found, one is able to describe and classify the physical phenomena that are playing a role in the occurrence of this effect.

## 1. INTRODUCTION

The phenomenon of surface plasmon resonance of metallic structures has been known and theoretically described for more than fifty years. Nowadays, the plasmonic nanostructures are used in a variety of applications, such as molecular sensing and imaging devices, optical devices and metamaterials. One of the materials in which plasmonic effects are encountered are metals that have a frequency dependant conductivity described by the Drude model. Many researchers investigated Drude model metals in the frequency domain in order to determine the optical behaviour of metals within a certain frequency range. However, to understand what phenomena are at work when the electromagnetic field interacts in the near vicinity with a Drude model metal, a thorough analysis in the space-time domain is needed. Once the space-time domain counterpart of the surface plasmon effect is identified, a description and classification of the physical phenomena that occur can be given. These results can be of great importance in the field of time-domain terahertz spectroscopy.

The pulsed electromagnetic radiation from a line source above a plane plasmonic half-space, modeled by a classical Drude model, is investigated. In order to obtain closed-form expressions for the reflected electric field anywhere above the plasmonic half-space, the reflection factor $R_E$ that corresponds to the Drude metal half-space is mathematically written in the form of a Laplace transform integral, enabling the application of the Cagniard-de Hoop method. In the Cagniard-de Hoop method the Laplace transform with respect to time is used, in which the Laplace parameter $s$ is kept real and positive in order to ensure causality in the space-time domain expressions for the electromagnetic field. Attention is paid to investigate which parts of the Laplace transform integral representation of $R_E$ give rise to the surface-plasmon effect by means of enhancement of the reflected electromagnetic field at the near vicinity of the Drude metal interface. The derived space-time domain expressions for the reflected electric field explicitly show the special time and space dependence that exists in the reflection factor at the interface of the plasmonic half-space. Consequently, leading to a better understanding of the physical effects, that occur at the interface of classical Drude metals in the space-time domain.

Finally, a suitable source signature is proposed, which enables, without loss of generality, a suitable determination of analytical closed-form inverse Laplace transforms of the distinct source functions. Numerical results that support the conclusions made in this paper will be presented at the Symposium.

## 2. DESCRIPTION OF THE CONFIGURATION

The reflected electromagnetic wavefield in the upper half-space of two homegeneous, isotropic, semi-infinite media is investigated. To specify the position in the configuration, we employ the coordinates $\{x_1, x_2, x_3\}$ with respect to a fixed, orthogonal, Cartesian reference frame with origin $\mathcal{O}$ and three mutually perpendicular base vectors $\{\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3\}$ of unit length each. In the indicated order the three base vectors form right-handed coordinate system. In accordance with previous papers, $\mathbf{i}_3$ points vertically upward towards the upper medium occupying the half-space $0 < x_3 < \infty$, which consists of vacuum with permittivity $\epsilon_0$ and permeability $\mu_0$. The lower medium occupying the half-space $-\infty < x_3 < 0$, consists of metal whose properties can be described by a Drude model, with permittivity $\epsilon_0$ and a conductivity relaxation function defined by

$$\mathbf{J}(\mathbf{x},t) = \frac{\eta}{\mu_0}\int_0^t \exp[-\beta(t-\tau)]\mathbf{E}(\mathbf{x},\tau)\,d\tau, \tag{1}$$

in which $\eta$ is defined by

$$\eta = \frac{\omega_p^2}{c^2}, \quad \text{with } \omega_p = \sqrt{\frac{4\pi n e^2}{m_e}}, \tag{2}$$

in which $c = (\epsilon_0\mu_0)^{-1/2}$, $\omega_p$ denotes the radial plasma frequency, $n$ is the density of free electrons, $e$ the electron charge and $m_e$ the effective electron mass. In Eq. (1) the parameter $\beta$ is the phenomenological relaxation constant.

## 3. METHOD OF SOLUTION

The electromagnetic field in the configuration is described in terms of the electric field $\mathbf{E}$ and the magnetic field $\mathbf{H}$. The action of the source is characterized by specifying the volume density of its electric current $\mathbf{J}_S$, which is defined by

$$\mathbf{J}_S = \mathbf{j}_{\|}(t)\delta(x_1,\, x_3 - h_T), \tag{3}$$

where $h_T$ with $h_T > 0$ denotes the vertical distance of the line source to the interface. In any domain where the field quantities are continuously differentiable, they satisfy the following electromagnetic field equations

$$\nabla\times\mathbf{H} = \epsilon_0\partial_t\mathbf{E} + \mathbf{J} + \mathbf{J}_S, \tag{4}$$
$$\nabla\times\mathbf{E} = -\mu_0\partial_t\mathbf{H} \tag{5}$$

together with,

$$\nabla\cdot(\epsilon_0\partial_t\mathbf{E} + \mathbf{J} + \mathbf{J}_S) = 0, \tag{6}$$
$$\nabla\cdot\mu_0\partial_t\mathbf{H} = 0, \tag{7}$$

in which $\mathbf{J}$ is defined by Eq. (1) in the plasmonic half-space and $\mathbf{J} = 0$ elsewhere. In our analysis we separate the total field into the incident field $\{\mathbf{E}^i,\,\mathbf{H}^i\}$, the reflected field $\{\mathbf{E}^r,\,\mathbf{H}^r\}$ and the transmitted field $\{\mathbf{E}^t,\,\mathbf{H}^t\}$ according to

$$\{\mathbf{E},\,\mathbf{H}\} = \left\{\mathbf{E}^i,\,\mathbf{H}^i\right\} + \{\mathbf{E}^r,\,\mathbf{H}^r\}, \text{ for } x_3 > 0, \tag{8}$$
$$\{\mathbf{E},\,\mathbf{H}\} = \left\{\mathbf{E}^t,\,\mathbf{H}^t\right\}, \text{ for } x_3 < 0, \tag{9}$$

where the incident field is the field that the source would generate if no boundary were present. Across the interface of the two media, we have the continuity of the tangential components $E_1$, $E_2$, $H_1$ and $H_2$ and we take into account the radiation condition, i.e., the fact that the incident field travels away from the source, while the reflected and transmitted field travel away from the interface. Next, we incorporate the two-dimensionality of the configuration. To this end, we decompose each field vector into a component that is parallel to the line source (denoted by the subscript $\|$) and a component in the plane perpendicular to it (denoted by by the subscript $\perp$). Taking into account that $\nabla = \nabla_\perp$ and $\mathbf{J}_S$ is parallel to the interface, we can rewrite Eqs. (4)–(7) in a form where only an $E$-polarized field is generated for which $\mathbf{E}_{\|} \neq 0$ and $\mathbf{H}_{\|} = 0$.

Finally, we subject the field quantities to a Laplace transformation with respect to time according to

$$\hat{\mathbf{E}}(\mathbf{x},s) = \int_{\tau=0}^{\infty} \mathbf{E}(\mathbf{x},\tau)\exp(-s\tau)\,d\tau, \tag{10}$$

and a spatial Fourier transformation with respect to $x_1$ according to

$$\tilde{\mathbf{E}}(\alpha,x_3,s) = \int_{-\infty}^{\infty} \exp(s\alpha x_1)\hat{\mathbf{E}}(\mathbf{x},s)\,dx_1, \;\; \alpha\in\mathcal{I}, \tag{11}$$

together with the inverse transformation

$$\hat{\mathbf{E}}(\mathbf{x},s) = -\frac{s}{2\pi i}\int_{-i\infty}^{i\infty} \exp(-s\alpha x_1)\tilde{\mathbf{E}}(\alpha,x_3,s)\,d\alpha, \tag{12}$$

in which $\mathcal{I}$ denotes the imaginary axis.

## 4. REFLECTED ELECTRIC FIELD REPRESENTATION

In this section, we determine the representation of the reflected electric field excited by the localized line source defined in Eq. (3). Application of the transformations (10) and (11) and solving Maxwell's equation for the generated $E$-polarized incident electric field yields

$$\tilde{\mathbf{E}}_{\|}^{i} = -\frac{\mu_0 \hat{\mathbf{j}}_{\|}(s)}{2\gamma} \exp(-s\gamma|x_3 - h_T|), \tag{13}$$

in which

$$\gamma = \left(c^{-2} - \alpha^2\right)^{1/2}, \ \ \mathrm{Re}\{\gamma\} \geq 0, \ \text{ with } \ c = (\epsilon_0\mu_0)^{-1/2}. \tag{14}$$

Similarly we obtain for the reflected electric field $\tilde{\mathbf{E}}_{\|}^{r}$,

$$\tilde{\mathbf{E}}_{\|}^{r} = R_E \tilde{\mathbf{E}}_{\|}^{i}(\alpha, 0, s)\exp(-s\gamma x_3), \ \text{ with } \ 0 < x_3 < \infty, \tag{15}$$

in which $R_E$ denotes the reflection factor which is given by

$$R_E = \frac{\gamma - \gamma_p}{\gamma + \gamma_p}, \ \text{ with } \ \gamma_p = \left(\gamma^2 + \frac{\eta}{s(s+\beta)}\right)^{1/2}, \ \ \mathrm{Re}\{\gamma_p\} \geq 0. \tag{16}$$

In Eq. (15), $h_T$ denotes the distance from the electric line source to the plasmonic interface and $x_3$ denotes the vertical position of the receiver above the interface. Next, we rewrite the reflection factor $R_E$ in the form of a Laplace transform integral in order to allow the application of the Cagniard-de Hoop technique [1]. First, we multiply both the numerator and the denominator of Eq. (16) with $\gamma - \gamma_p$, resulting in

$$R_E = -\left[\frac{2s(s+\beta)}{\eta}\gamma^2 + 1\right] + \frac{2\gamma}{\gamma_p}\left[\frac{s(s+\beta)}{\eta}\gamma^2 + 1\right]. \tag{17}$$

The expression $\gamma/\bar{\gamma}$ can be recognized as a Laplace transform integral according to [2], formula 29.3.55, as

$$\begin{aligned} \frac{\gamma}{\gamma_p} &= \sqrt{s(s+\beta)}\left[\frac{1}{\sqrt{(s+\beta/2)^2 + (\eta/\gamma^2 - \beta^2/4)}}\right] \\ &= \sqrt{s(s+\beta)}\int_0^\infty J_0\left(\frac{\kappa\beta}{2\gamma}\sqrt{4\eta/\beta^2 - \gamma^2}\right)\exp\left(-\beta\kappa/2\right)\exp(-s\kappa)\,d\kappa, \end{aligned} \tag{18}$$

in which $J_0$ denotes the Bessel function of order 0. Then, we can rewrite $R_E$ in Eq. (17) as

$$\begin{aligned} R_E &= -1 + s(s+\beta)\tilde{W}_1(\alpha) + s^{3/2}(s+\beta)^{3/2}\int_0^\infty \tilde{W}_2(\alpha,\kappa)\exp(-s\kappa)\,d\kappa \\ &\quad + s^{1/2}(s+\beta)^{1/2}\int_0^\infty \tilde{W}_3(\alpha,\kappa)\exp(-s\kappa)\,d\kappa, \end{aligned} \tag{19}$$

in which

$$\tilde{W}_1(\alpha) = -\frac{2\gamma^2}{\eta}, \tag{20}$$

$$\tilde{W}_2(\alpha,\kappa) = \frac{2\gamma^2}{\eta}J_0\left(\frac{\kappa\beta}{2\gamma}\sqrt{4\eta/\beta^2 - \gamma^2}\right)\exp(-\beta\kappa/2), \tag{21}$$

and

$$\tilde{W}_3(\alpha,\kappa) = 2J_0\left(\frac{\kappa\beta}{2\gamma}\sqrt{4\eta/\beta^2 - \gamma^2}\right)\exp(-\beta\kappa/2). \tag{22}$$

Now, we can rewrite the reflected electric field $\hat{\mathbf{E}}_{\|}^{r}(\mathbf{x}, s)$ in the form

$$\hat{\mathbf{E}}_{\|}^{r}(\mathbf{x}, s) = \hat{\mathbf{S}}_{\|,0}(s)\hat{g}_0(\mathbf{x}, s) + \hat{\mathbf{S}}_{\|,1}(s)\hat{g}_1(\mathbf{x}, s) + \hat{\mathbf{S}}_{\|,2}(s)\hat{g}_2(\mathbf{x}, s) + \hat{\mathbf{S}}_{\|,3}(s)\hat{g}_3(\mathbf{x}, s). \tag{23}$$

In Eq. (23), the source functions $\hat{\mathbf{S}}_{\|}$ are given by

$$\hat{\mathbf{S}}_{\|,0}(s) = \hat{\mathbf{j}}_{\|}(s), \tag{24}$$

$$\hat{\mathbf{S}}_{\|,1}(s) = s(s+\beta)\hat{\mathbf{j}}_{\|}(s), \tag{25}$$

$$\hat{\mathbf{S}}_{\|,2}(s) = s^{3/2}(s+\beta)^{3/2}\hat{\mathbf{j}}_{\|}(s), \tag{26}$$

$$\hat{\mathbf{S}}_{\|,3}(s) = s^{1/2}(s+\beta)^{1/2}\hat{\mathbf{j}}_{\|}(s), \tag{27}$$

and the Green's functions $\hat{g}(\mathbf{x}, s)$ are given by

$$\hat{g}_0(\mathbf{x}, s) = -\frac{\mu_0}{2\pi i}\int_{-i\infty}^{i\infty} \frac{1}{2\gamma}\exp\{-s[\alpha x_1 + \gamma(x_3 + h_T)]\}\, d\alpha, \tag{28}$$

$$\hat{g}_1(\mathbf{x}, s) = \frac{\mu_0}{2\pi i}\int_{-i\infty}^{i\infty} \frac{\tilde{W}_1}{2\gamma}\exp\{-s[\alpha x_1 + \gamma(x_3 + h_T)]\}\, d\alpha, \tag{29}$$

$$\hat{g}_{2,3}(\mathbf{x}, s) = \frac{\mu_0}{2\pi i}\int_{-i\infty}^{i\infty} \frac{\exp\{-s[\alpha x_1 + \gamma(x_3 + h_T)]\}}{2\gamma}\int_{\kappa=0}^{\infty} \tilde{W}_{2,3}\exp(-s\kappa)\, d\kappa\, d\alpha, \tag{30}$$

respectively. It is noted that $\hat{g}_0(\mathbf{x}, s)$ represents the total reflected Green's function in case $\eta \to \infty$, since $\hat{g}_1(\mathbf{x}, s)$, $\hat{g}_2(\mathbf{x}, s)$ and $\hat{g}_3(\mathbf{x}, s)$ vanish. In the next section we will show that the integrals at the right-hand side of Eqs. (28)–(30) can be transformed such that the space-time domain counterpart of the reflected wavefield in Eq. (23) can be expressed as

$$\begin{aligned}\mathbf{E}^r_{\|}(\mathbf{x}, t) = {}& U(t - R/c)\int_{R/c}^{t} \mathbf{S}_{\|,0}(t-\tau)g_0(\mathbf{x}, \tau) + \mathbf{S}_{\|,1}(t-\tau)g_1(\mathbf{x}, \tau) + \mathbf{S}_{\|,2}(t-\tau)g_2(\mathbf{x}, \tau) \\ & + \mathbf{S}_{\|,3}(t-\tau)g_3(\mathbf{x}, \tau)d\tau,\end{aligned} \tag{31}$$

in which $R = [x_1^2 + (x_3 + h_T)^2]^{1/2}$ denotes the distance from the image of the electric line source to the point of observation and $U$ denotes the Heaviside unit step function.

## 5. SPACE-TIME EXPRESSIONS

In order to obtain space-time domain expressions for the reflected electric field $\mathbf{E}^r_{\|}(\mathbf{x}, t)$ we apply the Cagniard-de Hoop technique. In the inverse spatial Fourier transformation with variable of integration $\alpha$ we introduce a parametrization along the variable $\tau$, which is real and positive. In view of subsequent deformation of the path of integration, we take $\mathrm{Re}(\bar{\gamma}) \geq 0$ not only on the imaginary $\alpha$-axis but everywhere in the complex $\alpha$-plane. This implies that branch cuts are introduced along $\{\alpha \in \mathcal{C} | c^{-1} < |\mathrm{Re}(\alpha)| < \infty, \mathrm{Im}(\alpha) = 0\}$. Now the path of integration in the complex $\alpha$-plane is deformed into a Cagniard-de Hoop contour defined by

$$\begin{aligned}\mathrm{Re}[\alpha x_1 + \bar{\gamma}(x_3 + h_T)] &= \tau, \\ \mathrm{Im}[\alpha x_1 + \bar{\gamma}(x_3 + h_T)] &= 0.\end{aligned} \tag{32}$$

Let $\alpha = \bar{\alpha}(\tau)$ denote the parametric representation of the solution of Eq. (32) in the upper half of the complex $\alpha$-plane, then the contour consists of $\bar{\alpha}$ together with its complex conjugate $\bar{\alpha}^\star$. We obtain

$$\bar{\alpha}(\tau) = \frac{x_1\tau}{R^2} + i\frac{x_3 + h_T}{R^2}\left[\tau^2 - T^2\right]^{1/2}, \tag{33}$$

for $T \leq \tau < \infty$, with $T = R/c$ and where the contour is a branch of a hyperbola. Along the contour we further have

$$\bar{\gamma}(\bar{\alpha}) = \frac{(x_3 + h_T)\tau}{R^2} - i\frac{r}{R^2}\left[\tau^2 - T^2\right]^{1/2}. \tag{34}$$

Taking into account the symmetry of the contour with respect to the real $\alpha$-axis, we obtain

$$g_0(\mathbf{x}, t) = -\frac{\mu_0 U(t - R/c)}{2\pi\left[t^2 - T^2\right]^{1/2}}, \tag{35}$$

$$g_1(\mathbf{x}, t) = \frac{\mu_0 U(t - R/c)\mathrm{Re}\left\{\bar{W}_1(\bar{\alpha}(t))\right\}}{2\pi\left[t^2 - T^2\right]^{1/2}}, \tag{36}$$

$$g_{2,3}(\mathbf{x}, t) = \frac{\mu_0 U(t - R/c)}{2\pi}\int_T^t \frac{\mathrm{Re}\left\{\bar{W}_{2,3}(\bar{\alpha}(\zeta), t-\zeta)\right\}}{\left[\zeta^2 - T^2\right]^{1/2}}\,d\zeta. \tag{37}$$

In Eqs. (36) and (37), $\bar{W}_1$ and $\bar{W}_{2,3}$ are obtained from Eqs. (20)–(22) by replacing $\gamma$ with $\bar{\gamma}$.

## 6. THE SOURCE SIGNATURE

In this section, we will present a source signature $\mathbf{j}_{\parallel}(t)$ that is chosen such that the time domain counterparts of Eqs. (24)–(27) can be found analytically. Since we are especially interested in space-time domain phenomena that give rise to the existence of surface wave like effects in the frequency domain, we can without loss of generality choose a suitable source signature that will simplify our computational efforts. If we take $\hat{\mathbf{j}}_{\parallel}(s) = s^{-3/2}(s+\beta)^{-3/2}$ to be our source signature, with [2], formula 29.3.50, we obtain the time domain counterpart

$$\hat{S}_0(s) = s^{-3/2}(s+\beta)^{-3/2} \overset{\mathcal{L}^{-1}}{\rightarrow} S_0(t) = t\exp(-\beta t/2)I_1(\beta t/2)U(t), \tag{38}$$

in which $I_1$ denotes the modified Bessel function of the first kind and order 1. Then the source functions $S_2(t) = \delta(t)$ and $S_1(t)$, $S_3(t)$ are found to be

$$\hat{S}_1(s) = s^{-1/2}(s+\beta)^{-1/2} \overset{\mathcal{L}^{-1}}{\rightarrow} S_1(t) = \exp(-\beta t/2)I_0(\beta t/2)U(t), \tag{39}$$

$$\hat{S}_3(s) = s^{-1}(s+\beta)^{-1} \overset{\mathcal{L}^{-1}}{\rightarrow} S_3(t) = \beta^{-1}(1 - \exp(-\beta t))U(t), \tag{40}$$

in which $I_0$ denotes the modified Bessel function of the first kind and order 0.

## 7. CONCLUSION

Numerical results of the space-time domain reflected electric field above and at the near vicinity of the interface will be presented at the Symposium. However, from the space-time expressions that we have derived in Eqs. (35)–(37) we can at this stage draw some preliminary conclusions. Since it is known that the enhancement of the electromagnetic field is mainly located close to the plasmonic interface, we can conclude that the surface plasmon effect is strongly space and time dependant. If we take a close look at the derived space-time expressions (35)–(37) to seek those parts that are responsible for the electromagnetic field enhancements at the near vicinity of the plasmonic half-space, we search for a connection between space and time that is only present in $\bar{W}_2$ and $\bar{W}_3$. Consequently, $g_{2,3}(\mathbf{x}, t)$ is the part of the space-time domain electric field that is responsible for the surface plasmon effect. Now, from Eq. (34) we observe that in the case the source and receiver locations are close to the plasmonic interface $\bar{\gamma}$ becomes imaginary, which changes the Bessel function $J_0$ in Eqs. (21) and (22) into a modified Bessel function $I_0$ with a real argument. Hence, when the receiver location is moved away from the plasmonic interface the Cagniard contour takes on the form of a branch of a hyperbola according to Eq. (33) and $\bar{\gamma}$ becomes complex, see (34), leading to a complex argument for the Bessel functions $J_0$ in Eqs. (21) and (22). We can conclude from this analysis that the enhancement of the electromagnetic field at the plasmonic interface is mathematically described by the Bessel function $J_0$ whose amplitude increases when the receiver location is moved towards the interface. These conclusions are supported by the numerical results that will be shown at the Symposium.

## REFERENCES


1. Kooij, B. J., "The transient electromagnetic field of an electric line source above a plane conducting earth," *IEEE Trans. on Electromagnetc Compatibility*, Vol. 33, No. 1, 19–24, 1991.
2. Abramowitz, M. and I. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Mineola, NY, 1965.

# Optical Activity in Organic Metamaterials

**N. Wongkasem and C. Kamtongdee**
Department of Electrical Engineering, Faculty of Engineering, Khon Kaen University
Khon Kaen 40002, Thailand

**Abstract**— Parameters which maneuver chirality and optical activity properties in organic metamaterials are investigated. Molecular optical rotation of a four-atom chain molecule and optical rotation constant of artificial organic structures are proposed to regulate in order to control chirality and optical activity. Artificial helix DNA is presented as a case study to monitor optical activity effects in organic metamaterials. Circular polarization effect, electromagnetic parameters and index of refraction are recorded. Once chirality and optical activity can be controlled at molecular level, designed organic metamaterials can be used in electromagnetic applications at high frequencies, up to x-rays regimes.

## 1. INTRODUCTION

Most natural macromolecules e.g., organic molecules, solid materials, as well as spin-polarized gases are chiral and optically active in the optical frequency regime. Chirality and electromagnetic parameters i.e., permittivity and permeability are the main factors controlling the refractive indices of the materials and the rotations of the out-going waves. The chirality parameter, $\kappa$, is related to optical rotary dispersion (ORD) and circular dichroism (CD), the two unique electromagnetic properties of the media. ORD or optical activity creates the rotation of the polarization plane, resulting from the splitting waves [1] propagating inside the media with different phase velocities; whilst CD causes a change of polarization ellipticity, according to the differential absorption of the left and right circularly polarized waves. Both CD and ORD are direct proportional to an optical rotation angle, the specific angle through which plane-polarized light is rotated on passing through an enantiomerically enriched solution. The optical angle can be used as a quantitative probe of molecular chirality [2].

In this paper, we study parameters which influence the chirality and optical activity in organic metamaterials. Rotation constant and dihedral angle are recommended to adjust in order to design any desired chirality and optical activity. We also discuss how optical activity affects artificial helix DNA, our case study. Circular polarization effect, electromagnetic (EM) parameter and index of refraction are observed.

## 2. OPTICAL ROTATION VS. OPTICAL ACTIVITY

An organic compound is a large class of chemical compounds whose molecules contain carbon. The relation between the optical activity and geometrical structure of organic compounds has been analyzed [3, 4]. The molecular optical rotation of a four-atom chain, as shown in Figure 1, depends on the dihedral angle sine formed with the two terminal atoms and sum of each of those parameters [3]. It was proved that the sign and the magnitude of the molecular optical rotation of a substance could be related to the structure, conformation and absolute configuration [4]. Optical rotation angle ($\varphi$) and ellipticity ($\theta$) signify optical rotatory dispersion (ORD) and circular dichroism (CD), correspondingly. Note that in chemistry, molar rotation and molar ellipticity, both in unit of (deg cm$^2$ dmol$^{-1}$), are often used to represent ORD and CD. The Kramers-Kronig (KK) relationship holds between ORD and CD, which are a real and imaginary part of complex optical rotation ($\Phi$) or complex optical angle, as shown in Equation (1).

$$\Phi = \varphi + i\theta \tag{1}$$

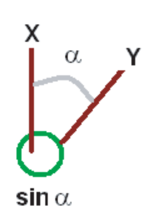



Figure 1: $X$-$C$-$C$-$Y$, dihedral angle.

ORD ($\varphi$) or optical activity is related to the optical rotatory parameter ($\beta$) as presented in Equation (2).

$$\varphi = \frac{1}{3} f^2 \left(n^2 + 2\right) N \mu_0 \beta \tag{2}$$

where $f$ is the frequency (rad/s). $(n^2 + 2)$ is a very large number (in $10^8$ range), obtained from Ab initio calculation. $N$ is an inverse volume unit ($\mathrm{m}^{-3}$) and $\mu_0$ is free space permeability. It can be seen that $\beta$ is a main variable to control ORD. A specific optical rotation $[\alpha]_D$, calculated from molecular optical rotation of a substance $[M]_D$, is direct proportional to $\beta$. $[M]_D$, $\mathrm{deg \cdot cm^2 \cdot dmol^{-1}}$, representing molecular optical rotation of a substance, can be calculated from Equation (3) [5], where $f_i$ is the population of the conformer $i$. The dihedral angle $\alpha_k$ is formed by four consecutive atoms for each type of dihedral angle, as shown in Figure 1, while $K_{XY}$ is a rotation constant.

$$[M]_D = \sum_i f_i \sum_k K_{XY} \sin \alpha_k \tag{3}$$

Several parameters are linked to chirality and optical activity; however, there are only some which can be manipulated in order to control both $\kappa$ and $\varphi$ effectively. Here, we recommend focusing on the rotation constant ($K_{XY}$) and dihedral angle ($\alpha_k$). An example of an artificial organic structure will be shown in the presentation.

## 3. CASE STUDY: ARTIFICIAL HELIX DNA

DNA or deoxyribonucleic acid belongs to the group of organic compounds known as nucleic acids. It is formed in double helix shape. An example of DNA structure is shown in Figure 3(a). Three main parts, i.e., sugar phosphate backbone, base and hydrogen bonds, form DNA structure; however, in most artificial helix studies, only the backbone part is taken into consideration. Helix is known as one of the most promising three-dimensional chiral media. Artificial single helical structure, introduced as broadband circular polarizer, has been proposed [6]. Physical geometry, i.e., helix pitch, helix radius, two-dimensional lattice constant, wire radius, number of helix pitch, and angle of incidence of single helix was studied. It was found that the helix optical properties are controlled by resonances of the individual helices as well as interaction effects [7].

In this session, we are investigating optical activity in artificial helix DNA, focusing on its circular polarization (CP) and electromagnetic responses. Helices are designed using DNA dimensions,

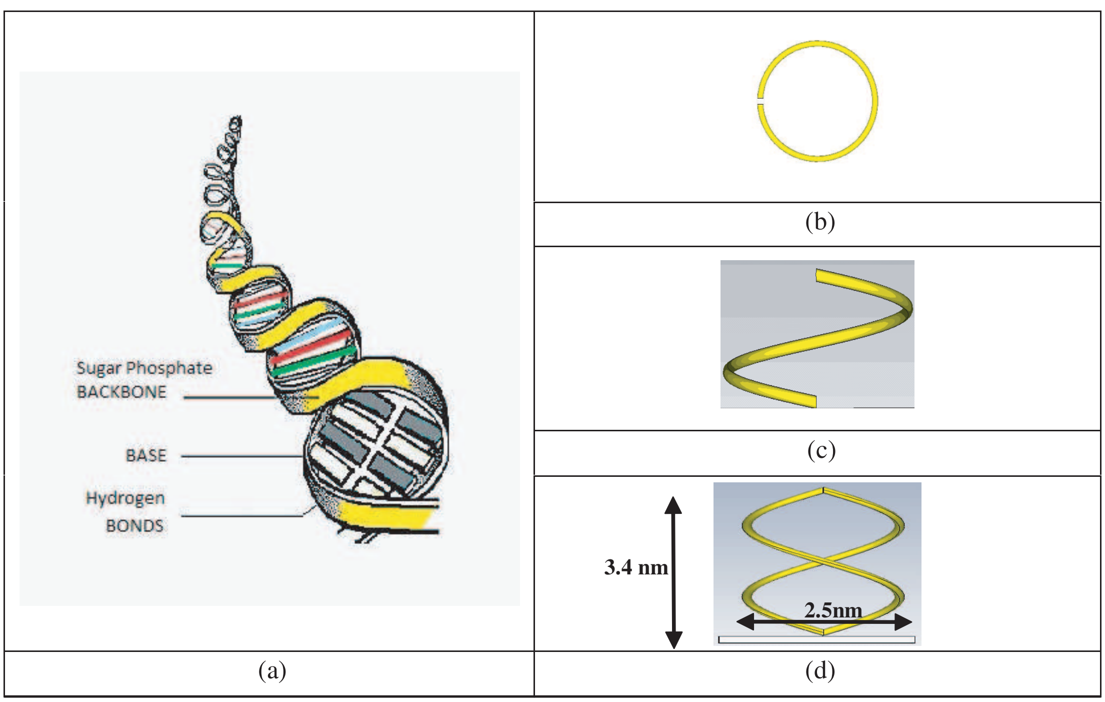



Figure 2: (a) DNA structure (double helix), (b) split ring resonator (SRR), (c) single helix and (d) double helix with DNA dimension: 2.5 nm (diameter) × 3.4 nm (helix turn).

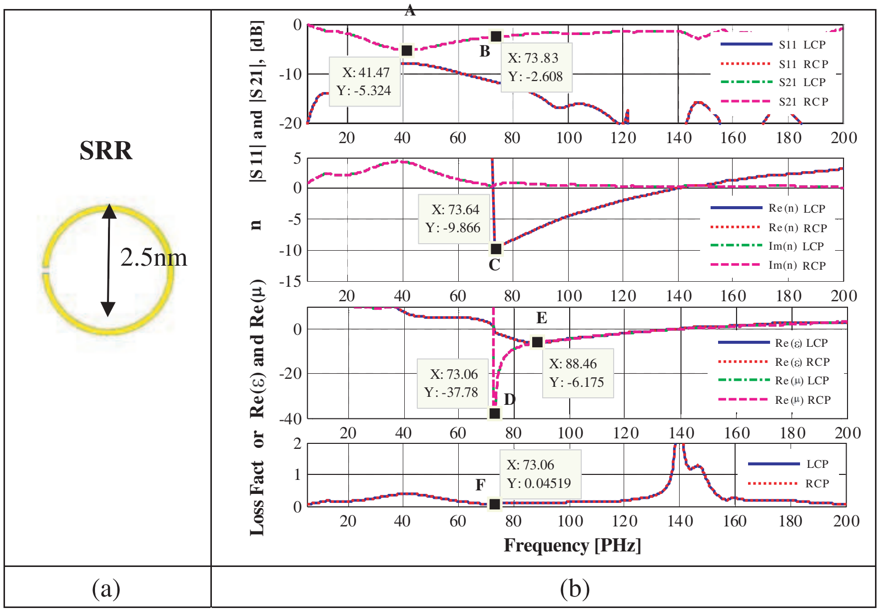



Figure 3: (a) Split ring resonator (SRR) and (b) the comparison of scattering and material parameters from LCP and RCP excitation.

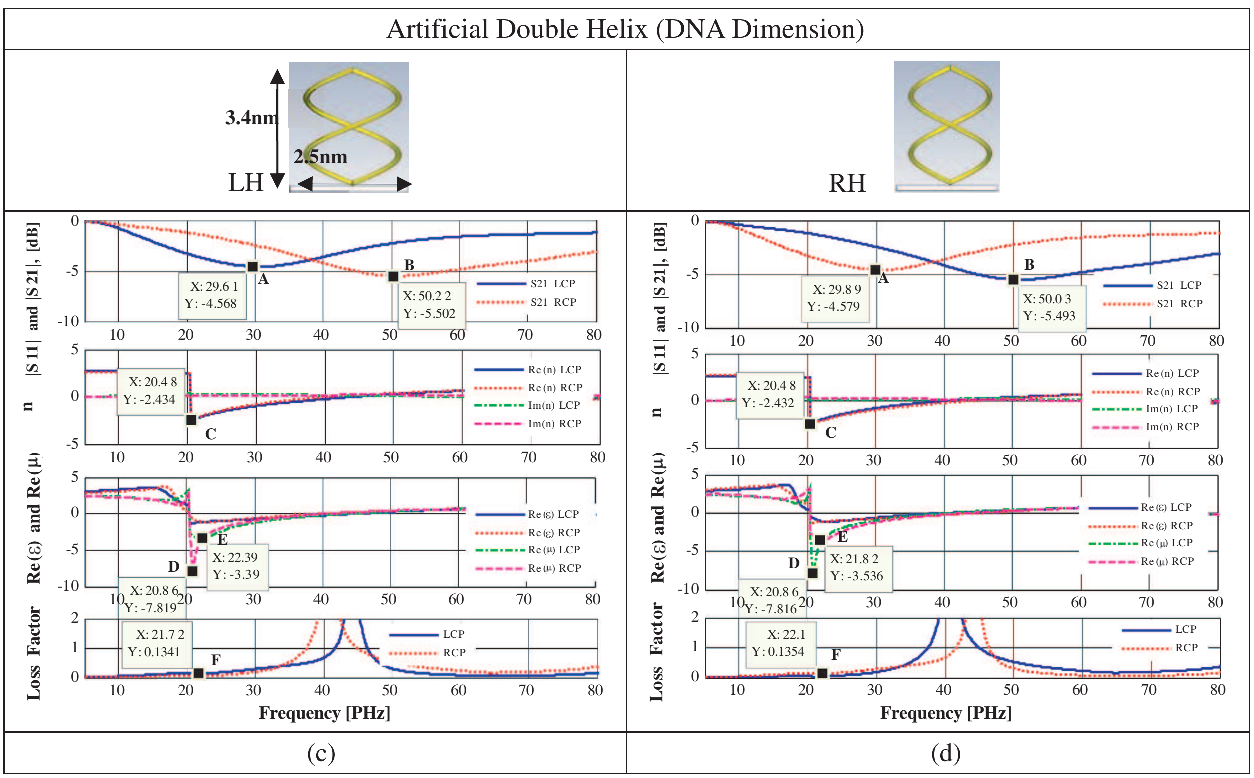



Figure 4: (a) Clockwise (right-handed) helix DNA and (b) the comparison of scattering and material parameters from LCP and RCP excitation. (c) Counter clockwise (left-handed) helix DNA and (d) the comparison of scattering and material parameters from LCP and RCP excitation.

i.e., 2.5 nm (diameter) × 3.4 nm (helix turn), without considering the base and hydrogen bonds. Left and right circular polarized waves are individually launched in the same direction as the DNA (or helix) axis. Transmission properties, as well as EM material parameters, i.e., refractive index, permittivity and permeability from each excitation are recorded. The helix is twisted in clockwise (right handed or RH) and counter clockwise (left handed or LH) to generate the structure's handedness. The properties will be compared with those from a well known bi-anisotropic planar split ring resonator (SRR), shown in Figure 2(b). Single helix and double helix with DNA dimension: 2.5 nm (diameter) $\times$ 3.4 nm (helix turn) are illustrated in Figures 2(c) and (d). CST Microwave Studio® software is used to observe the transmission properties and to determine scattering parameters, which afterward are used to extract EM material parameters.

Considering the normal excitation onto the SRR plane, this structure does not contain any chirality or handedness; therefore, the scattering parameters, both transmission coefficient ($S_{21}$) and reflection coefficient ($S_{11}$) of SRR from both LCP and RCP waves are identical. The simulation results are presented in Figure 3(b). $S_{21}$ from both eigenwaves generates a resonance at 41.47 PHz (point A). The (-)$n$ passband with low loss (4.52% shown at point F) begins around 73 PHz. This structure produces high (-)$\mu$ (peaks at $-37.78$) at point D; while the highest (-)$\varepsilon$ is at point E.

On the other hand, circular polarization effect, resulting from optical activity, is observed in three-dimensional double helical structures, as shown in Figure 4. Opposite handedness structures produce reverse transmission results, i.e., left-handed helical (counter clockwise) structure in Figure 4(c) shows a potential resonance at point A ($-4.57$ dB at 29.61 PHz) and point B ($-5.50$ dB at 50.22 PHz) when it is excited by LCP and RCP plane wave, respectively, and vice versa. Within the resonance band, both LCP and RCP waves generate broad (-)$n$, from 20 PHz to 50 PHz, with low loss (loss factor $\sim 0$, point F). There is some difference between $\mu$ and $\varepsilon$ from each excitation, as illustrated, as point D and E. Note that LH helix with RCP excitation provides quite similar polarization and EM properties to those of the RH helix with LCP excitation. The properties of LH helix with LCP and RH helix with RCP are also nearly identical.

## 4. CONCLUSION

A dihedral angle of a four-atom chain molecule, main structure of organic compounds, and optical rotation constant, discussed among other influenced parameters, are proposed to regulate to be able to control chirality and optical activity of artificial organic structures. Polarization and electromagnetic properties, resulting from optical activity are investigated in artificial helix DNA. The DNA based helices are shown to have well-built reverse circular polarization effect.

## ACKNOWLEDGMENT

This work was supported by the Coordinating Center for Thai Government Science and Technology Scholarship Students (CSTS), National Science and Technology Development Agency (NSTDA) and the Faculty of Engineering, Khon Kaen University, Khon Kaen, Thailand. We also wish to thank TRIDI, the Telecommunications Research and Industrial Development Institute for funding the Optical and Wireless Communications Research Laboratory, particularly CST Microwave Studio® software used in simulation study.

## REFERENCES

1. Wongkasem, N. and A. Akyurtlu, "Light splitting effects in chiral metamaterials," *J. of Optics*, 035101, 2010.
2. Kondru, R. K., P. Wipf, and D. N. Beratan, "Atomic contribution to the optical rotation angle as a quantitative probe of molecular chirality," *Science*, Vol. 282, 2247–2250, 1998.
3. Whiffen, D. H., "Optical rotation and geometrical structure," *Chem. Ind.*, No. 479, 964–968, 1956.
4. Brewster, J. H., "A useful model of optical activity," *J. Am. Chem. Soc.*, Vol. 81, 5475, 5483, 5493, 1959.
5. Sastre, J. A. L., R. N. Miquel, A. G. Ramos, and D. Galisteo, "Empirical relationship between structure and molecular optical rotation in six-membered cyclic compounds," *J. of Molecular Structure*, Vol. 526, 1–15, 2000.
6. Gansel, J. K., M. Thiel, M. S. Rill, M. Decker, K. Bade, V. Saile, G. V. Freymann, S. Linden, and M. Wegener, "Gold helix photonic metamaterial as broadband," *Science*, Vol. 325, No. 5947, 1513–1515, 2009.
7. Gansel, J. K., M. Wegener, S. Burger, and S. Linden, "Gold helix photonic metamaterials: A numerical parameter study," *Optics Express*, Vol. 18, No. 2, 1059–1069, 2010.