An Efficient 3D Integral Equation Method for Computation of Electromagnetic Wavefields in a Layered Configuration Containing Inhomogeneous Objects

P. M. van den Berg[1], **A. Abubakar**[2], **and T. M. Habashy**[2]

[1]Delft University of Technology, The Netherlands
[2]Schlumberger-Doll Research, Cambridge, USA

Abstract— This paper is concerned with the source-type of integral equation to compute the electromagnetic scattering by an inhomogeneous 3D object in a planar layered medium in the frequency domain. By decomposing the scattered field into a particular and a general constituent, the structure of the integral operator of the integral equation is constructed. The particular constituent represents the scattered field inside the layer that embodies the contrasting object, due to the presence of virtual contrast sources inside the inhomogeneous object, while the general constituent represents the interaction with the other layers due to the presence of source distributions on each side of the layer that embodies the contrasting object. The particular constituent has a convolution structure in all spatial directions. The general constituent consists of two terms; one has again a convolution structure with respect to all spatial coordinates, while the other has a convolution structure with respect to the horizontal coordinates and a correlation structure in the vertical coordinates. These properties facilitate a fast and efficient computation of the integral operator with the help of Fast Fourier Transforms. In view of numerical efficiency, it is desirable to keep the spatial derivatives outside the Fourier integral, rather than to consider them as spectral multiplications with the wave vector inside the Fourier integral. The method is applied to simulate the geophysical low-frequency electromagnetic problem, i.e., the controlled-source electromagnetic (CSEM) method.

1. INTRODUCTION

We consider the electromagnetic scattering by an inhomogeneous 3D object in a planar layered medium in the frequency domain. The inhomogeneous object, with a bounded supporting domain $\mathbb{D}$, is located in the homogeneous layer $n = 0$ of a horizontally layered medium (see Fig. 1). The vectors $\boldsymbol{x} = (x_1, x_2, x_3)$ and $\boldsymbol{x}' = (x_1', x_2', x_3')$ denote the vectorial position in $\mathbb{R}^3$. We use the complex time factor $\exp(-\mathrm{i}\omega t)$. In each homogeneous layer, $n = -M, \ldots, 0, \ldots, N$, the medium is characterized by a constant conductivity σ_n. We assume that the frequency is low enough to neglect dielectric displacement currents. The permeability μ is constant in whole space. Then the wave number is given by $k_n = \omega(\sigma_n\mu)^{\frac{1}{2}}$. We define the incident wavefield, $\{\boldsymbol{E}^{\mathrm{inc}}, \boldsymbol{H}^{\mathrm{inc}}\}$, as the field in the layered background medium in absence of the inhomogeneous object in $\mathbb{D}$, and the scattered field in the layered background as $\{\boldsymbol{E}^{\mathrm{sct}}, \boldsymbol{H}^{\mathrm{sct}}\} = \{\boldsymbol{E}, \boldsymbol{H}\} - \{\boldsymbol{E}^{\mathrm{inc}}, \boldsymbol{H}^{\mathrm{inc}}\}$. For an object in a homogeneous embedding, the scattering problem at hand can be formulated as a domain integral equation over the object domain $\mathbb{D}$. By a weak formulation, in which the spatial derivatives are kept outside the domain integral, a very efficient computational method is arrived at [1]. In this paper we extend this weak formulation to a horizontally layered embedding.

2. CONSTRUCTION OF THE PARTICULAR AND GENERAL SOLUTIONS

The scattered electromagnetic field in the layered configuration (including the object) is written as a superposition of a particular solution and a general solution, viz.,

$$\{\boldsymbol{E}^{\mathrm{sct}}, \boldsymbol{H}^{\mathrm{sct}}\} = \{\boldsymbol{E}^{\mathrm{prt}}, \boldsymbol{H}^{\mathrm{prt}}\} + \{\boldsymbol{E}^{\mathrm{gen}}, \boldsymbol{H}^{\mathrm{gen}}\}. \tag{1}$$

We assume that the particular solution vanishes outside the layer $x_{3,0} \leq x_3 \leq x_{3,1}$, while inside this layer it satisfies Maxwell's equations. Inside this layer the particular solution is given by the following contrast source integral representation

$$\boldsymbol{E}^{\mathrm{prt}}(\boldsymbol{x}) = (k_0^2 + \nabla\nabla\cdot) \int_{\mathbb{D}} G^{(0)}(\boldsymbol{x} - \boldsymbol{x}')\chi(\boldsymbol{x}')\boldsymbol{E}(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}', \tag{2}$$

where the contrast is given by $\chi(\boldsymbol{x}') = \sigma(\boldsymbol{x}')/\sigma_0 - 1$. Note that $\chi\boldsymbol{E}$ can be considered as the contrast source that generates the the scattered wavefield. With the notation $\boldsymbol{x} = \boldsymbol{x}_T + x_3\boldsymbol{i}_3$ and $\boldsymbol{x}_T = \{x_1, x_2\}$, the Green function may be written as a Fourier Bessel integral,

$$G^{(0)}(\boldsymbol{x} - \boldsymbol{x}') = \frac{\exp(\mathrm{i}k_0|\boldsymbol{x} - \boldsymbol{x}'|)}{4\pi|\boldsymbol{x} - \boldsymbol{x}'|} = \frac{\mathrm{i}}{4\pi}\int_0^\infty \frac{\kappa}{\gamma_0} J_0(\kappa|\boldsymbol{x}_T - \boldsymbol{x}'_T|)\exp(\mathrm{i}\gamma_0|x_3 - x'_3|)\mathrm{d}\kappa, \tag{3}$$

where $\gamma_0 = (k_0^2 - \kappa^2)^{\frac{1}{2}}$. The Green function in this domain integral representation for the particular solution has a convolution structure both in the horizontal direction and in the vertical direction.

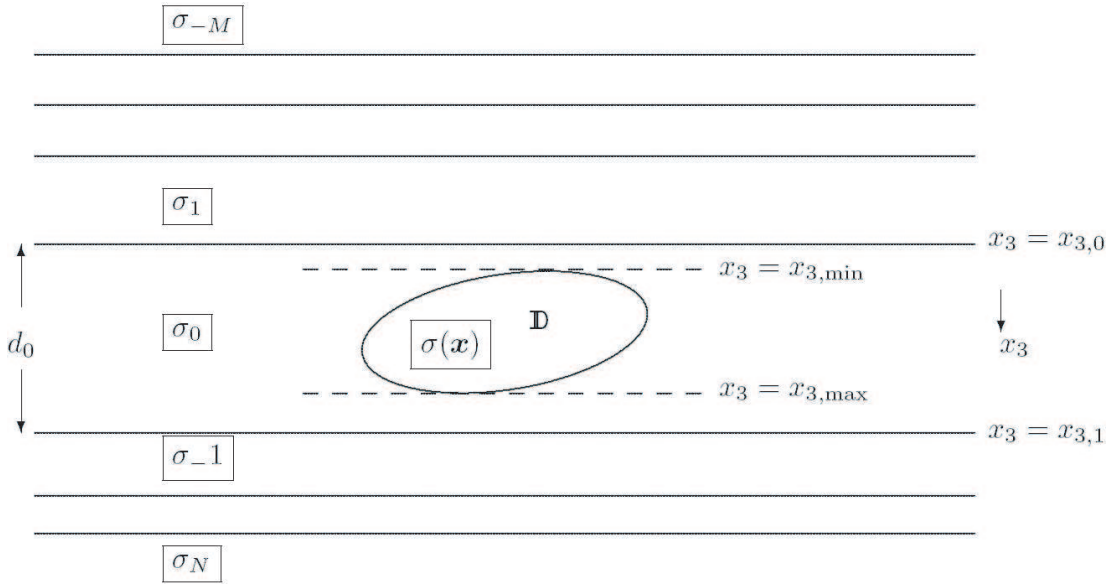


Figure 1: The object domain $\mathbb{D}$ in a homogeneous layer of a horizontally layered medium.

Let us assume that $x_{3,\max}$ and $x_{3,\min}$ are the maximum and minimum extensions of the object $\mathbb{D}$ in the x_3-direction (see Fig. 1). We observe that for $x_3 > x_{3,\max}$ the particular wave field consists of a down-going plane wave spectrum and for $x_3 < x_{3,\min}$ it consists of an up-going wave spectrum. The particular constituent of our field solution does satisfy the Maxwell equations and satisfies all the interface conditions, except at $x_3 = x_{3,0}$ and $x_3 = x_{3,1}$, where the particular solution jumps. Hence, at these levels extra surface sources occur, both of the electric type and of the magnetic type. These surface sources generate the general constituent of our field solution, being a superposition of a field with zero vertical electric field (TE-wavefields) and a field with zero vertical magnetic field (TM-wavefields). After decomposition of the particular wavefield in a TE-part and a TM-part, the general wavefield in up- and down-going wavefields in each layer, and application of the appropriate continuity conditions at the interfaces between the layers, (see [2, 3]), the general wavefield constituent is obtained as $\boldsymbol{E}^{\mathrm{gen}} = \boldsymbol{E}^{\mathrm{gen},(1)} + \boldsymbol{E}^{\mathrm{gen},(2)}$, where

$$\begin{aligned}
\boldsymbol{E}^{\mathrm{gen},(1)}(\boldsymbol{x}) \;=\; & k_0^2\int_{\mathbb{D}} G^{(1)}_{\mathrm{TE}}(\boldsymbol{x}_T - \boldsymbol{x}'_T,\, x_3 + x'_3)\chi(\boldsymbol{x}')\boldsymbol{E}_T(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}' \\
& + k_0^2\int_{\mathbb{D}} G^{(1)}_{\mathrm{TM}}(\boldsymbol{x}_T - \boldsymbol{x}'_T,\, x_3 + x'_3)\chi(\boldsymbol{x}')\boldsymbol{E}_3(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}' \\
& - (\nabla\bar{\nabla}\cdot)\int_{\mathbb{D}} G^{(1)}_{\mathrm{TM}}(\boldsymbol{x}_T - \boldsymbol{x}'_T,\, x_3 + x'_3)\chi(\boldsymbol{x}')\boldsymbol{E}(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}' \\
& + (\nabla_T\nabla_T\cdot)\int_{\mathbb{D}} G^{(1)}_{\mathrm{EM}}(\boldsymbol{x}_T - \boldsymbol{x}'_T,\, x_3 + x'_3)\chi(\boldsymbol{x}')\boldsymbol{E}_T(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}',
\end{aligned} \tag{4}$$

and

$$\begin{aligned}\boldsymbol{E}^{\mathrm{gen},\,(2)}(\boldsymbol{x}) \;=\; & k_0^2\int_{\mathbb{D}} G^{(2)}_{\mathrm{TE}}(\boldsymbol{x}_T-\boldsymbol{x}'_T,\,x_3-x'_3)\chi(\boldsymbol{x}')\boldsymbol{E}_T(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}' \\ & +k_0^2\int_{\mathbb{D}} G^{(2)}_{\mathrm{TM}}(\boldsymbol{x}_T-\boldsymbol{x}'_T,\,x_3-x'_3)\chi(\boldsymbol{x}')\boldsymbol{E}_3(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}' \\ & +(\nabla\nabla\cdot)\int_{\mathbb{D}} G^{(2)}_{\mathrm{TM}}(\boldsymbol{x}_T-\boldsymbol{x}'_T,\,x_3-x'_3)\chi(\boldsymbol{x}')\boldsymbol{E}(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}' \\ & +(\nabla_T\nabla_T\cdot)\int_{\mathbb{D}} G^{(2)}_{\mathrm{EM}}(\boldsymbol{x}_T-\boldsymbol{x}'_T,\,x_3-x'_3)\chi(\boldsymbol{x}')\boldsymbol{E}_T(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}', \end{aligned} \tag{5}$$

in which we have made the decomposition of the electric field vector $\boldsymbol{E}$ and the differential operator ∇ in a transverse part and a vertical part, viz., $\boldsymbol{E}=\boldsymbol{E}_T+\boldsymbol{E}_3$ and $\nabla=\nabla_T+\boldsymbol{i}_3\partial_3$. The overbar above the nabla-operator indicates that we have to change the sign of the vertical derivative, i.e., $\bar{\nabla}=(\partial_1,\,\partial_2,\,-\partial_3)$. This is directly related to the fact that the general solution $\boldsymbol{E}^{\mathrm{gen},\,(1)}$ corresponds to wavefields being generated by secondary sources at the interfaces at $x_3=x_{3,\,0}$ and $x_3=x_{3,\,1}$ and reflected an odd number of times by these latter interfaces. The Green functions in the domain integral representations for this part of the general solution have a convolution structure in the horizontal directions and a correlation structure in the vertical direction. The general solution $\boldsymbol{E}^{\mathrm{gen},\,(2)}$ correspond to wavefields being generated by secondary sources at the interfaces at $x_3=x_{3,\,0}$ and $x_3=x_{3,\,1}$ and reflected an even number of times by these latter interfaces. The Green functions in the domain integral representations for this part of the general solution have a convolution structure both in the horizontal directions and in the vertical direction. Further the various Green functions are given as Fourier Bessel integrals:

$$\begin{aligned}
G^{(1)}_{\mathrm{TE}}(\boldsymbol{x}_T,\,x_3) =& \frac{\mathrm{i}}{4\pi}\int_0^\infty \frac{\kappa}{\gamma_0}J_0(\kappa|\boldsymbol{x}_T|)\Big\{R^{(0,\,-1)}_{\mathrm{TE}}\exp[\mathrm{i}\gamma_0(x_3-2x_{3,\,0})]+R^{(0,\,1)}_{\mathrm{TE}}\exp[\mathrm{i}\gamma_0(2x_{3,\,1}-x_3)]\Big\}\mathrm{d}\kappa,\\
G^{(1)}_{\mathrm{TM}}(\boldsymbol{x}_T,x_3) =& \frac{\mathrm{i}}{4\pi}\int_0^\infty \frac{\kappa}{\gamma_0}J_0(\kappa|\boldsymbol{x}_T|)\Big\{R^{(0,\,-1)}_{\mathrm{TM}}\exp[\mathrm{i}\gamma_0(x_3-2x_{3,\,0})]+R^{(0,\,1)}_{\mathrm{TM}}\exp[\mathrm{i}\gamma_0(2x_{3,\,1}-x_3)]\Big\}\mathrm{d}\kappa,\\
G^{(1)}_{\mathrm{EM}}(\boldsymbol{x}_T,x_3) =& \frac{\mathrm{i}}{4\pi}\int_0^\infty \frac{\kappa}{\gamma_0}J_0(\kappa|\boldsymbol{x}_T|)\Big\{R^{(0,\,-1)}_{\mathrm{EM}}\exp[\mathrm{i}\gamma_0(x_3-2x_{3,\,0})]+R^{(0,\,1)}_{\mathrm{EM}}\exp[\mathrm{i}\gamma_0(2x_{3,\,1}-x_3)]\Big\}\mathrm{d}\kappa,\\
G^{(2)}_{\mathrm{TE}}(\boldsymbol{x}_T,x_3) =& \frac{\mathrm{i}}{4\pi}\int_0^\infty \frac{\kappa}{\gamma_0}J_0(\kappa|\boldsymbol{x}_T|)R^{(0,\,-1|0,\,1)}_{\mathrm{TE}}2\cosh(\gamma_0 x_3)\exp(2\mathrm{i}\gamma_0 d_0)\mathrm{d}\kappa,\\
G^{(2)}_{\mathrm{TM}}(\boldsymbol{x}_T,x_3) =& \frac{\mathrm{i}}{4\pi}\int_0^\infty \frac{\kappa}{\gamma_0}J_0(\kappa|\boldsymbol{x}_T|)R^{(0,\,-1|0,\,1)}_{\mathrm{TM}}2\cosh(\gamma_0 x_3)\exp(2\mathrm{i}\gamma_0 d_0)\mathrm{d}\kappa,\\
G^{(2)}_{\mathrm{EM}}(\boldsymbol{x}_T,x_3) =& \frac{\mathrm{i}}{4\pi}\int_0^\infty \frac{\kappa}{\gamma_0}J_0(\kappa|\boldsymbol{x}_T|)R^{(0,\,-1|0,\,1)}_{\mathrm{EM}}2\cosh(\gamma_0 x_3)\exp(2\mathrm{i}\gamma_0 d_0)\mathrm{d}\kappa,
\end{aligned} \tag{6}$$

where

$$\begin{aligned}
R^{(0,\,-1)}_{\mathrm{EM}} &= k_0^2\frac{R^{(0,\,-1)}_{\mathrm{TE}}+R^{(0,\,-1)}_{\mathrm{TM}}}{\kappa^2}, \quad R^{(0,\,1)}_{\mathrm{EM}} = k_0^2\frac{R^{(0,\,1)}_{\mathrm{TE}}+R^{(0,\,1)}_{\mathrm{TM}}}{\kappa^2},\\
R^{(0,\,-1|0,\,1)}_{\mathrm{EM}} &= k_0^2\frac{R^{(0,\,-1|0,\,1)}_{\mathrm{TE}}-R^{(0,\,-1|0,\,1)}_{\mathrm{TM}}}{\kappa^2}.
\end{aligned} \tag{7}$$

and

$$\begin{aligned}
R^{(0,\,-1)}_{\mathrm{TE}} &= \frac{\tilde{R}^{\mathrm{TE}}_{0,\,-1}}{1-\tilde{R}^{\mathrm{TE}}_{0,\,-1}\tilde{R}^{\mathrm{TE}}_{0,\,1}\exp(2\mathrm{i}\gamma_0 d_0)}, \qquad R^{(0,\,-1)}_{\mathrm{TM}} = \frac{\tilde{R}^{\mathrm{TM}}_{0,\,-1}}{1-\tilde{R}^{\mathrm{TM}}_{0,\,-1}\tilde{R}^{\mathrm{TM}}_{0,\,1}\exp(2\mathrm{i}\gamma_0 d_0)},\\
R^{(0,\,1)}_{\mathrm{TE}} &= \frac{\tilde{R}^{\mathrm{TE}}_{0,\,1}}{1-\tilde{R}^{\mathrm{TE}}_{0,\,-1}\tilde{R}^{\mathrm{TE}}_{0,\,1}\exp(2\mathrm{i}\gamma_0 d_0)}, \qquad R^{(0,\,1)}_{\mathrm{TM}} = \frac{\tilde{R}^{\mathrm{TM}}_{0,\,1}}{1-\tilde{R}^{\mathrm{TM}}_{0,\,-1}\tilde{R}^{\mathrm{TM}}_{0,\,1}\exp(2\mathrm{i}\gamma_0 d_0)},\\
R^{(0,\,-1|0,\,1)}_{\mathrm{TE}} &= \frac{\tilde{R}^{\mathrm{TE}}_{0,\,-1}\tilde{R}^{\mathrm{TE}}_{0,\,1}}{1-\tilde{R}^{\mathrm{TE}}_{0,\,-1}\tilde{R}^{\mathrm{TE}}_{0,\,1}\exp(2\mathrm{i}\gamma_0 d_0)}, \quad R^{(0,\,-1|0,\,1)}_{\mathrm{TM}} = \frac{\tilde{R}^{\mathrm{TM}}_{0,\,-1}\tilde{R}^{\mathrm{TM}}_{0,\,1}}{1-\tilde{R}^{\mathrm{TM}}_{0,\,-1}\tilde{R}^{\mathrm{TM}}_{0,\,1}\exp(2\mathrm{i}\gamma_0 d_0)}.
\end{aligned} \tag{8}$$

The reflection coefficients $\tilde{R}^{\mathrm{TE}}_{0,-1}$ and $\tilde{R}^{\mathrm{TM}}_{0,-1}$ denote the generalized reflection coecients of the stack of layers above the interface at $x_3 = x_{3,0}$, relating the amplitude of the down-going TE- and TM-wavefields to the up-going wavefields in the region above the object in layer $n = 0$. Similarly, the reflection coefficients $\tilde{R}^{\mathrm{TE}}_{0,1}$ and $\tilde{R}^{\mathrm{TM}}_{0,1}$ denote the generalized reflection coefficients of the stack of layers below the interface at $x_3 = x_{3,1}$, relating the amplitude of the up-going TE- and TM-wavefields to the down-going wavefields in the region below the object in layer $n = 0$. These coefficients can easily be computed by a recursive scheme [4].

An integral equation for the total electric field is obtained by observing that

$$\boldsymbol{E}(\boldsymbol{x}) = \boldsymbol{E}^{\mathrm{inc}}(\boldsymbol{x}) + \boldsymbol{E}^{\mathrm{prt}}(\boldsymbol{x}) + \boldsymbol{E}^{\mathrm{gen},\,(1)}(\boldsymbol{x}) + \boldsymbol{E}^{\mathrm{gen},\,(2)}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{D}. \tag{9}$$

Substituting Eqs. (2), (4) and (5) in the above equation we arrive at

$$\begin{aligned}
\boldsymbol{E}(\boldsymbol{x}) \;=\; & \boldsymbol{E}^{\mathrm{inc}}(\boldsymbol{x}) + (\nabla\nabla_T\cdot) \int_{\mathbb{D}} \left[G^{(0)} - G^{(1)}_{\mathrm{TM}} + G^{(2)}_{\mathrm{TM}}\right](\boldsymbol{x},\,\boldsymbol{x}')\chi(\boldsymbol{x}')(\boldsymbol{x}')\boldsymbol{E}_T \mathrm{d}\boldsymbol{x}' \\
& + (k_0^2 + \nabla\partial_3) \int_{\mathbb{D}} \left[G^{(0)} + G^{(1)}_{\mathrm{TM}} + G^{(2)}_{\mathrm{TM}}\right](\boldsymbol{x},\,\boldsymbol{x}')\chi(\boldsymbol{x}')E_3(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}' \\
& + k_0^2 \int_{\mathbb{D}} \left[G^{(0)} + G^{(1)}_{\mathrm{TE}} + G^{(2)}_{\mathrm{TE}}\right](\boldsymbol{x},\,\boldsymbol{x}')\chi(\boldsymbol{x}')\boldsymbol{E}_T(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}' \\
& + (\nabla_T\nabla_T\cdot) \int_{\mathbb{D}} \left[G^{(1)}_{\mathrm{EM}} + G^{(2)}_{\mathrm{EM}}\right](\boldsymbol{x},\,\boldsymbol{x}')\chi(\boldsymbol{x}')\boldsymbol{E}_T(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}'
\end{aligned} \tag{10}$$

which shows that, once the spatial 3D Fourier transforms of the Green functions have been computed, the computation of the electric FIeld in the object involves three forward 3D Fourier transforms and seven inverse 3D Fourier transforms.

In view of numerical efficiency, it is desirable to keep the spatial derivatives outside the domain integrals, rather than to consider them as spectral multiplications with the wave vector inside the Fourier Bessel integrals. The discretization of the integral equation at hand runs along similar lines as in [1]. The spatial derivatives are replaced by a finite difference rule. The convolution and correlation properties of the Green function facilitate a fast and efficient computation of the integral operator with the help of Fast Fourier Transforms. Hence, the total computational time of this approach will be proportional to $20 N_{\mathrm{iter}} N \log(N)$, where N_{iter} is the total number of iterations of a BiCGSTAB type of solver and N is the number of discretization points of the 3D rectangular domain that encloses the object domain $\mathbb{D}$. The Fourier Bessel integrals in the Green functions are truncated and computed with a trapezoidal rule. Furthermore, since the integral operator acts as a filter for the product of the fields and the contrast, the number of Fourier components in the Green functions, that are needed to accurately calculate the integral operator, can be reduced to a number that is related to the sampling theorem.

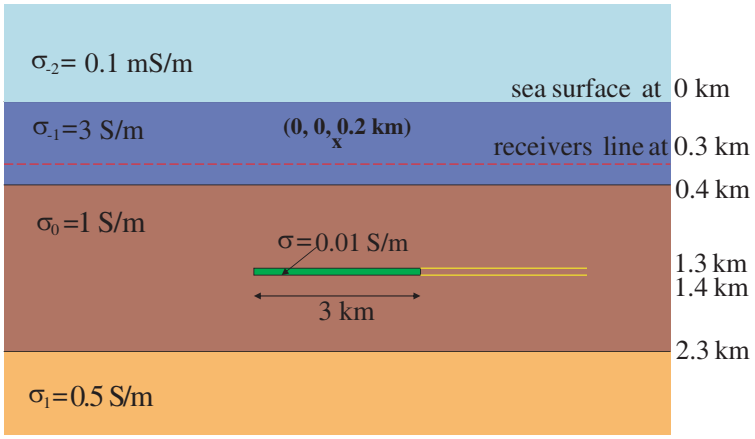


Figure 2: Configuration in the $(x_1,\, x_3)$-plane. The source at depth of 200 m operates at 0.25 Hz. The length of the object in the x_2-direction is 1 km. The receivers are at $x_2 = 0$, $x_3 = 300$ m, with a total extent of 20 km in the x_1-direction.

3. NUMERICAL EXAMPLE

Our domain integral method is used for simulating the geophysical low-frequency electromagnetic problem, i.e., the controlled-source electromagnetic (CSEM) method in marine geophysics. The 3D object $\mathbb{D}$ of interest is a rectangular domain of 3 km by 0.1 km by 1 km, in which an oil-based medium with an electrical conductivity of 0.01 S/m is present (see Fig. 2). It is embedded in a four-layer configuration. The sea surface is located at $x_3 = 0$. Above the sea surface we have a semi-infinite air layer with a conductivity of 0.1 mS/m. Below the sea surface we have a water layer with a conductivity of 3 S/m. In this water layer we assume that we have a vertical electric dipole at position (0, 0, 0.2 km). We have 21 receivers distributed equidistantly along the the receiver line at $-10\,\text{km} \leq x_1 \leq 10\,\text{km}$, $x_2 = 0$ and $x_3 = 0.4\,\text{km}$. The object is embedded in a layer with conductivity of 1 S/m. Below this layer a semi-infinite layer is present with conductivity of 0.5 S/m. The frequency of operation is 0.25 Hz. The object domain is discretized in a grid of 120 by 40 by 4. The grid size is 25 m by 25 m by 25 m. The BiCGSTAB iterative solver reaches the error criterion of 10^{-3} in 13 iterations. The results for the vertical electric field at the receiver locations are presented as solid blue lines in Fig. 3. For comparison we also plot the results by using an integral-equation based preconditioned finite-difference method [5]. In the latter method, we deal with a grid of 400 by 40 by 120 and a grid size of 50 m by 50 m by 25 m. The results are plotted as dashed red lines and we observed that the agreement is excellent. The computational time of the domain integral equation method is about 148 s on a PC with 3.04 GHz processor. It is noted that the computer time of the present method can be further reduced when we store the Bessel functions in the Fourier Bessel integrals for various arguments in the computer memory. Presently, these Bessel functions are repeatedly computed for the same arguments.

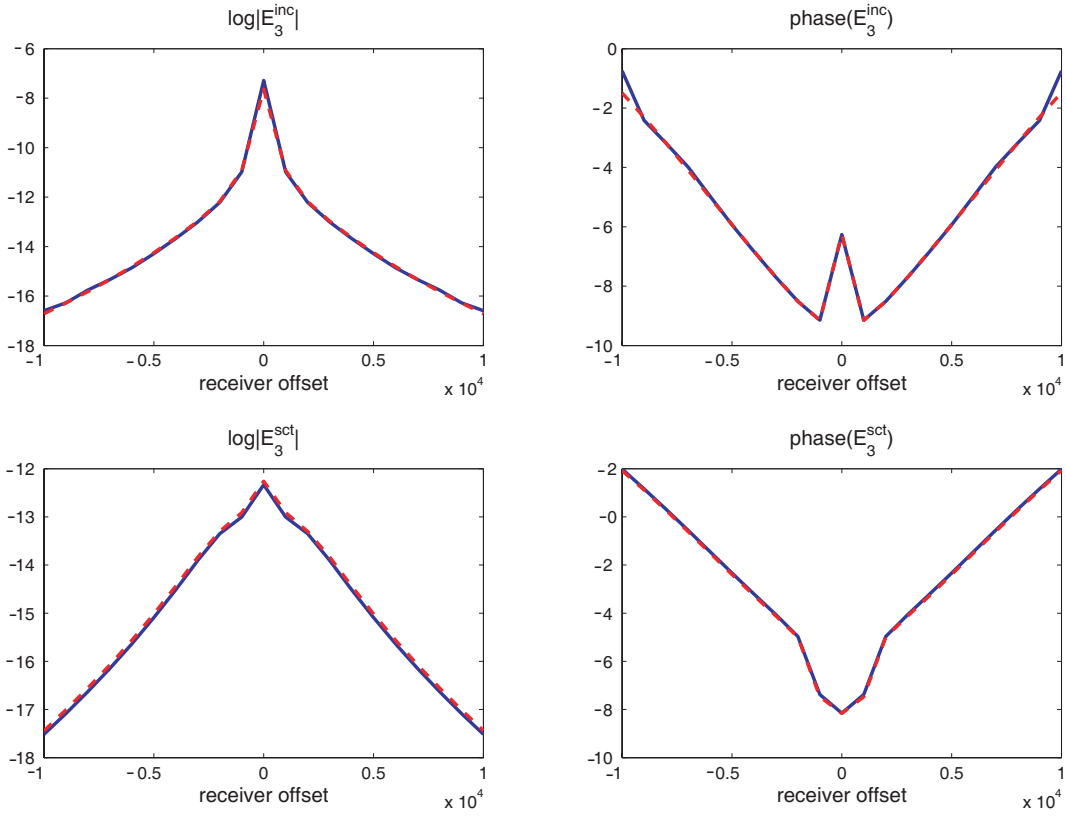


Figure 3: The logarithm of the amplitude (*left*) and the phase (*right*) of the vertical electric field component $\hat{E}_3$ at the receiver locations, generated by a vertical electric dipole in the x_3-direction: incident field (*top*), scattered field (*bottom*). Solid blue lines: integral equation method. Dashed red lines: finite difference method.

4. CONCLUSIONS

We have developed an efficient method to compute the scattering by an object in a planar layered medium. The method extend the weak form of the source-type domain integral equation of the scattering problem of an object in a homogeneous embedding to a similar one for a layered embedding. Specifically the low-frequency (diffusive) electromagnetic problem has been dealt with. For an example representative for the geophysical CSEM configuration the efficiency and the accuracy has been shown.

REFERENCES

1. Abubakar, A. and P. M. van den Berg, "Iterative forward and inverse algorithms based on domain integral equations for three-dimensional electric and magnetic objects," *Journal of Computational Physics*, Vol. 195, 236–262, 2004.
2. Van den Berg, P. M., "On the source type of integral equations for scattering by an object in a layered medium," *Wave Motion*, Vol. 41, 194–210, 2005.
3. Van den Berg, P. M., "On the source type of integral equations for scattering by an object in a layered medium," *Wave Motion*, Vol. 41, 285–287, 2005.
4. Chew, W. C., *Waves and Fields in Inhomogeneous Media*, Chapter 2.1.3, IEEE Press, New York, 1995.
5. Zaslavsky, M., S. Davydycheva, V. Druskin, L. Knizhnerman, A. Abubakar, and T. M. Habashy, "Finite-difference solution of the 3D EM problem using integral equation type preconditioners," *Progress In Electromagnetics Research Symposium Abstracts*, 225, Cambridge, MA, USA, March 26–29, 2006.

A Parallel, Fourier Finite-Element Formulation with an Iterative Solver for the Simulation of 3D LWD Measurements Acquired in Deviated Wells

D. Pardo[1], **M. J. Nam**[1], **C. Torres-Verdin**[1], **and M. Paszynski**[2]

[1]The University of Texas at Austin, USA
[2]AGH University of Science and Technology in Krakow, Poland

Abstract— We describe a new method to simulate resistivity measurements acquired with induction logging instruments in deviated wells. The method combines: (1) a highly efficient iterative solver, (2) a parallel implementation, (3) a Fourier Finite-Element (FFE) formulation in a non-orthogonal system of coordinates, and (4) a 2D *hp*-Finite Element (FE) goal-oriented self-adaptive grid-refinement strategy.

We apply the new method to simulate measurements acquired with a logging-while-drilling (LWD) instrument operating at 1.75 MHz in a 55-degree deviated well. Numerical results confirm the high-accuracy and efficiency of the method. An error level below 1% is achieved in reservoirs with high-contrast in electrical resistivity using an average CPU time of 1–3 minutes per logging position.

1. INTRODUCTION

Resistivity logging instruments are routinely used by oil-companies to quantify the spatial distribution of electrical conductivity in the vicinity of boreholes. Conductivity of the rock formation is utilized to asses the material properties of the subsurface, and estimate the volume of hydrocarbons (oil and gas) existing in a reservoir. To improve the interpretation of results obtained with resistivity logging instruments, and thus, to better quantify and determine existing subsurface materials and increase hydrocarbon recovery, diverse numerical methods have been developed to perform computer simulations.

Numerical methods developed by the oil industry for simulation of resistivity logging measurements include fast 1D and 2D axial-symmetric simulators such as, for example, those described in [6, 9, 11]. Due to the dimensionality reduction of these simulators, they are unable to solve problems involving deviated wells. 3D algorithms such as, for example, those described in [1, 2, 4, 7, 13] are capable of simulating resistivity measurements acquired in deviated wells. However, they are not widely used by the logging industry because the accuracy of these methods is compromised and/or the CPU time required for simulations exceeds reasonable limits (several hours per logging position).

In [8] and [12], we described a FFE formulation in a non-orthogonal system of coordinates for simulation of Direct Current (DC) and Alternating Current (AC) problems, respectively. We also demonstrated the suitability and advantages of this formulation for simulating various resistivity logging measurements in deviated wells.

The main contribution of this work is the integration of the formulation described in [12] with a new iterative solver specially designed for the FFE formulation. In addition, we also describe a new simple but powerful parallel implementation of the entire method that is suitable for both shared-memory and distributed-memory machines. The resulting numerical method provides an improved performance and drastically minimizes the memory requirements associated with the use of direct solvers. The method is applied to the simulation of LWD measurements in a 55-degree deviated well.

2. MODEL PROBLEM: LWD MEASUREMENTS

Figure 1 describes the assumed LWD instrument operating at 1.75 MHz in a 55-degree deviated well in a borehole environment. The LWD instrument is composed of a metallic mandrel with resistivity equal to $10^{-6}\,\Omega\cdot\mathrm{m}$ that incorporates magnetic buffers used to reduce the inductive coupling between antennas and the mandrel. The antennas are embedded within the indented mandrel, thereby minimizing mechanical friction due to drilling. Formation is composed of six different layers with high-contrast resistivities, varying from $0.1\,\Omega\cdot\mathrm{m}$ to $10{,}000\,\Omega\cdot\mathrm{m}$.

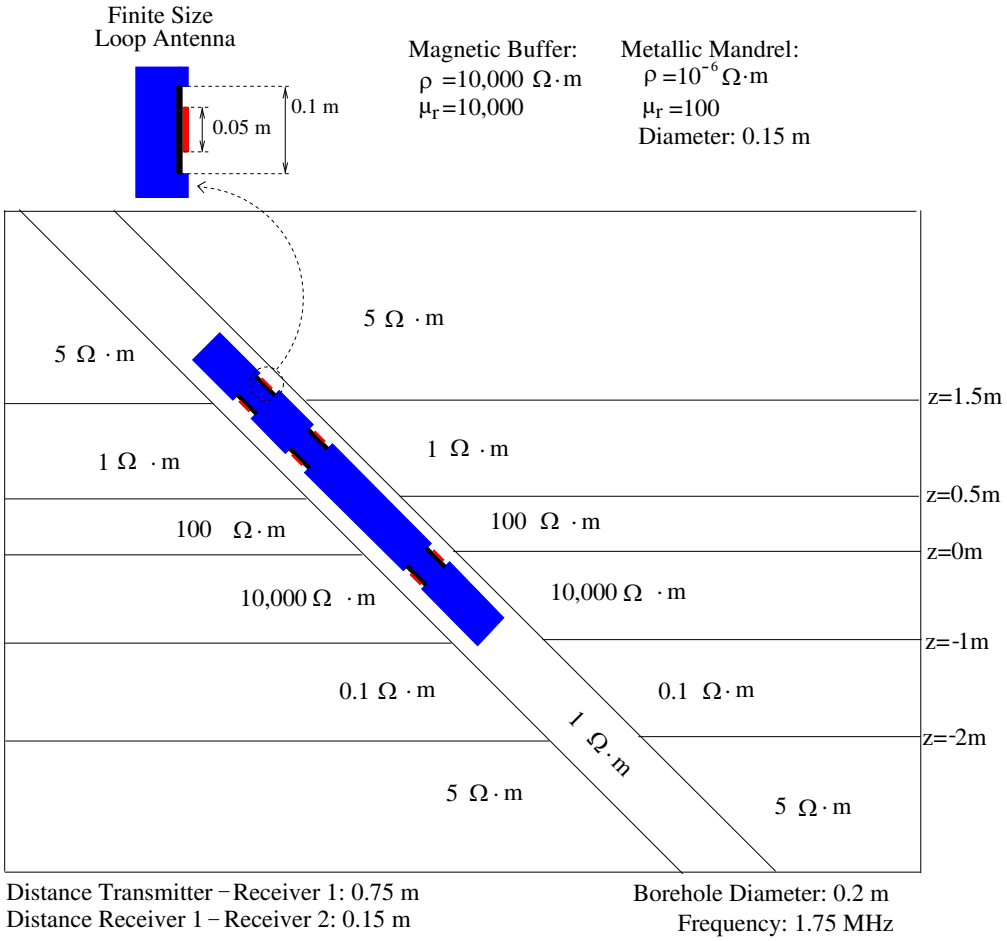


Figure 1: Cross-section of a LWD instrument in a 55-degree deviated well. Formation has large spatial variations in electrical conductivity.

3. METHOD

In problems with axial-symmetric materials, it is customary to consider a cylindrical system of coordinates (ρ, ϕ, z), and then utilize a Fourier series expansion in terms of the azimuthal variable ϕ to reduce the spatial dimension of the problem. For general 3D sources, the resulting formulation for electromagnetic problems consists of a sequence of independent 2D problems. If sources are axial-symmetric, this sequence of 2D problems further reduces to a single 2D problem.

When simulating resistivity logging problems in deviated wells, the axial-symmetry condition is violated, and a dimensionality reduction of the problem becomes elusive. However, we noted in [8, 12] that we may consider a quasi-cylindrical system of coordinates ($\tilde{\rho}$, $\tilde{\phi}$, $\tilde{z}$) and employ a Fourier series expansion with respect to variable $\tilde{\phi}$. The resulting formulation for resistivity logging problems in deviated wells consists of a sequence of coupled 2D problems. Although a sequence of coupled 2D problems constitutes (by definition) a 3D problem, using this formulation we obtain a very special interaction among the various 2D problems. Namely, each 2D problem only interacts (couples) with a maximum of five 2D problems, which results in a penta-diagonal structure for the associated stiffness matrix. Furthermore, this interaction among five different 2D problems only takes place in part of the borehole. In other areas of the computational domain, such as the formation, the associated stiffness matrix is tri-diagonal, while in the subdomain occupied by the logging instrument, the stiffness matrix is simply diagonal. The sparsity of the stiffness matrix becomes a major advantage of this formulation with respect to more traditional 3D formulations.

We construct one common grid for all (coupled) 2D problems by using a 2D *hp*-FE goal-oriented self-adaptive grid-refinement strategy. h indicates the element size, and p the polynomial order of approximation, both varying locally throughout the grid. The goal-oriented *hp*-FE method delivers exponential convergence rates in terms of the error in the quantity of interest versus the number of unknowns and CPU time. A detailed description of the *hp*-FE method and its exponential

convergence properties can be found in [3]. The outstanding performance of the *hp*-FE method for simulating diverse resistivity logging measurements has been documented in [9, 11]. We refer to [10] for technical details on the goal-oriented adaptive algorithm applied to electrodynamic simulation problems.

Due to the existing interaction among the 2D problems resulting from our formulation, we need to solve the final system of linear equations by using a 3D solver. Direct solvers require large amounts of memory and CPU time. Iterative solvers minimize those computational requirements. However, most iterative solvers diverge in presence of high-contrast materials or grids containing elongated elements. In this paper, we consider a LWD problem with up to ten order of magnitude contrast in electrical conductivity. In addition, the resulting optimal *hp*-grid contains elongated elements with different orders of approximation, as shown in Fig. 2 (right panel). Under these conditions, traditional iterative solvers will diverge.

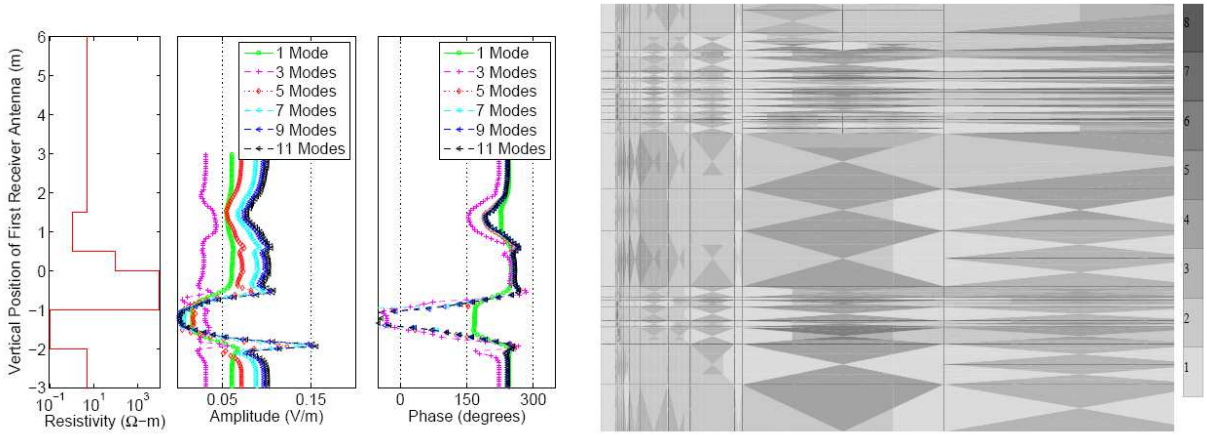


Figure 2: Left panel: Simulated LWD measurements as we move the logging instrument along the well trajectory. Different curves indicate the use of a different number of Fourier modes. Right panel: Amplification of the final *hp*-grid when the first receiver antenna is located at $z = 0$. Dimensions: 2.5 m (horizontal) $\times$ 2 m (vertical). Different colors indicate different polynomial orders of approximation, from $p = 1$ (lighter) to $p = 8$ (darker).

To overcome the above difficulties, we design a new iterative solver by utilizing a block-Jacobi preconditioner, where the size of each block corresponds to a full 2D problem. Thus, the preconditioner solves exactly all high-contrast materials and elongated elements present within each 2D problem. A Krylov subspace optimization method is then used to resolve the interactions occurring among different 2D problems, where material coefficients are smooth.

Finally, we implement a parallel version to speedup computations and minimize the memory used per processor. A traditional approach to design parallel implementations for FE methods consists of decomposing the original computational domain Ω into subdomains $\{\Omega_i\}_{i=1}^N$, and assigning one subdomain to each processor. In presence of self-adaptive grid-refinement algorithms, it is necessary to *reconciliate* the grid after refinements, in order to guarantee compatibility of all grid-refinements occurring in different processors. For the case of *hp*-FE methods, this grid-reconciliation procedure becomes technical and difficult to implement.

In this paper, we resolve the above difficulty by storing a copy of the entire grid in all processors. Decision about grid refinements is performed in parallel, while actual refinements are executed in all processors simultaneously, making unnecessary the grid-reconciliation step. With this new approach, implementation of the parallel version of the software is drastically simplified, while the overhead associated with having an entire copy of the 2D grid in all processors remains at a level below 10% of the total memory and CPU time cost for a moderate number of processors (below 200).

4. NUMERICAL RESULTS

Figure 2 (left panel) displays the simulated measurements as we move the LWD instrument along the well trajectory. Different curves correspond to the use of different numbers of Fourier modes. As we increase the number of Fourier modes, we observe the convergence of the method both in

terms of the amplitude and phase at the receiver antennas. Simulated results confirm a strong sensitivity of the LWD response with respect to the electrical conductivity of the formation.

Figure 2 (right panel) describes a section of the final 2D *hp*-grid obtained with seven Fourier modes and the receiver located at $z = 0$. We observe the need for elongated elements with varying orders of approximation.

Table 1 describes the total CPU time and memory used by six steps of the self-adaptive goal-oriented *hp*-adaptive strategy as we increase the number of processors. For these computations, we have employed the parallel direct solver [5]. Results indicate the substantial gaining that is achieved by using a moderate number of processors. We note that the degeneration of the scalability is related to a poor mesh-redistribution of work due to the small size of the problem.

Table 1: Maximum CPU time and memory (per processor) for the parallel implementation. We have simulated LWD measurements with the first receiver located at $z = 0\,\mathrm{m}$.

Nr. Processors	CPU Time (sec)	Memory (MB)
1	640	6300
2	413	3319
4	274	2251
16	153	851
64	130	459

Finally, in Fig. 3 we compare the time and memory utilized by our new iterative solver and the direct solver MUMPS [5], when using a 2D *hp* grid containing 12,896 unknowns with order of approximation $p = 2$. The relative difference between the two solutions is below 0.05%, that is, the iterative solver has converged to the exact solution delivered by the direct solver. Results confirm the superior scaling expected from iterative solvers. In particular, when using 19 Fourier modes, the iterative solver reduces the CPU time by a factor of 2, and the memory requirements by a factor of 6. These savings increase as we augment the number of Fourier modes and/or the polynomial order of approximation p.

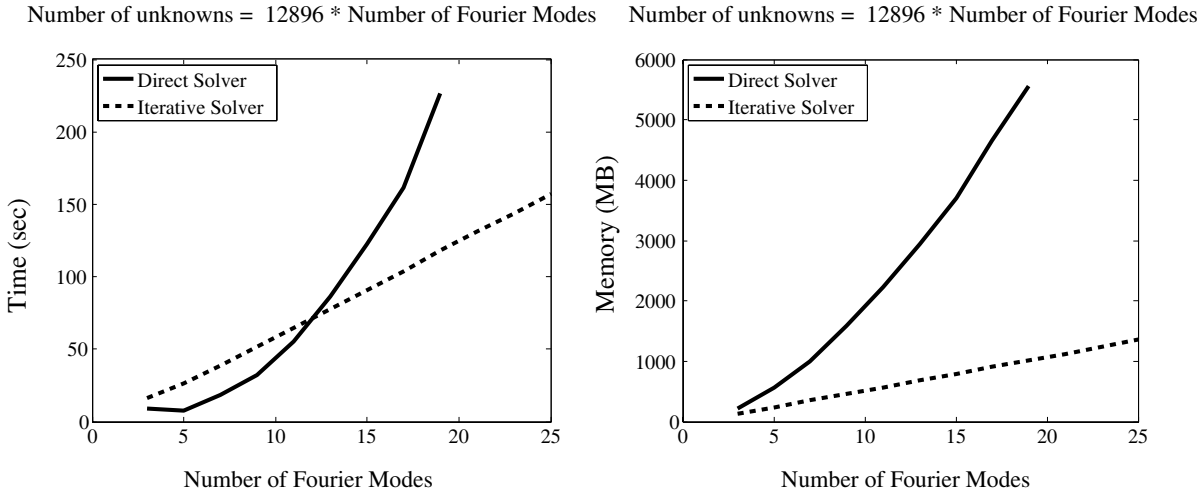


Figure 3: Time (left) and memory (right) used by the direct and iterative solvers, respectively, as a function of the number of Fourier modes. We have simulated LWD measurements with the first receiver located at $z = 0\,\mathrm{m}$. For the simulations, we have employed a machine equipped with a 2.0 GHz processor and 8 GB of RAM.

5. CONCLUSIONS

We have employed a FFE formulation for the simulation of LWD measurements acquired in deviated wells. A simple parallel implementation for a moderate number of processors (below 200) has enabled us to consider more Fourier modes, thereby providing additional accuracy. In addition, we

have designed a new iterative solver for the FFE formulation that further reduces the CPU time and memory requirements.

ACKNOWLEDGMENT

The work presented in this paper was supported by The University of Texas at Austin's *Joint Industry Research Consortium on Formation Evaluation* sponsored by Anadarko, Aramco, Baker Atlas, British Gas, BHP-Billiton, BP, Chevron, Conoco-Phillips, ENI E&P, ExxonMobil, Halliburton, Marathon, Mexican Institute for Petroleum, Hydro, Occidental Petroleum, Petrobras, Schlumberger, Shell E&P, Statoil, TOTAL, and Weatherford International, Ltd. The work of the fourth author has been partially supported by the Foundation for Polish Science under Homming Programme. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing high performance computing resources that have contributed to the research results reported in this paper.

REFERENCES

1. Avdeev, D. B., A. V. Kuvshinov, O. V. Pankratov, and G. A. Newman, "Three-dimensional induction logging problems, part 1: An integral equation solution and model comparisons," *Geophysics*, Vol. 67, 413–426, 2002.
2. Davydycheva, S., V. Druskin, and T. Habashy, "An efficient finite-difference scheme for electromagnetic logging in 3D anisotropic inhomogeneous media," *Geophysics*, Vol. 68, 1525–1536, 2003.
3. Demkowicz, L., *Computing with hp-adaptive Finite Elements, Volume I: One and Two Dimensional Elliptic and Maxwell Problems*, Chapman and Hall, 2006.
4. Druskin, V. L., L. A. Knizhnerman, and P. Lee, "New spectral Lanczos decomposition method for induction modeling in arbitrary 3-D geometry," *Geophysics*, Vol. 64, 701–706, 1999.
5. http://graal.ens lyon.fr/MUMPS/, MUMPS: A Multifrontal Massively Parallel Sparse Direct Solver, 2008.
6. Lu, X. and D. L. Alumbaugh, "One-dimensional inversion of three-component induction logging in anisotropic media," *SEG Expanded Abstract*, Vol. 20, 376–380, 2001.
7. Newman, G. A. and D. L. Alumbaugh, "Three-dimensional induction logging problems, part 2: A finite-difference solution," *Geophysics*, Vol. 67, 484–491, 2002.
8. Pardo, D., V. Calo, C. Torres-Verdin, and M. J. Nam, "Fourier series expansion in a non-orthogonal system of coordinates for the simulation of 3D DC borehole resistivity measurements," *Computer Methods in Applied Mechanics and Engineering*, Vol. 197, Nos. 1–3, 1906–1925, Apr. 2008.
9. Pardo, D., L. Demkowicz, C. Torres-Verdin, and M. Paszynski, "Simulation of resistivity logging-while-drilling (LWD) measurements using a self-adaptive goal-oriented *hp*-finite element method," *SIAM Journal on Applied Mathematics*, Vol. 66, No. 6, 2085–2106, 2006.
10. "A goal oriented *hp*-adaptive finite element strategy with electromagnetic applications, Part II: electrodynamics," *Computer Methods in Applied Mechanics and Engineering*, Vol. 196, 3585–3597, 2007.
11. Pardo, D., C. Torres-Verdin, and L. Demkowicz, "Feasibility study for two-dimensional frequency dependent electromagnetic sensing through casing," *Geophysics*, Vol. 72, No. 3, F111–F118, 2007.
12. Pardo, D., C. Torres-Verdin, M. J. Nam, M. Paszynski, and V. Calo, "Fourier series expansion in a non-orthogonal system of coordinates for the simulation of 3D AC borehole resistivity measurements," *Computer Methods in Applied Mechanics and Engineering*, in press, 10.1016/j.cma.2008.03.007, Mar. 2008.
13. Wang, T. and J. Signorelli, "Finite-difference modeling of electromagnetic tool response for logging while drilling," *Geophysics*, Vol. 69, 152–160, 2004.

Evaluation of the Sacttering Matrix of Flat Dipoles Embedded in Multilayer Structures

S. J. S. Sant'Anna[1, 2], **J. C. da S. Lacava**[2], **and D. Fernandes**[2]

[1]Instituto Nacional de Pesquisas Espaciais, Brazil
[2]Instituto Tecnológico de Aeronáutica, Brazil

Abstract— The expression for the scattering matrix of a dipole embedded in a multilayer planar structure is derived analytically. The structure is excited by an elliptically polarized plane wave incident at an oblique angle. Calculation of the electromagnetic fields present in a four-layer structure is performed in the spectral domain. The scattering matrices of electric and magnetic flat dipoles printed on the interface between two confined layers are evaluated using two polarimetric features.

1. INTRODUCTION

Interest in retrieving parameters from natural targets out of microwave remote sensing data has increased lately. Full polarimetric SAR data make the analysis feasible, especially when polarimetric decomposition techniques are utilized. These techniques are based on the characterization of radar scattering from complex targets in terms of a linear combination of scattering from simpler ones, constitute an effective tool for the analysis of polarimetric data. From this point of view, knowledge of the scattering matrix from simpler targets is crucial for extracting information gathered from polarimetric SAR data. Besides, the closer to the exact solution the determination of the scattering matrix is, the more precise the scattering analysis will be. From polarimetric SAR images and using the knowledge of scattering matrix it is possible, for instance, to develop mathematical models for natural targets and to derive several polarimetric target descriptors. It is also possible to improve the classification accuracy of land cover and use, to identify the scattering mechanisms intrinsic on an ensemble of pixels, and to build up a sensor calibration processes.

In this context, the scattering matrix computation plays a central role in the polarimetric radar remote sensing imagery. Therefore, the aim of this work is to develop an analytical expression for the scattering matrix of flat dipoles embedded in a multilayer planar structure illuminated by an elliptically polarized plane wave at oblique incidence. The theory behind the computation of the electromagnetic fields present in the structure is briefly described in Section 2. In Section 3, the scattering matrices for an electric and a magnetic flat dipoles printed on the interface $z = \ell_1$ (see Fig. 1) are developed and compared.

2. THEORY

The structure under investigation, as depicted in Fig. 1, is composed of four isotropic linear homogenous layers stacked up in the z direction, that is: two confined layers located between free space (the upper layer) and ground (the lower layer). The lower layer, of complex permittivity ε_g and complex permeability μ_g, occupies the negative-z region. Each of the confined layers is characterized by thickness ℓ_n, complex permittivity ε_n and complex permeability μ_n, where $n = 1, 2$. Perfect electric and magnetic conductors of infinitesimal thickness that act as scattering elements are printed on each layer interface. The planar interface $z = \ell_1 + \ell_2$ separates layer 2 from free space. The layers are assumed to be unbounded along the transversal x and y directions. The development is based on a global rectangular coordinate system located on top of the ground layer (interface $z = 0$) lying on the xy-plane.

The electromagnetic fields in a multilayer structure are determined through the methodological approach described in [1]. According to this methodology the structure is treated as a boundary value problem. The analysis is carried out by employing the spectral domain full-wave technique. The induced electric and magnetic surface current densities on the conductors are the virtual sources of the scattered fields. Firstly, the wave equations for every layer are solved in the Fourier domain, which leads to a system of differential equations. Application of the proper electromagnetic boundary conditions at each interface yields a set of twelve equations with an equal numbers of unknowns. The analytical solution of this system leads to the spectral Green's functions in simple, closed form. Combination of the Green's functions and the transformed surface current densities

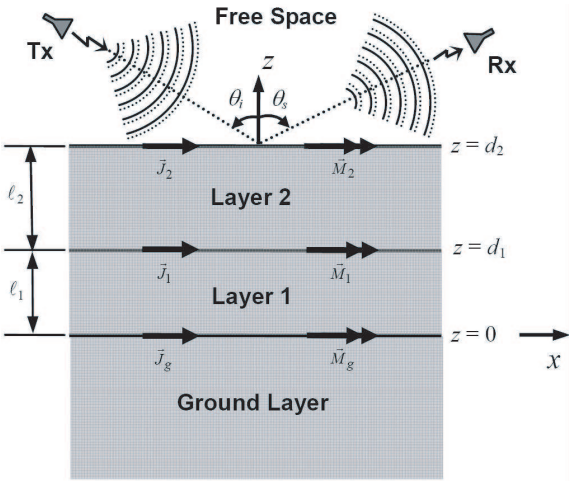


Figure 1: Lateral view of a planar structure with two confined layers.

allows the determination of the transformed fields at any point of the multilayer structure. Finally, the electromagnetic fields in spatial domain are obtained from the inverse double Fourier transform.

The stationary phase method [2] is used to obtain asymptotic expressions for the far electromagnetic fields scattered by the multilayer structure. From these asymptotic expressions and the knowledge of the current densities, the scattering matrix elements are completely determined for any incidence and scattering direction. Using this method the far scattered electric field is expressed by:

$$\boldsymbol{E}_0(r,\,\theta,\,\phi) \cong -\frac{ik_0}{2\pi}\frac{e^{-ik_0 r}}{r}\cot\theta\{\hat{\boldsymbol{\theta}}\mathbb{e}_{0z}(k_{xe},\,k_{ye}) - \hat{\boldsymbol{\phi}}\eta_0\mathbb{h}_{0z}(k_{xe},\,k_{ye})\}, \tag{1}$$

in the spherical coordinate system, where bold face letters represent vectors, η_0 is the intrinsic impedance of free space, $k_{xe} = k_0\sin\theta\cos\phi$ and $k_{ye} = k_0\sin\theta\sin\phi$ are the stationary phase points, k_0 is the wave number of the exciting wave, r is the distance between the receiving antenna and the target, and $\mathbb{e}_{0z}$ and $\mathbb{h}_{0z}$ are the longitudinal components of the spectral electromagnetic fields. As a note, the current densities induced on the conductors could be determined by the moment method, or other techniques.

A particular structure containing electric and magnetic dipoles printed only on the interface $z = \ell_1$ is analyzed next. The dipole dimension along the x-axis of a standard rectangular coordinate system is considered much larger than that along the y-axis, such that the y-component of the surface current density can be neglected. For this case, the longitudinal components $\mathbb{e}_{0z}$ and $\mathbb{h}_{0z}$, on the stationary phase points, are:

$$\mathbb{e}_{0z}(k_{xe},\,k_{ye}) = \frac{4\omega^2}{\Delta_e}\{k_{xe}\Phi_0^e j_{1xe}(k_{xe},\,k_{ye}) + k_{ye}\Theta_0^e m_{1xe}(k_{xe},\,k_{ye})\}, \tag{2}$$

$$\mathbb{h}_{0z}(k_{xe},\,k_{ye}) = \frac{4\omega^2}{\Delta_m}\{-k_{ye}\Phi_0^h j_{1xe}(k_{xe},\,k_{ye}) + k_{xe}\Theta_0^h m_{1xe}(k_{xe},\,k_{ye})\}, \tag{3}$$

where $j_{1xe}(k_{xe},\,k_{ye})$ and $m_{1xe}(k_{xe},\,k_{ye})$ are respectively the transformed electric and magnetic surface current densities, and the following factors together with the spectral variables represent the spectral Green's functions of free space.

$$\Phi_0^e = -\varepsilon_1\gamma_1\gamma_2(\varepsilon_1\gamma_g\cos\alpha_1 + i\varepsilon_g\gamma_1\sin\alpha_1), \tag{4}$$

$$\Theta_0^e = \omega\varepsilon_1\varepsilon_2\gamma_2(\varepsilon_g\gamma_1\cos\alpha_1 + i\varepsilon_1\gamma_g\sin\alpha_1), \tag{5}$$

$$\Phi_0^h = -\omega\mu_1\mu_2\gamma_2(\mu_g\gamma_1\cos\alpha_1 + i\mu_1\gamma_g\sin\alpha_1), \tag{6}$$

$$\Theta_0^h = \mu_2\gamma_1\gamma_2(\mu_1\gamma_g\cos\alpha_1 + i\mu_g\gamma_1\sin\alpha_1), \tag{7}$$

$$\begin{aligned}
\Delta_e &= -4\omega^3 e^{-i\alpha_0}\{\varepsilon_1\gamma_2(\varepsilon_2\gamma_0\cos\alpha + i\varepsilon_0\gamma_2\sin\alpha)(\varepsilon_g\gamma_1\cos\alpha_1 + i\varepsilon_1\gamma_g\sin\alpha_1) \\
&\quad +\varepsilon_2\gamma_1(\varepsilon_0\gamma_2\cos\alpha + i\varepsilon_2\gamma_0\sin\alpha)(\varepsilon_1\gamma_g\cos\alpha_1 + i\varepsilon_g\gamma_1\sin\alpha_1)\}, \qquad (8) \\
\Delta_m &= 4\omega^3 e^{-i\alpha_0}\{\mu_1\gamma_2(\mu_2\gamma_0\cos\alpha + i\mu_0\gamma_2\sin\alpha)(\mu_g\gamma_1\cos\alpha_1 + i\mu_1\gamma_g\sin\alpha_1) \\
&\quad +\mu_2\gamma_1(\mu_0\gamma_2\cos\alpha + i\mu_2\gamma_0\sin\alpha)(\mu_1\gamma_g\cos\alpha_1 + i\mu_g\gamma_1\sin\alpha_1)\}, \qquad (9) \\
\gamma_m &= (-1)^\tau\sqrt{\omega^2\mu_m\varepsilon_m - (k_{xe}^2 + k_{ye}^2)} \qquad \mathbb{I}\mathrm{m}\{\gamma_m\} \le 0, \qquad (10)
\end{aligned}$$

with $m \in \{0,\, 1,\, 2,\, g\}$, $\alpha_0 = \gamma_0 d_2$, $\alpha_1 = \gamma_1 d_1$, $\alpha_2 = \gamma_2 d_2$, $\alpha_3 = \gamma_2 d_1$ and $\alpha = \alpha_2 - \alpha_3$. The τ variable, which defines the wave propagation direction, can take values of either 1 or 2. For the free space $\tau = 1$, i.e., a wave propagating in the positive z-direction. For the inner layers, τ takes both values, and for the ground layer $\tau = 2$ to represent a wave propagating in the negative z-direction.

3. SCATTERING MATRICES

The scattering matrix can be seen as a mathematical characterization of target scattering. This matrix relates the electric fields of the wave scattered by a target to the ones of the incident wave. It is the most important parameter in polarimetric SAR analysis, since it provides complete information about the scattering mechanism; all polarimetric features that describe the target scattering can be derived from it. In order to determine the scattering matrix of a target in a multilayer structure, the procedure defined in [3] will be followed. This procedure directly relates the elements of the scattering matrix to the components of the scattered electric field; a standard spherical coordinate system (r, θ, ϕ) is chosen to coincide with a coordinate system (k, v, h) defined in terms of horizontal and vertical linear polarization components, as stated in [4].

The incident wave sets up currents on the scattering element (dipole), which in turn re-radiates a scattered wave. The currents induced on the dipole depend on the incident wave polarization. Therefore, from the reciprocity theorem, relation (11) between the components of the transformed far electric field in free space due to each linear polarization

$$\mathbb{e}_{0z}^{(h)}(k_{xe},\, k_{ye}) = -\eta_0 \mathbb{h}_{0z}^{(v)}(k_{xe},\, k_{ye}), \qquad (11)$$

can be obtained, where the superscripts (v, h) are associated to polarization of the incident wave. Consequently, for an electric planar dipole the induced currents due to horizontal and vertical linear polarizations are, then, related by

$$j_{1xe}^{(h)}(k_{xe},\, k_{ye}) = -\mathbb{C}_{\mathbb{1}}\left(\frac{k_{ye}}{k_{xe}}\right) j_{1xe}^{(v)}(k_{xe},\, k_{ye}). \qquad (12)$$

This relation leads to the following scattering matrix

$$[S_{ed}] = \frac{\mathbb{C}_2}{\cos\phi_s}\begin{bmatrix} (\cos\phi_s)^2/\mathbb{C}_{\mathbb{1}} & \sin\phi_s\cos\phi_s \\ \sin\phi_s\cos\phi_s & \mathbb{C}_{\mathbb{1}}(\sin\phi_s)^2 \end{bmatrix}, \qquad (13)$$

where the constants $\mathbb{C}_{\mathbb{1}}$ and $\mathbb{C}_2$ are given by:

$$\mathbb{C}_{\mathbb{1}} = \frac{\Delta_e\Phi_0^h}{\Delta_m\Phi_0^e}\eta_0 \quad \text{and} \quad \mathbb{C}_2 = -\frac{i2\omega^2 k_0^2\eta_0\Phi_0^h}{\pi\Delta_m}\cos\theta_s j_{1xe}^{(v)}(k_{xe},\, k_{ye}).$$

Similarly for the magnetic dipole, the relation between the induced currents and its corresponding scattering matrix are expressed, respectively, by:

$$\begin{aligned}
m_{1xe}^{(h)}(k_{xe},\, k_{ye}) &= -\mathbb{C}_{\mathbb{3}}\left(\frac{k_{xe}}{k_{ye}}\right) m_{1xe}^{(v)}(k_{xe},\, k_{ye}), \qquad (14) \\
[S_{md}] &= \frac{\mathbb{C}_4}{\sin\phi_s}\begin{bmatrix} (\sin\phi_s)^2/\mathbb{C}_{\mathbb{3}} & -\sin\phi_s\cos\phi_s \\ -\sin\phi_s\cos\phi_s & \mathbb{C}_{\mathbb{3}}(\cos\phi_s)^2 \end{bmatrix}, \qquad (15)
\end{aligned}$$

where the constants $\mathbb{C}_{\mathbb{3}}$ and $\mathbb{C}_{\mathbb{4}}$ are represented by:

$$\mathbb{C}_{\mathbb{3}} = \frac{\Delta_e\Theta_0^h}{\Delta_m\Theta_0^e}\eta_0 \quad \text{and} \quad \mathbb{C}_{\mathbb{4}} = -\frac{i2\omega^2 k_0^2\eta_0\Theta_0^h}{\pi\Delta_m}\cos\theta_s m_{1xe}^{(v)}(k_{xe},\, k_{ye}).$$

Note that the constants $\mathbb{C}_2$ and $\mathbb{C}_4$, which are directly related to the currents induced on the dipoles, affect equally all the elements of the scattering matrices.

The ability to distinguish different scattering mechanisms improves the accuracy of polarimetric SAR image classification. Several polarimetric features aiming at the characterization of the scattering mechanism are presented in the literature. One of these is the α-angle, which is derived from the eigenvalue/eigenvector decomposition [5]. Its physical interpretation is associated to the type of scattering mechanism and it can represent a wide variety of scatters. The α-angle for the electric and magnetic dipoles is computed from (16).

$$\alpha = \cos^{-1}\left\{\frac{|S_{xx}+S_{yy}|}{\sqrt{|S_{xx}+S_{yy}|^2+|S_{xx}-S_{yy}|^2+4|S_{xy}|^2}}\right\}, \quad \text{with} \quad [S]=\begin{bmatrix} S_{xx} & S_{xy} \\ S_{xy} & S_{yy}\end{bmatrix}. \tag{16}$$

Even though the dipoles are of a different electromagnetic nature the α-angle for both is 45°, a typical value of dipole scatter. Another feature used for the comparison between the scattering matrices of the electric and magnetic dipoles is a similarity measurement developed in [6]. It is based on square of a correlation coefficient and it is also independent of the spans of the scattering matrices and of the target orientation. The similarity measure is defined as:

$$\rho([S_1],\,[S_2]) = \frac{|\boldsymbol{k}_1^H \cdot \boldsymbol{k}_2|^2}{\|\boldsymbol{k}_1\|_2^2\|\boldsymbol{k}_2\|_2^2}, \tag{17}$$

where $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ correspond respectively to the modified Pauli-scattering vectors of the scattering matrices $[S_1]$ and $[S_2]$, as defined in [5], the superscript H denotes complex conjugate transpose, and $\|\|_2$ denotes the Euclidean norm of the vector. In the specific case of the scattering matrices of the electric and magnetic dipoles

$$\rho([S_{ed}],\,[S_{md}]) = \left\{\frac{(\mathbb{C}_{\mathbb{1}}^*\mathbb{C}_{\mathbb{3}}-1)(\mathbb{C}_{\mathbb{1}}\mathbb{C}_{\mathbb{3}}^*-1)(\sin\phi_s\cos\phi_s)^2}{[|\mathbb{C}_{\mathbb{1}}|^2(\sin\phi_s)^2+(\cos\phi_s)^2]\,[|\mathbb{C}_{\mathbb{3}}|^2(\cos\phi_s)^2+(\sin\phi_s)^2]}\right\}^2. \tag{18}$$

It can be noted that this coefficient carries the information of the scattering mechanism taking into account the electromagnetic parameters of the layers and the type of dipole. In particular, when the effect of the layers is neglected (i.e., the four layers are considered to be free space) both scattering matrices are non-similar, since the constants $\mathbb{C}_{\mathbb{1}}$ and $\mathbb{C}_{\mathbb{3}}$, are reciprocal real values, which yields $\rho([S_{ed}],\,[S_{md}]) = 0$.

4. CONCLUSIONS

The full-wave technique in the spectral domain was used to derive the electromagnetic fields scattered by a multilayer planar structure. In the analysis, the structure was excited by an elliptically polarized plane wave at oblique incidence angle. From these fields, analytical expressions for the scattering matrices of electric and magnetic flat dipoles embedded in the structure were derived. Two polarimetric features were used as metrics to compare both scattering matrices: the α-angle and the correlation coefficient. Based on the former it was not possible to differentiate between the two dipoles, despite their different electromagnetic nature. On the other hand, it was shown that the latter feature can be used as a discriminating measure between these two scattering mechanisms.

ACKNOWLEDGMENT

The authors are grateful to FINEP-CAPTAER project for the partial support.

REFERENCES

1. Lacava, J. C. S., A. V. Proaño De la Torre, and L. Cividanes, "A dynamic model for printed apertures in anisotropic stripline structures," *IEEE Trans. Microwave Theory Tech.*, Vol. 50, No. 1, 22–26, 2002.
2. Collin, R. E. and F. J. Zucker, *Antenna Theory: Part 1*, McGraw-Hill Book Company, New York, 1969.
3. Sant'Anna, S. J. S., J. C. S. Lacava, and D. Fernandes, "Closed form expressions for scattering matrix of simple targets in multilayer structures," *Proceedings of International Geoscience and Remote Sensing Symposium*, 714–717, Barcelona, Spain, July 2007.

4. Ulaby, F. T. and C. Elachi, *Radar Polarimetry for Geoscience Applications*, 364, Artech House, Norwood, 1990.
5. Cloude, S. R. and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sens.*, Vol. 35, No. 1, 68–78, 1997.
6. Yang, J., Y. N. Peng, and S. M. Lin, "Similarity between two scattering matrices," *Electronics Letters*, Vol. 37, No. 3, 193–194, 2001.

Radio Frequency Metrology for Mobile Atmospheric Pressure Plasma Devices

V. J. Law, **N. O'Connor**, **and S. Daniels**
National Center of Plasma Science and Technology, Dublin City University
Collins Avenue, Glasnevin, Dublin 9, Dublin, Ireland

Abstract— Real-time time-domain and frequency-domain measurement of mobile atmospheric pressure plasma is described. The measurements are non-invasive and deployed on the main power-line to the plasma device.

1. INTRODUCTION

The atmospheric pressure coaxial dielectric barrier discharge (DBD) has been established for non-thermal plasma processing for more then 20 years [1, 2]. Their potential for direct impact on society in the guise of surface modification of engineering materials and the destruction of microbial pathogens on contaminated surfaces are now becoming apparent. Recently atmospheric pressure glow discharges in the form of: plasma pencil [3] needle [4, 5], jet [6] and torch [7] have been reported to have low electrical power consumption. The common feature of these devices is that they have a volume discharge that acts as a source of ion species and an expanding plume that is driven by the flow of gas passing through the chamber volume. The drive circuitry typically employs a variety of sinusoidal or pulse power technology; some of which have fixed frequencies (13.56 MHz [5, 6]) or variable frequencies (7 kHz and 1 MHz [7–9]). Within this body of work and within the water and food purification sectors [10], it is commonly believed that continuous power switching technology has better power conversion efficiencies than sinusoidal excitation.

This paper presents the application of Radio Frequency (RF) metrology to atmospheric pressure plasma devices [11, 12]. In particular, to semiconductor power switching circuitry and Flyback transformer that drives an atmospheric discharge. Section 2 describes active RF metrology deployed on the power-line of the Flyback driven DBD. Section 3 describes passive RF metrology on the power-line. Section 4 describes the combination of time-domain and frequency-domain passive RF metrology of a sinusoidal driven plasma. Finally Section 5 provides a conclusion to this work.

2. ACTIVE RF METROLOGY

The salient points of the helium DBD have been presented in [12, 13]. Here only the RF measurement is described. The drive circuitry and the Rohde & Schwarz FSL3 scalar network analyzer is shown in Figure 1. Figure 2 shows the equivalent electrical model (EEM) of the measurement. The EEM comprises a 7 element Chebyshev filter front-end probe, a 1.2 m length of 50-Ohm transmission-line and a variable resistor termination (R_3) that represents the plasma. Figure 3 shows the measured frequency dependant modulus of the plasma as a function of power: the resolution bandwidth (RBW) is set to 30 kHz. Here it can be seen that the modulus is plasma dependent at 165 MHz.

A standard linear software package is used to simulate these measurements and after a few parameter iterations the transmission-line responds to the plasma termination at 165 MHz, see Figure 4. At this frequency the EEM yields a resistance termination of 10–21 Ohm with a transmission length of $l = 1.3$ m and a velocity factor of 0.62. *N.B. a relative offset of 23 dB is used to normalize the dynamic gain and account for the insertion of the Chebychev filter. The additional $l = 0.1$ m and reduced velocity factor may also be attributed to the additional line length through filter.* These simulations show the depth of the node increasing with applied power, thus indicating the helium DBD is becoming more conductive as the power is increased. From a plasma point-of-view this is expected as the ion/electron density increases with applied power.

3. PASIVE RF METROLOGY

In this section the same helium DBD is integrated using passive RF metrology to measure the drive oscillator phase noise and frequency pulling as a function of plasma plume-surface interaction [12]. The measurement is made on the power-line as shown in Figure 1. However, in this case the spectrum analyzer is protected from the high voltage input level by a capacitive clamp front-end probe which has a voltage coupling factor of 0.001.

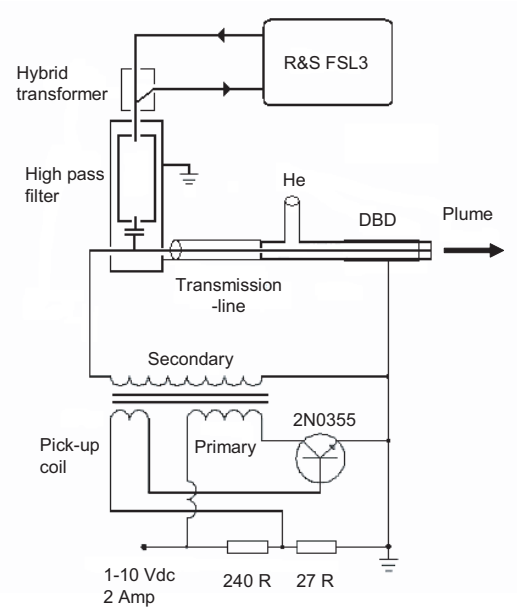


Figure 1: Schematic of SNA and DBD.

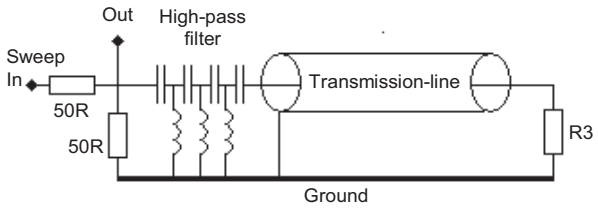


Figure 2: EEM of DBD circuit: 50 Ohm transmission = 1.5 m in length.

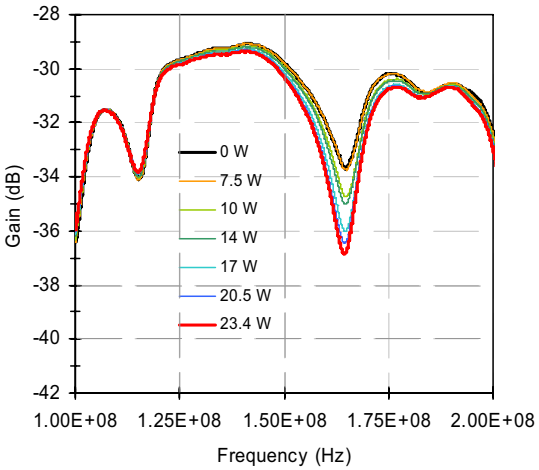


Figure 3: Reflection modulus of DBD as a function of power.

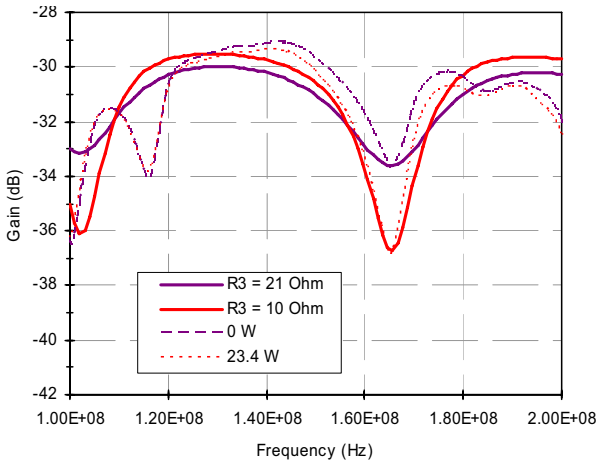


Figure 4: EEM simulation of DBD. l = 1.3 m, velocity factor = 0.62, and R3: 10 and 21 Ohm.

To exemplify this measurement, Figures 5 and 6 shows the helium DBD driven at 177 KHz with an applied dc power of 11 W under stable and unstable modes of operation due to flow rate conditions. The terms stable and unstable are used to classify visual instabilities within the plasma plume. Here the only operating difference between the stable and unstable mode is the gas flow rate: 2 standard liters per minute (slm) for Figure 5 and 5 slm for Figure 6. These results show that the spectrum analyzer captures the flow rate instability as a 2 kHz modulation of the drive carrier frequency. This instability is in the audible range of the human ear and can be clearly heard and captured on a microphone.

In the second example the oscillator drive frequency is measured as a function of plasma plume interaction with two different surfaces (copper connected to ground and glass) at a distance of 16 mm, and plume expanding into free-space. The helium flow for these measurements is 2 slm at a drive frequency of 172.9 kHz and applied dc power of 10.5 W. The results for the copper (dotted-line) glass (dashed-line) and air (solid-line) are shown in Figure 7. The result shows frequency pulling of $\sim$ 700 Hz from free space to metal coupling with the plume. For example: 172.9 kHz without the work-piece; 172.6 with the glass slide and; 172.2 kHz with the copper strip. This equates to a frequency pulling ratio of $f_{\max}$ / $f_{\min}$ of 1.004. This result indicates that the plume-surface

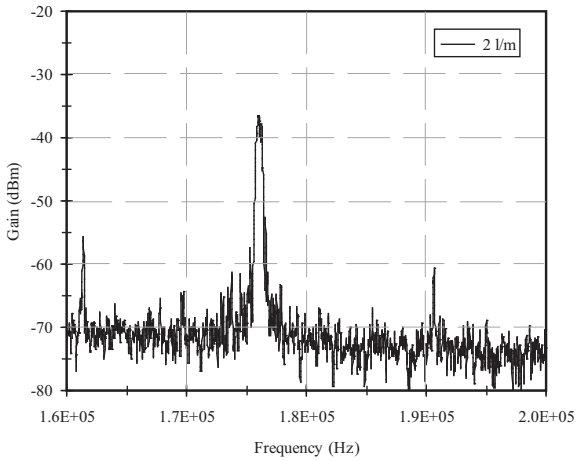


Figure 5: Stable operating mode.

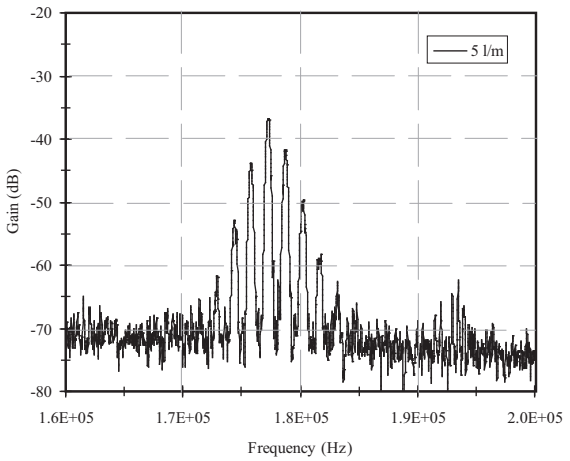


Figure 6: Unstable operating mode.

interaction is adding a reactive component to the circuit. It is thought that this affect may have many processing applications.

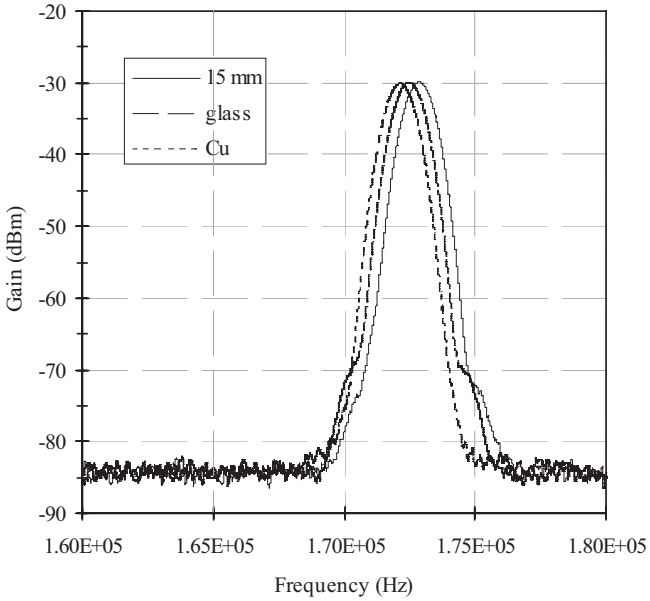


Figure 7: Oscillator frequency pulling due plasma plume-surface interaction. RBW = 100 Hz.

4. PULSE CURRENT WIDTH MODULATION

In this section a Dow Corning plasma solutions Labline^{TM} reel-to-reel helium atmospheric pressure plasma system [13] is monitored using passive RF metrology. The time-domain voltage and current waveforms are captured and compared to the frequency-domain current signal. The purpose of these measurements is to identify the effective pulse of the current drawn by the plasma in each half of the driving frequency period. The voltage and current waveforms for a 1400 W helium plasma is shown in Figures 8 and 9. The spectrum analyzer frequency-domain envelope measurement was optimized by setting the RBW to be larger than the carrier frequency, but significantly smaller than the effective current pulse width, see Figure 10.

A comparison between time-domain and frequency-domain measurements reveals a correlation between the envelope nulls and the current pulse width. For example the nulls have a frequency span (t) of ~ 140 KHz which approximate to the current pulse width ($\sim 7\,\mu$s). In effect the current pulses drawn appears to be pulse width modulating the carrier. If this is the case, the micro-scale events (Figure 6: 0.1 µs with current amplitudes of ~ 0.15 A) may have an influence on the effective current and hence the power dissipated in the plasma.

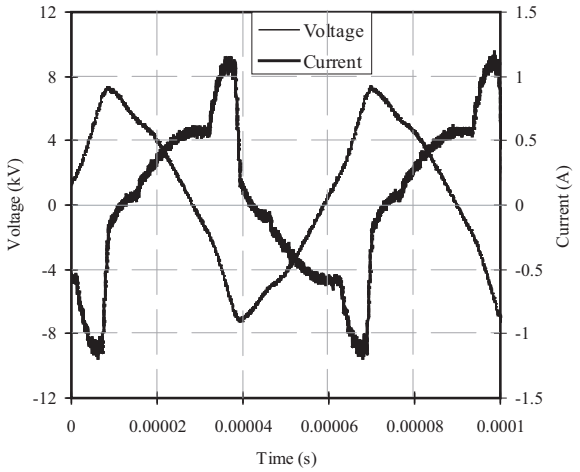


Figure 8: Voltage and current waveforms.

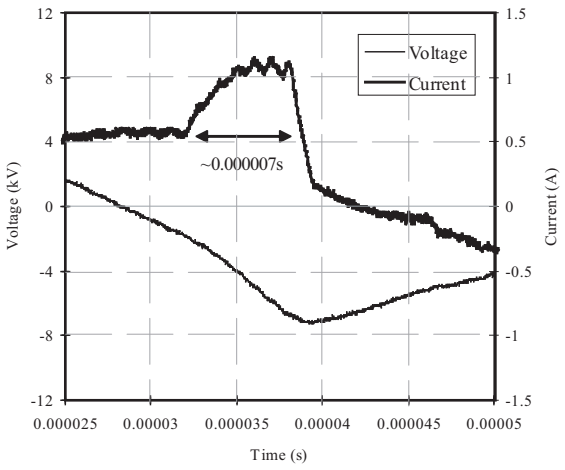


Figure 9: Single current pulse.

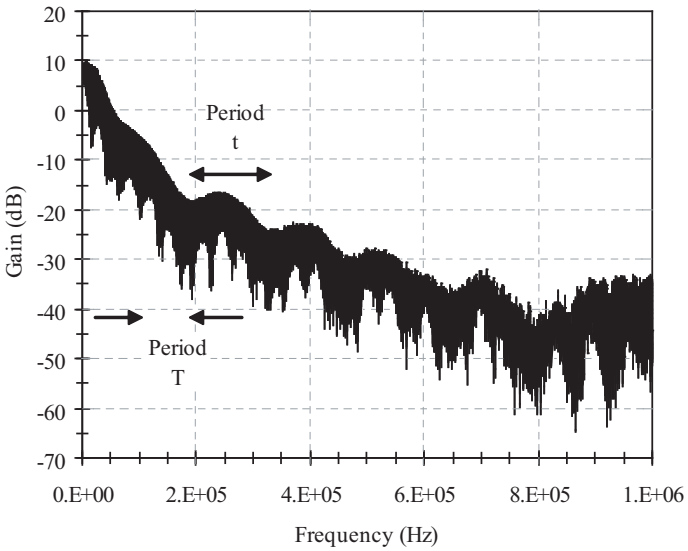


Figure 10: Frequency-domain envelope measurement of current waveform using a 30 kHz RBW.

5. CONCLUSION

This paper has presented active and passive RF metrology measurement of atmospheric pressure plasmas. These measurement are non-invasive (they are only placed on the power-line). Hence they are not limited by the size of the plasma device: This being an important factor when the plasma device is miniaturized for mobile hand-held applications.

The active RF Metrology approach provides an indication of the DBD electrical performance, and when coupled with an EEM the scalar measurement provides information on plasma termination. To gain access to the plasma impedance a vector network measurement needs to be implemented. Passive RF metrology measurements do not require an EEM to interpret the results in that they can be directly correlated to gas flow and plume-surface interactions. The reduction in instrumentation complexity when compared to active RF metrology makes this approach attractive. The frequency-domain measurements provide both temporal (instabilities) and spatial (plume-surface interaction) information. Both of these attributes are necessary for monitoring plasma-surface interactions. The combination of time and frequency-domain measurements allows the current drawn by the plasma to be interrogated. Micro-scale plasma events in the current waveform are likely to modify the effective current pulse width and hence the power developed within the plasma. The origin and timing of these micro-scale events are of great interest as they may disclose the underlying plasma physics and chemistry.

ACKNOWLEDGMENT

This work is supported by Enterprise Ireland under grant number CFTD/7/IT/304.

REFERENCES

1. Eliasson, B., M. Hirth, and U. Kogelschatz, "Ozone synthesis form oxygen in dielectric barrier discharge," *J. Phys. D: Appl. Phys.*, Vol. 20, No. 11, 1421–1437, 1987.
2. Chen, Q., Y. Zhang, E. Han, and Y. Ge, "Atmospheric pressure DBD gun and its application in ink printability," *PSST*, Vol. 14, No. 4, 670–675, 2005.
3. Laroussi, M., "Nonthermal decontamination of biological media by atmospheric pressure plasma: Review, analysis, and prospects," *IEEE Trans. Plasma Sci.*, Vol. 30, No. 4, 1409–1415, 2002.
4. Kieft, I. E., E. P. Van den Laan, and E. Stoffels, "Electrical and optical characterisation of the plasma needle," *New J. Phys.*, Vol. 6, 149, 2004.
5. Stoffels, E., I. E. Kieft, R. E. J. Sladek, L. J. M. Van den Bedem, E. P. Van den Laan, and M. Steinbuch, "Plasma needle for in vivo medical treatment: Recent developments and perspectives," *PSST*, Vol. 15, No. 4, S169–S180, 2006.
6. Walsh, J. L., J. J. Shi, and M. G. Kong, "Contransting characteristics of pulsed and sinusoidal cold atmospheric plasma jets," *Appl. Phys. Letts.*, Vol. 88, No. 17, 171501–3, 2006.
7. Anghel, S. D. and A. Simon, "Measurement of electrical characteristics of atmospheric pressure non-thermal He plasma," *Meas. Sci. Technol.*, Vol. 18, 2642–2648, 2007.
8. Panousis, E., F. Cléments, J.-F. Loiseau, N. Spyrou, B. Held, M. Thomachot, and L. Marlin, "An electrical comparative study of two atmospheric pressure dielectric barrier discharge reactors," *PSST*, Vol. 15, No. 4, 828–839, 2006.
9. Kong, M. G. and X. T. Dang, "Electrically efficient production of a diffuse nonthermal atmospheric plasma," *IEEE Trans. Plasma Sci.*, Vol. 31, No. 1, 7–18, 2003.
10. Rowan, N. J., S. J. MacGregor, J. G. Anderson, D. Cameron, and O. Farish, "Inactivation of mycobacterium paratuberculosis by pulsed electric fields," *Appl. Environ. Microbiol*, Vol. 67, No. 6, 2833–2836, 2001.
11. Law, V. J., F. J. Lawler, and S. Daniels, "Non-invasive VHF Injection signal monitoring in atmospheric pressure plasma and DC Magnetron," *Vacuum*, Vol. 82, No. 5, 514–520, 2008.
12. Law, V. J., "Process induced oscillator frequency pulling and phase noise within plasma systems," *Vacuum*, 82, doi:10.1016/j, vacuum.2007.10.001.
13. Twomey, B., D. P. Dowling, G. Byrne, L. O'Neill, and L. O'Hare, "Properties of siloxane coatings deposited in a reel-to-reel atmospheric pressure plasma system," *Plasma Process and Polymers*, Vol. 4, 450–454, 2007.

Monitoring of Satellite Thermal Pattern in Ocean Front Evolution

S. Nakamura
Kyoto University, Japan

Abstract— Monitoring of satellite thermal pattern of ocean front evolution is studied. The author has obtained a set of satellite thermal pattern on the ocean to find spacial structure of ocean front and its time evolution. A typical pattern is introduced in order to get a dynamical understanding of the ocean front evolution.

1. INTRODUCTION

The author has monitored satellite thermal pattern to find ocean front evolution. For this purpose, a simple antenna of cross-bar type is used for directly receiving satellite signal. The signal was processed by a simple processing system to get satellite thermal pattern. The ocean front between the coastal water and the offshore ocean water can be found approximately by tracing the maximum thermal gradient on the sea surface. It is introduced a schematic pattern of the ocean front evolution in order to see what process is found in the scope of geophysical hydrodynamics.

2. MONITORING

A monitoring system settled on the coast facing an ocean is introduced. A set of data has been obtained after directly receiving satellite signal using the system for the sea surface thermal pattern. Several examples of the typical patterns were already studied by this time to find the ocean front evolution in the northwestern Pacific, though this pattern should be essentially similar to the other area of the ocean, for example, in a case of the ocean front around the Gulf stream in the Atlantic Ocean.

3. SIMPLIFIED MODEL

A simplified model of an ocean front is introduced first as seen in Fig. 1. A strait line of the model ocean front is separating the coastal water and the offshore ocean water. Usually the ocean water is understood circulating in a global scale though only a local area of the interested ocean is considered in this work.

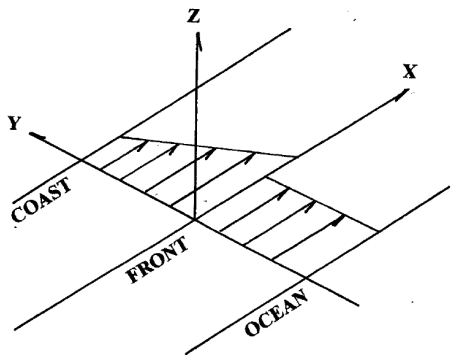


Figure 1: Ocean from model.

4. MOTION OF SEA WATERS

The ocean current is driven by the effect of the densimetric property of the sea water and by the effect of the earth's rotation. The global pattern of the ocean current can be seen after geophysical hydrodynamics. At this time, it is necessary to see that the sea water density is determined by the temperature, salinity and water depth. In this case, thermal pattern of the sea surface is approximately determined by the sea water temperature.

In the actual ocean, the ocean front separates the coastal water and the ocean water. The ocean water is governed by the earth's motion though the coastal water is a small scale motion. In this

work, a simple schematic pattern is assumed for a convenience, and it should be noted that a part of the motion in the sea surface layer is driven by the wind stress on the sea surface.

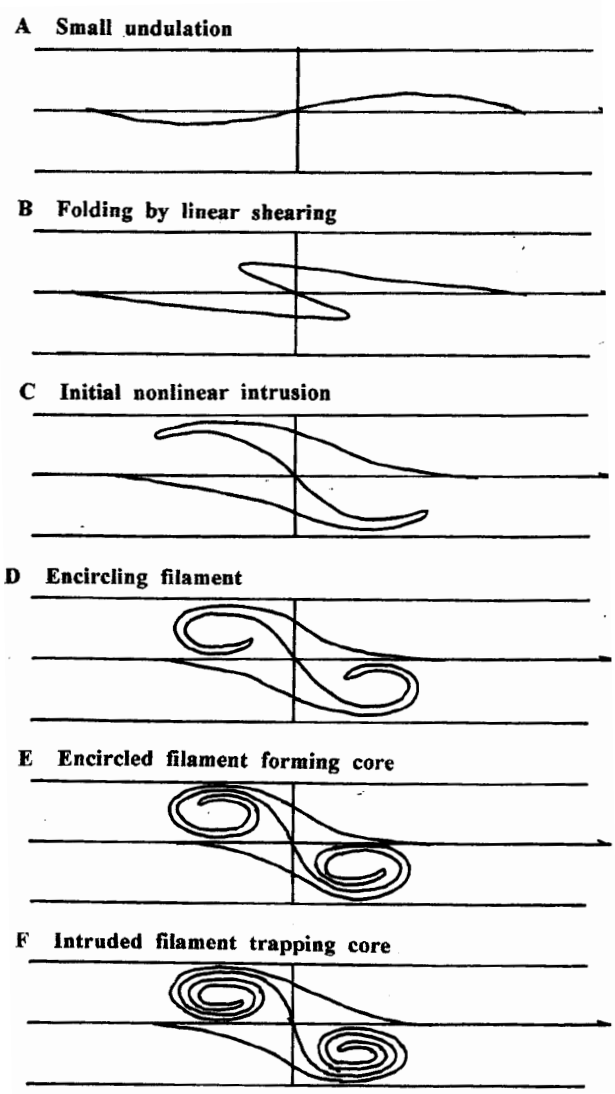


Figure 2: Ocean front evolution.

5. OCEAN FRONT EVOLUTION

Actual ocean front is extended on the ocean and is evolving time to time. The pattern of the ocean front is usually complicated in fact. Then, the author here introduces a simplified model ocean for convenience. The ocean front is found on the sea surface, nevertheless, the contact face of the ocean water and the coastal water is actually in a three-dimensional form. In this work, the author considers a two-dimensional ocean. With some understanding of "hydrodynamics", it can be understood that what process is seen to form a case of complicated case.

Now, consider about an ocean front evolution which can be found when monitoring the satellite thermal pattern of the sea surface.

When a straight front is undulated in the ocean, there should be a shearing effect of the coastal and ocean waters to the ocean front.

When the considering horizontal undulation is small enough, ocean front evolution may be appeared as if it were a linear process. For example, a schematic pattern of this pattern can be seen as shown the pattern A in Fig. 2.

Now, several patterns are introduced which has been found at the direct monitoring of the satellite thermal pattern of the ocean front. The most important factor of ocean front is sea surface

temperature. Salinity is also effective but a minor factor. The motion of the ocean water is in the surface layer of about 200 meters thick. The coastal water is not necessarily higher than the ocean water. Then, an ocean front evolution can be schematically shown as shown in Fig. 2. That is, The first stage "A" usually evolves as those of "B" by the horizontal shear effect.

An ocean front evolution may be seen the stage "A" to "F" with time elapse (cf. Fig. 2).

A) Small undulation — It look as if the undulation were sinusoidal in a linear process.

B) Folding by linear shearing — This could be seen the above undulation in a system of an obliquely crossing Y-co-ordinate about the X-co-ordinate.

C) Initial nonlinear intrusion — This process of intrusion is not well solved by theory. Because the sea water is not an ideal fluid, the intrusion process is not so simple. Nevertheless, during the satellite thermal pattern of the ocean front, the intrusion takes this pattern at the initial stage to develop to form a filament or eddy.

D) Encircling filament — The folding of the ocean front evolves to form a filament of ocean water intruding into the coastal water, and vise versa.

E) Encircled filament forming core — Encircled filament to form a warm core.

F) Intruded filament trapping core — A warm core is established.

This model helps us to see the satellite thermal pattern in a scope of hydrodynamics.

6. SATELLITE MONITORING AND SHIP SURVEY

Before the age of the satellite monitoring, the information of the ocean had been seen by a direct survey along a scheduled legs using the survey ships. The survey data is discrete and the data on the leg requires a long time, for example, several days, or several months.

A schematic example in Fig. 3 shows a relation between the leg P, Q, R, S and T and the matured stage F in Fig. 2 as seen by monitoring satellite thermal pattern of the ocean front.

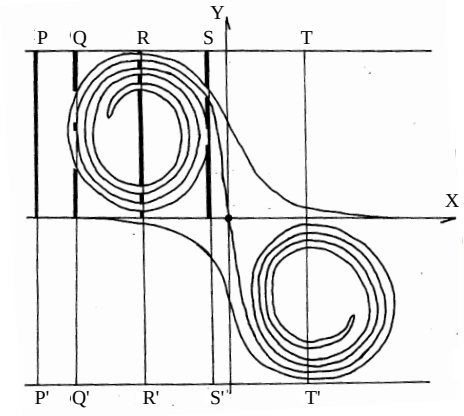


Figure 3: Survey legs for ocean surface observation in relation to satellite thermal pattern of ocean front monitored directly.

In the case of the pattern for $x < 0$ shown in Fig. 3, it can be taken as extensive example of the case shown in Fig. 2. (1) As for the survey leg P to P' stretched the coastal zone to the offshore ocean zone, the segment of the thick line (P–P' for $y > 0$) corresponds to the coastal water zone and the rest of the segment corresponds to the ocean water. (2) As for the leg Q to Q' ($y > 0$), an filament of ocean water intrude to separate the coastal zone into three sub-segments. (3) As for the leg R to R' ($y > 0$), the intruded ocean water force to make 6 coastal water zone. (4) As for the leg S to S' ($y > 0$), three zones of the coastal water are found after intruding the ocean water. In the case for $x > 0$ in Fig. 3, reversed patterns can be seen.

When the coastal water is cooler than that of the ocean water, the front acts as the face of thermal exchange to form a thermal transitional zone between the two waters. At the tip of the intruding filament, final stage of the thermal exchange is seen as a cold core eddy with a warm ring

must be formed in the coastal zone. Similar way, it can be seen a warm core eddy with a cold ring in the ocean water zone.

The author has his understand that there can not be any discontinuous process of the ocean front, even though some scientists name it as a jump on the bases of the discrete data on the survey legs of the survey ships or a coarse scheme for detecting the maximum of satellite thermal gradient. Nevertheless, it is essentially hard to see any discontinuity of the ocean front in the ocean in a scope of hydrodynamics.

The author has understanding that the maximum of a faint satellite thermal gradient might be taken as a lost ocean front or as a discontinuity of the ocean front.

This apparent discontinuity led to a concept of jump apparently appeared during the period of sounding and measurement for the data of the extended legs on the ocean can only be taken as the reliable data if no information of any satellite thermal pattern is given or is referred for relating the surveyed data and satellite information.

Now, it is the time to use the data from the existing satellites and the data from the survey ship.

The model introduced in this work may give us a key to join the information from the survey ship on the leg and the satellite thermal pattern.

7. CONCLUSION

In relation to monitoring satellite thermal pattern of ocean front, a simplified model is introduced in order to illustrate a process of a ocean front evolution. A small undulation can be evolves to a matured warm core in a cold front ring or vise versa. Then, it is shown us that a simple model suggests what is a way to join the discrete data on the legs of the survey ship and a satellite thermal pattern of the sea surface for our dynamical understanding of the ocean front updated.

A New Approach to Mars Ionosphere Characterisation

M. Iorio[1], F. Fois[1], R. Mecozzi[1], G. Picard[2], R. Seu[2], and E. Flamini[3]

[1]Thales Alenia Space Italy, BU Observation Systems & Radar
Via Saccomuro 24, Rome 00131, Italy
[2]Infocom Department, University of Rome "La Sapienza", Via Eudossiana 18, Rome 00184, Italy
[3]ASI, Agenzia Spaziale Italiana, Viale Liegi 26, Rome 00189, Italy

Abstract— Recently in order to investigate the distribution of water, liquid and solid, in the upper portions of the Martian's crust, two different instruments are operatives, the Mars Advanced Radar for Subsurface and Ionosphere Sounding (MARSIS) and the SHAallow RADar (SHARAD). The secondary objective of the MARSIS instrument is the Mars Ionosphere characterisation through the Active Ionosphere Sounding (AIS) sensor. This operative mode presents the frequency agility capability in order to estimate the Mars ionosphere spectral properties. This method allows the sounding of the ionosphere first layer only. However another method to characterise the Mars ionosphere can be used. Here will be presented a new ionosphere data inversion approach using the ground penetrating (GPR) data in order to remove the data instability and the solutions uncertainly. After the analytical data inversion formulation will be evaluated the performances of the method and its applicability. Finally, will be presented an application to the experimental data in order to test the proposed ionosphere data inversion method.

1. INTRODUCTION

The Mars Advanced Radar for Subsurface and Ionosphere Sounding (MARSIS) and the SHAallow RADar (SHARAD) are a low frequency (1.8–5 MHz for MARSIS and 20 MHz for SHARAD) Ground Penetrating Radar (GPR) in altimeter configuration which uses the synthetic aperture technique. The manly differences between the two instruments are the spatial resolution (150 m vertical resolution, 5–10 Km along track resolution for MARSIS and 15 m vertical resolution, 300–500 m along track resolution for SHARAD) and the penetrating capability (about 5 Km for MARSIS and 1 Km for SHARAD). One of the operative modes of the MARSIS instrument is the Active Ionosphere Sounding (AIS). The basic objective of the AIS mode is the Mars ionosphere characterisation, in particular the ionosphere plasma frequency measure in order to obtain information about the ionosphere total electron content (TEC) versus the solar zenith angle (SZA). According to the plasma frequency trend versus the altitude sketched in the Fig. 1, this method allows the measure of the plasma frequency shape in the first layer of the ionosphere and the determination of the maximum plasma frequency value. By these two parameters can be derived the complete plasma frequency profile with a best fit procedure, the first result of the AIS experiment are reported in [1]. Even though AIS is the basic instrument to ionosphere characterisation, to this scope there is an alternative approach based on the GPR data. As illustrated in [2], the impact of the ionosphere on the GPR signal appears relevant despite the frequency of the signal is higher to maximum plasma frequency. The GPR signal passes thorough the ionosphere layer and after the surface reflection is received by the instrument, during this path the signal is distorted by the ionosphere presence. To optimise the GPR data this distortion shall be estimated to compensate the signal. The information of the compensation function coefficients can be used to extrapolate the ionosphere property. Clearly this is an indirect approach respect to AIS mode, however is equally valid as presented in [3], although in [3] one of the coefficient used to estimation procedure is the signal delay introduced by the ionosphere that is a parameter with a large sensibility to the surface roughness. To this reason in the present paper a more robust procedure will be illustrated in order to reduce the estimation accuracy of the method illustrated in [3].

2. IONOSPHERE MODEL

As well known [4] the refraction index of a ionized medium is a function of the frequency:

$$n(f) = \sqrt{1 - \left(\frac{f_P(z)}{f}\right)^2} \tag{1}$$

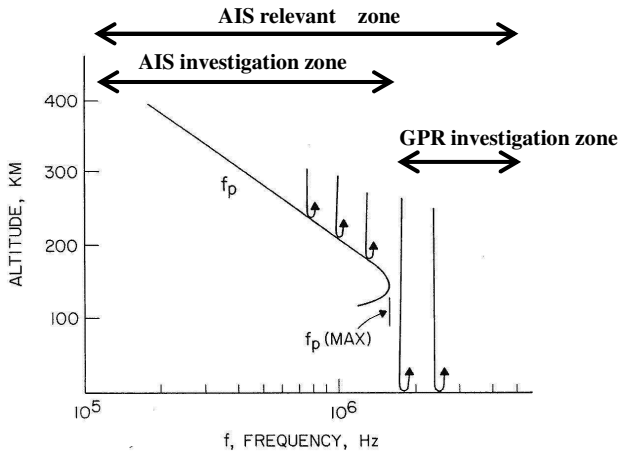


Figure 1: Representative profile of the electron plasma frequency.

Here $n(f)$ is the refraction index, $f_p(z)$ is the plasma frequency vertical profile, z is the generic altitude coordinate and f is the frequency. During the propagation through the ionosphere the signal subject to a extra-phase shift which can be expressed by the following:

$$\Delta\phi(f) = \frac{4\pi}{c} f \int_0^L \left[\sqrt{1 - \left(\frac{f_P(z)}{f} \right)^2} - 1 \right] dz \tag{2}$$

being c the speed of light in vacuum and L the thickness of medium. As seen from (2) the phase shift as function of the frequency, it depends by in the variation of the plasma frequency along the dispersive layer. The plasma frequency can be related to the plasma local electron density by the relationship:

$$f_P(z) = 8.98\sqrt{n_e(z)} Hz \cdot m^{3/2} \tag{3}$$

In order to quantify the distortion effects on the signal impulse due to the ionosphere propagation, can be introduced the following "gamma" model [3] to characterize the distribution of the plasma frequency vs. height:

$$f_P(z) = f_{P\max} \frac{z - h_0}{b} e^{1 - \frac{z-h_0}{b}} \delta_{-1}(z - h_0) \tag{4}$$

According to this model the ionosphere plasma frequency is zero for $z < h_0$ (below the altitude $h_0 \sim 100$–150 Km the ionosphere profile is zero), and assume its maximum value $f_{p,\max}$ for $z = h_0 + b$ and then tends smoothly to zero for increasing z. The model (4) appears in perfect accordance with the experimental data measured profiles of Viking mission [4]. In the following the model (2) will be used to evaluate the inpact of the ionosphere phase distorction on the GPR singnal.

3. IONOSPHERE INVERSION APPROACH

In order to reduce the estimation complexity an alternative ionosphere model can be performed. This model is based on an equivalent ionosphere model characterized by a constant plasma frequency $f_{P,EQ}$ and a thickness L_{EQ}, so that the Equation (2) becomes (uniform model):

$$\Delta\phi_{EQ}(f) = \frac{4\pi f L_{EQ}}{c} \left(\sqrt{1 - \left(\frac{f_{P,EQ}}{f} \right)^2} - 1 \right) = 2\pi\tau_0 \left(\sqrt{f^2 - f_{P,EQ}^2} - f \right) \tag{5}$$

Where $\tau_0 = 2L_{EQ}/c$. In [5] is demonstrated the good agreement of the uniform model vs. the gamma model when the value of factor b minimize the error due to particular L_{EQ} assumption. Using a numeric regressive technique it is possible to demonstrate the linear relationship between the equivalent and maximum plasma frequency [6]. This equation can be expressed as:

$$f_{P\max} = k_p f_{P,EQ} \tag{6}$$

Where k_p is a constant respect to carrier frequency that can be fixed to $k_p = 1.18$. The uniform model and the relationship (6) can be used to simplify the data inversion procedure. As above mentioned, in [3] is illustrated an application to MARSIS GPR data of the uniform model to estimate the equivalent plasma frequency of the Mars ionosphere. However in this case is used as a parameter the time delay measured by the instrument in order to extract the ionosphere time delay component. This aspect introduces a very high inaccuracy due to dependence of the time delay by the hardware path delay and the surface roughness. For this reason in [3] it is derived only a bandwidth of the ionosphere plasma frequency. A way to reduce the ionosphere plasma frequency estimation accuracy can be based on a second order estimation technique. This procedure is based on the demonstration of the ionosphere two frequency quadratic coefficients ratio invariant property respect to the ionosphere shape factor. By the Equations (1) and (5) the ionosphere phase coefficients can expressed as:

$$a_1 = 2\pi\frac{2}{c}\int_0^L \left(\frac{f_0}{\sqrt{f_0^2 - f_p^2(s)}} - 1\right) ds, \quad a_2 = -2\pi\frac{2}{c}\int_0^L \left(\frac{f_p^2(s)}{2\left(f_0^2 - f_p^2(s)\right)^{3/2}}\right) ds \tag{7}$$

for the gamma model and:

$$a_1 = 2\pi\frac{2L_{EQ}}{c}\left(\frac{f_0}{\sqrt{f_0^2 - f_{P,EQ}^2}} - 1\right), \quad a_2 = -2\pi\frac{2L_{EQ}}{c}\left(\frac{f_{P,EQ}^2}{2\left(f_0^2 - f_{P,EQ}^2\right)^{3/2}}\right) \tag{8}$$

in the case of the uniform model. Considering the ratio of the phase quadratic coefficients (8) at two different carrier frequencies by the Equation (8) is simple to conclude that this ratio appears independent by the equivalent thickness LEQ. This property can be demonstrated in the general case using the gamma model, considering the Equation (7) the quadratic coefficients ratio can expressed as:

$$\frac{a_2|_{f_{01}}}{a_2|_{f_{02}}} = \frac{\sum_{k=0}^{\infty}(-1)^k \binom{-3/2}{k} \frac{e^{2k+2}(2k+2)!}{(2k+2)^{2k+3}} \left(\frac{f_{P\max}}{f_{01}}\right)^{2k+2}}{\sum_{k=0}^{\infty}(-1)^k \binom{-3/2}{k} \frac{e^{2k+2}(2k+2)!}{(2k+2)^{2k+3}} \left(\frac{f_{P\max}}{f_{02}}\right)^{2k+2}} = Fun\left(f_{01}, f_{02}, f_{P\max}\right) \tag{9}$$

Where f_{01} and f_{02} are the two different carrier frequencies. Remembering the relationship (8) and (9) the maximum ionosphere plasma frequency can be calculated by the relation:

$$f_{P\max} = k_p\sqrt{\frac{f_{02}^2 - f_{01}^2\left(a_2|_{f_{01}}/a_2|_{f_{02}}\right)^{2/3}}{1 - \left(a_2|_{f_{01}}/a_2|_{f_{02}}\right)^{2/3}}} \tag{10}$$

In the (10) the quadratic phase coefficients can be determined by the application of the methodologies illustrated in [7] and [8] using the amplitude contrast, Shannon entropy or peak detection

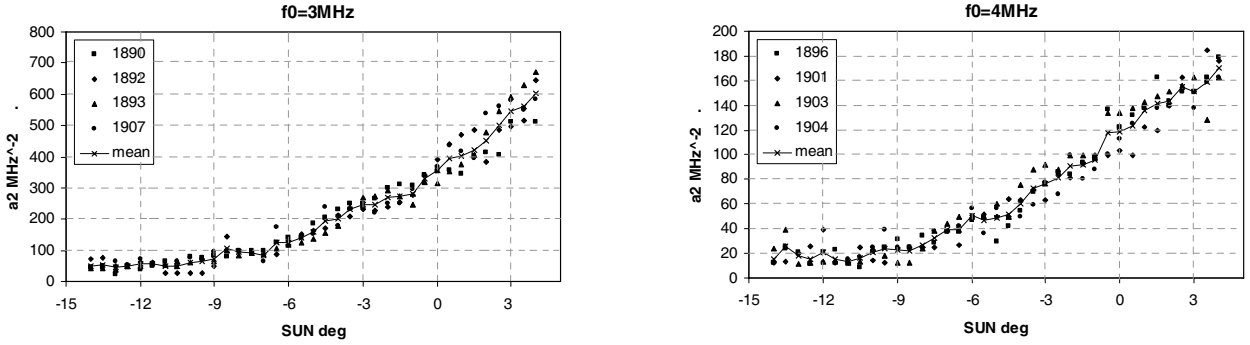


Figure 2: Quadratic coefficient at carrier frequency of 3 and 4 MHz.

method. Although the inversion approach based on the relationship (10) is a second order estimation technique with an apparently lower sensibility respect to the first order approach presented in [3], this methodologies is very simple and roust respect to surface roughness property. Clearly the accuracy of the estimation depend by the estimation of the ionosphere quadric phase coefficients, these details are illustrated in [5] and [7].

4. EXAMPLE OF APPLICATION AND CONCLUSIONS

In this section will be applied the Mars ionosphere plasma frequency estimation methodology illustrated in the previous section using the MARSIS and SHARAD available data. In the case of the MARSIS instrument the GPR data are collected at the carrier frequency of 3 MHz (orbit 1890,1892,1893 and 1907) and 4 MHz (orbit 1896, 1901, 1903 and 1904). The estimated quadratic phase coefficients, as a function of the sun elevation angle, using the amplitude contrast method are reported in Fig. 2.

In order to reduce the dispersion of the quadratic coefficients it is necessary operate an opportunely average (see Fig. 2) and successively can be evaluated the ratio between the coefficients (see Fig. 3).

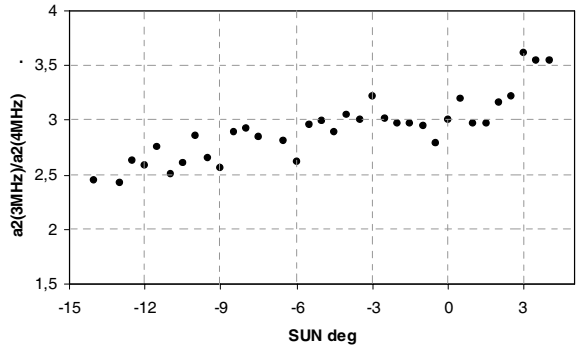


Figure 3: Quadratic phase coefficients ratio.

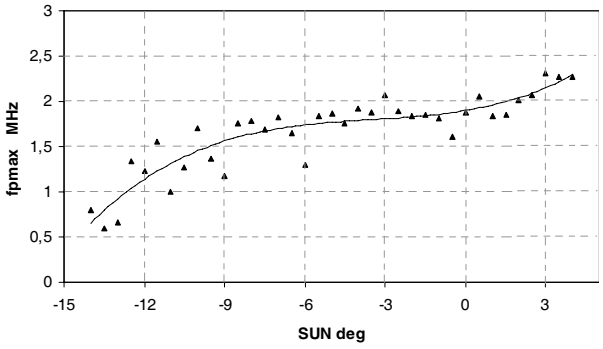


Figure 4: Estimated maximum plasma frequency.

The accuracy of the estimated plasma frequency can be investigated using the results obtained in [5, 7] and the relationship (11). Assumed $\sigma_{a2} = 15\,\mathrm{MHz}^{-2}$ [5] (take into account that the principal effect is due from the speckle and noise) and being the value of the quadratic phase ratio ~ 3 for a $SUN = -5deg$ the total variation of the ratio can be approximated as:

$$\sigma^{\%}_{a_2}(f_{01}) + \sigma^{\%}_{a_2}(f_{02}) = \frac{\sigma_{a_2}(f_{01})}{a_2(f_{01})} + \frac{\sigma_{a_2}(f_{01})}{a_2(f_{02})} \cong 0.4 \tag{11}$$

By the Fig. 4 the accuracy of the plasma frequency results $\sigma^{\%}_{fp} \sim 0.1$ for $SUN = -5deg$ the plasma frequency can be assumed in the range $1.67 \div 2.05\,\mathrm{MHz}$. The estimated plasma frequency

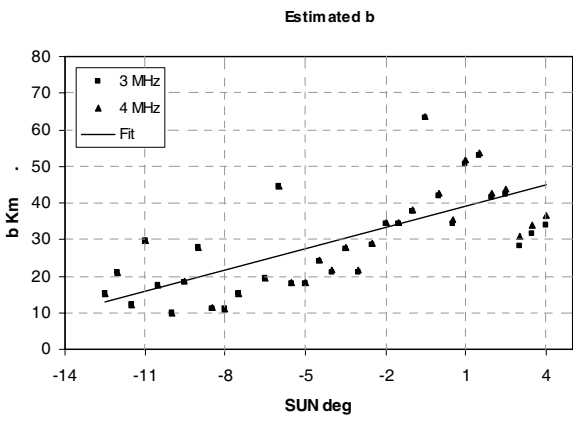


Figure 5: Estimated ionosphere shape factor.

at $SUN = -5deg$ is 1.86 MHz, this value is in the range of the expected accuracy. Applying the Equation (8) at different carrier frequencies it is possible to evaluate the parameter b as reported in Fig. 5, in conclusion, in this case can be assumed a mean value of $b \sim 30$ Km. The analogous procedure can be applied to SHARAD GPR data, in this case the carrier frequencies are more highs and the sensibility is lower respect to the MARSIS data, an example of this estimation is reported in Fig. 6. In Fig. 6 are compared the Mars ionosphere maximum plasma frequency estimated using the MARSIS and SHARAD data vs. the data available in literature [4]. As can be observed the estimated ionosphere plasma frequency results in accordance with the older data, in corresponding of a solar zenith angle of $100deg$ the estimated value are more highs, this trend is confirmed by the AIS operative mode and is due by the ionosphere ionization reduction delay that can be associated to the ionosphere inertial characteristic.

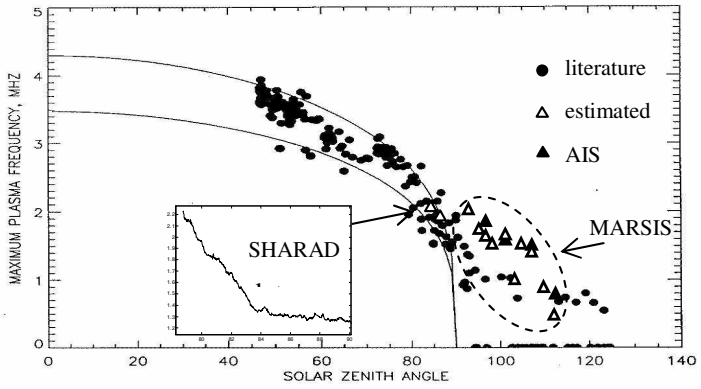


Figure 6: Comparison between the estimated ionosphere maximum plasma frequency.

REFERENCES

1. Gurnet, D. A., et al., "Radar soundings of the ionosphere of Mars," *Science*, Vol. 30, November 2005.
2. Safenili, A., W. Kofman, J.-F. Nouvel, A. Herique, and R. L. Jordan, "Impact of Mars ionosphere on orbital radar sounder operation and data processing," *Planetary and Space Science*, Vol. 51, No. 7, June 2003.
3. Picardi, G., et al., "Mars ionosphere data inversion by MARSIS subsurface signals analysis," *Geophysical Research*, Vol. 8, 2006.
4. Hansot, W. B., et al., "The martian ionosphere as observed by the Viking retarding potential analyzers," *Journal of Geophysical Research*, Vol. 82, 1977.
5. Picardi, G. and S. Sorge, "Adaptive compensation of Mars ionosphere dispersion," Technical Report, University of Rome 'La Sapienza', 1999.
6. Iorio, M., "GPR sensors: Modelling and inversion approach," PhD Thesis, University of Rome 'La Sapienza', 2004.
7. Biccari, D., G. Picardi, and R. Seu, "Adaptive compensation of Mars ionosphere dispersion in the MARSIS experiment," *IEEE*, 2001.
8. Ilyushin, Y. A., V. E. Kunitsyn, and T. Hagfors, "A new algorithm for ionosphere correction in the MARSIS experiment," Technical Report, Moscow State University, 2000.

Representation of Einstein's Relativity by Smith's Chart

C. F. Chen[1] and Y. F. Tung[2]

[1]Department of Electronic Engineering, The Chinese University of Hong Kong
Shatin, N. T., Hong Kong, China
[2]Department of Translation, The Chinese University of Hong Kong
Shatin, N. T., Hong Kong, China

Abstract— The first author is an engineer and the second a humanist. They co-author this paper starting from two completely different points of view: one through a mathematical reasoning, the other the pictorial approach on the Smith Chart. Eventually, a new representation of Einstein's special theory of relativity is obtained.

P. A. M. Dirac's well-known pamphlet, *General Theory of Relativity*, uses the following set of notations [1]

$$t = x^0, \quad x = x^1, \quad y = x^2, \quad z = x^3, \quad \text{or written as } x^\mu, \ \mu = 0, 1, 2, 3$$

where t and x, y, z are the four coordinates of the time t and three space coordinates x, y, z, while Einstein's original work is to use

$$t = t, \quad x = x, \quad y = 0, \quad z = 0$$

He lets $y = 0$, $z = 0$, just for simplicity.

In his book *Relativity* [2], Einstein states that the space-time continuum of the special theory of relativity considered as a Euclidean Continuum introduces Minkowski's work for description of the four-dimensional space-time continuum. For these systems, the four coordinates t, x, y, z which determine an event or a point of the four dimensional continuum, are designed physically in a simple manner.

The validity of the Lorentz transformation follows from the following condition

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2$$

which belongs to two adjacent points of the four dimensional space time continuum and has the same value for all selected reference-bodies.

Now, using $y = z = 0$, we have

$$ds^2 = d(ct)^2 + dX^2 \tag{1}$$

where ct and X are the time in terms of distance and the space respectively. And S is two adjacent points of the space time continuum. According to Einstein, c is the speed of light, regardless of how fast the source is moving.

Assuming v is the speed of a uniform-motion-object, we have

$$X = vt \tag{2}$$

then Equation (1) can be written as

$$S^2 = \left(c^2 + v^2\right) t^2 \tag{1a}$$

as far as

$$v \leq c. \tag{3}$$

In other words, combining (1a) and (3), we can write

$$S^2 = \left(c^2 - v^2\right) t^2. \tag{4}$$

However, if $v > c$, (4) does not hold.

Rewrite (4):

$$S^2 = c^2 t^2 - X^2 \tag{4a}$$

That implies in general

$$S^2 = (ct)^2 - (X^2 + Y^2 + Z^2) \tag{4b}$$

Equation (4b) is equivalent to the idea derived from Minkowski in 1908.
Minkowski, moreover, using the symbol $j = \sqrt{-1}$, yields

$$(jS)^2 = (jct)^2 + (X^2 + Y^2 + Z^2)$$

or

$$S^2 = -(jct)^2 - X^2 - Y^2 - Z^2$$

Equation (4a) will be used to explain the theory of relativity on the Smith Chart.
Now let us review the basic construction principle of Smith Chart.
There are two complex planes $\gamma = \alpha + j\beta$, $w = u + jv$
Through the following bilinear transformation:

$$w = \frac{\gamma - 1}{\gamma + 1}, \quad \text{or} \quad \gamma = \frac{1 + w}{1 - w}$$

we obtain

$$\alpha = \frac{1 - u^2 - \nu^2}{(1-u)^2 + \nu^2} \quad \text{and} \quad \beta = \frac{2v}{(1-u)^2 + \nu^2}$$

Rearranging gives

$$\left(u - \frac{\alpha}{1+\alpha}\right)^2 + \nu^2 = \left(\frac{1}{1+\alpha}\right)^2 \tag{5}$$

$$(v-1)^2 + \left(\nu - \frac{1}{\beta}\right)^2 = \left(\frac{1}{\beta}\right)^2 \tag{6}$$

Equation (5) is the equation for a circle having a radius $\frac{1}{1+\alpha}$ and a center $(\frac{\alpha}{1+\alpha}, 0)$. Different values of α yield circles of different radii with centers at different points on the u axis. A family of α circles are shown by solid lines in Fig. 1. Equation (6) will give a family of β circles as shown by dotted lines in the same figure.

The Chart was originally designed for using in impedance matching in Transmission-line [3, 4], and then has been applied to Waveguides studies [5], and later on, with some modification to feedback systems design and other applications by the first author of this paper [6–8].

The second author, however, simply sees the Smith's Chart as a picture without knowing anything about its mathematic background. She is very knowledgeable about non-mathematic literature about Einstein because she has translated many non-mathematical books and written one concerning Einstein [9]. When, incidentally, the first author was working on the subject of relativity on the Smith Chart, she asked: "Is the Smith Chart the whole universe of time-space of Einstein's relativity?"

Surprisingly inspired by her question, the first author answered: "Perhaps you are considering in the right direction. If you are standing on the leftist point of the Chart, which is the zero point of the circle, and go along the upper half circle to the rightist point of the Chart, then that is your t variable, and the horizontal line is your x variable. The Smith Chart indeed can be considered as the whole picture of your universe.

She then immediately quoted Bertrand Russell's *The ABC of Relativity* [10] as follows:

> Everybody knows that Einstein has done something astonishing, but very few people know exactly what it is that he has done. It is generally recognized that he has revolutionised our concept of the physical world, but his new conception are wrapped up in mathematical technicalities. What is demanded is a change in our imaginative picture of the world ..... A change in our imagination is always difficult, especially when we are no longer young.
>
> In exploring the surface of the earth, we make use of all our senses, more particularly of the senses of touch and sight. In measuring lengths we use "foot" which is a part of the human body. For longer distances, we think of time it takes to walk from one place to another. We gradually learn to judge distances roughly by the eye but rely upon touch for accuracy.

In studying the heavens we are debarred from all senses except sight ..... It has turned out that much of what we learned from the sense of touch was unscientific prejudice, which must be rejected if we are to have a true picture of the world.

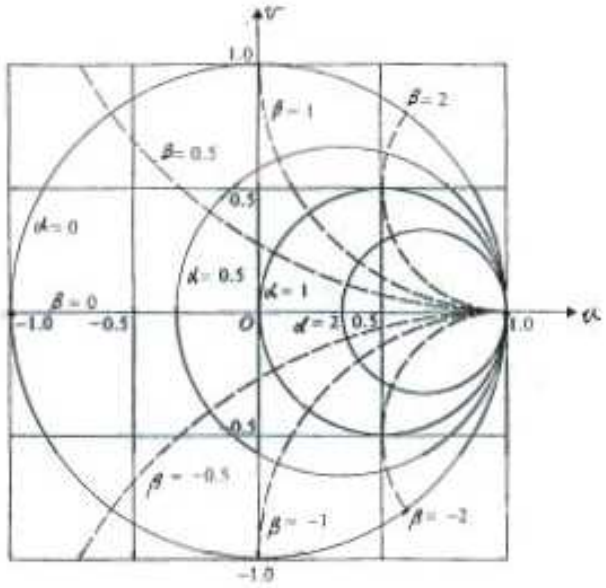


Figure 1: Smith Chart with rectangular coordinates.

Based on Russell's point of view, the surface of the Smith Chart is a surprisingly good device for explaining the theory of relativity. For example, let us stand on the leftist point and look to the right hand side of the chart. The horizontal axis of Fig. 2 shows exactly as we are looking at uniformly distributed trees along a straight long road. These trees form a line progressing in a logarithmic scale to infinity.

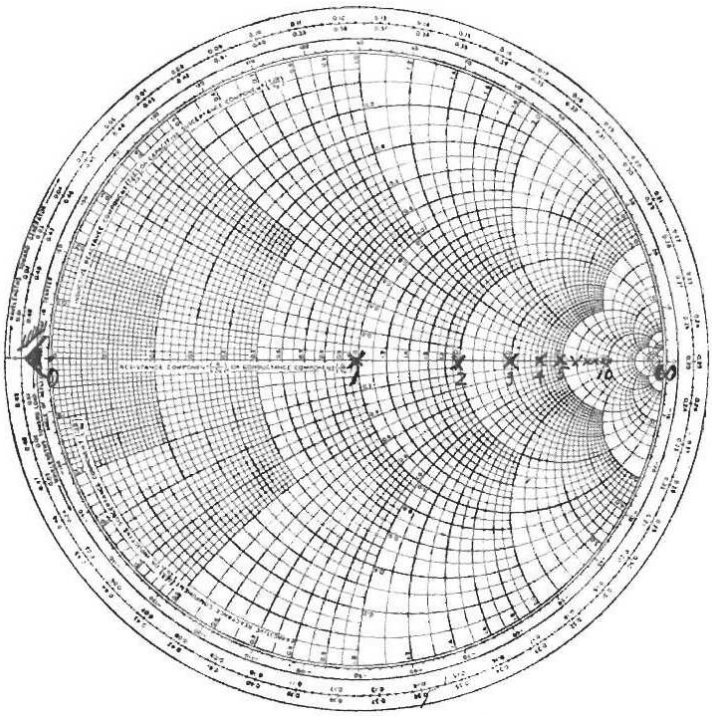


Figure 2: Details of the Smith Chart.

Secondly, let us check the idea of light cones. The light cone is defined by the rectangular coordinates using the horizontal axis as space and the vertical axis as time. We then have two cones: the inverted cone is for the future and the regular cone for the past while the origin as the observer stands at the present time. These mean that we imagine the universe in the sense of touch. Not only our geometry and physics but our whole conception of what exists outside us, is based upon the sense of touch. In studying "heavens" we cannot use the sense of touch at all. We cannot touch the sun or travel to it. However, scientists have unhesitatingly applied the geometry which they found serviceable on the surface of the earth, and which they had based upon touch and travel. In doing so, they brought down trouble on their heads which it has been left for Einstein to clear up. It has turned out that much of what we learned from the sense of touch was unscientific prejudice, which must be rejected if we are to have a true picture of the world.

Those cones can be mapped on the Smith Chart by connecting $\alpha = \beta$ as shown in Fig. 3. This means that we imagine the picture of the universe in the sense of sight rather than that of touch. Between the arc drawn from the Chart and the horizontal axis is the region of "absolute elsewhere", a term coined by Sir Arthur Stanley Eddington. Among other ideas in the special relativity, some of which can already be explained on the Smith Chart while others may also be done in the course of time.

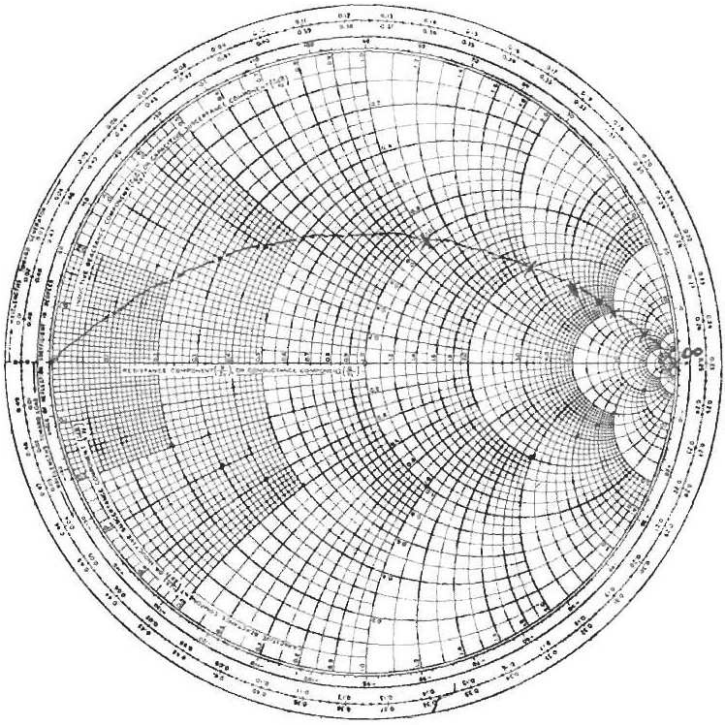


Figure 3: The Smith Chart with the light-cone boundary.

In conclusion, the research on the Smith Chart which is frequently used in the electrical engineering and particularly in the microwave technology has been extended to the representation problem of relativity theory. The cross points of α, β circles of the γ plane depend exclusively on sight. The background is the u–v rectangular coordinates of the w plane. The two complex planes are related by linear fractional transformation. This transformation has the most beautiful property: Any linear fractional transformation transforms circles and lines into circles and lines. In doing so, we are getting a true universe in the sense of sight instead of a universe in the sense of touch.

REFERENCES

1. Dirac, P. A. M., *General Theory of Relativity*, Wiley, New York, 1975.
2. Einstein, A. and translated by R. W. Lawson, Chapters 25 and 26, *Relativity — The Special and the General Theory — A Popular Exposition*, 96–107, Wing Books, New York, 1961.
3. Smith, P. H., "Transmission-line calculation," *Electronics*, Vol. 12, 29, January 1939.

4. Smith, P. H., "An improved transmission-line calculator," *Electronics*, Vol. 17, 130, January 1944.
5. Cheng, D., *Field and Wave Electromagnetics*, Second Edition, Addison and Wesley, Reading, Mass., 1989.
6. Chen, C. F. and I. J. Haas, *Elements of Control Systems Analysis: Classical and Modern Approaches*, Prentice Hall, Englewood Cliffs, N. J., 1968.
7. Chen, C. F., *Digital Filters Design*, Lecture Notes, The Chinese University of Hong Kong, Hong Kong, 1978–1984.
8. Chen, C. F. and D. W. C. Shen, "A new chart relating the open loop and closed loop frequency responses of linear control system," *American Institute of Electrical Engineers Transactions on Applications and Industry*, 252–255, September 1959.
9. Tung, Y. F.(translator), A Chinese Translation of Lightman, A., *Einstein's Dreams*, Pantheon Books, New York, 1993 and Renn, J. and R. Schulmann, editors, *Albert Einstein/Mileva Marić: The Love Letters*, Princeton University Press, Princeton, N. J., 1992.
10. Russell, B., "Touch and sight: The earth and the heavens," (One Chapter from *The ABC of Relativity*, Harper & Brothers, New York, 1925), *Selected Papers of Bertrand Russell*, The Modern Library, New York, 1927.

A Toroidal Harmonic Representation of the Yukawa-potential Kernel for a Circular Cylindrical Source

J. P. Selvaggi, S. J. Salon, and M. V. K. Chari
Rensselaer Polytechnic Institute, Troy, NY 12180-3590, USA

Abstract— A true cylindrical series expansion of the Yukawa- or screened Coulomb-potential kernel is developed for a finite circular cylindrical source through the application of a toroidal harmonic expansion. The Yukawa kernel is separated into a singular and nonsingular part. The singular part is expanded in terms of the associated toroidal harmonics and the nonsingular part is expanded in terms of an elementary binomial expansion.

1. INTRODUCTION

The Yukawa or screened Coulomb potential has applications in fields such as plasma physics [1–3], molecular and biophysics [4–6], nanoscience [7] and many others. The Yukawa potential also plays a significant role in the study of classical and quantum gravity [8–12]. An exact series expansion is developed which employs the associated toroidal harmonics of the second kind and are applicable for a finite circular cylindrical source with a rectangular cross section.

A problem that one encounters when working with the Yukawa potential, in circular cylindrical coordinates, is trying to find an expansion of $K(\boldsymbol{r},\boldsymbol{r}') = \frac{e^{-\lambda|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}$ valid for all regions of space. However, the application of toroidal functions of the second kind offers a solution to this problem. In fact, the toroidal functions of the second kind also have some definite numerical advantages [13]. The associated toroidal harmonics of the second kind are represented in the literature by $Q^l_{m-\frac{1}{2}}(\xi)$ where $\xi = \frac{\rho^2+\rho'^2+(z-z')^2}{2\rho\rho'} > 1$. Another very important feature of these functions is that they are used to represent the inverse distance function in circular cylindrical coordinates and all its higher order derivatives using one family of functions, namely–$Q^l_{m-\frac{1}{2}}(\xi)$. These functions are sometimes called ring functions [14] or Legendre functions of the second kind and of half-integral degree and of order l [15].

2. THE YUKAWA-POTENTIAL KERNEL OF A CIRCULAR CYLINDRICAL SOURCE

The Yukawa or screened Coulomb potential, due to a general circular cylindrical source, at some arbitrary field or observation point $P(\rho,\phi,z)$ shown in Figure 1 below is given by

$$\Phi_P(\rho,\phi,z) = \int_{V'} \zeta_v(\boldsymbol{r}') \frac{e^{-\lambda|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3r', \tag{1}$$

where ζ_v is the forcing function(not necessarily a constant), $\boldsymbol{r}$ is the position vector of the observation point, $\boldsymbol{r}'$ is the position vector of differential volume associated with source point and the kernel function is given by

$$K(\boldsymbol{r},\boldsymbol{r}') = \frac{e^{-\lambda|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{2}$$

For an electrostatic source, ζ_v represents a volume charge density. For a gravitational source, ζ_v characterizes the relative strength of the non-Newtonian interaction compared to that of the Newtonian one. For applications in non-Newtonian gravity, ζ_v is generally a constant but it need not be. If Eq. (1) represents the screened Coulomb potential then λ represents the Debye screening factor. In non-Newtonian gravity, λ^{-1} represents a length scale and this provides a means for knowing when the Yukawa potential becomes relevant.

The reciprocal distance, in circular cylindrical coordinates, between the source point, $P'(\rho',\phi',z')$, and an arbitrary observation point, $P(\rho,\phi,z)$, is given by

$$\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} = \frac{1}{\sqrt{\rho^2+\rho'^2+(z-z')^2-2\rho\rho'\cos(\phi-\phi')}}. \tag{3}$$

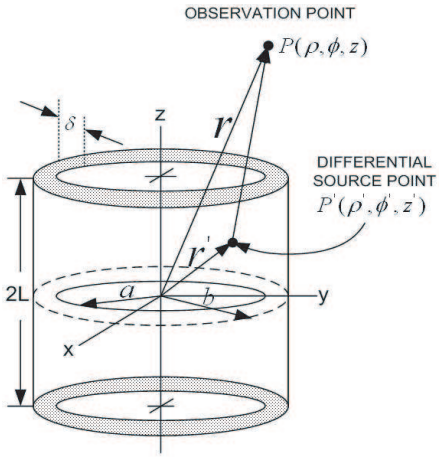


Figure 1: Finite and hollow circular cylinder source.

Figure 1 illustrates a simple model of a hollow circular cylindrical source with a finite thickness of δ.

This hollow cylinder is a general source since it leads to all the limiting cases one may encounter. This may be a solid finite circular cylinder, a thin flat disk, a thin ring, a thin hollow tube, a very thin and hollow finite circular cylinder or infinite cylinders. The main purpose is to find a general expansion for the Yukawa-potential kernel, in terms of cylindrical coordinates, which can be employed to handle any case that one may encounter when dealing with circular cylindrical geometry.

The Yukawa kernel given by Eq. (2) can be written as

$$K(\boldsymbol{r},\boldsymbol{r}')=\frac{1}{R}\sum_{l=0}^{\infty}\frac{(-1)^l\lambda^l R^l}{l!}=\frac{1}{R}\left\{\sum_{l=0}^{\infty}\frac{\lambda^{2l}R^{2l}}{(2l)!}-\sum_{l=0}^{\infty}\frac{\lambda^{2l+1}R^{2l+1}}{(2l+1)!}\right\}=\sum_{l=0}^{\infty}\frac{\lambda^{2l}R^{2l-1}}{(2l)!}-\sum_{l=0}^{\infty}\frac{\lambda^{2l+1}R^{2l}}{(2l+1)!}, \quad (4)$$

where $R=|\boldsymbol{r}-\boldsymbol{r}'|$. The first sum in Eq. (4) has a singularity at $R=0$ but the second sum does not. One can employ the not-so-well-known associated toroidal harmonic expansion found in Snow [16] and MacRobert [14] (appendix) to write the R^{2l-1} in first sum of Eq. (4) as follows:

$$R^{2l-1}=\frac{1}{\pi\sqrt{\rho\rho'}}2^{2l}(2l-1)!!(\rho\rho')^l(\xi^2-1)^{\frac{l}{2}}\sum_{m=0}^{\infty}\frac{(2m-2l-1)!!}{(2m+2l-1)!!}Q^l_{m-\frac{1}{2}}(\xi)\cos[m(\phi-\phi')], \quad (5)$$

Where $Q^l_{m-\frac{1}{2}}(\xi)$ are called the associated Legendre functions of the second kind and half-integral degree and of order l [15] or the associated toroidal functions of the second kind and of order l [17]. Its argument $\xi=\frac{\rho^2+\rho'^2+(z-z')^2}{2\rho\rho'}>1$. Equation (5) is valid for any integer l. The Neumann factor ε_m [18] is 1 for $m=0$ and 2 otherwise. The R^{2l} given in the second sum of Eq. (4) can easily be represented using a binomial expansion and is expressed as

$$R^{2l}=(2\rho\rho')^l\left[\xi-\cos(\phi-\phi')\right]^l=(2\rho\rho'\xi)^l\sum_{k=0}^{l}(-1)^k\binom{l}{k}\xi^{-k}\cos^k(\phi-\phi'). \quad (6)$$

Employing Eqs. (5) and (6) allows one to write Eq. (4) as

$$\begin{aligned}K(\boldsymbol{r},\boldsymbol{r}') &= \frac{1}{\pi\sqrt{\rho\rho'}}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\varepsilon_m\lambda^{2l}C(m,l)(\rho\rho')^l(\xi^2-1)^{\frac{l}{2}}Q^l_{m-\frac{1}{2}}(\xi)\cos[m(\phi-\phi')]\\ &\quad-\sum_{l=0}^{\infty}\sum_{k=0}^{l}\lambda^{2l+1}D\,(l,k)\left(\rho\rho'\right)^l\xi^{l-k}\cos^k(\phi-\phi'), \quad (7)\end{aligned}$$

where $C(m,l) = \frac{(2l-1)!(2m-2l-1)!!2^{2l}}{(2l)!(2m+2l-1)!!}$ and $D(l,k) = \frac{(-1)^k l! 2^l}{(2l+1)!k!(l-k)!}$. Also, the identities $\frac{(2l-1)!!}{(2l)!} = \frac{1}{2^l l!}$ and $\left(\begin{array}{c} l \\ k \end{array} \right) = \frac{l!}{k!(l-k)!}$ have been employed. Equation (7) represents an exact expansion, in terms of the associated toroidal harmonics, valid for an arbitrary point in space not coincident with the source. If one considers using the same terminology found in antenna theory, then taking the "static limit" ($\lambda \to 0$) in Eq. (7) yields:

$$\frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} = \frac{1}{\pi\sqrt{\rho\rho'}} \sum_{m=0}^{\infty} \varepsilon_m Q^0_{m-\frac{1}{2}}(\xi) \cos[m(\phi - \phi')]. \tag{8}$$

The right hand side of Eq. (8) is the zeroth order toroidal harmonic expansion for the inverse distance function in circular cylindrical coordinates [19–25] and can be derived directly from the free-space Green's function in circular cylindrical coordinates [26, 27].

Figures 2 and 3 are plots which illustrate a comparison between Eqs. (2) and (7).

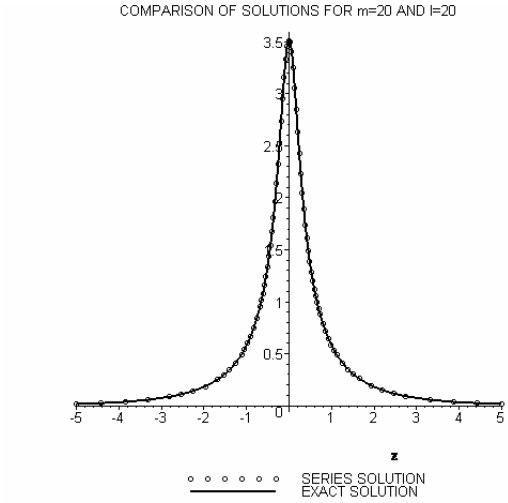


Figure 2: $\lambda = 0.5$.

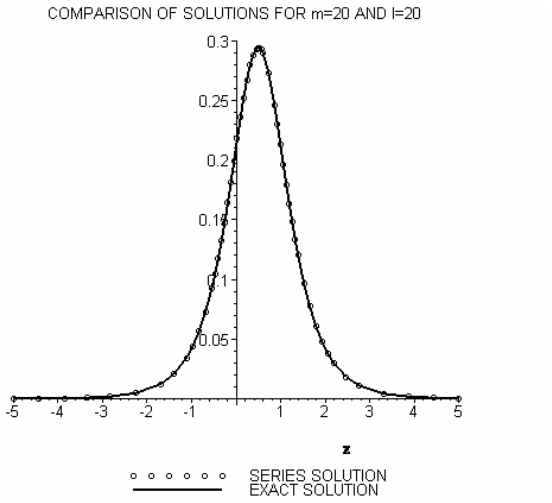


Figure 3: $\lambda = 1.5$.

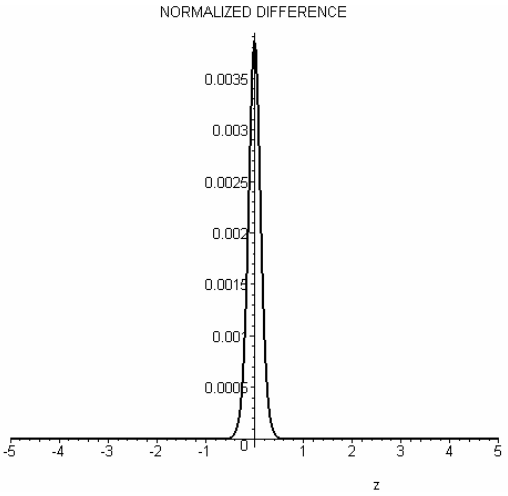


Figure 4: $N_d^{\text{Figure 2}} = \frac{K(\boldsymbol{r},\boldsymbol{r}')_{\text{exact}} - K(\boldsymbol{r},\boldsymbol{r}')_{\text{series}}}{K(\boldsymbol{r},\boldsymbol{r}')_{\text{exact}}}$.

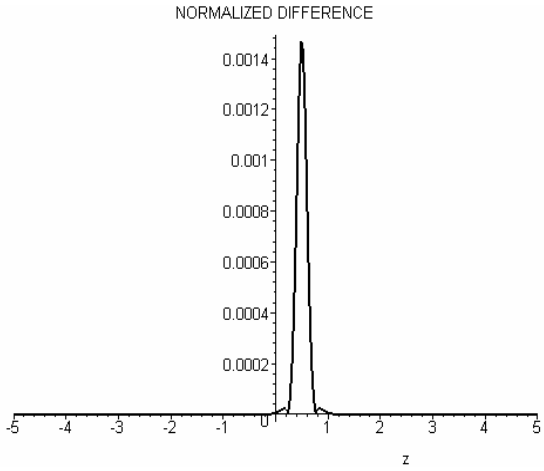


Figure 5: $N_d^{\text{Figure 3}} = \frac{K(\boldsymbol{r},\boldsymbol{r}')_{\text{exact}} - K(\boldsymbol{r},\boldsymbol{r}')_{\text{series}}}{K(\boldsymbol{r},\boldsymbol{r}')_{\text{exact}}}$.

In Figure 2, $\rho' = a = 1.0\,\text{m}$, $\phi = \phi' = 0$, and $z' = 0$. Also, an observation radius of $1.25\,\text{m}$ was chosen for the plot as well as a screening factor of $\lambda = 0.5\,\text{m}^{-1}$. The axial variation ranges from $-5 \leq z \leq 5$. In Figure 3, the same source radius, observation radius and observation angle were chosen, but the source angle is now $\phi' = \frac{\pi}{4}$. Also in Figure 3, the axial source variable is

now $z' = 0.5\,\mathrm{m}$ and the screening factor is $\lambda = 1.5\,\mathrm{m}^{-1}$. Figures 4 and 5 illustrate the normalized difference between the exact kernel and its equivalent series solution for Figures 2 and 3 respectively. The authors defined the normalized difference in the captions of Figures 4 and 5. All plots were created with Maple [28].

Both the plot of the exact kernel given by Eq. (2) and its series representation given by Eq. (7) yield equivalent results. The authors chose $m = 20$ and $l = 20$ for both plots. These values of m and l were chosen by a simple trial-and-error process. As $m, l \to$ large, $N_d \to 0$. For a more detailed numerical study on the toroidal functions of zeroth order, one should refer to Cohl and Tohline [13]. Research on the numerical properties of the *associated* toroidal functions is ongoing. However, a very useful report written by Segura and Temme [29] highlights many interesting numerical features of the toroidal functions and gives a number of new methods for computing them.

3. DISCUSSION AND CONCLUSION

A new exact series expansion is given for the Yukawa-potential kernel. This expansion is an associated toroidal harmonic expansion and is valid at an arbitrary point in space. The toroidal harmonics of the zeroth order, $Q^0_{m-\frac{1}{2}}(\xi)$, have received some recent interest [13, 19–25] where a number of applications in electrostatics, electromagnetics and gravitation are discussed. The zeroth order toroidal harmonic expansion is used to represent the equivalent series expansion of the inverse distance function, $|\boldsymbol{r} - \boldsymbol{r}'|^{-1}$, in circular cylindrical coordinates. However, the associated toroidal harmonics increase the scope of applicability by handling the inverse distance function and all its higher order derivatives using one family of functions, namely–$Q^l_{m-\frac{1}{2}}(\xi)$. In other words, the associated toroidal harmonic expansion is a true cylindrical expansion of $|\boldsymbol{r} - \boldsymbol{r}'|^{2l-1}$, where $-N \leq l \leq N$ and $N = 0, 1, 2, 3, \ldots$ etc. This allows one to tackle problems exhibiting circular cylindrical symmetry in non-Newtonian gravity, electromagnetic radiation, eddy current analysis and other formulations where the kernel function can be written as that given in Eq. (2). One may notice that if λ in Eq. (2) is replaced with $i\lambda$ then one can, using a similar expansion, develop an expansion for the radiation kernel.

APPENDIX: DERIVATION OF EQ (5)

The derivation of Eq. (5) can be derived by starting from the known Bessel expansion given by

$$e^{\frac{\alpha}{2}\left(t-\frac{1}{t}\right)} = \sum_{m=-\infty}^{\infty} t^m J_m(\alpha), \tag{A.1}$$

where $J_m(\alpha)$ is a Bessel function of the first kind. Let $\alpha = i\chi$ and $t = -ie^{i\beta}$ in Eq. (A.1). With a little algebra, one obtains

$$e^{\chi\cos(\beta)} = \sum_{m=0}^{\infty} \varepsilon_m I_m(\chi)\cos(m\beta), \tag{A.2}$$

where $I_m(\chi)$ is the modified Bessel function and where one will need the identities $J_m(i\chi) = e^{\frac{m\pi i}{2}} I_m(\chi)$ and $I_m(\chi) = I_{-m}(\chi)$ in order to derive Eq. (A.2) from Eq. (A.1). Multiply both sides of Eq. (A.2) by $e^{-\chi\xi}\chi^{n-\frac{1}{2}}$ and integrate on χ from 0 to ∞ to obtain

$$\int_0^\infty e^{-\chi[\xi-\cos(\beta)]}\chi^{n-\frac{1}{2}}d\chi = \sum_{m=0}^{\infty}\varepsilon_m\cos(m\beta)\int_0^\infty I_m(\chi)e^{-\chi\xi}\chi^{n-\frac{1}{2}}d\chi. \tag{A.3}$$

The integral in (A.3) is known [30] and, using our notation, it is given by

$$\int_0^\infty I_m(\chi)e^{-\chi\xi}\chi^{n-\frac{1}{2}}d\chi = \sqrt{\pi}\frac{(2m+2n-1)!!}{m!2^{2m+n}\xi^{m+n+\frac{1}{2}}}F\left(\frac{2m+2n+1}{4}, \frac{2m+2n+3}{4}; m+1; \frac{1}{\xi^2}\right) \text{ for } \xi > 1. \tag{A.4}$$

One can now make use of the known expression for toroidal function of the second kind [31] give by

$$Q^n_{m-\frac{1}{2}}(\xi) = \frac{\pi}{\sqrt{2}}\frac{(-1)^n(2m+2n-1)!!}{m!2^{2m+n}}\frac{\left(\xi^2-1\right)^{\frac{n}{2}}}{\xi^{m+n+\frac{1}{2}}}F\left(\frac{2m+2n+1}{4}, \frac{2m+2n+3}{4}; m+1; \frac{1}{\xi^2}\right) \text{for } \xi > 1. \tag{A.5}$$

Employing Eqs. (A.3), (A.4), and (A.5) yields the following:

$$\sum_{m=0}^{\infty} \varepsilon_m Q_{m-\frac{1}{2}}^{n}(\xi)\cos(m\beta) = (-1)^n \sqrt{\frac{\pi}{2}}\left(\xi^2-1\right)^{\frac{n}{2}} \int_0^{\infty} e^{-\chi[\xi-\cos(\beta)]}\chi^{n-\frac{1}{2}}d\chi$$
$$= \sqrt{\frac{\pi}{2}}\frac{(-1)^n\Gamma\left(n+\frac{1}{2}\right)\left(\xi^2-1\right)^{\frac{n}{2}}}{[\xi-\cos(\beta)]^{n+\frac{1}{2}}}. \tag{A.6}$$

Letting $n \to -l$, $\beta \to \phi - \phi'$ and employing the identity $Q_{m-\frac{1}{2}}^{-l}(\xi) = \frac{(2m-2l-1)!!}{(2m+2l-1)!!}Q_{m-\frac{1}{2}}^{l}(\xi)$ [31] in Eq. (A.6) yields:

$$R^{2l-1} = \frac{1}{\pi\sqrt{\rho\rho'}}2^{2l}(2l-1)!!(\rho\rho')^l(\xi^2-1)^{\frac{l}{2}}\sum_{m=0}^{\infty}\frac{(2m-2l-1)!!}{(2m+2l-1)!!}Q_{m-\frac{1}{2}}^{l}(\xi)\cos\left[m(\phi-\phi')\right] \tag{A.7}$$

which is Eq. (2). For $l = 0, \pm 1$ and ± 2, for example, one obtains a few representative expansions.

$$R^{-1} = \frac{1}{\pi\sqrt{\rho\rho'}}\sum_{m=0}^{\infty}\varepsilon_m Q_{m-\frac{1}{2}}^{0}(\xi)\cos\left[m(\phi-\phi')\right], \tag{A.8}$$

$$R^{-3} = \frac{1}{\pi(\rho\rho')^{\frac{3}{2}}\sqrt{\xi^2-1}}\sum_{m=0}^{\infty}\varepsilon_m Q_{m-\frac{1}{2}}^{1}(\xi)\cos\left[m(\phi-\phi')\right], \tag{A.9}$$

$$R^{-5} = \frac{1}{3\pi(\rho\rho')^{\frac{5}{2}}\left(\xi^2-1\right)}\sum_{m=0}^{\infty}\varepsilon_m Q_{m-\frac{1}{2}}^{2}(\xi)\cos\left[m(\phi-\phi')\right], \tag{A.10}$$

$$R^{1} = \frac{4}{\pi}\sqrt{\rho\rho'}\sqrt{\xi^2-1}\sum_{m=0}^{\infty}\varepsilon_m\frac{1}{4m^2-1}Q_{m-\frac{1}{2}}^{1}(\xi)\cos\left[m(\phi-\phi')\right], \tag{A.11}$$

$$R^{3} = \frac{48}{\pi}\left(\rho\rho'\right)^{\frac{3}{2}}\left(\xi^2-1\right)\sum_{m=0}^{\infty}\varepsilon_m\frac{1}{(4m^2-9)(2m-1)}Q_{m-\frac{1}{2}}^{2}(\xi)\cos\left[m(\phi-\phi')\right]. \tag{A.12}$$

REFERENCES

1. Liu, B. and J. Goree, "Phonons in a one-dimensional Yukawa chain: Dusty plasma experiment and model," *Phys. Rev.*, Vol. 71, 046410–12, 2005.
2. Henning, C., H. Baumgartner, A. Piel, P. Ludwig, V. Golubnichiy, M. Bonitz, and D. Block, "Ground state of a confined Yukawa plasma," *Phys. Rev.*, Vol. 74, 056403–6, 2006.
3. Khrapak, S. A., A. V. Ivlev, G. E. Morfill, S. K. Zhdanov, and H. M. Thomas, "Scattering in the attractive Yukawa potential: Application to the ion-drag force in complex plasmas," *IEEE Trans. on Plasma Sci.*, Vol. 32, April 2004.
4. Harreis, H. M., C. N. Likos, and H. Lowen, "Azimuthal frustration and bundling in columnar DNA aggregates," *Biophys. Jour.*, Vol. 84, 3607–1378, June 2003.
5. Rowan, D. G., J. P. Hansen, and E. Trizac, "Screened electrostatic interactions between clay platelets," *Mol. Phys.*, Vol. 98, No. 18, 1369–1378, 2000.
6. Wagner, K., E. Keyes, T. Kephart, and G. Edwards, "Analytical Debye-huckel model for electrostatic potentials around dissolved DNA," *Biophys. Jour.*, Vol. 73, July 21–30, 1997.
7. Gillespie, A., R. Paul, and S. J. Paddison, "Electrical potential distribution in nanopores," *Nanotech*, Vol. 3, 497–499, 2003.
8. Bordag, M., B. Geyer, G. L. Klimchitskaya, and V. M. Mostepanenko, "New constraints for non-Newtonian gravity in the nanometer range from the improved precision measurement of the Casimir force," *Phys. Rev. D*, Vol. 62, 011701(R)-5, June 2000.
9. Smullin, S. J., A. A. Geraci, D. M. Weld, and A. Kapitulnik, "Testing gravity at short distances," *SLAC Summer Institute on Particle Physics (SSI04)*, Aug. 2-13, 2004.
10. Quinn, T. J., C. C. Speake, S. J. Richman, R. S. Davis, and A. Picard, "A new determination of G using two methods," *Phys. Rev. Let.*, Vol. 87, No. 11, 111101-1-111101-4, Sept. 2001.
11. Shapiro, I. L., "Asymptotic behavior of effective Yukawa coupling constants in quantum R2 gravity with matter," *Class. Quantum Grav.*, Vol. 6, 1197–1201, 1989.

12. Stacey, F. D., G. J. Tuck, and G. I. Moore, "Quantum gravity: Observational constraints on a pair of Yukawa terms," *Phys. Rev. D (Particles and Fields)*, Vol. 36, Issue 8, 2374–238015, Oct. 1987.
13. Cohl, H. S. and J. E. Tohline, "A compact cylindrical Green's function expansion for the solution of potential problems," *Astrophys. Jour.*, Vol. 527, 86–101, Dec. 1999.
14. MacRobert, T. M., *Spherical Harmonics*, 209–211, Pergamon Press, Oxford, 1967.
15. Lebedev, N. N., *Special Functions and Their Applications*, 184–199, Prentice-Hall, New Jersey, 1965.
16. Snow, C., "Hypergeometric and legendre functions with applications to integral equations of potential theory," *Natl. Bur. Stand. App. Math. Ser.*, Vol. 19, 56–82, 228–238, Washington D. C., 1952.
17. Abramowitz, M. and I. Stegun, *Handbook of Mathematical Functions*, 336-337, Dover Publications Inc., New York, 1972.
18. Morse, P. M. and H. Feshbach, *Methods of Theoretical Physics*, Parts I and II, 744 and 1301–1304, McGraw Hill, New York, 1953.
19. Cohl, H. S., "Portent of Heine's reciprocal square root identity," *3D Stellar Evolution, ASP Conference Proceedings*, Vol. 293, University of California Davis, Livermore, California, USA, July 22–26, 2002.
20. Cohl, H. S., J. E. Tohline, and A. R. P. Rau, "Developments in determining the gravitational potential using toroidal functions," *Astron. Nachr.*, 321, 363–372, 2000.
21. Cohl, H. S., A. R. P. Rau, J. E. Tohline, D. A. Browne, J. E. Cazes, and E. I. Barnes, "Useful alternative to the multipole expansion of $1/r$ potentials," *Phys. Rev. A*, Vol. 64, 052509-1-052509-5, Oct. 2001.
22. Selvaggi, J., S. Salon, O. Kwon, and M. V. K. Chari, "Calculating the External magnetic field from permanent magnets in permanent-magnet motors-an alternative method," *IEEE Trans. Mag.*, Vol. 40, No. 5, 3278–3285, Sept. 2004.
23. Selvaggi, J., S. Salon, and M. V. K. Chari, "An application of toroidal functions in electrostatics," *Am. J. Phys.*, Vol. 75, No. 8, 724–727, Aug. 2007.
24. Selvaggi, J., S. Salon, O. Kwon, and M. V. K. Chari, "Computation of the external magnetic field, near-field or far-field, from a circular cylindrical magnetic source," *IEEE Trans. Mag.*, Vol. 43, No. 4, 1153–1156, April 2007.
25. Selvaggi, J., S. Salon, O. Kwon and M. V. K. Chari, "Computation of the three-dimensional magnetic field from solid permanent-magnet bipolar cylinders by employing toroidal harmonics," *IEEE Trans. Mag.*, Vol. 43, No. 10, 3833–3839, Oct. 2007.
26. Jackson, J. D., *Classical Electrodynamics-3rd ed.*, 125–127, John Wiley & Sons, New York, 1999.
27. Smythe, W. R., 187–189 and 204, Static and Dynamic Electricity-3rd ed., McGraw-Hill, New York, 1968.
28. Maple, *Maple*, Version 10.06, Waterloo Maple Inc., Waterloo, Ontario, Canada, 2004.
29. Gil, A., J. Segura, and N. M. Temme, "Computing toroidal functions for wide ranges of the parameters," Report MAS-R0014, CWI, National Research Institute for Mathematics and Computer Science, P.O. Box 94079 GB Amsterdam, Netherland., May 31, 2000.
30. Grobner, W. and N. Hofreiter, *Integraltafel Zweiter Teil Bestimmte Integrale*, 199, Springer-Verlag, Wein, integral (6a), 1966.
31. Wang, Z. X. and D. R. Guo, *Special Functions*, 254–255, World Scientific, Singapore, 1989.

GL Acoustic Mechanical and EM Uniform Coupled Modeling and Inversion

Ganquan Xie, Jianhua Li, and Feng Xie
GL Geophysical Laboratory, USA

Abstract— In this paper, an Acoustic and Mechanical (ME) and Electromagnetic (EM) field uniform equation system and coupled equation are presented. We propose GL Acoustic and ME and EM uniform coupled field modeling and inversion (GLAMEEM). Their domain is a finite inhomogeneous medium which is embedded into the infinite homogeneous domain. The domain is divided to a set of sub domains. The acoustic, displace velocity, stress, magnetic and electric field are presented as explicit recursive sum of the integrals in the sub domains. The analytical Acoustic and ME and EM field and the Green's tensor in the background domain are called initial Global field. The Global field is updated by Local scattering field successively sub by sub domain. Once all sub domains are scattered, the GL Acoustic and ME and EM field solution are obtained. Unlike FEM and FD, GL method directly assembles inverse matrix for solution and does not need to solve lager matrix equation. The GL modeling is great faster and more accurate than the FEM and FD method. Moreover, it does not need artificial boundary for truncating infinite domain. A GL nonlinear inversion is presented as an explicit recursive form for acoustic, mechanical electric or mechanical magnetic materials. Also the GL Acoustic and EM and ME uniform modeling GLAMEEM modeling and inversion can be used in the geophysical exploration and earthquake exploration and fabrication of the non icing high voltage line.

1. INTRODUCTION

The Electromagnetic (EM) equation was derived by Faraday and Maxwell. In particularly, Maxwell introduced the displaced current term and completed the EM Maxwell equation. The Maxwell equation was derived from experiments, instead of Newton's principle. The elastic mechanical (ME) equation was derived partly from Newton mechanical principle, partly from experiment and Hook's constitutive law. The mechanics and electromagnetics are developed separately in sciences, engineering and education. The ME field equation and EM equation were developed independently in history as mention above. However, they have closed relationship. Maxwell had working in the elastic mechanics before he developed Maxwell equation. In this paper, we proposal new uniform form to present acoustic, ME and EM equation and their coupled equation. A GL modeling in frequency domain and time domain for solving acoustic, EM, elastic, flow, and quantum field have been proposed by authors [1–3] since 2002. We also proposed GL inversion in frequency domain and time domain to recover the EM and other above wave field parameters in various applications [4, 5]. These GL methods in frequency domain are based on EM, elastic, acoustic frequency domain integral equations, respectively [1–3, 6–9]. We propose GL Acoustic, Mechanical and Electromagnetic field coupled modeling and inversion in time and space domain in this paper. We derive acoustic ME and EM integral equation in time domain. Based on these time domain uniform field integral equations, we construct GL ME and EM coupled modeling (GLAMEEM) in time space domain. The GLEMEMM method does not need to solve any large matrix, only 16×16 matrices are needed to solve. Moreover, the artificial boundary and absorption condition are unnecessary in the GLAMEEM GL modeling method.

The presenting plan of this paper is as follows. The introduction has been presented in Section 1. We derive 3D ACOUSTIC and ME and EM field uniform equation in time domain in Section 2. In Section 3, The 3D Acoustic and ME and EM field time domain integral equation. In Section 4, we propose 3D GLAMEEM coupled modeling. In Section 5, we propose 3D GL GLAMEEM coupled inversion. The advantages of the GLAMEEM time domain modeling and inversion are described in Section 6. We describe several applications and simulation in Section 7. In Section 8, we make conclusions.

2. 3D EM AND ME AND ACOUSTIC FIELD UNIFORM EQUATION IN TIME DOMAIN

We propose a uniform field equation for the acoustic, elastic, and electromagnetic equation as follows.

$$A+\begin{bmatrix} \dfrac{\partial}{\partial t} & 0 \\ 0 & \dfrac{\partial}{\partial t} \end{bmatrix} B \begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} 0 & L \\ L^T & 0 \end{bmatrix} \begin{bmatrix} U \\ V \end{bmatrix} + \begin{bmatrix} S_1 \\ S_2 \end{bmatrix}, \tag{1}$$

where A and B are coefficient matrix with corresponding dimension, L is space differential matrix, S_1 and S_2 are source term.

2.1. The Acoustic Wave Equation by Uniform Equation (1)

(2.1.1) Let

$$A=0, \quad B=\begin{bmatrix} \dfrac{1}{c^2} & \\ & I \end{bmatrix}, \quad L=\nabla\cdot \tag{2}$$

B is 4×4 matrix, the right low corner of B is 3×3 unit matrix, c is scale wave velocity, $L=\nabla\cdot$ and $L^T=\nabla$, by substituting A, B, and L into the uniform Equation (1), then the acoustic wave equation

$$\frac{1}{c^2}\frac{\partial^2 U}{\partial t^2}=\nabla^2 U+S, \tag{3}$$

is obtained.

(2.1.2) Let

$$A=0, \quad B=\begin{bmatrix} 1 & 0 \\ 0 & \dfrac{1}{K}I \end{bmatrix}, \quad L=\nabla\cdot \tag{4}$$

B is 4×4 matrix, the right low corner of B is 3×3 matrix $\frac{1}{K}I$, K is square of scale wave velocity, $L=\nabla\cdot$ and $L^T=\nabla$, by substituting A, B, and L into the uniform Equation (1), then the acoustic wave equation

$$\frac{\partial^2 U}{\partial t^2}=\nabla\cdot(K\nabla U)+S, \tag{5}$$

is obtained.

2.2. The Maxwell Electromagnetic Wave Equation by Uniform Equation (1)

Upon substituting

$$A=\begin{bmatrix} \sigma I & \\ & 0 \end{bmatrix}, \quad B=\begin{bmatrix} \varepsilon I & \\ & \mu I \end{bmatrix}, \quad L=\nabla\times, \quad U=E, \quad V=H \tag{6}$$

into the uniform Equation (1), we obtain the Maxwell EM equation in the space time domain.

$$\begin{bmatrix} \sigma+\varepsilon\dfrac{\partial}{\partial t} & 0 \\ 0 & \mu\dfrac{\partial}{\partial t} \end{bmatrix} \begin{bmatrix} E \\ H \end{bmatrix} = \begin{bmatrix} 0 & \nabla\times \\ -\nabla\times & 0 \end{bmatrix} \begin{bmatrix} E \\ H \end{bmatrix} - \begin{bmatrix} J \\ M \end{bmatrix}. \tag{7}$$

where the I in the formula (6) is 3×3 unit matrix, also $L^T=-\nabla\times$, $U=E$ is the electric field, $V=H$ is magnetic field in the (6).

2.3. The Elastic Mechanical Wave Equation by Uniform Equation (1)

In the uniform Equation (1), let $U=\begin{bmatrix} U_x & U_y & U_z \end{bmatrix}^T$ be displacement velocity, the stress be $V=\begin{bmatrix} \sigma_{xx} & \sigma_{yy} & \sigma_{zz} & \sigma_{xy} & \sigma_{yz} & \sigma_{xz} \end{bmatrix}$, $A=0$, B is 9×9 matrix, its left up corner is 3×3 ρI, I is unit matrix, and its right corner is 6×6 inverse elastic matrix by Hook's law,

$$L=\begin{bmatrix} \dfrac{\partial}{\partial x} & & & \dfrac{\partial}{\partial y} & & \dfrac{\partial}{\partial z} \\ & \dfrac{\partial}{\partial y} & & \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial z} & \\ & & \dfrac{\partial}{\partial z} & & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial x} \end{bmatrix}. \tag{8}$$

Up substituting above A, B, L, and S into the uniform Equation (1), we obtain 3D elastic mechanical wave equation

$$\sigma_{ij,\,j} - \rho \frac{\partial^2 U}{\partial t^2} = S, \tag{9}$$

The uniform Equation (1) governs the acoustic, elastic, and EM etc equation. The uniform Equation (1) is important to study the relationship between the acoustic, EM, and elastic field and their couple properties.

3. 3D ACOUSTIC, EM AND ELASTIC FIELD INTEGRAL EQUATION IN TIME DOMAIN

We have proposed the 3D EM integral equation in frequency domain in PIERS 2005 in Hangzhou of China and PIERS 2006 in Cambridge [1]. By using the field uniform Equation (1) and its Green field uniform equation, we propose 3D acoustic, EM and ME integral equation in a uniform form in this section.

3.1. An Acoustic Wave Field Integral Equation in Time Domain

In this section, we propose the acoustic wave field integral equation in time domain as follows,

$$u\left(r,t\right) = u_b\left(r,t\right) - \int\limits_{\Omega} \left(\frac{1}{c_b^2\left(r\right)} - \frac{1}{c^2\left(r\right)}\right) G\left(r',r,t\right) *_t \frac{\partial^2}{\partial t^2} u_b\left(r',t\right) dr', \tag{10}$$

where r is space variable $r = (x, y, z)$, $r \in R^3$, $c\left(r\right)$ is the coefficient, so called wave velocity, which is inhomogeneous media function in a finite bounded domain $\Omega \subset R^3$, $c\left(r\right) = c_b$ in $r \in R^3 \backslash \Omega$, c_b is a constant, t is time, $u\left(r,t\right)$ is the unknown wave field, If the $c\left(r\right) = c_b$ in whole space R^3, the solution $u_b(r,t)$ of the wave equation is called incident wave in the background media, $u_b\left(r,t\right) = \delta\left(t - \frac{|r-r_s|}{c_b}\right) / \left|r - r_s\right| / \left(4\pi\right)$, The solution of the wave equation satisfies the following radiation condition,

$$\lim_{r\to\infty} r\left(\frac{\partial u}{\partial n} - \frac{\partial}{c\partial t} u\right) = 0,$$

and $ru(r,t)$ is bounded in infinite, $G\left(r',r,t\right)$ is Green's function of the wave equation, $*_t$ is convolution with respect to time t.

3.2. An EM Integral Equation in Time Domain

In this section, we propose the EM integral equation in time domain as follows

$$\left[\begin{array}{c} E\left(r,t\right) \\ H\left(r,t\right) \end{array}\right] = \left[\begin{array}{c} E_b\left(r,t\right) \\ H_b\left(r,t\right) \end{array}\right] - \int\limits_{\Omega} G_{E,\,H}^{J,\,M}\left(r',r,t\right) *_t \left[D_b - D\right] \left[\begin{array}{c} E_b\left(r',t\right) \\ H_b\left(r',t\right) \end{array}\right] dr', \tag{11}$$

where

$$G_{E,\,H}^{J,\,M}\left(r',r,t\right) = \left[\begin{array}{cc} E^J\left(r',r,t\right) & H^J\left(r',r,t\right) \\ E^M\left(r',r,t\right) & H^M\left(r',r,t\right) \end{array}\right], \tag{12}$$

$$D_b = \left[\begin{array}{cc} \sigma_b + \varepsilon_b \dfrac{\partial}{\partial t} & \\ & -\mu_b \dfrac{\partial}{\partial t} \end{array}\right], D = \left[\begin{array}{cc} \sigma + \varepsilon \dfrac{\partial}{\partial t} & \\ & -\mu \dfrac{\partial}{\partial t} \end{array}\right], \tag{13}$$

$E\left(r,t\right)$ is the electric field, $H\left(r,t\right)$ is the magnetic field, $E^J\left(r',r,t\right)$ is the electric Green's field function excited by the impulse current, $H^J\left(r',r,t\right)$ is the magnetic Green's field function excited by the impulse current, $E^M\left(r',r,t\right)$ is the electric Green's field function exited by the impulse magnetic moment, $H^M\left(r',r,t\right)$ is the magnetic Green's field function excited by the impulse magnetic moment, $E^J\left(r',r,t\right)$, $H^J\left(r',r,t\right)$, $E^M\left(r',r,t\right)$, and $H^M\left(r',r,t\right)$ are EM Green's function field in the electric conductivity σ, dielectric ε, and magnetic permeability μ media, $E_b\left(r,t\right)$ and $H_b\left(r,t\right)$ are EM wave field in the background σ_b, ε_b, and μ_b media, $*_t$ is convolution with respect to time t.

3.3. An Elastic Mechanical Integral Equation in Time Domain

A 3D mechanical field integral equation is proposed in this section

$$\begin{bmatrix} \sigma(r,t) \\ V(r,t) \end{bmatrix} = \begin{bmatrix} \sigma_b(r,t) \\ V_b(r,t) \end{bmatrix} - \int_\Omega G^{s,d}_{\sigma,V}(r',r,t) *_t [D_b - D] \begin{bmatrix} \sigma_b(r',t) \\ V_b(r',t) \end{bmatrix} dr', \tag{14}$$

$$D = \begin{bmatrix} [E,\nu]\dfrac{\partial}{\partial t} & \\ & \rho\dfrac{\partial}{\partial t} \end{bmatrix}, D_b = \begin{bmatrix} [E,\nu]_b\dfrac{\partial}{\partial t} & \\ & \rho_b\dfrac{\partial}{\partial t} \end{bmatrix}, \tag{15}$$

$\sigma(r,t)$, is strees field, $V(r,t)$ is the displacement velocity field, $G^{s,d}_{\sigma,V}(r',r,t)$ is the displacement velocity and stress Green's tensor, its component $G^s_\sigma(r',r,t)$ is stress Green's field function excited by the stress impulse, $G^s_V(r',r,t)$ is the displace velocity Green's field function excited by the stress impulse, $G^d_\sigma(r',r,t)$ is the stress Green's field function exited by the velocity impulse, $G^d_V(r',r,t)$ is the displace velocity Green's field function excited by the velocity impulse. $G^{s,d}_{\sigma,V,b}(r',r,t)$ is background Green tensor.

4. GLAMEEM COUPLED MODELING IN TIME DOMAIN

In his section, we propose the new GLAMEEM coupled modeling in time domain as follows.

(4.1) The domain Ω is divided into set of the N sub domains $\{\Omega_k\}$, such t hat $\Omega = \bigcup_{k=1}^{N} \Omega_k$, where the division can be mesh or meshless.

(4.2) Suppose that, when $k = 0$, $[\ U\quad \sigma\quad V\quad E\quad V\ ]_0(r,t) = [\ U\quad \sigma\quad V\quad E\quad V\ ]_b(r,t)$, is the global analytical time domain acoustic, elastic mechanical, and EM wave field and $G^{q,s,d,J,M}_{u,\sigma,V,E,H,0}(r',r,t)$ $= G^{q,s,d,J,M}_{u,\sigma,V,E,H,b}(r',r,t)$ is the global acoustic, mechanical and EM analytic wave Green's tensor function in the background medium with constant medium, or half space medium, or waveguide medium, or multiple layered medium. By induction, suppose that the time domain acoustic, ME, and EM field $[\ U\quad \sigma\quad V\quad E\quad V\ ]_{k-1}(r,t)$, and the acoustic, mechanical and EM wave Green's tensor function $G^{q,s,d,J,M}_{u,\sigma,V,E,H,k-1}(r',r,t)$ have been calculated in the $(k-1)$th step.

(4.3) In Ω_k, we solve the acoustic, mechanical and EM wave Green's tensor function integral equation system in time domain which is based on the Equations (10)–(15). By dual operation, the integral equation are reduced into 16×16 matrix equations. By solving the 16×16 equations, We obtain time domain the acoustic, mechanical and EM wave coupled Green tensor field $G^{q,s,d,J,M}_{u,\sigma,V,E,H,k}(r',r,t)$.

(4.4) We improve the global acoustic, elastic mechanical, and EM wave coupled field $[\ U\quad \sigma\quad V\quad E\quad V\ ]_k(r,t)$ by the local time domain coupled scattering filed in Ω_k,

$$\begin{bmatrix} U_k(r,t) \\ \sigma_k(r,t) \\ V_k(r,t) \\ E_k(r,t) \\ H_k(r,t) \end{bmatrix} = \begin{bmatrix} U_{k-1}(r,t) \\ \sigma_{k-1}(r,t) \\ V_{k-1}(r,t) \\ E_{k-1}(r,t) \\ H_{k-1}(r,t) \end{bmatrix} - \int_\Omega G^{a,s,d,J,M}_{U,\sigma,V,E,H,k}(r',r,t) *_t [D_b - D] \begin{bmatrix} U_{k-1}(r,t) \\ \sigma_{k-1}(r,t) \\ V_{k-1}(r,t) \\ E_{k-1}(r,t) \\ H_{k-1}(r,t) \end{bmatrix} dr', \tag{16}$$

$k = 1, 2, \ldots N$, successively. The $[\ U\quad \sigma\quad V\quad E\quad V\ ]_N(r,t)$ is GL acoustic, elastic mechanical and EM wave coupled field modeling solution.

5. 3D GLAMEEM COUPLED INVERSION

We present a 3D GLAMEEM acoustic, elastic, and EM parameters coupled inversion as follows,

$$\delta \begin{bmatrix} U_k(r,t) \\ \sigma_k(r,t) \\ V_k(r,t) \\ E_k(r,t) \\ H_k(r,t) \end{bmatrix} = \int_\Omega G^{a,s,d,J,M}_{U,\sigma,V,E,H,k}(r',r,t) *_t \delta[D] \begin{bmatrix} U_{k-1}(r,t) \\ \sigma_{k-1}(r,t) \\ V_{k-1}(r,t) \\ E_{k-1}(r,t) \\ H_{k-1}(r,t) \end{bmatrix} dr', \tag{17}$$

6. THE ADVANTAGES OF THE GL GLAMEEM COUPLED MODELING AND INVERSION IN TIME DOMAIN

6.1. GL Method Assembles to Build Inverse Matrix and Does Not Need to Solve Any Big Matrix

The Finite Element Method (FEM) and implicit Finite Difference (FD) method is to build matrix and then to solve matrix, For large scale scientific and engineering problem, the large FEM matrix solving is massive storage and run time cost. In contrast of FEM and FD method, our GL time domain and frequency domain modeling assemble to build inverse matrix and does not need to save big matrix. Therefore, the GL modeling overcome the difficulty of FEM and FD method for solving big matrix. The GL method is more fast than FEM and implicit FD method. The Green's function is inverse operator of the differential equation. Its discretization is more accurate than the inversion of FEM or FD matrix. The GL modeling for EM, elastic, and acoustic is more accurate than FEM and FD method.

6.2. The Artificial Boundary to Truncate Infinite Domain Is Unnecessary in the GL Modeling

When the FEM and FD etc numerical method is used to solve PDE in the infinite domain, the artificial boundary and on which the absorption boundary condition are necessary to truncate the infinite domain. From Sommerfeld's radiation condition, the research for constructing the absorption condition on the artificial boundary has been 80 years. There are many papers and books to report the construction of the absorption condition on the artificial boundary. In interntional mathematical congress, there are 45 minute speaker for the absorption boundary condition. However, there is refelection in all absorption condition from the corner of the artificial boundary. The error refelection degrade the accuracy of the modeling solution, in particular, the error reflections of the absorption condition from the artificial boundary prapagate into the internal domain in inverse iteration that will be a unremoveable noise to damage the inversion. In contrast of all effort for studying absorption condition, the GL method removed the artificial bounday for infinite domain. Because the Sommerfeld's radiation condition in the far field in the infinite domain, we derive the EM, elastic, and acoustic field equation are equivalent to their corresponding integral equation, respectively, the volumical integral equation does not include any boundary condition term, therefore, The GL method based on these integral equation does not need any artificial boundary and absorption condition for the infinite domain. The GL method overcome the FEM and FD's difficulty of the reflction errors from their artificial boundary and absorption condition for the infinite domain.

6.3. The Analytical Method and Numerical Method Is Consistently Combining in the GL Method

The Analytical and numerical methods are two main methods for solving the differential equation. They have been developing separately in the history. The analytical method is suitable for the whole space, half space, multiple layered media, spherical, cylindrical and regular geometry domain, but it is not available for solving differential equation in the irregular and complex geometry and inhomogeneous media domain. The FEM and FD etc numerical method is available for solving differential equation in the irregular and complex geometry and inhomogeneous media, but it is not suitable for the whole space, half space, multiple layered media infinite domain. In the numerical method, the artificial boundary and the absorption condition on it is necessary that produces reflection errors to degrade the accuracy of forward solution and damage inversion. The GL method combines the analytical method and numerical method consistently together. The GL method is either available for whole space, half space, multiple layered media, spherical, cylindrical and regular geometry domain, also is suitable to solve the differential equation in the irregular and complex geometry and inhomogeneous media domain without artificial boundary.

6.4. The GL Method Is Natural Parallel Algorithm and Multiple Scale Parallel Algorithm in Periodic Space Time Domain

The GL modeling method is natural parallel algorithm. The GL EM acoustic, and elastic wave field dispersion modeling are useful to construct global periodic to local periodic, global periodic to local aperiodic, global aperiodic to local periodic multiple scale parallel algorithms.

6.5. The GL Method Overcome the Coordinate Singularity Difficulty in FEM and FD Method

There exist the coordinate singularity in the pole of the differential equation in cylindrical coordinate pole, south pole and north pole of the spherical coordinate system. The cylindrical pole singularity

and spherical poles' singularity are difficulties of FEM and FD method. The GL method overcomes the coordinate singularity difficulty.

6.6. The Accurate GL Method vs. Inaccurate Born Approximation

The Born approximation is sample iteration but it is inaccurate method and only suitable for low frequency and low contrast. The GL method can available for all frequency from low to high and high contrast medium anomalous.

6.7. There is No Complexity for High Order Differential Term in the GL Method, If the Inhomogeneous Medium is Located in the Low Order Term

6.8. GL Method Displays Methodology from Inversion to Forward Modeling

6.9. The Solution of GL Time Domain Converges to Exact Solution and Has Super Convergence

We prove that the solution of GL time domain modeling converges to exact wave field. Provided the middle point and trapezoid integral is used in the integrals in GL method, the convergent rate is $O(h^2)$. If the high accurate 20 nodes curve isoparameter element [19] and Gaussion integral is used in the integrals in GL method, then GL method has super convergent rate $O(h^4)$ [17, 18]. In paper [19], the author use a novel method to construct the 20 nodes curve isoparameter element and develop 3D elastic FEM and its software first in China, also author used the 3D isoparameter FEM software to calculate all dams and underground structures displacement and stress of China. In particular, author Xie and Li discovered superconvergence of the isoparameter element first of the world that the discover has been recognized by Brandts, J. and M. Krizek's paper "History and Future of Superconvergence in Three Dimensional Finite Element Methods" [20] and Michael Krizek' paper "Superconvergence Phenomena on Three Dimenional Meshes" [21]. By using the 20 nodes curve isoparameter element [19] and Gaussion integral, GL method products pointwise superconvergence $O(h^4)$.

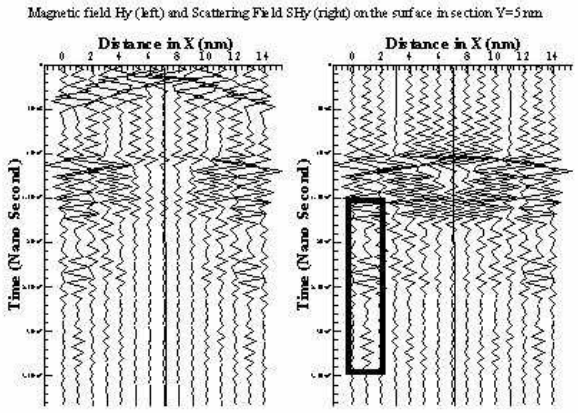


Figure 1: Hy in the transverse section $Y = 5\,\text{nm}$ and on surface, Hy is in left, scattering field SHy is in right. A quanta is shown in rectangle frame.

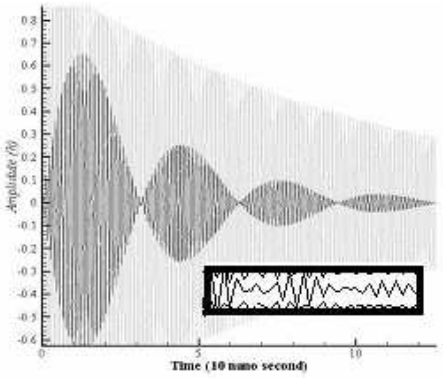


Figure 2: GLEMQUAN quanta in the rear of the scattering magnetic field Hy which is shown in black rectangle frame in Figure 1.

7. APPLICATIONS AND SIMULATIONS

The 3D and 2D GL modeling for EM field dispersion in periodic space time domain is useful to simulate EM dispersion in the macro and nanometer materials, optical materials, and dispersion engineering of photonic crystals. In paper [3], we present coupled GL EM and quantum mechanical modeling for studying nanometer material properties. The nanometer material is mixed periodic photonic crystal and other aperiodic structure. The GL modeling to simulate coupled EM dispersion in periodic domain and for EM field in aperiodic domain and obtained very reasonable imaging.

The $20\,\text{nm} \times 20\,\text{nm} \times 20\,\text{nm}$ Bragg crystal optical sensor material is embedded into the metal bulk cheap with $5\,\mu\text{m} \times 5\,\mu\text{m} \times 5\,\mu\text{m}$. The X ray propagate through the periodic and aperiodic composite material. The optical material is divided 4096 unit cell lattices. The outside gasket metal bulk is divided into 1024 domains. The initial quantum field is ψ_0 which satisfies Schrödinger equation with Hamiltonian H_0. We use the GLEMQUAN EM and Quantum field global aperiodic

to local periodic modeling to calculate the EM field and quantum wave field function. The Hy in the transverse section $Y = 5\,\text{nm}$ and on surface of the bulk is shown in Figure 1, Hy is shown in left graphic, scattering field SHy is shown in right graphic which includes dispersion wave packets. A quanta dispersion wave package is obviously shown in rectangle frame in rear of the scattering wave. The quanta dispersion wave package is an EM-quantum coupled effect which is match to the X ray observation. The GLEMQUAN quanta dispersion wave package is shown in the Figure 2, it is different from the rear part of macro EM field in which the scattering wave vanished to zero from peak. From the GLEMQUAN EM field and quantum wave field, the quantum fiber strain is calculated. The deformation of the nanometer material crystal due to the quantum fiber strain is shown in the Figure 3. The electric microscope image of the deformation of the nanometer material is shown in Figure 4. The similarity between images in Figure 3 and Figure 4 shows that the GLEMQUAN modeling can obtain high resolution image of the nanometer fiber deformation. Figure 5 and Figure 6 show other crystal deformation. The simulations show that the GLEMQUAN modeling is fast and accurate [3].

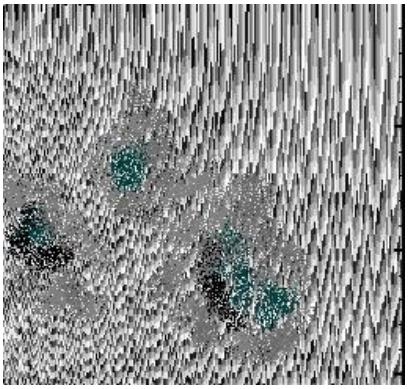


Figure 3: The deformation of the nanometer material crystal due to the quantum fiber strain.

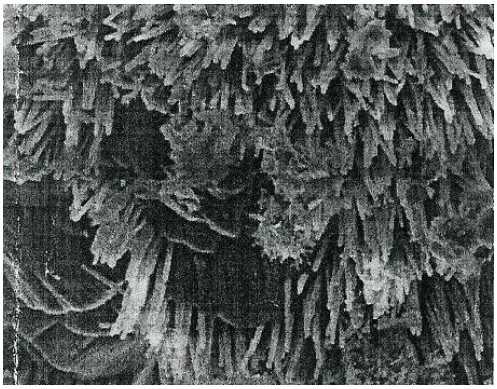


Figure 4: The scan electric microscope SEM image of the deformation of the nanometer material crystal.

8. CONCLUSION

By checking the research references, it is appear that our uniform Equation (1) is lack in exist research references. The uniform equation not only governs the acoustic, mechanical, and EM equations, in particular, but also it is useful to make their physical coupling, and predict new physical field. We try to use (1) to couple gravity and elastic and EM wave in earth and air space. That is useful to simulate Wenchuan earthquake [22] and icing around the high voltage line [23], EM stirring, mevhanical electric, piezoelectric material property and flow and EM in earth center. Simulations show and validate that the GL acoustic, and elastic and EM aperiodic and periodic space time domain modeling are fast, accurate, stable high resolution and have the advantages over FEM and FDTD method [24–27]. These advantages are described in the context of the paper. GL time domain modeling for EM, acoustic and elastic wave field with dispersion have wide applications in nanometer materials, optical materials, dispersion engineering of photonic crystals, porous material, nondestructive testing, desctructive mechanics, and geophysical exploration. The GL EM, acoustic, elastic quantum periodic lattice and aperiodic space time domain modeling algorithms and software are developed and patented by authors in GL Geophysical laboratory and all rights are reserved by authors in GLGEO.

REFERENCES

1. Xie, G., F. Xie, L. Xie, and J. Li, "New GL method and its advantages for resolving historical diffculties," *Progress In Electromagnetics Research*, PIER 63, 141–152, 2006.
2. Xie, G., J. Li, L. Xie, F. Xie, and J. Li, "The 3D GL EM-flow-heat-stress coupled modeling," *PIERS Online*, Vol. 3, No. 4, 411–417, 2007.
3. Xie, G., J. Li, F. Xie, and L. Xie, "3D GL EM and quantum mechanical coupled modeling for the nanometer materials," *PIERS Online*, Vol. 3, No. 4, 418–422, 2007.
4. Xie, G., J. Li, L. Xie, and F. Xie, "GL metro carlo EM inversion," *Journal of Electromagnetic Waves and Applications*, Vol. 20, No. 14, 1991–2000, 2006.

5. Li, J., G. Xie, L. Xie, and F. Xie, "A 3D GL EM modeling and inversion for forest exploration and felling," *PIERS Online*, Vol. 3, No. 4, 402–410, 2007.
6. Xie, G. and Y. M. Chen, "A modified pulse spectrum technique for solving inverse problem of two-dimensional elastic wave equation," *Appl. Numer. Math.*, Vol. 1, No. 3, 217–237, 1985.
7. Xie, G. and J. Li, "New parallel SGILD modeling and inversion," *Physics D*, Vol. 133, 477–487, 1999.
8. Xie, G., J. Li, E. Majer, D. Zuo, and M. Oristaglio, "3-D electromagnetic modeling and nonlinear inversion," *Geophysics*, Vol. 65, No. 3, 804–822, 2000.
9. Xie, G. and J. Li, "New parallel GILD-SOR modeling and inversion for E-O-A strategic simulation," *IMACS Series Book in Computational and Applied Math.*, Vol. 5, 123–138, 1999.
10. Xie, G., J. Li, L. Xie, and F. Xie, "GL EM mechanical and acoustic field time domain modeling for materials and exploration with dispersion," to be appear to *Progress In Electromagnetics Research Symposium Abstract*, 2008.
11. Xie, G., J. Li, and C.-C. Lin, "New SGILD EM modeling and inversion in geophysics and nano-physics," *Three Dimensional Electromagnetics*, Vol. 2, 193–213, (SCI), 2002.
12. Xie, G., C.-C. Lin, and J. Li, "GILD EM modeling in geophysics and nanometer materials using a magnetic field integral equation," *Three Dimensional Elctromagnetics III*, Vol. 3, 11–12, 2003.
13. Pustai, D. M., C. Chen, A. Sharkawy, S. Shi, J. Murakowski, and D. W. Prather, "Dispersion engineering of photonic crystal devices," *Mat. Res. Soc. Symp. Proc.*, Vol. 797, 6.8.1–6.8.6, 2004.
14. Wei, S., Y. Zhao, Y. Bao, S. Li, Z. Zhang, and T. Xu, "Numerical simulation and analysis on mode property of photonic crystal fiber with high birefringence by fast multipole method," *PIERS Online*, Vol. 3, No. 6, 386–841, 2007.
15. Lin, C.-C., G. Xie, J. Li, and B.-J. Lin, "Deformation analysis for materials using gild mechanical modeling," (Invited paper), *The Chinese Journal of Mechanics-series A*, Vol. 19, No. l, 207–221, (A special issue dedicated to Prof. T. C. T. Ting in celebration of his seventieth birthday), (SCI & EI), 2003.
16. Xie, G., C.-C. Lin, and J. Li, "New 3D gild mechanical modeling and inversion," *Mathematica Applicata*, Vol. 16, No. 2, 149–156, 2003.
17. Li, J., G. Xie, and L. Xie, "3D and 2D GL modeling for EM mechanical and heat field with dispersion in mixed aperiodic and periodic space time domain for material sciences and exploration," GL Geophysical Laboratory Technological Patent Report, GL2007-1, 2007,
18. Xie, G., J. Li, and F. Xie, "GL time domain modeling based integral equation for EM acoustic and elastic wave field with dispersion in crystal and porous material," GL Geophysical Laboratory Technological Patent Report, GL2007-2, 2007.
19. Xie, G., "Three dimensional finite element method in the elastic problem," *Mathematical Practice and Knowledge*, Vol. 1, 28–41, (Chinese), 1975.
20. Brandts, J. and M. Krizek, "History and future of superconvergence in three dimensional finite element methods," *Mathematical Sciences and Applications*, Vol. 15, 24–35, 2001.
21. Krizek, M., "Superconvergence phenomena on three dimenional meshes," *International Journal of Numerical Analysis and Modeling*, Vol. 2, No. 1, 43–56, 2005.
22. Xie, G., J. H. Li, M. Oristaglio, and X. Zhou, "A GLEMFHS coupled modeling for Wenchuan earthquake," to be published in *PIERS Proceedings*, Cambridge, July 2–6, 2008.
23. Xie, G., J. H. Li, F. Xie, and T. Kang, "A GLEMFCS coupled modeling and inversion for icing disaster on high voltage lines," to be published in *PIERS Proceedings*, Cambridge, July 2–6, 2008.
24. Xie, G., J. H. Li, L. Xie, and F. Xie, "GL EM mechanical and acoustic field time domain modeling for materials and exploration with dispersion," *PIERS Proceedings*, 354–364, Hangzhou, March 24–29, 2008.
25. Li, J., G. Xie, L. Xie, and F. Xie, "GL time domain modeling for em acoustic and elastic wave field with dispersion in crystal and porous material," *PIERS Proceedings*, 655–663, Hangzhou, March 24–29, 2008.
26. Xie, G., J. H. Li, and F. Xie, "GL mechanical and EM coupled modeling and inversion," *XXII ICTAM*, 25–29, Adelaide, Australia, August 2008.
27. Xie, G., J. Li, L. Xie, and F. Xie, "GL method for solving equation in math-physics and engineering," *Acta Mathematicae Applicatae Sinica, English Series*, Vol. 24, No. 2, 291–404, 2008.

An Improved Forward Scattering Simulation Technique for Microwave Breast Imaging

B. Babu and M. Condon

RINCE, School of Electronic Engineering, Dublin City University, Dublin 9, Ireland

Abstract— Microwave imaging is a promising alternative to conventional mammography methods. At microwave frequencies, normal and malignant tissues show high contrasts in their electrical properties. Microwave Imaging (MWI) systems can be used to construct three-dimensional profiles of the electrical properties of the body part that is being examined. MWI systems illuminate the body part with electromagnetic radiation of a suitable frequency. Using the measured scattered field at the surface of the body, inverse scattering algorithms reconstruct profiles of the electrical properties of the target. It is therefore of the uttermost importance that the forward scattering set-up is correct so that an inversion algorithm can create accurate profiles of electrical properties. We propose an improvement over the existing integral equation based forward scattering simulation techniques for microwave breast cancer imaging. Early detection of breast cancer is crucial. At this stage, the size of malignant tissue can be in the order of millimeters. For imaging involving such a small malignancy, one must use high-frequency radiation. At such frequencies, in order to overcome the relaxation effect of complex permittivity, we use the Debye model. For solving the forward scattering problem, we use the stabilized bi-conjugate gradient fast Fourier transform method (BI-CGSTAB-FFT). For the scattering domain, we apply the so-called cyclic boundary condition. This reduces the number of FFTs involved thus saving time and memory. For the BI-CGSTAB-FFT iteration method, we choose the initial value of the total field to be the incident electric field. This choice yields better convergence than a random selection of initial condition.

1. INTRODUCTION

Current imaging techniques for the early detection of breast cancer include X-rays, MRI and ultrasound imaging. X-ray mammography is the most common method employed at present. However, it poses certain health risks owing to the ionizing nature of the radiation. In addition, the method has relatively high false detection rates [1, 2]. Recently, microwave imaging has been proposed as a painless alternative to conventional X-ray mammography. Its non-ionizing nature makes it a safe probing method for the patient. In addition, the relatively lower manufacturing and maintenance cost makes it a viable economic replacement for conventional mammography machines. Studies have shown that at microwave frequencies, normal and malignant tissues show reasonable contrast in their electrical properties [3]. A recent large-scale study [4] reveals more insight into the electrical properties of the breast tissues. According to [4], there is no significant difference among the dielectric properties of breast tissue samples obtained from a large number of patients. This finding strengthens the case for using the MWI for identifying malignant regions as for all patients, regions with high contrast in dielectric properties will indicate abnormalities in normal tissues. Accurate simulation of the underlying scattering problem is of immense importance. These simulations facilitate the design and development of suitable imaging chambers, the selection of better antennas and efficient image reconstruction algorithms.

Simulation of electromagnetic problems at microwave frequencies is challenging as the size of governing matrix equations grows with frequency. Any improvement in algorithms that reduces the computation time or memory usage is of immense benefit. In this paper, we include cyclic boundary conditions [7] in the discretization procedure which reduces overall computation time and memory requirements while the accuracy is not affected. At high frequencies, in order to overcome the relaxation effect of complex pemittvity, we use the single-pole Debye dispersion expression [9].

2. SIMULATION SETUP

MWI involves the scattering of microwave radiation by an inhomogeneous body and the reconstruction of three-dimensional (3-D) maps of the dielectric properties of the target using the measured scattered field. In this study, we use an inhomogeneous sphere of radius 0.064 m as shown in Fig. 1 as the scatterer to mimic the breast. The forward scattering problem is modelled with the volume electric field integral equation approach of scattering theory. The source of microwave radiation

is an electric dipole placed at $z = -0.08\,\mathrm{m}$ as shown in Fig. 1. In the spherical polar coordinate system, the expression for the electric field due to an electric dipole is given by [5]:

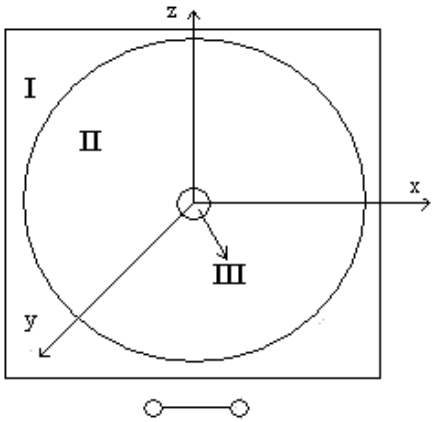


Figure 1: Simulation setup.

$$\begin{aligned} E(r) &= \frac{[I]L\cos\theta}{2\pi\epsilon_0}\left(\frac{1}{cr^2}+\frac{1}{j\omega r^3}\right) \\ E(\theta) &= \frac{[I]L\sin\theta}{4\pi\epsilon_0}\left(\frac{j\omega}{c^2 r}+\frac{1}{cr^2}+\frac{1}{j\omega r^3}\right) \\ E(\phi) &= 0 \end{aligned} \tag{1}$$

In Eq. 1, L is the length of the dipole, r is the distance of any point in the scattering domain from the dipole centre and f is the frequency of operation. The angular frequency is given by $\omega = 2\pi f$. $[I]$ is the retarded current and c is the velocity of light. Throughout our simulations $f = 6\,\mathrm{GHz}$, $L = 0.01\,\mathrm{m}$ and $[I] = 10\,\mathrm{A}$. The total field $\mathbf{E}$ will be measured on the surface of scattering domain at $z = 0.07\,\mathrm{m}$. The scattered field $\mathbf{E}^s$ is given by $\mathbf{E}^s = \mathbf{E} - \mathbf{E}^i$ where $\mathbf{E}^i$ is the incident field.

At high frequencies, to overcome the relaxation effect of complex permittivity, the single-pole Debye dispersion expression [9] is used for representing complex permittivity:

$$\tilde{\epsilon}_r(\omega) = \epsilon_r^{'}(\omega) - j\epsilon_r^{''}(\omega) = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + j\omega\tau} - j\frac{\sigma_s}{\omega\epsilon_0} \tag{2}$$

Here, ϵ_∞ is the relative permittivity at infinite frequency, ϵ_s is the static relative permittivity, σ_s is the static conductivity and τ is the relaxation time constant. The scattering domain in Fig. 1 has three regions with different dielectric properties as given in Table 1 [9]. Region I is outside the sphere and acts as a coupling liquid between the breast phantom and radiation source. Region II acts as the normal breast phantom and region III acts as the malignant region. Throughout the scattering domain, the permeability $\mu = \mu_0$ is constant.

Table 1: Debye parameters (in SI units) for the breast phantom in Fig. 1 at 6 GHz.

	Dielectric properties of the breast phantom					
	ϵ_∞	ϵ_s	σ_s	τ	ϵ_r	σ
Region I	4.0	10.0	0.1	7.0e-12	10	0.1
Region II	6.57	16.29	0.23	7.0e-12	15.66	1.03
Region III	3.99	54.00	0.70	7.0e-12	50.74	4.82

3. VOLUME ELECTRIC FIELD INTEGRAL EQUATION (VEFIE)

The electric field inside the scattering domain can be represented by a Fredholm integral equation of the second kind [6], given by:

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^i(\mathbf{r}) + (k_b^2 + \nabla\nabla.) \int_V G(|\mathbf{r} - \mathbf{r}'|)\chi(\mathbf{r}')\mathbf{E}(\mathbf{r}')d\mathbf{r}' \tag{3}$$

Here, $\mathbf{E}^i$ is the incident electric field and $\mathbf{E}$ is the total electric field in the scattering domain V. k_b is the wavenumber in the coupling medium given by $k_b = \omega\sqrt{\mu_b\tilde{\epsilon}_b}$, χ is the contrast function given by $\chi = \frac{k^2}{k_b^2} - 1$. $G(|\mathbf{r} - \mathbf{r}'|)$ is the scalar Green's function. Note that the harmonic time dependence $e^{-j\omega t}$ is omitted in Eq. (3) and in all equations henceforth.

4. DISCRETIZATION OF VEFIE USING CYCLIC BOUNDARY CONDITIONS

The magnetic vector potential is defined [6] as:

$$\mathbf{A} = -j\omega\mu_b\tilde{\epsilon}_b \int_V G(|\mathbf{r} - \mathbf{r}'|)\chi(\mathbf{r}')\mathbf{E}(\mathbf{r}')d\mathbf{r}' \tag{4}$$

Then the integral Eq. (3) can be written as:

$$\mathbf{E}(\mathbf{r}) - j\omega\mathbf{A} - \mathbf{E}^{irr} = \mathbf{E}^i(\mathbf{r}) \tag{5}$$

In Eq. (5), $\mathbf{E}^{irr}$ is the irrotational part of scattered electric field and is denoted as in Eq. (6).

$$\mathbf{E}^{irr} = \frac{j\omega}{k_b^2}\nabla\nabla\cdot\mathbf{A} \tag{6}$$

The scattering domain V is discretized into $N \times N \times N$ cubic cells of volume $dv = dx \times dy \times dz$. In our simulations, N is set to be equal to 21. The cell centres are located at:

$$\mathbf{r}_{j,k,l} = [j \times dx, k \times dy, l \times dz] \tag{7}$$

To mimic a continuous scattering domain, we include the cyclic boundary condition [7] in the discretization procedure. According to this condition, if the scattering domain has N cells in each direction, the Nth cell in the x-direction will be considered as the left neighbour of first cell in that direction and the first cell will be considered as the right neighbour of the Nth cell. It is likewise in the y and z directions. Testing Eq. (5) with the delta function $\delta(\mathbf{r} - \mathbf{r}_{j,k,l})$ gives the field values at $\mathbf{r}_{j,k,l}$ given by Eq. (8), where $j \in [1, J]$, $k \in [1, K]$, $l \in [1, L]$.

$$\mathbf{E}^i_{j,k,l} = \mathbf{E}_{j,k,l} - j\omega\mathbf{A}_{j,k,l} - \mathbf{E}^{irr}_{j,k,l} \tag{8}$$

The magnetic vector potential in (4) can be expressed as:

$$\mathbf{A}_{j,k,l} = -j\omega\mu_b\tilde{\epsilon}_b dv \sum_{j'=1}^{J}\sum_{k'=1}^{K}\sum_{l'=1}^{L} G_{j-j',k-k',l-l'}\chi_{j',k',l'}\mathbf{E}_{j',k',l'} \tag{9}$$

where, $j \in [1, J]$, $k \in [1, K]$, $l \in [1, L]$. The Green's function [10] is given by:

$$G_{j,k,l} = \begin{cases} \frac{(1-\frac{1}{2}jk_0dr)\exp(\frac{1}{2}jk_0dr)-1}{\frac{1}{6}\pi k_0^2 dx^3} & \text{if } |r| = 0 \\ \frac{\exp(jk_0|r|)\left[\frac{\sinh(\frac{1}{2}jk_0dr)}{\frac{1}{2}jk_0dr)} - \cosh(\frac{1}{2}jk_0dr)\right]}{\frac{1}{3}\pi(k_0dr)^2|r|} & \text{if } |r| > 0 \end{cases} \tag{10}$$

The convolution in Eq. (9) may be efficiently evaluated by applying the Fast Fourier Transform, i.e.,

$$\mathbf{A}_{j,k,l} = C\left[FFT^{-1}[FFT[G_{j',k',l'}] \cdot FFT[\chi_{j,k,l}\mathbf{E}_{j,k,l}]\right] \tag{11}$$

where $C = -j\omega\mu_b\tilde{\epsilon}_b dv$. In our simulation, the inclusion of the cyclic boundary condition reduces the number of FFTs involved by a factor $3(N+2)^3 - 3N^3$. The components of $\mathbf{E}^{irr}_{j,k,l}$ in Eq. (8) can be computed as in [6]. Eq. (8) may now be written as a matrix equation of the form

$$\mathbf{E}^i = \mathfrak{L}[\mathbf{E}] \tag{12}$$

In Eq. (12), $\mathfrak{L}$ represents the linear operations in Eq. (8) and 11 and in [6] acting on the total field $\mathbf{E}_{j,k,l}$.

5. NUMERICAL RESULTS

In order to find the unknown electric field, Eq. (12) needs to be solved numerically. Direct solution involves the inversion of a large matrix and this is prohibitive from a computational viewpoint. Consequently, iterative methods are employed for the solution of Eq. (12). We intend to use a Krylov subspace method for iteratively solving this equation. In addition, it should be noted that because of the probable non-symmetric location of tumour tissues in the breast phantom, Eq. (12) will not be symmetric. To tackle this asymmetry, we use the BI-CGSTAB-FFT method [8] to solve the matrix equation in (12). The residual error in satisfying Eq. (12) at the nth iteration step is given by $\mathbf{e}_n = \mathbf{E}^i - \mathfrak{L}[\mathbf{E}_n]$. The rate of convergence of the BI-CGSTAB-FFT method is measured as:

$$\delta_n = log\left(\frac{||\mathbf{e}_n||_2}{||\mathbf{e}_1||_2}\right) \tag{13}$$

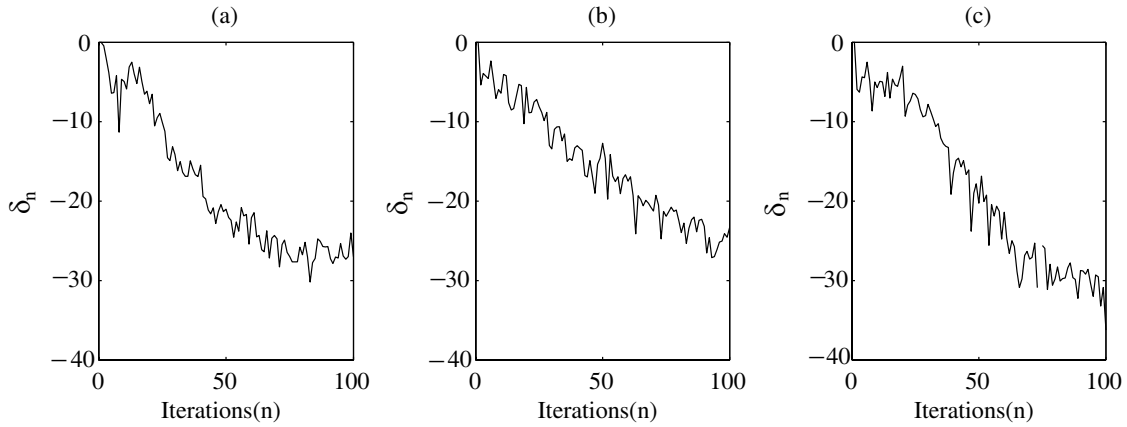


Figure 2: Rate of convergence of the BI-CGSTAB-FFT method measured as in Eq. (13), (a) with a random initial value for $\mathbf{E}$, (b) with the incident field $\mathbf{E}^i$ as the initial value for $\mathbf{E}$, (c) with the incident field $\mathbf{E}^i$ as initial value for $\mathbf{E}$ and the cyclic boundary condition.

Figure 2(a) shows the convergence rate δ_n of the BI-CGSTAB-FFT implementation when a random initial value is selected for $\mathbf{E}$. When the incident field $\mathbf{E}^i$ is used as the initial value for $\mathbf{E}$, better convergence is achieved as evident from Fig. 2(b). Fig. 2(c) shows the convergence rate when the cyclic boundary condition is employed. Note also that for Fig. 2(c), $\mathbf{E}^i$ is chosen as the initial value for $\mathbf{E}$. Comparing the figures, it is obvious that the cyclic boundary condition does not affect the convergence rate. However, in our case, it reduces the number of FFTs involved in the simulation by 23% and saves a considerable amount of memory.

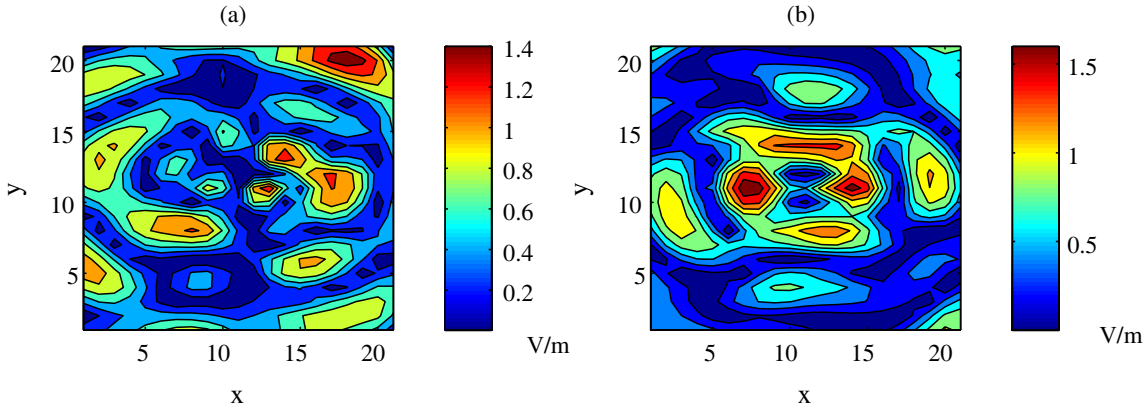


Figure 3: Scattered field $\mathbf{E}^s$ measured on the surface ($z = 0.07$ m) of the scattering domain, (a) without the tumour and (b) with the tumour.

Figure 3 shows the scattered field measured on the surface ($z = 0.07$ m) of the scattering domain, (a) without the tumour and (b) with the tumour. From Fig. 3, it is clear that the measured scattered

field on the surface of the scattering domain is different when an inhomogeneity is present inside the scatterer. The scattered field can then be used with an inversion algorithm to identify the location of the tumour.

6. CONCLUSION

Simulation of electromagnetic scattering problems at microwave frequencies is computationally challenging because of the large size of governing matrix equations. Any improvement in the formulation and algorithm that reduces the memory requirements and computation times is of immense importance. Microwave imaging is a promising alternative for mammography and other kinds of medical imaging. Simulation of these systems is computationally expensive. We have examined the inclusion of a cyclic boundary condition in the discretization procedure which indeed mimics a continuous scattering domain and reduces the overall computation by a considerable amount compared to previous works. Furthermore, in the BI-CGSTAB iteration method, we proposed the selection of the incident electric field itself as the initial condition for solving for the total field. Results confirm that it gives better convergence when compared to the standard choice of initial condition.

ACKNOWLEDGMENT

This work was funded by Science Foundation Ireland

REFERENCES

1. MARIBS study group, "Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: A prospective multicentre cohort study (MARIBS)," *The Lancet*, Vol. 365, 1769–1778, May 16, 2005.
2. Khor, W. C., et al., "An ultra wideband microwave imaging system for breast cancer detection," *IEICE Transactions on Communication*, Vol. E90-B, No. 9, 2376–2380, 2007.
3. Chaudhary, S. S., et al., "Dielectric properties of normal and malignant human breast tissues at radiowave and microwave frequencies," *Indian J. Biophys*, Vol. 21, No. 10, 76–79, 1984.
4. Lazebnik, M., et al., "A large-scale study of the ultrawideband microwave dielectric properties of normal, benign, and malignant breast tissues obtained from cancer surgeries," *Physics in Medicine and Biology*, Vol. 52, 6093–6115, 2007.
5. Krauss, J. D., *Antennas 2ed.*, Mc Graw-Hill, New York, 1998.
6. Zhang, Z. Q., et al., "Microwave breast imaging: 3-D forward scattering simulation," *IEEE Transactions on Biomedical Engineering*, Vol. 50, No. 10, 1180–1188, 2003.
7. Johnson, H. J. and G. E. Christensen, "Consistent landmark and intensity-based image registration," *IEEE Transactions on Medical Imaging*, Vol. 21, No. 5, 450–461, 2002.
8. Van der Vorst, H. A., "BI-CGSTAB: A fast and smoothly converging variant of BI-CG for the solution of nonsymmetric linear systems," *SIAM J. Sci. Statist. Comput.*, Vol. 13, 631–644.
9. Converse, M., et al., "A computational study of ultrawideband versus narrowband microwave hyperthermia for breast cancer treatment," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 54, 2169–2180, 2006.
10. Zwarmborn, P. and P. van den Berg, "The three dimensional weak form of the conjugate gradient FFT method for solving scattering problems," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 40, 1757–1766, 1992.

Extending MAS/TSA Technique for Conducting Environments to Enhance Underwater UXO Discrimination

D. Kakulia[1], **G. Ghvedashvili**[1], **and F. Shubitidze**[2]

[1]Chavchavadze Ave. 3, Tbilisi State University, Tbilisi 0128, Georgia
[2]Thayer School of Engineering, Dartmouth College, HB 8000, Hanover, NH 03755, USA

Abstract— The main objective of the paper is to demonstrate the applicability of the combined Method of Auxiliary Sources (MAS) and thin skin depth approximation (TSA) (MAS/TSA) technique for underwater environment and to investigate mutual coupling between highly conducting and permeable metallic objects and their surrounding conductive medium in marine environments. The theoretical bases and numerical realization of the MAS/TSA technique for conducting medium is demonstrated. Namely, the numerical results are designed to for illustrating the underlying physics of EMI scattering phenomena for metallic objects that are placed inside a conducting medium, and the interaction effects between the object and it surrounding conducting medium are analyzed.

1. INTRODUCTION

There are approximately one million acres of underwater lands at Department of Defense (DOD) and Department of Energy (DOE) sites that are highly contaminated with unexploded ordnance (UXO) and land mines. The detection and disposal of Underwater Military Munitions are more expensive than excavating the same targets on land. Therefore, innovative detection and discrimination systems with 0% false alarm rate that can reliably discriminate between hazardous UXO and innocuous items are required. Recently, electromagnetic induction (EMI) sensing has been considered as a potential candidate for underwater detection. In order to explore the potential of various EMI sensing technologies for underwater detection and discrimination and to achieve a high ($\sim$100%) probability of detection, and to distinguish UXO from non-UXO items accurately and reliably, first the underlying physics of the EM scattering phenomena in underwater environment needs to be investigated in detail. This could be achieved by using an accurate 3D numerical code, such as the MAS/TSA approach that has been developed previously for land based UXO detection and discrimination [1–3]. Here the combined Method of Auxiliary Sources and thin skin depth approximation procedure are extended for solving a boundary problem when a highly conductive and permeable object is illuminated with a quasi-static magnetic field and placed in electrically conductive environment. By performing comparative analysis in the EMI frequency range between magnetic field responses from objects placed in free space and conducting medium, interaction effects between the object and its surrounding are illustrated. The numerical analysis are done for different canonical shaped geometries, such as sphere and prolate spheroid, which could be considered as as prototypes of the real ordinances or land-mines. The studies of interaction effects reported here for the simple objects are expected to be applicable for actual objects and under real field conditions.

2. MSA/TSA EXTENSION

The MSA/TSA algorithm for free space is described in great data in [1, 2]. Here the method will be extended for under water environments. The MSA/TSA algorithm was developed for frequency domain to solve boundary problem in low frequency range for highly conductive and permeable object placed in quasi-static magnetic field. Notice that the low frequency range implies to frequency range from several Hz up to several hundred's of kilo Hertz. This covers the working frequencies of magnetic sensors. Usually, the magnetic sensors detect the disturbance of the magnetic field responses from metallic objects in their vicinity, thus investigate interaction effects between object and its surrounding medium has a great interests for subsurface objects detection and discrimination.

In the original MSA/TSA algorithm, the scattered field is represented as a summation of magnetic fields produced by set of magnetic charges, which are distributed inside on a fictitious surface inside the object. This assumption is valid only for cases when object's surrounding medium is free space. To over this problem here the auxiliary magnetic charges are replaced with auxiliary

magnetic dipoles with corresponding dyadic Green's function. Namely, the scattered magnetic and electric fields produced by a magnetic dipole with magnetic moment $\vec{M}$ are:

$$\vec{H} = \frac{e^{i(kr-\omega t)}}{4\pi\mu_r\mu_0}\left(\frac{3\cdot\vec{r}\cdot\left(\vec{r}\cdot\vec{M}\right)}{r^5} - \frac{\vec{M}}{r^3} - i\cdot k\left(\frac{3\cdot\vec{r}\cdot\left(\vec{r}\cdot\vec{M}\right)}{r^4} - \frac{\vec{M}}{r^2}\right) - \frac{k^2\vec{r}\times\left(\vec{r}\times\vec{M}\right)}{r^3}\right) \quad (1)$$

$$\vec{E} = \frac{ikW_0e^{i(kr-\omega t)}}{4\pi\mu_0}\cdot\left(\left(\frac{1}{r^3} - \frac{ik}{r^2}\right)\cdot\left(\vec{M}\times\vec{r}\right)\right) \quad (2)$$

Here $\vec{r}$ vector between magnetic dipole and observation points. $r = |\vec{r}|$, μ_r is relative magnetic permeability of medium, μ_0 is magnetic permeability of free space. ω is cyclic frequency of magnetic dipole's oscillation. t is time. W_0 is wave impedance of free space. k is wave number in medium, so for conductive medium it is complex number. It can be expressed as

$$k = \frac{\omega}{c}\sqrt{\varepsilon_r}\sqrt{\mu_r} \quad (3)$$

where c is velocity of light in free space. ε_r is relative electric permittivity of medium. In case of conductive medium ε_r is complex number with imaginary part $\frac{\sigma}{\omega\varepsilon_0}$, where σ is electrical conductivity and ε_0 is electrical permittivity of free space.

3. TESTING OF MSA/TSA EXTENSION

In order to validate the numerical code, first the extended MAS/TSA algorithm is compared with magnetic charge based MAS/TSA algorithm for case when surrounding medium is free space. The test problem simulated here is 10 cm radius conductive sphere with electrical conductivity 10^7 S/m and relative permeability 10 placed in homogeneous quasi-static magnetic fields with unit strength. The magnetic field is polarized along z-axis. The observation point is at 5 cm above the sphere on z-axis. The magnetic field response at this point has only z component. Corresponding results of comparative calculation are given at Figure 1.

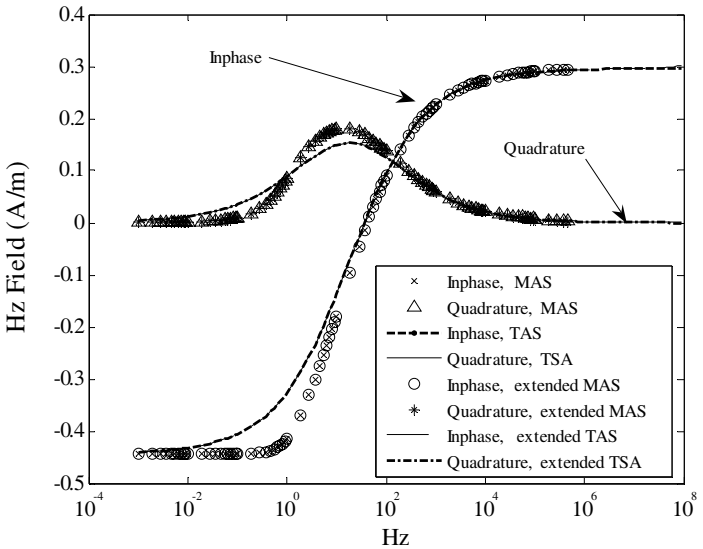


Figure 1: The scattered magnetic field versus frequency for a sphere.

The comparisons depicted on Figure 1 show that the extended MAS/TSA algorithm produces similar result as the original MAS/TSA algorithm. It can be note that TSA algorithm gives proper results at the upper frequency range and MAS algorithm numerically fails.

4. NUMERICAL RESULTS

In order to investigate an interaction effects between object and its surrounding medium, we apply extended MSA/TSA algorithm to previously examined problem for a sphere, cylindrical object and

prolate spheroid. Here a surrounding medium is a low conductive which electrical conductivity of 4 S/m, that corresponds a salt water's conductivity.

The Figure 2 shows results of calculation that is performed by extended MSA/TSA algorithm for case of sphere in conductive medium and its comparison with results of Figure 1. Interaction effect between object and conductive medium becomes significant for upper frequencies. In this frequency range, TSA algorithm is suitable and allows observing the effect that significantly distorts response from object. The Figure 3 illustrates in XOZ slice Y components distribution of current density. These currents are induced by scattered field of object in surrounding medium. Current's density is calculated by formula $\vec{j} = \sigma\vec{E}$, where σ is conductivity and $\vec{E}$ electric field induced in medium. So, It gives real picture of physical processes having place in surrounding medium.

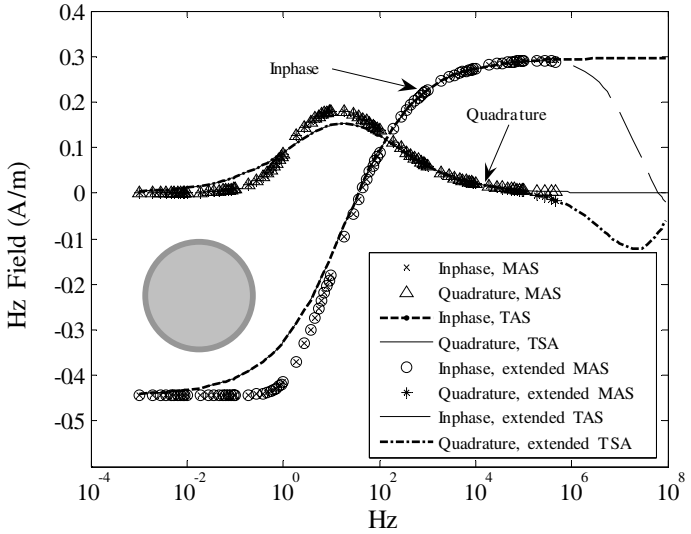


Figure 2: The scattered magnetic field versus frequency for a sphere, for free space and conductive medium.

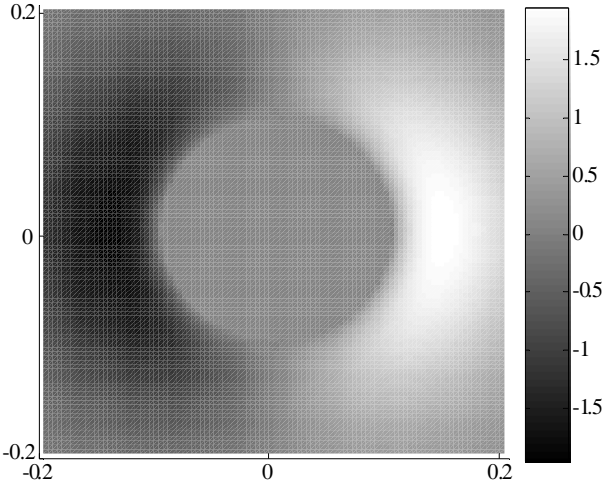


Figure 3: Distribution of induces currents in XOZ slice at frequency 10 MHz.

In addition, we consider cases with cylindrical object and prolate spheroid, which are completely analogical to the problem with sphere. Geometrical data of cylindrical object are height 20 cm, radius 10 cm. It is rounded on the bottom and top as depicted on Figure 4. The radius of the rounding is 3.5 cm. Object's axis of symmetry is coincided with z-axis.

Corresponding results of simulation are given on Figure 4. Its field response also is observed at 5 cm above object. In cases of sphere and cylindrical object, behavior of frequency characteristics is similar. But for second case interaction effect is stronger. It caused by object's geometry, by

currents induced of the cylinder's top, which induce stronger currents in surrounding medium. Following this, we observe interference of fields of both currents in observation point.

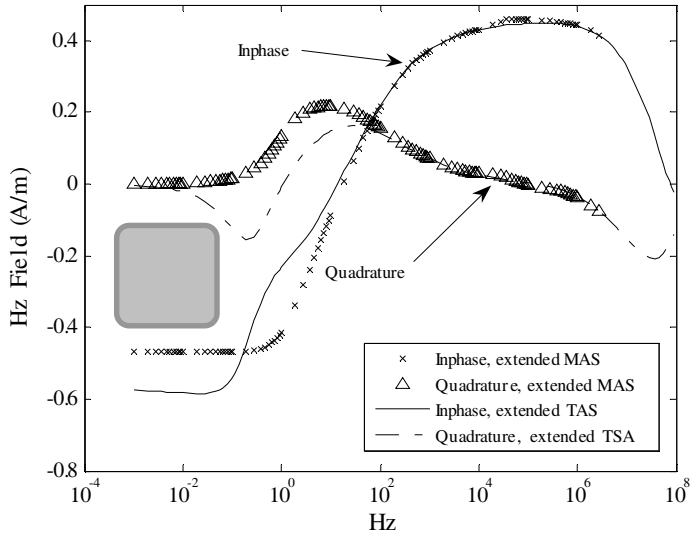


Figure 4: The scattered magnetic field versus frequency for cylindrical object, for conductive medium.

The spheroid is prolate along z-axis. Its minor axes $a = b = 10\,\text{cm}$, major axis $c = 40\,\text{cm}$. Its field response also is observed at 5 cm above it. Figure 5 is very similar to Figure 2. In case of spheroid, we can see same interaction effects, which appear at upper frequencies. Here interaction effect is weaker than in cases of sphere and cylindrical object. It caused also by object's geometry and can be explained by same reasoning as in case of cylindrical object.

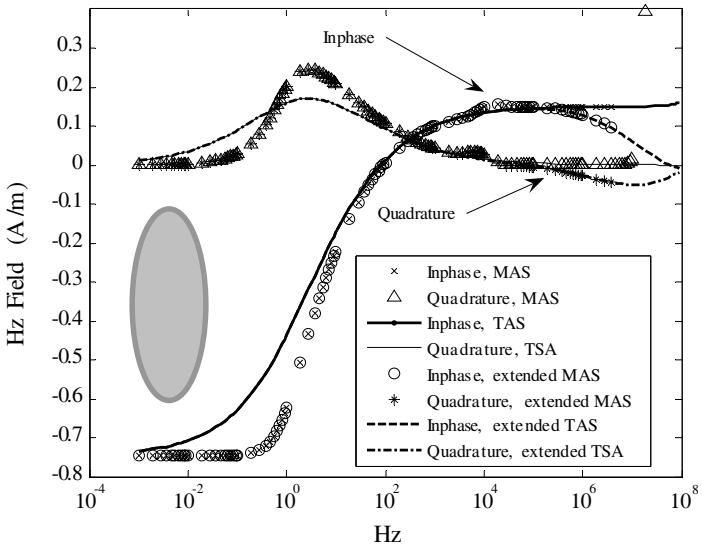


Figure 5: The scattered magnetic field versus frequency for a spheroid, for free space and conductive medium.

5. CONCLUSION

Validation of MSA/TSA algorithm's extension is performed. Interaction effects between the object and its surrounding medium are illustrated. A physical phenomenon of observed interaction effects is analyzed. Results of numerical analysis underline an importance of effects caused by medium; prevision of it improves objects discrimination capabilities.

ACKNOWLEDGMENT

To GRDF/CRDF for financial support of performed research.

REFERENCES

1. Shubitidze, F., et al., "A hybrid full MAS and combined MAS/TSA algorithm for electromagnetic induction sensing," *Applied Computational Electromagnetic Society Journal*, 112–126, Mar. 2004.
2. Shubitidze, F., et al., "Interaction between highly conducting and permeable metallic objects in the EMI frequency range," *Proceedings of 19th Annual Review of Progress in Applied Computational Electromagnetics*, 625–631, Monterey, CA, Mar. 24–28, 2003.
3. Sun, K., et al., "Theoretical analysis of TSA formulation and its domain of validity," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 42, No. 9, 1871–1881, Sept. 2004.

An Analysis of the Landau-Lifshitz Reaction Term in Classical Electrodynamics

G. Ares de Parga
Dpto. de Física, Escuela Superior de Física y Matemáticas
Instituto Politécnico Nacional
U.P. Adolfo López Mateos, Zacatenco, C.P. 07738, México D.F., México

Abstract— Since Dirac obtained the so named Lorentz-Dirac equation [LD] as the equation of motion for a charged point particle, it has focused many discussions about its validity. Indeed, the runaway solutions, the preaccelerations, the renormalization of the electron's mass and the violation of physical causality by the use of the advanced solutions of the Maxwell equations, are the main reasons for the long historical discussion about the LD equation. This unsatisfactory situation is evidenced by the continued appearance of new equations of motion in the literature.

By using an approximation of first order of the LD equation, Landau and Lifshitz obtained an equation in the frame of classical electrodynamics, the Landau-Lifshitz equation [LL]. Spohn has claimed that the LL equation can be obtained with the same degree of accuracy than the LD equation. Rohrlich has noticed that LL equation is a second order differential equation which doesn't permit runaway solutions or preaccelerations. It is important to note that Ares de Parga have proposed a physical deduction of the LL equation which implies a change in the concept of the radiation rate of energy; that is: the regular LD reaction term is substituted by the LL reaction term. We analyze the different situations where the LL reaction term vanishes. In these cases, the LL equation of motion coincides with the Lorentz equation. We propose some physical interpretations in order to understand the absence of radiation rate of energy in such situations.

1. INTRODUCTION

The Landau-Lifshitz equation is considered by many authors [1, 2] as the correct equation which describes the motion of a point charged particle. Nevertheless, it apparently presents some unphysical solutions. Indeed, for the case when the electric field is constant in the same direction of the motion, it has been proved that the motion of the particle coincides with the one that will follow the Lorentz path [3–5]; that is: the Landau-Lifshitz equation and the Lorentz equation coincide for this case. On the other hand, Baylis [6] has showed that there exists a divergent radiated energy when pulses are considered and that we have to mix both theory, the Lorentz-Dirac equation and the Landau-Lifshitz equation, in order to obtain a consistency. Due to this reasons we can always think that since the correct equation of motion will be no longer the Lorentz-Dirac equation, it will be possible that the Larmor formula will be no longer too the correct expression for the radiated power. We will focus our attention on the Landau-Lifshitz radiation term (Γ^μ); that is:

$$\Gamma^\mu = \tau_o \left[\frac{q}{c} \frac{dF^{\mu\nu} v_\nu}{d\tau_L} + \frac{q^2}{c^4 m} F^2 v^\mu \right], \tag{1}$$

where $\tau_o = \frac{2q^2}{3c^3 m}$, q, c and m represent the characteristic time of the charge, the charge, the speed of light and the mass of the charge. $F^{\mu\nu}$ is the Faraday electromagnetic tensor (strength tensor) and $d\tau_L$ corresponds to a differential of the proper time but following the Lorentz path. This means that $\frac{dF^{\mu\nu} v_\nu}{d\tau_L}$ represents the derivative of the force along a path described by the Lorentz equation [7, 8] (see Fig. 1).

Then, if the Landau-Lifshitz radiation term (Γ^μ) vanishes, the Landau-Lifshitz equation coincides with the Lorentz equation and after a reasoning it is easy to see that it also happens with the Lorentz-Dirac equation; that is:

$$\Gamma^\mu = 0 \Rightarrow \dot{a}^\mu + \frac{a^2}{c^2} v^\mu = 0. \tag{2}$$

We will analyze the cases where Eq. (2) is possible and we will try to understand the radiation in those cases.

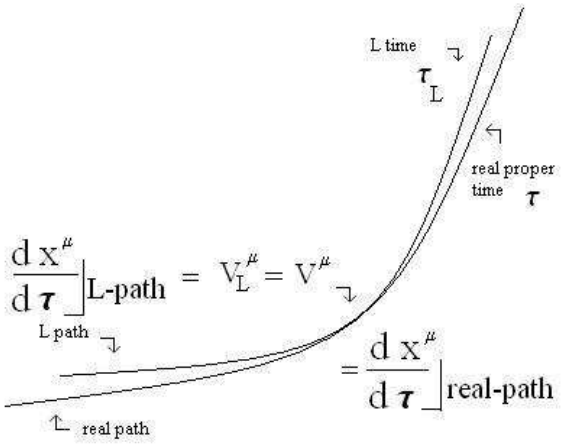


Figure 1: L-time.

2. VANISHING RADIATION

Let us see some examples.

2.1. A. The Constant Force

Ares de Parga et al. [3] demonstrated that when the particle is submitted to a constant force in the same direction of its motion the Landau-Lifshitz term vanishes.

In this case we can just consider two dimension: ct, x. Therefore the equations of motion are

$$\begin{aligned} \ddot{ct} &= w\dot{x} + \tau w^2 \dot{ct} + \frac{\tau}{c^2}\left[\dot{x}^2 - \dot{ct}^2\right]\dot{ct} = w\dot{x} \\ \text{and} \quad \ddot{x} &= w\dot{ct} + \tau w^2 \dot{x} + \frac{\tau}{c^2}\left[\dot{x}^2 - \dot{ct}^2\right]\dot{x} = w\dot{ct}, \end{aligned} \tag{3}$$

where the dot $\cdot$ represents the derivative with respect to the proper time and $w = \frac{qE}{cm}$ where E is the constant electric field in the x-axis. Since Eq. (3) is equivalent to the one that would appears for the Lorentz equation. It is obvious that for this case, Eq. (2) is satisfied.

2.2. B. The Central Field

In the example above, the initial conditions were hidden, since we constrained the motion to the direction of the field. Now if we consider a more complicated situation as a central force, Coulomb force for example, and we consider that $\Gamma^\mu = 0$. Then,

$$a^\mu = -\frac{q}{cm} F^{\mu\upsilon} v_\upsilon, \tag{4}$$

where $F^{\mu\upsilon}$ represents the strenght tensor and for the case of the Coulomb electric field,

$$F^{\mu\upsilon} = \frac{q}{r^2}\begin{bmatrix} 0 & -\hat{r} \\ \hat{r} & 0 \end{bmatrix}. \tag{5}$$

It is easy to demonstrate that:

$$\ddot{ct} = -\frac{w}{r^2}\dot{r} \quad \text{and} \quad \ddot{\vec{r}} = -\frac{w}{r^2}\hat{r}\dot{ct}. \tag{6}$$

By using the first equation, we arrive at:

$$\dot{ct} - \frac{w}{r} = \lambda = const \quad \text{or} \quad mc^2\gamma - \frac{q^2}{r} = mc^2\lambda = E, \tag{7}$$

where λ represents a constant and E the constant energy. Therefore,

$$\vec{\Gamma} = \vec{0} \Rightarrow \dddot{\vec{r}} + \frac{a^2}{c^2}\dot{\vec{r}} = 2\frac{w}{r^3}\left(\lambda + \frac{w}{r}\right)\dot{r}\hat{r} - \frac{w}{r^2}\left(\lambda + \frac{2w}{r}\right)\dot{\theta}\hat{\theta} = \vec{0}. \tag{8}$$

We arrive at two possible solutions

$$\text{for } \dot{r} \neq 0 \Rightarrow \lambda = -\frac{w}{r} \quad \text{and for} \quad \dot{\theta} \neq 0 \Rightarrow \lambda = -\frac{2w}{r}. \tag{9}$$

That is:

$$\text{for } \dot{r} \neq 0 \Rightarrow v^2 = 0 \quad \text{and for} \quad \dot{\theta} \neq 0 \Rightarrow c\dot{t} = -\frac{w}{r}. \tag{10}$$

This means that for the Coulomb case, there is no motion that satisfied both Lorentz equation and vanishing Landau-Lifshitz radiation term.

3. CONSTANT MAGNETIC FIELD

If we consider a constant magnetic field in the z-axis, putting $\Omega = \frac{qB}{cm}$

$$c\ddot{t} = \tau\Omega^2(\dot{t}^2 - 1)c\dot{t}. \tag{11}$$

Consequently, we will not have a Lorentz motion unless

$$\dot{t}^2 = 1 \Rightarrow \vec{v} = \vec{0}. \tag{12}$$

4. CONCLUSION

By using the Landau-Lifshitz equation of motion, we analyze tree types of motion. The constant force on a direction, the central field and the constant magnetic field. The only case where the radiation term vanishes, was the constant force with initial conditions on the same direction of the force. The other two cases predict that there exist non vanishing radiation terms. Then, the Landau-Lifshitz equation of motion presents two fundamental problems: the first one is related with the divergence of the radiated power for a step function [6], and the other physical problem comes from the case where the constant force which is parallel to its velocity. For Ares de Parga [7], both problem can be explained by a physical rearrangement of the energy and by considering that the radiated power is expressed by

$$P = -\tau\frac{q^2}{c^4 m}F^2 v^o. \tag{13}$$

ACKNOWLEDGMENT

This work was partially supported by C.O.F.A.A and E.D.I., I.P.N.

REFERENCES

1. Spohn, H., "The critical manifold of the Lorentz-Dirac equation," *Europhys. Lett.*, Vol. 50, 287–292, 2000.
2. Rohrlich, F., "The correct equation of motion of a classical point charge," *Phys. Lett. A*, Vol. 283, 276–279, 2001.
3. Ares de Parga, G., R. Mares, and S. Domínguez, "Landau-Lifshitz equation of motion in classical electrodynamics revisited," *Ann. Fond. Louis de Broglie*, Vol. 30, 283–288, 2005.
4. Ares de Parga, G. and S. Domínguez, *Rev. Mex. Fís*, Vol. 52, No. 2, 139–142, 2006.
5. Spohn, H., *Dynamics of Charged Particles and Their Radiation Field*, Chap. 9, Cambridge University Press, Cambridge, 2004.
6. Baylis, W. E. and J. Huschilt, "Energy balance with the Landau-Lifshitz equation," *Phys. Lett. A*, Vol. 301, 7–12, 2002.
7. Ares de Parga, G., *Foundations of Physics*, Vol. 36, No. 10, October 2006.
8. Ares de Parga, G., M. Ortiz-Domínguez, and R. Mares, "A deduction of the landau-lifshitz equation of motion without mass renormalization in classical electrodynamics," Submitted to *Physica Scripta*, 2008.

Convergence of Krylov Solvers and Choice of Basis and Weighting Set of Functions in the Moment Method Solution of Electrical Field Integral Equation

G. Angiulli and S. Tringali
DIMET, Univ. Mediterranea, 89100 Reggio Calabria, Italy

Abstract— In Computational Electromagnetics, iterative techniques for solving algebraic linear system of equations are of fundamental importance, since actual problems give rise to linear systems too large to be practically solved by direct methods. In this work we investigate as performances of the major Krylov subspace iterative solver (i.e., GMRES), is affected by different choice of these set of functions. Specifically, we consider the algebraic linear system of equations obtained by reducing the electrical field integral equation (EFIE) from the TM_z scattering of a plane wave by a metallic strip. It can be observed that exists a critical threshold Δ_0 such that, whenever either the basis or the weight pulses are given with an amplitude greater than Δ_0, then the total number of internal loops necessary for taking the relative residual under a definite tolerance $\epsilon > 0$ increases all of a sudden, in such a dramatic way that it can even prevent the process at all from convergence. We try to explain this numerical behavior by inquiring the relationship between the MoM matrix condition number and the number of overall iterations necessary to numerical convergence.

1. INTRODUCTION

It is not easy to give a comprehensive treatment of so many issues involved by analytical and numerical methods in the field of Computational Electromagnetics, and specifically by the method of moments, for solving integral equations relevant to electromagnetic theory. These include a very large amount of mathematical elements and capabilities, ranging from numerical integration and differentiation to the feasibility of converting operator equations to a discrete system of algebraic equations, from the choice of suitable sets of basis and weight functions to a deep knowledge of linear algebra, matrix manipulations, algorithms and actual implementations of iterative and direct solvers and so on going [1]. The lack of analytical solutions provides a further item of difficulty when trying in formal terms to expound mathematical clues intended to be potentially notable from a practical point of view. Thus it is common to consider simple (i.e., typically scalar) problems for preliminary tests and then to extrapolate general ideas to be applied to more complex cases. In fact, this is the approach we follow in the present paper, where we deal with the scalar problem of computing the current density induced on the surface of an ideal metallic strip when struck by a planar wave of a given wavelength λ which is supposed to be small or large with respect to the width $2w$ of the strip section. A harmonic time variation $\exp(i\omega t)$ is assumed and suppressed throughout. The problem serves as an useful and somehow simple example in electromagnetics for tackling basic concepts and discussing practical issues for numerical methods. Specifically, we investigate and try to characterize the behavior of iterative techniques for solving algebraic linear system of equations. Nowadays these are of basic importance, since actual problems give rise to linear systems that are too large to be practically solved by direct methods. There exists much study regarding the numerical aspects of such subjects, but it's not clear at all how the efficiency of iterative solvers are conditioned by the choosing of the basis and weighting functions in the Moment Method (MM) [2, 3]. In this sense, we study as performances of the Generalized Minimun Residual (GMRES) Krylov iterative solver [4] is affected by different choice of these set of functions.

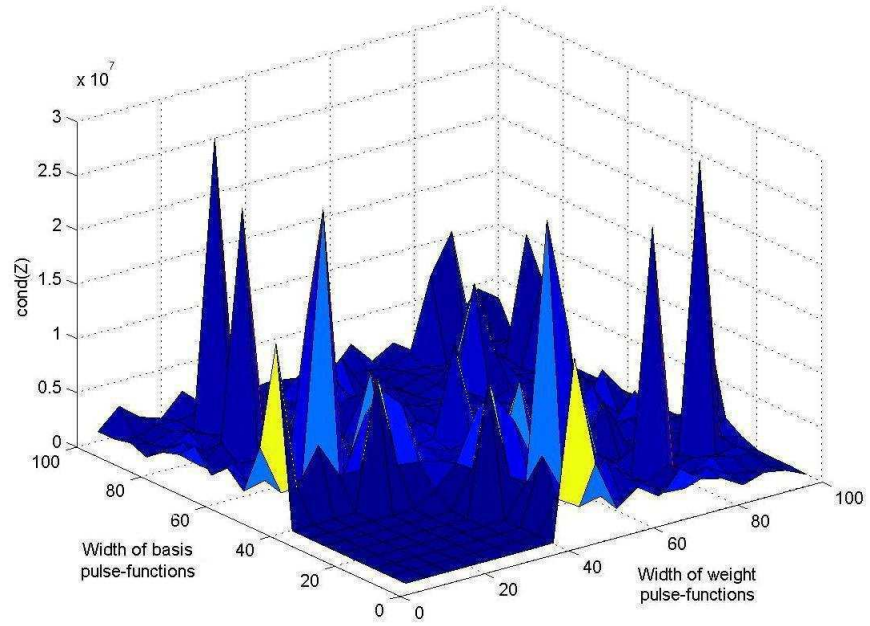
2. A WIDTH-VARYING PULSE-PULSE MOMENT METHOD SCHEME

An implementation of the moment method exploiting width-varying rectangular pulses for both the basis and the weight sets of functions is adopted for solving the EFIE

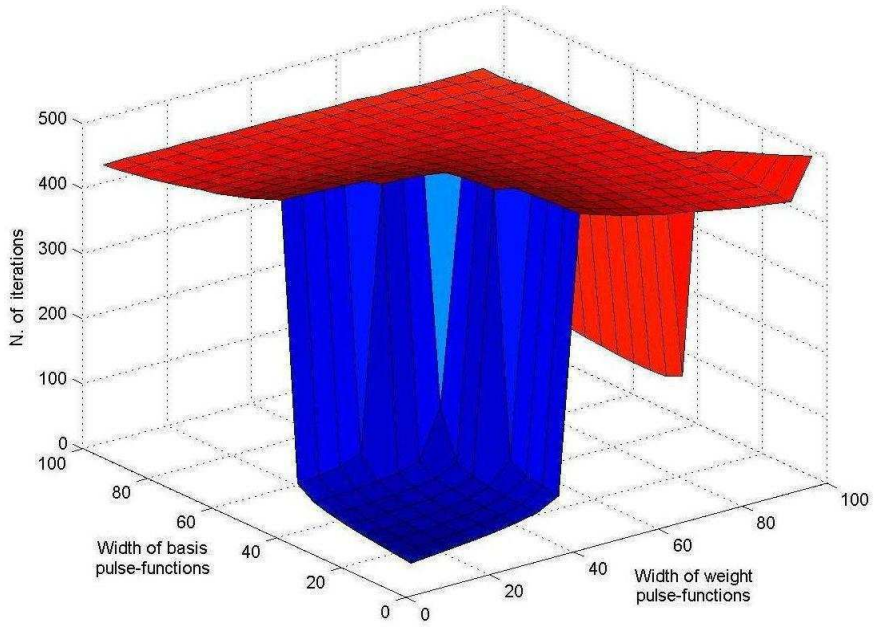
$$E_0 e^{i\beta}\left(x\cos\phi_i + y\sin\phi_i\right) = \frac{1}{4}\beta\eta_0 \int_0^{2w} H_0^{(2)}\left(k|x-\xi|\right) J(\xi) d\xi \tag{1}$$

arising from the plane wave TM_z scattering from a planar metallic strip, supposed to have finite width $2w$ and to be indefinitely extended in the direction of the z-axis [5]. The medium is the free

space, with constitutive parameters (ϵ_0, μ_0) and intrinsic impedance η_0. Equation (1) is indeed a standard Fredholm integral equation of the first kind in the unknown current density $J(\xi)$. Since there exists no chance to express its solution in a simple closed form, we have discretized it through the Method of Moments [6] and solved the subsequent system of linear equations by means of the GMRES. First we spend some words about the descent scheme in order to make clear our plan. As well-known, provided $\mathcal{U} = \{u_p(\cdot)\}_{p=1}^{\infty}$ is some Hilbert complete set of vectors for the unitary space $\mathcal{L}^2(a, b)$ on the complex field of all functions being square-summable on the segment $(0, 2w)$



(a)



(b)

Figure 1: Actual behavior of some numerical parameters as two-variable functions depending on the widthes Δ_u and Δ_w of our width-varying pulse-pulse implementation of the moment method. (a) Total number of iterations to 2-convergence with GMRES. (b) Condition number surface.

according to Lebesgue (even if in a generalized sense), we can expand the unknown current through

$$J(\xi) = \sum_{p=1}^{\infty} I_p u_p(\xi), \quad 0 < \xi < 2w \tag{2}$$

and then we form the moment linear system of equations

$$\mathbf{ZI} = \mathbf{V}, \tag{3}$$

by letting the (p, q)-entry in the impedance matrix $\mathbf{Z} \in \mathbb{C}^{n,n}$ be $\langle Au_q, w_p \rangle$ and arresting the process at a given level n of discretization. Here, $\langle \cdot, \cdot \rangle$ denotes the usual inner product in $\mathcal{L}^2(a, b)$ and $\mathcal{W} = \{w_q(\cdot)\}_{q=1}^{\infty}$ is another Hilbert basis of the same space (eventually, $\mathcal{W} = \mathcal{U}$). Furthermore, $A : \mathcal{L}^2(a, b) \mapsto \mathcal{L}^2(a, b)$ is the Hilbert-Schmidt linear operator realizing the mapping

$$u(\cdot) \to \int_0^{2w} H_0^{(2)}(\beta|x - \xi|)u(\xi)d\xi, \quad 0 < x < 2w \tag{4}$$

In fact, in the framework of our numerical experiments, we have set both $u_q(\cdot)$ and $w_p(\cdot)$ to be rectangular pulse functions centered in the points x_p and x_q, of width Δ_u and Δ_w, respectively. Here, $x_1, x_2, \ldots, x_n$ are the middle points of an uniform division of the real segment $(0, 2w)$ into n internally disjoint subintervals of constant width $\Delta = 2w/n$, so that $x_k = \frac{1}{2}(2k - 1)\Delta$ for any $k = 1, 2, \ldots, n$; and $0 < \Delta_u$, $\Delta_w \leq \Delta$. For completeness, we remark the rectangular pulse of width $T > 0$ centered in the point t_0 is the function

$$\mathbb{R} \mapsto \mathbb{R} : t \to \text{rect}\left(\frac{t - t_0}{T}\right) = \begin{cases} 1 & \text{if } |t - t_0| \leq \frac{1}{2}T \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

3. NUMERICAL RESULTS

Operatively, we have studied the way some numerical properties related to the system (3) are dependent on the actual choice of the basis and weight sets $\mathcal{U}_n = \{u_p(\cdot)\}_{p=1}^n$ and $\mathcal{W}_n = \{w_q(\cdot)\}_{q=1}^n$, by tuning the widthes Δ_u and Δ_w of their elements, respectively. In detail, it can be observed (see Fig. 1(a)) there is a critical threshold $\Delta_0 > 0$ such that, whenever Δ_u or Δ_w is greater than Δ_0, then the condition number κ (as a two-variable function of Δ_u and Δ_w) presents an abrupt growing, getting even many orders of magnitude larger than in the square $0 < \Delta_u$, $\Delta_w < \Delta_0$ and describing an irregular surface in the three dimensional space, whereas it maintains a very flat (practically constant) and relatively low profile as far as $0 < \Delta_u$, $\Delta_w < \Delta_0$. In principle, this wouldn't be particularly significant by itself, since the pick value attained by κ is still too small for resulting into trouble and causing a significant loss of numerical precision. This not standing, it is notable that, for the very same values of Δ_u and Δ_w corresponding to the entry into the instability region for the condition number κ, the total number of iterations necessary for the GMRES to pull down the residual vector 2-norm under a pre-fixed tolerance $\epsilon > 0$ (in our tests, $\epsilon = 1\text{e} - 6$) grows dramatically, provided that $\max(\Delta_u, \Delta_w) > \Delta_0$, in such a way that the *tout-court* convergence process is indeed very slow (see Fig. 1(b)) and maintains constant not depending on the actual values of the condition number. This seems to be related to the wide spreading of the spectrum through the complex plane, that can be observed entering the instability region, but we feel this is not enough to explain the numerical behavior of the GMRES, suggesting that deeper investigations are necessary for giving a truly satisfactory answer.

4. CONCLUSIONS

Our numerical experiments reveal that, even in the hypothesis of an uniform dissection of the real segment $[0, 2w]$ into n subintervals of amplitude Δ there exists a critical threshold Δ_0 such that, whenever either the basis or the weight pulses are given with an amplitude greater than Δ_0, then the total number of internal loops necessary for taking the relative residual under a definite tolerance $\epsilon > 0$ increases all of a sudden, in such a dramatic way that it can even prevent the process at all from convergence. Finally, we try to explain this numerical behavior by inquiring into the spectral properties of both the starting integral equation and the impedance matrix $\mathbf{Z}$ focusing our attention on its condition number. In conclusion, i) it seems that the condition number solely cannot be adopted as a reliable quantity to measure the effectiveness of the numerical solutions coming out

from the application of the method of moments (3) to the problems relevant with electromagnetics; ii) the choice of the basis and weight sets $\mathcal{U}$ and $\mathcal{W}$ is critical for ensuring the method can eventually result into a faithful discrete representation of the analytical operator equations.

REFERENCES

1. Chew, W. C., J. M. Jin, E. Michielssen, and J. Song, *Fast and Efficient Algorithms in Computational Electromagnetics*, Artech House, 2001.
2. Klein, C. A. and R. Mittra., "The effect of different testing functions in the moment method solution of thin-wire antenna problems," *IEEE Trans. on Antennas and Propagation*, 258–261, March 1975.
3. Peterson, A. F., C. F. Smith, and R. Mittra, "Eigenvalues of the moment-method matrix and their effect on the convergence of the conjugate gradient algorithm," *IEEE Trans. on Antennas and Propagation*, Vol. 36, No. 8, 1177–1179, 1988.
4. Saad, J., *Iterative Methods for Sparse Linear Systems*, SIAM, 2003.
5. Balanis, C. A., *Advanced Engineering Electromagnetics*, Wiley, 1989.
6. Harrington, R. F., *Field Computation by Moment Methods*, Krieger, Malabar, FL, 1982.

Simulation of Relevant Process Variables for Electrochemical Etching

R. Neugebauer, H. Knüpfer, K. Wolf, and H.-J. Roscher
Fraunhofer Institute for Machine Tools and Forming Technology IWU
Chemnitz, Germany

Abstract— The process of electrochemical etching requires the existence of an electrolytic bath and two electrodes with a certain distance in between, which is called the gap width. For an intensive etching process the gap between the electrodes should vary over short periods. The increase of the gap provides the possibility to rinse the removed material. Material which is ablated by the process of electrochemical etching in an electrolytic bath must be removed by rinsing. A periodic expanding of the working gap between electrode and work piece is essential for an intensive etching process. The duration of the rinsing process should be kept short for the reason of efficiency. Small working gaps lead to high contour accuracy during the etching process itself. The arising process forces reach to kilo Newton range. Thereby, a short circuit between electrode and work piece must be avoided. By the usage of new high dynamic actuators instead of powered drive solutions a profile of the electrode movement could be realized, which is better adapted to the process requirements. This article introduces a Simulink® model (under simplifying assumptions) for the determination of relations between the movement of the electrode and the arising forces in the working gap. The results are fundamental input data for dimensioning of actuating elements.

1. INTRODUCTION

Electrochemical etching, a kind of Electro Chemical Machining (ECM) is a machining technique by which metal removal is achieved by means of controlled dissolution process. This is done by means of a salt solution and a current supply. To connect the product to the positive pole and the electrode to a conductive negative pole. See Figure 1. When the electrons flow from the minus through the salt by means of the ions to the plus, the metal (work piece) will dissolute.

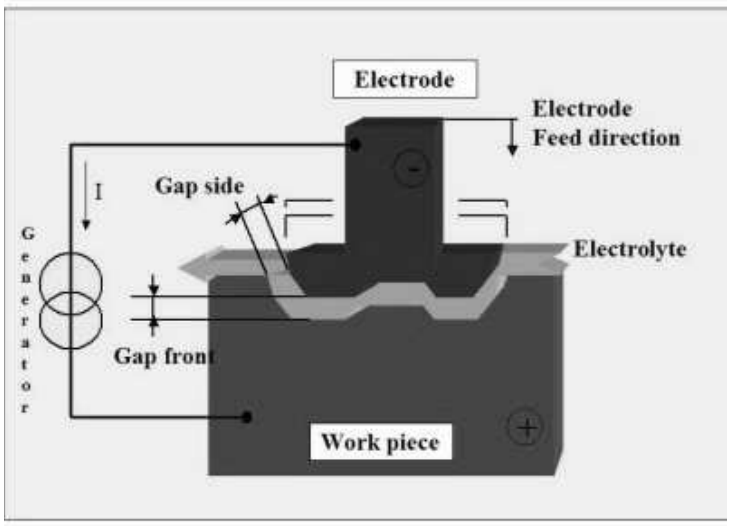


Figure 1: Electro Chemical Machining (ECM) — The process.

State-of-the-art for the adjustment of the gap size is to use a powered drive. A motor-driven tappet provides a sinusoidal curve of the gap width over time. A rectangular-shaped gap over short periods would be optimal. Moreover, the motor-driven approach requires large electrical engine power and high effort for the positioning accuracy of the tiny etching gaps. To overcome these problems high dynamic actuators, for example piezo-stacks, can be used because they can act with relatively high frequency and high accuracy while providing high reaction forces and the necessary stroke.

To prove the usability of different actuators simulations were done, which are introduced in the first section. This section also introduces the physical model of the system from which the differential equations and the Simulink® model are derived. The behavior of the system and conclusions concerning dimensioning of certain input parameters are discussed in the following section. Finally, the last section provides conclusions and an outlook.

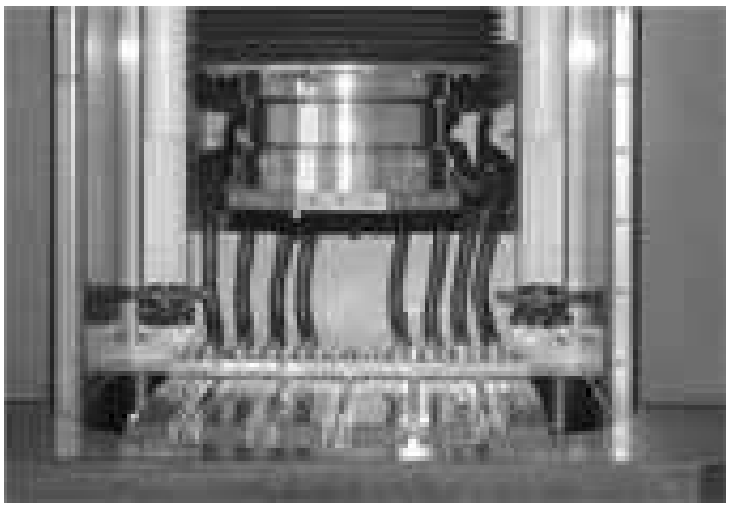


Figure 2: Electro Chemical Etching machine [4].

2. VIBRATION DRIVE — MODEL AND SOLUTION

2.1. Model of the Piezo-based Vibration Drive

The piezo-based vibration drive consists of a linear drive with large mass, the piezo actuator with stiffness c and the electrode with an additional mass m (see Figure 3). The piezo-stack performs a sinusoidal vibration which is superposed by the movement of the linear drive z_0 resulting in the total excitation

$$z(t) = z_0 + \hat{z}_p \cdot \sin(2 \cdot \pi \cdot f \cdot t). \tag{1}$$

Since the mass of the linear drive is large compared to the rest of the system, z_0 can be assumed to be constant.

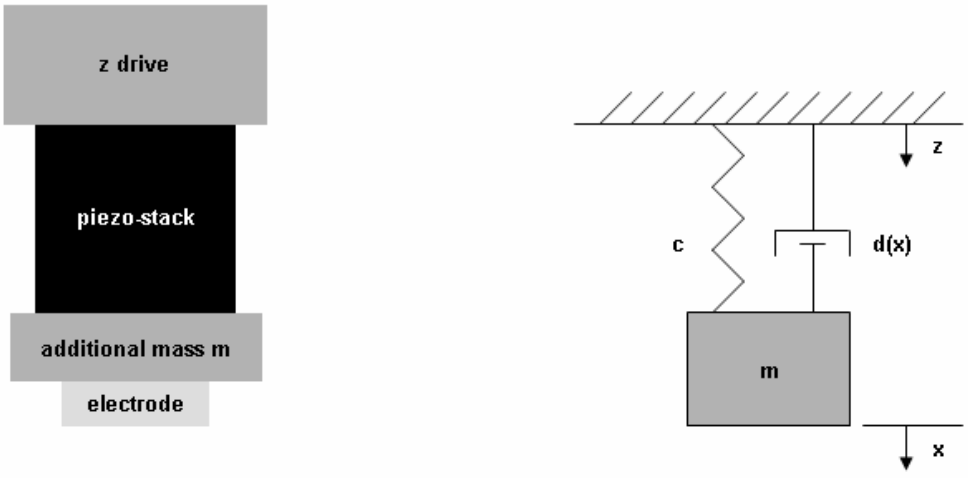


Figure 3: Model of the piezo-based vibration drive.

Inert forces of the mass m and squeezing forces $F_{squeeze}$ of the fluid in the working gap compress or stretch the spring c in the spring-mass-system by the difference of the excited movement z and

the movement x of the mass m to which the electrode is firmly connected:

$$c \cdot (z - x) = F_{squeeze} + m \cdot \frac{d^2x}{dt^2}. \tag{2}$$

The squeezing force acts as highly nonlinear damping and can be approximated by the term

$$F_{squeeze} = \frac{1}{2} \cdot \eta \cdot A^2 \cdot \frac{1}{x^3} \cdot \frac{dx}{dt}, \tag{3}$$

in which A denotes the shell area of the electrode and η the dynamic viscosity of the fluid (see [1]). The shell area, which is dependent on the gap width but considered to be constant, marks an additional flow resistance.

The following differential equation can be derived from Equations (1) to (3):

$$\frac{d^2x}{dt^2} + \frac{K}{m} \cdot \frac{1}{x^3} \cdot \frac{dx}{dt} + \frac{c}{m} \cdot x = \frac{c}{m} \cdot z \left(K = \frac{1}{2} \cdot \eta \cdot A^2 \right). \tag{4}$$

On basis of Equation (4) a Simulink® model, which can be seen in Figure 4, was built.

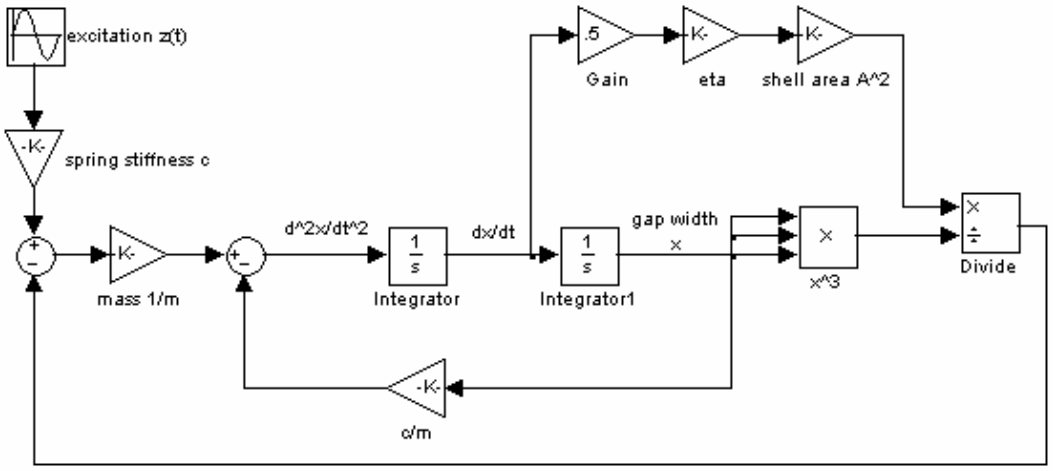


Figure 4: Simulink® model of the piezo-based vibration drive.

2.2. Results

For the first simulation the following parameters are set.

- Amplitude of the sinusoidal excitation of $\hat{z}_p = 100\,\mu\text{m}$
- Frequency of the excitation of $f = 50\,\text{Hz}$
- Starting position of $z_0 = 110\,\mu\text{m}$
- Shell surface area equals the electrode surface area of $A = 141.4\,\text{mm} \times 141.4\,\text{mm} = 20000\,\text{mm}^2$
- Dynamic viscosity of the fluid of $\eta = 0.001\,\text{Pa·s}$ (water)
- Mass of $m = 50\,\text{kg}$
- Spring stiffness of the piezo actuator $c = 300\,\text{N} \cdot \mu\text{m}^{-1}$

The simulation results can be seen in Figure 5.

Looking at the graph of the gap width a rectangular-like shape seems to be possible. The minimal working gap is about 40 micrometer and needs to be decreased. The system forces, i.e., the spring force which is driven by the inert force and the squeezing force, are also depicted in

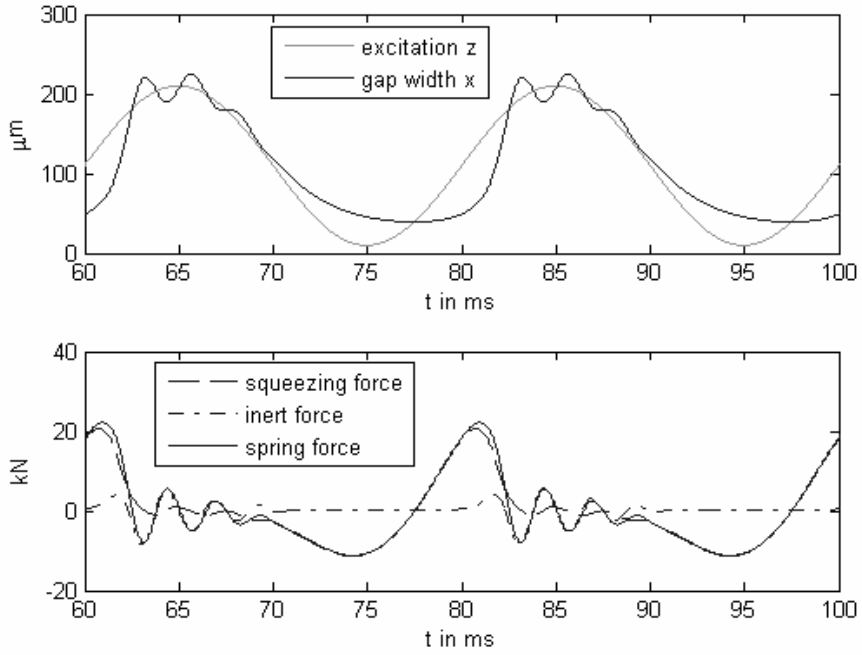


Figure 5: Simulation results of the system with the first settings.

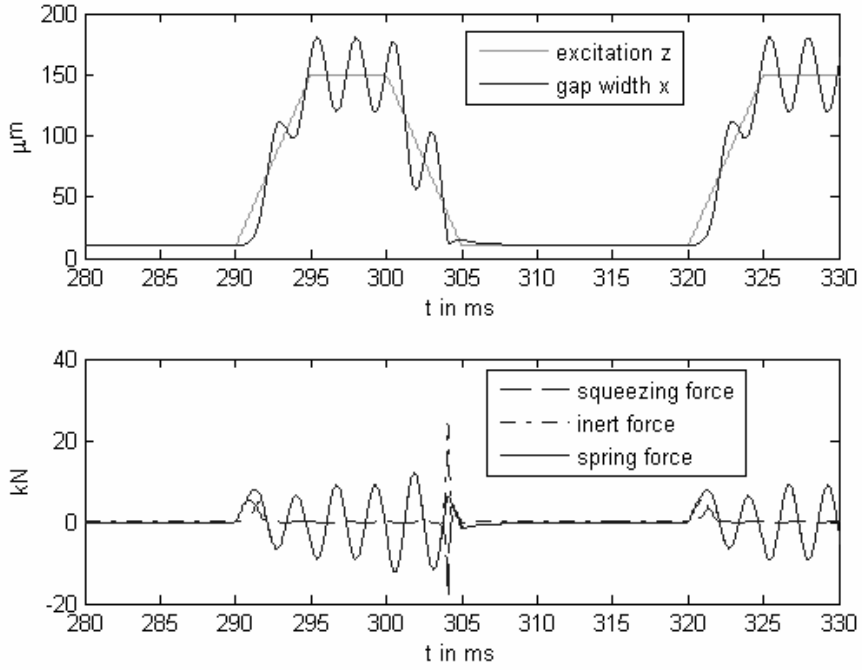


Figure 6: Simulation results of the system with small shell area and trapezoid excitation.

Figure 3. Because of the small mass, the inert force is relatively small and the spring force is vibrating quite fast.

To improve the system behavior a parameter variation was done. Varying the spring stiffness has a significant influence on the gap width. To increase the spring stiffness means to decrease the gap width and to decrease the vibration during the opening phase. The change of the frequency of the excitation does not result in a significant gap reduction but the decrease of the shell area does. An increase of mass leads to an increased kinematic energy. The large inert force leads to

less vibration of the spring force. The influence of the mass is much more severe in the opening phase because with the inert force also the squeezing force increases when the electrodes converge.

A shell area of 1600 mm^2 and a trapezoid excitation with periodic time of 30 ms result in a gap width which holds for about 17 milliseconds at about 10 micrometer and opens to average 150 micrometer (see Figure 6).

3. CONCLUSIONS AND OUTLOOK

As the simulations show, the realization of a piezo-based vibration drive is very promising. Since the estimated spring stiffness of the piezo actuator as well as the electrode mass are realistic values, they are retained. The working gap can be forced to be relatively constant caused by the nonlinear system properties and by applying a drive frequency which is advantageous to the resonance frequency. Additionally, a trapezoid excitation performed by the piezo actuator is reasonable to achieve the desired gap shape. The large reaction forces limit the minimal working gap to approximately 10 micrometer at relatively small electrode shell areas. The resulting opened gap should be large enough for the rinsing.

The next step is to verify the simulation results by doing measurements of the system. On basis of the results the model can be improved by considering the masses of the piezo-stack and the linear drive as well as its stiffness. Additionally, the shell area can be calculated depending on the gap width and the real dynamic viscosity of the liquid should be used.

REFERENCES

1. Kurth, S., "Theoretische und experimentelle Verhaltensbeschreibung kinetischer mikromechanischer Komponenten mit elektrostatischer Feldkopplung," Dissertation TU Chemnitz, D, 1995.
2. Griffin, W. S., H. H. Richardson, and S. Yamanami, "A study of fluid squeeze-film damping," *ASME J. Basic Eng.*, 451–456, 1966.
3. Mc Geough, J., *Principles of Electrochemical Machining*, Chapman & Hall, London, UK, 1974.
4. www.pemtec.de.

Singular Analytical Integration for Efficient Volume Integral Equation Implementation

F. J. Pérez Soler[1], **F. D. Quesada Pereira**[1], **A. Alvarez Melcon**[1], **and L. Peregrini**[2]
[1]Technical University of Cartagena, Campus Muralla del Mar s/n, Cartagena 30202, Spain
[2]University of Pavia, Strada Nuova 65, Pavia 27100, Italy

Abstract— We present a novel efficient implementation of the Volume Integral Equation formulation, for the analysis of printed circuits with finite size dielectric objects. The singularity of the Green's functions is extracted from the kernel of the integral equation, as static terms. These terms are then evaluated analytically for coincident rectangular cells in two and three dimensions. A semi-analytical formulation is also derived for the interactions involving adjacent cells. Results show the gain in efficiency and accuracy that can be obtained in the frame of the Volume Integral Equation formulation, when the new analytical techniques are used.

1. INTRODUCTION

The Volume Integral Equation technique (VIE) is a widely used formulation for the analysis of a large class of problems containing dielectric objects of finite dimensions and complex shapes. In this type of problems, Green's functions can not account for the dielectric objects in a semi-analytic fashion, and one has to resort to the use of the well known free-space Green's functions. With this formulation the dielectric objects are modeled with the aid of polarization currents defined on their volumes. The subsequent application of the Method of Moments (MoM) requires the computation of overlapping integrals between the free space Green's functions and the basis and testing functions. An interesting feature when developing this approach, is that surface charge distributions appear at the outer surfaces of homogeneous bodies. As a result, several combinations of surface and volume integrals need to be computed, before filling up the complete MoM matrix system. The whole situation is even more complicated, due to the singular behavior of the Green's functions used inside the Kernel of the integral equations.

In the above frame, the authors have derived an efficient implementation of the Volume Integral Equation (VIE) approach, by following different singularity extraction techniques, combined with analytical integration. First, a Taylor series expansion is applied to the free-space Green's functions, splitting the whole contribution in static and dynamic parts. The dynamic part thus obtained is not anymore singular, and it can be integrated numerically without effort. What remains, then, is the calculation of the static terms, containing the singular behavior of the Kernel. Now, the main difficulty in the implementation of this formulation is the integration of coincident basis and testing cells for these static terms. This is because of the well known singularity of the potential Green's functions. To overcome this problem, integration techniques based on polar and spherical coordinates have been developed in the past [1].

In this paper, we propose to apply analytical integration to these singular terms. The analytical integration increases accuracy, and reduces considerably the computational cost. The analytical integration is based on the dyadic identities reported in [2], which are extended for the first time to account for volume integration, and mixed surface-volume interactions, in rectangular and hexahedral domains. Besides, the authors have also treated for the first time the quasi-singular case of adjacent cells. For these adjacent cells, the fast variations due to the quasi-singular behavior of the kernel may introduce important errors during numerical integration. To overcome this problem, we propose in this paper to evaluate analytically the source contribution of the integrals for this quasi-singular situation.

The included results show the efficiency and accuracy of the new approach for the treatment of the singularity, in comparison with direct numerical integration. The presented results show big improvements in efficiency and accuracy during the analysis of practical devices using the VIE approach.

2. THEORY

The formulation under consideration is a general VIE technique including metallic and dielectric objects [3]. First, the static terms of the Green's functions are extracted following a Taylor series

expansion [4]. A standard VIE formulation leads to the following volume and surface integrals for the static terms:

$$I_1 = \int_S \int_{S'} \frac{1}{R_S} dS' dS \quad \text{On metallic surfaces} \tag{1}$$

$$I_2 = \int_S \vec{f_t} \int_{S'} \frac{1}{R_S} \vec{f_b} dS' dS \quad \text{On metallic surfaces} \tag{2}$$

$$I_3 = \int_V \int_{V'} \frac{1}{R_V} dV' dV \quad \text{In dielectric objects} \tag{3}$$

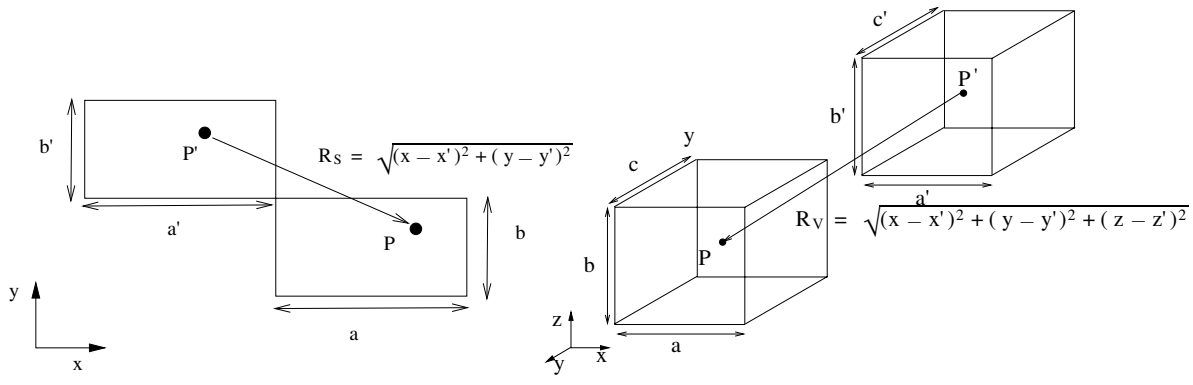
$$I_4 = \int_V \vec{f_t} \int_{V'} \frac{1}{R_V} \vec{f_b} dV' dV \quad \text{In dielectric objects} \tag{4}$$

$$I_5 = \int_V \int_{S'} \frac{1}{R_V} dS' dV \quad \text{On the surfaces of dielectric objects} \tag{5}$$

where $\vec{f_b}$ and $\vec{f_t}$ are the basis and testing functions employed in the MoM implementation. Also, R_S is the spatial distance between observation and source points, considering two dimensions (surface case), whereas R_V is the volumetric counterpart of this distance. Integrals (1) and (3) are those corresponding to the static MoM contributions of the electric scalar potential in two and three dimensions, respectively. On the other hand, (2) and (4) refer to the contributions of the magnetic vector potential in the same cases. Finally, (5) corresponds to a mixed surface-volume interaction due to the external charges distributed in the outer surfaces of homogeneous bodies.

Integrals I_1 to I_5 can be performed numerically using standard quadrature rules, when source and observation integration domains do not coincide. For the coincident cases $S = S'$ and $V = V'$, a different strategy must be adopted in order to avoid numerical problems and serious inaccuracies. A useful approach consists in transforming the source and observation points into polar coordinates. It is well known that these transformations mathematically cancel the term $(1/R)$. The main disadvantage of this method is that many integration points must be considered to assure convergence in the numerical integrals.

To overcome this problem, a new strategy is proposed in this paper, based on the analytical integration of terms of the form $(1/R)$, for coincident surface and volume domains. The procedure for triangular surfaces was presented in [2]. Based on this work, the authors have adpated the procedure to rectangular domains, for both surface and volume cases (as illustrated in Fig. 1).



(a) 2D case: rectangular cells (b) 3D case: hexahedral cells

Figure 1: General geometry of the integration domains considered in this paper.

2.1. Identities and Integrals for Rectangular Domains

In [2], the authors started from a set of useful algebraic identities which helped to reduce the complexity of the MoM integrals. In two dimensions, and for rectangular surfaces, these identities

have the form:

$$\frac{1}{R_S} = -\nabla_S \cdot \nabla'_S R_S \tag{6}$$

$$\frac{\bar{I}}{R_S} = \nabla_S \times \nabla'_S \times \bar{I} R_S - \nabla_S \cdot \nabla'_S R_S \tag{7}$$

$$\left(\nabla'_S \times \bar{I} R_S \cdot \vec{f}_b\right) = \nabla'_S \times \left(R_S \cdot \vec{f}_b\right) \tag{8}$$

$$\vec{f}_t \cdot \left(\nabla_S \nabla'_S R_S\right) \cdot \vec{f}_b = \nabla_S \cdot \left(\nabla'_S \cdot R_S \left[\vec{f}_t \vec{f}_b + \frac{1}{3}\left(\vec{f}_t \vec{R}_S - \vec{R}_S \vec{f}_b\right) - \frac{1}{9}\bar{I} R_S^2\right]\right) \tag{9}$$

Here, the subscript S indicates that the spatial distance and the differential operators are expressed in two-dimensional coordinates (case of Fig. 1(a)). These are the same expressions that appear in [2] for triangular domains, except for the case of (9), where the constants $\frac{1}{3}$ and $\frac{1}{9}$ change depending on the definition of the basis functions employed (rectangular for our case).

These identities can now be extended to the 3D case of tetrahedral cells, resulting into:

$$\frac{2}{R_V} = -\nabla_V \cdot \nabla'_V R_V \tag{10}$$

$$\frac{2\bar{I}}{R_V} = \nabla_V \times \nabla'_V \times \bar{I} R_V - \nabla_V \cdot \nabla'_V R_V \tag{11}$$

$$\left(\nabla'_V \times \bar{I} R_V \cdot \vec{f}_b\right) = \nabla'_V \times \left(R_V \cdot \vec{f}_b\right) \tag{12}$$

$$\vec{f}_t \cdot \left(\nabla_V \nabla'_V R_V\right) \cdot \vec{f}_b = \nabla_V \cdot \left(\nabla'_V \cdot R_V \left[\vec{f}_t \vec{f}_b + \frac{1}{4}\left(\vec{f}_t \vec{R}_S - \vec{R}_S \vec{f}_b\right) - \frac{1}{12}\bar{I} R_V^2\right]\right) \tag{13}$$

Now, the subscript V denotes that operations are performed in three dimensions (Fig. 1(b)). It can be noticed that these new identities maintain the same form as the previous ones, differing only in some constants that appear due to the additional spatial dimension. Finally, because we are also interested in mixed surface-volume interactions, an additional identity must be added to complete the whole set, namely:

$$\frac{2}{R_V} = -\nabla_S \cdot \nabla'_S R_V + \frac{\partial^2}{\partial z^2} R_V \tag{14}$$

Using these identities, in combination with the well known Gauss, Stokes and Curl theorems, the integrals expressed in (1) to (5) can be transformed into:

$$I_1 = -\int_C \int_{C'} R_S \vec{u} \vec{u}' dl' dl \tag{15}$$

$$I_2 = -\int_C \int_{C'} R_S \vec{u} \cdot \left[\vec{f}_t \vec{f}_b - \vec{f}_b \vec{f}_t + \frac{1}{3}\left(\vec{f}_t \vec{R}_S - \vec{R}_S \vec{f}_b\right) + \bar{I}\left(\vec{f}_t \cdot \vec{f}_b - \frac{1}{9} R_S\right)\right] \vec{u}' dl' dl \tag{16}$$

$$I_3 = -\frac{1}{2}\int_S \int_{S'} R_V \vec{u} \vec{u}' dS' dS \tag{17}$$

$$I_4 = -\frac{1}{2}\int_S \int_{S'} R_V \vec{u} \cdot \left[\vec{f}_t \vec{f}_b - \vec{f}_b \vec{f}_t + \frac{1}{4}\left(\vec{f}_t \vec{R}_V - \vec{R}_V \vec{f}_b\right) + \bar{I}\left(\vec{f}_t \cdot \vec{f}_b - \frac{1}{12} R_V\right)\right] \vec{u}' dl' dl \tag{18}$$

$$I_5 = \frac{1}{2}\left(-\int_0^c \int_S \int_{S'} R_V \vec{u} \vec{u}' dS' dS + \int_S \int_{S'} \frac{c}{R_V|_{(z-z')=c}} dS' dS\right) \tag{19}$$

In above expressions, C and C' denote the linear contours of the integration cells, and the vectors $\vec{u}$, $\vec{u}'$ are the outward vectors normal to the observation and source integration domains (linear or surface depending on the case), respectively. It is interesting to note that, due to the application of the identities, the term $(1/R)$ has changed to (R), which is not singular for $(R \to 0)$. Also, the complexity of the integration domains have been reduced in one dimension for each case. Hence, it is possible to evaluate analytically the expressions (15) to (18). Although the integration domains are simple, the final analytical results for all the integrals require very long

calculus and algebraic manipulations, specially in the case of the volumetric cells. These expressions are composed of simple logarithmic and inverse of trigonometric functions, depending only on the size cell dimensions. They can be evaluated, therefore, very fast with a computer. They are not included here for the sake of space.

Apart from the singularity cases already explained, other important sources of numerical errors are the MoM interactions between very close surface cells. Specially in the cases of adjacent cells (one or two common nodes), the numerical techniques applied to (1) and (2) do not behave properly, due to the fast variations of the Kernel for small source-observer distances. In order to minimize this effect, we propose to evaluate this kind of interactions following a semi-analytical procedure. The strategy consists in computing analytically the source contribution of I_1 and I_2, while the integral extended to the observer cell is computed numerically. The analytical integration on the source cell effectively eliminates the fast variations of the integral equation Kernel. Consequently, the numerical integration on the observer cell converges very fast.

3. RESULTS

In this section we will show the advantages and improvements of the new proposed approach. The calculation of the integrals has been performed for the MoM analysis of a practical microstrip device following the VIE formulation. The geometry under consideration is a metallic strip line of dimensions $0.8 \times 40\,\text{mm}$ printed over a $1\,\text{mm}$ substrate of Alumina ($\varepsilon_r = 9.8$), and shielded in a metallic box of dimensions $40 \times 40 \times 4\,\text{mm}$.

We present in Fig. 2 the error obtained for the volume interactions computed numerically with the transformation to spherical coordinates, as compared to the closed-form solution obtained with the analytical expressions derived in this work. It can be noticed in the figure that errors below 10^{-3} are obtained when more than 10^3 points are used in the numerical integration. When 14^3 points are used, the total static MoM matrix needs 7.2 minutes to be computed, whereas only 17.1 seconds are needed when using the proposed analytical approach. This shows the bottleneck problem of the numerical method, when computing the singular interactions with high accuracy.

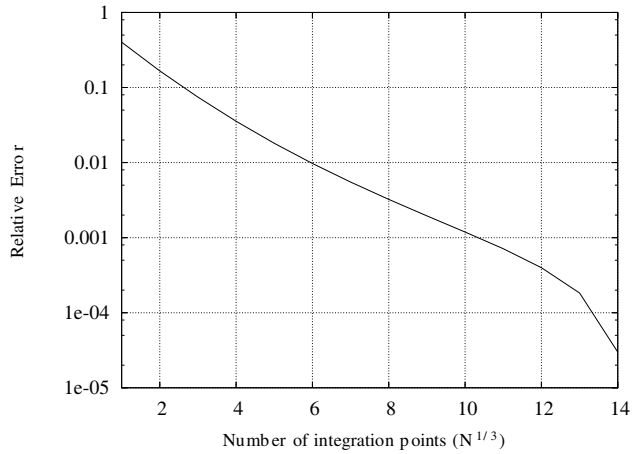


Figure 2: Relative error between numerical and analytical values of I_3 as a function of the number of quadrature integration points.

Now, we will examine the behavior of the semi-analytical approach for adjacent interactions. In Fig. 3, the efficiency of this new approach is compared to a standard numerical quadrature integration. First, Fig. 3(a) shows the relative errors obtained between the numerical and the semi-analytical values of I_1 and I_2, as a function of the distance between the centers of the source and observation cells. The number of quadrature points employed in the two methods has been optimized according to this distance. It can be observed that, as the distance decreases, the relative error grows, until significant values are obtained for the case of two adjacent cells (the first three points in the graphic). The discontinuity peaks that appear in the curves occur when the observation cell changes from one row to another row inside the mesh structure. As it can be noticed, the error is larger for the I_2 integral, where the linear variations of the basis functions are present. This is because these linear variations contribute to increase the fast variations of the Kernel for small source-observer distances.

On the other hand, Fig. 3(b) shows the effects of the number of quadrature points employed in the evaluation of the integrals using the two methods (pure numerical and semi-analytical). For this test we have considered the most critical case of two adjacent cells. The relative errors are taken with respect to the value obtained with a sufficiently large number of quadrature points (60) for all cases. It is interesting to observe how the analytical integration of the source contribution in I_1 allows to reduce the error several orders of magnitude with respect to the standard case of the double numerical integration. This fact helps to efficiently reduce the maximum number of quadrature points, and still achieve high accuracy.

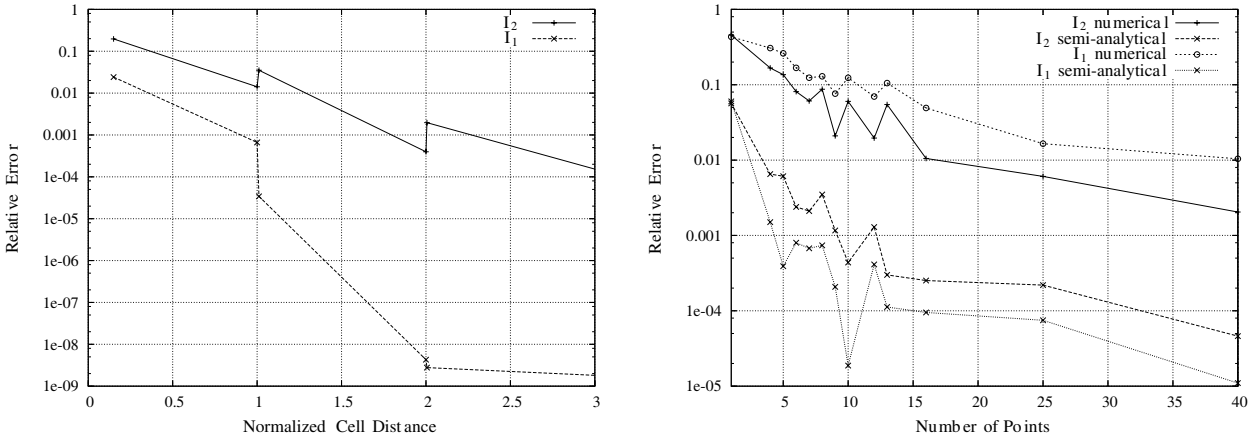


Figure 3: Comparison of numerical and semi-analytical approaches for the calculation of I_1 and I_2. Sizes of the cells are $a = 0.4\,\text{mm}$, $b = 2.66\,\text{mm}$, (a) Relative error as a function of the distance between cells, (b) Relative error as a function of the number of quadrature integration points.

4. CONCLUSIONS

In this paper we have presented a new efficient technique for evaluating the singular contributions in the general frame of a volume integral equation (VIE) formulation solved by MoM. First, the singular behavior is extracted from the free space Green's functions, and it is isolated in static terms for the vector and scalar potentials. A novel analytical integration in rectangular domains is then applied to these static terms. It is shown that the new technique improves the accuracy and reduces the computational cost, as compared to pure numerical integration. Also, the quasi-singular case of adjacent cells has been evaluated semi-analytically, reducing in this way the errors introduced by standard numerical integration. Useful numerical results are presented to confirm the practical value of the new approach.

ACKNOWLEDGMENT

The authors thank Ministerio de Educación y Ciencia of Spain, which has supported this work under Grant TEC2007-67630-C03-02/TCM.

REFERENCES

1. Pereira, F. Q., J. L. G. T. D. C. Rebenaque, J. P. Garcia, and A. A. Melcon, "Analysis of finite microstrip structures using an efficient implementation of the integral equation technique," *Radio Science*, Vol. 2004rs003036, January 2005, in print.
2. Arcioni, P., M. Bressan, and L. Perregrini, "On the evaluation of the double surface integrals arising in the application of the boundary integral method to 3-d problems," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 3, 436–439, 1997.
3. Thomas, J. B., "An integral equation approach to the analysis of finite microstrip antennas: Volumen/surface formulation," *IEEE Transactions on Antennas and Propagation*, Vol. 38, No. 3, 305–312, March 1990.
4. Soler, F. P., F. Q. Pereira, A. A. Melcon, V. B. Esbert, and B. G. Martinez, "Efficient software tool for the analysis of planar based metamaterial structures," *EuCAP, European Conference on Antennas and Propagation*, 1–7, IET, Edinburgh, UK, November 11–16, 2007. Fr2.2.10.

A Study on the Interference in Single Frequency Network and On-Channel Repeater

Sung-Woong Choi and Heon-Jin Hong
ETRI-Radio & Broadcasting Technology Lab., Radio Technology Department
161 Kajong-Dong, Yusong-Gu, Taejon 305-350, Korea

Abstract— In digital TV, SFN (Single Frequency Network) and OCR (On-Channel Repeater) are often considered for the efficiency of frequency allotment. Both of them broadcast using the same frequency. In this paper, we discuss the performance and evaluate some coverage criterions for SFN and OCR. Also, we propose method and simulation for coverage planning and estimation.

1. INTRODUCTION

Broadcasting had been used during long period by efficient method that can carry the large quantity of information to many people. To carry much information and satisfy various consumer's desire, various digital broadcasting techniques are introduced and studied recently. Because such new broadcasting system must co-exist existing broadcasting system during some period, frequency assignment that is different from existing frequency is required.

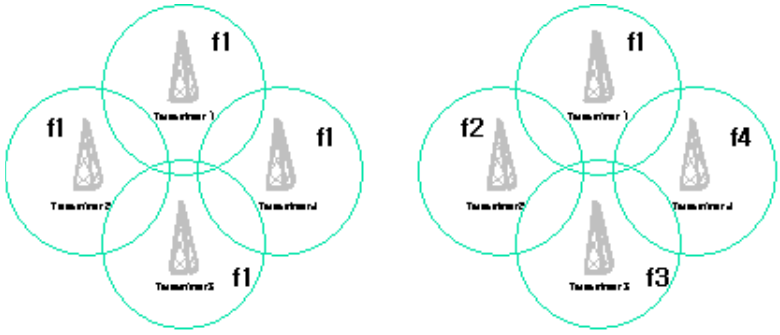
Because of limited frequency resource, most digital broadcasting system is supplying technique of SFN (Single Frequency Network) or OCR (On-Channel Repeater) etc. to use minimum frequency resources. In this paper, we analyze the scheme to apply SFN of DVB-T [1, 2] and OCR of ATSC [3, 6] in domestic, and propose to simulate coverage area by position of relay station and transmitter.

2. THE CONCEPT OF THE SFN AND OCR

2.1. The SFN Concept

The analog terrestrial broadcasting utilizes MFN (Multi-Frequency Network), a transmission scheme that uses a different transmitting frequency in each service area. Hundreds of stations are established for national broadcasting service, therefore much frequencies are used.

From late 1980s, using OFDM in Europe or Japan, technology that operate on the same frequency in adjacent service area had been examined, this is SFN. Single Frequency Network is two or more transmissions that operate on the same frequency, broadcast the same program so as to fill gaps within or extend the service provided by a single transmitter. All transmitters in an SFN must be synchronized so that their broadcasts are frequency identical and bit identical. SFN operation therefore requires special equipment in the primary distribution network to ensure this synchronization using a universal time and frequency reference such as that available from the GPS satellite system. Considering an SFN, the distance between two adjacent transmitter stations determines the necessary length of the guard interval. Figure 1 shows the SFN and MFN.



(a) SFN (b) MFN

Figure 1: Single Frequency Network and Multi Frequency Network.

2.2. The OCR Concept

The OCR allows rebroadcast of a TV signal, without frequency shifting, into an area previously unable to receive the originally transmitted signal. OCR consists of non-regenerative repeater and regenerative repeater.

Non-regenerative repeater receives signal of transmitting station, simply amplifies, filters and retransmit the signal. The received signal in the repeater service area consist of the signal of main transmitting station and signal of repeater itself, and each is regarded as delayed multi-path signal. The regenerative repeater demodulates perfectly the received signal of the transmitting station, re-modulates, and re-transmit the signal. OCR signal is not coherent with the primary signal. New signal acts like co-channel interference. In case of the co-channel interference reception is possible in situation that D/U ratio is more than 15 dB, In case of multi-path, because the equalizer of receiver compensate this, reception is possible in situation that D/U ratio is up to 5 dB.

Usually Equalizer's performance in the VSB receiver is decided according to filter Tap number. Forward transverse filter removes the Pre-ghost and decision feedback filter removes Post-ghost. Sampling rate of the VSB receiver equalizer is 10.762 MHz ($T = 0.09292\,\mu s$). So it is easily calculated by following equation.

- Pre-ghost = 64-tap $\times\, T = 5.947\,\mu s \approx 5\,\mu s$
- Post-ghost = 192-tap $\times\, T = 17.84\,\mu s \approx 18\,\mu s$

As see in Figure 2, early model (Chip P) of Equalizer can handle Pre-ghost about 5 μs, and could handle Post-ghost about 18 μs, but present model can handle about 40 μs.

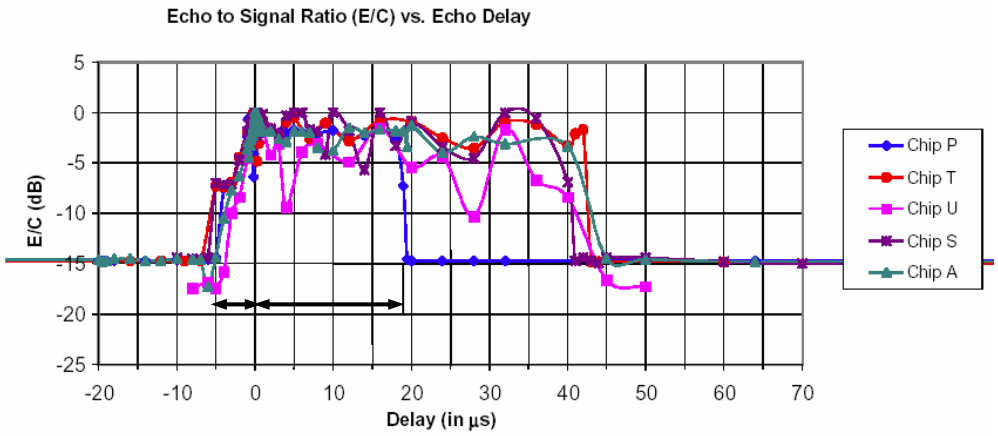


Figure 2: Receiver performance according to the chipset.

3. SIMULATION OF SFN AND OCR

3.1. SFN Simulation

In single frequency network, we must forecast reception area according to the distance between transmitter and receiver, modulation, code rate, guard interval. In case that arrival time difference of the receiving signals is within the guard interval, reception area is simply depended on path difference between transmitter and receiver. But, in case that arrival time difference exceeds the guard interval, two signals are acted by noise each other. So reception area should be decided by C/N ratio by modulation, encoding rate, channel characteristic etc.

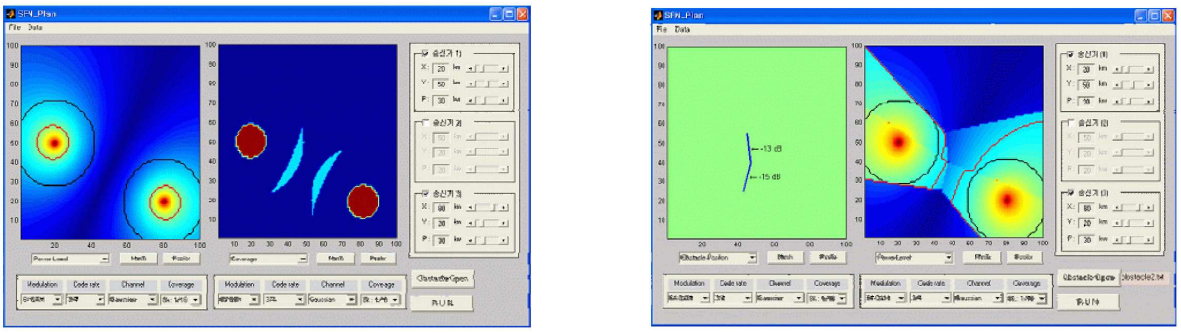
First, the position and transmit power are established, and modulation mode, the encoding rate, channel, guard interval are selected. And we simulate using position of two transmitters and attenuation value when obstacles are added. The following parameters for SFN simulation were described in detail.

Input

- Transmitter: Up to 3, position, transmit power
- Transmission mode: Mode 1/4, 1/8, 1/16, 1/32 guard interval

- modulation: QPSK, 16-QAM, 64-QAMs
- The encoding rate: 1/2, 2/3, 3/4, 5/6, 7/8
- Channel: Gaussian, Ricean, Rayleigh

In case that two transmitting stations transmit the equal power of 30 kW with Figure 3, we can see the distribution of received power difference from two transmitting stations. Figure 3(b) shows when obstacle exists.



a) power difference and reception area b) obstacle added-reception area

Figure 3: SFN simulation (two station of 30 kW).

3.2. OCR Simulation

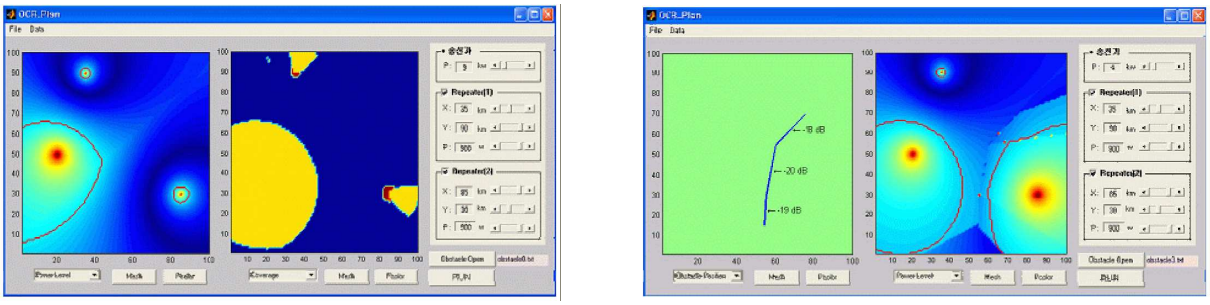
In OCR we must forecast reception area according to the distance between the main transmitter and repeater and repeater isolation. Because the regenerative repeater demodulates the received signal of the transmitting station, re-modulates and transmit this, much times are delayed. So this paper considers the non-regenerative repeater. In this paper, we designed to simulate the reception area of OCR Considering mentioned item in front. The following parameters for OCR simulation were described in detail.

Input

- Transmitter: Position, transmit power
- Repeater: Up to 2, position, transmit power

Output

- Distance and power from transmitter and repeater
- Power difference from transmitter and repeater
- Area of Pre-ghost and Post-ghost



a) power difference and reception area b) obstacle added-reception area

Figure 4: OCR simulation (one main transmitter of 9 kW and two repeaters of 900 W).

As see in Figure 4, when one main transmitter of 9 kW and two repeaters of 900 W transmit, the Power difference is big around the main transmitter.

4. CONCLUSION

In this paper we analyze the SFN and OCR that can efficiently use limited domestic frequency resources and developed simulator to forecast reception coverage. In the developed simulator, SFN simulator considered channel required C/N ratio, different guard interval and modulation, and OCR simulator considered distance between the transmitter and repeater, isolation.

Present proposed simulator supposed simple model that do not consider the effect of obstacle, using the basic propagation model. But we have plan to extend the simulator to consider Non-Line of Sight or simple topography factor.

REFERENCES

1. EN 300 744 V1.4.1, "Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television," 2001.
2. TR 101 190 V1.1.1, "Digital Video Broadcasting (DVB); Implementation guidelines for DVB terrestrial services; Transmission aspects," 1997.
3. ATSC Revision 1.0, "Performance Assessment of the ATSC Transmission System, Equipment and Future Directions," 2001.
4. ETRI, "DTV on channel repeater," 2000.
5. ETRI, "The examination result terrestrial regenerative OCR," 2000.
6. ATSC, "Guide to the use of the ATSC Digital Television Standard," Doc. A/54, 1995.

A Novel Ultra-wideband Bandpass Filter

I-Tseng Tang[1], **Ding-Bing Lin**[2], **Chi-Min Li**[3], **and Min-Yuan Chiu**[3]

[1]Department of Environment and Energy, Nation University of Tainan, Tainan, Taiwan
[2]Institute of Computer and Communication Engineering
National Taipei University of Technology, Taipei, Taiwan
[3]Department of Communications and Guidance Engineering
Nation Taiwan Ocean University, Keelung, Taiwan

Abstract— A novel ultra-wideband (UWB) band-pass filter (BPF) with a fractional bandwidth (FBW) of about 110% is proposed using a microstrip structure. The proposed UWB band-pass filter structure using five short-circuited stubs with reverse double U-shape defected ground structure (RDU-DGS), which employ a very wide bandwidth from 3.1 GHz to 10.6 GHz and very wide notch band in the undesired band. The new proposed filter has a good performance, including an ultra-wideband bandpass (3.1–10.6 GHz), a small size, a low insertion loss, and sharp rejection. The filter also demonstrated an UWB reject band from 12.31 GHz to more than 20 GHz at −20 dB.

1. INTRODUCTION

In 2002, the Federal Communications Commission (FCC) approved the use of ultra-wideband (UWB) (range of 3.1–10.6 GHz) for commercial purposes [1]. Recently, the development of UWB enabling high data transmission rates and low power consumption, and simple hardware configuration in communication applications such RFID devices, sensor networks, radar and location tracking, have received attention. This rise in the field of UWB technology has pushed for the development of new UWB filters has increased via different methods and structures [2–6]. In order to accurate, filter design already from lumped elements (resistance, capacitance and inductance) to microstrip components. Lumped-element filter design is generally unpopular due to the difficulty of its use at microwave frequencies along with the limitations of lumped-element values [7]. Hence, conventional microstrip filters are often used.

The new proposed filter design based on quarter wavelength short-circuited stub. In order to reduce filter size, bending connecting line and let five short-circuited stubs via the same hole. Consequently, hexagon UWB filter was modeled. Nevertheless, spurious mode also excited out-of-band, for dispose of spurious band that used reverse double U-shape defected ground structure (RDU-DGS) to eliminate spurious bands. Cascade different dimension of RDU-DGS can act different attenuation poles in the different frequencies, which consist of a wider band-notch.

Here, a novel UWB BPF structure with wide upper stopband is proposed. The structure is also compact in microstrip line direction than those reported in [8–10].

2. FILTER DESIGN

The filter using five short-circuited stubs was designed for an optimum distributed high-pass filter whose connecting lines are non-redundant ($\theta_C = 2\theta_S$) [11]. Thus, the filter can exhibit a frequency selectivity equivalent to that of a conventional 9-pole Chebyshev filter. But an optimum distributed high-pass filter requires a greater area. In order to reduce the filter, the connecting line had been

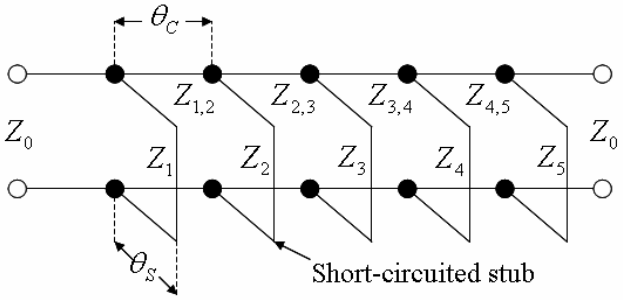


Figure 1: Conventional of five short-circuited stubs filter.

reduced $\theta_C = \theta_S$ is shown in Fig. 1. Thus, let 9-pole decrease to 5-pole, even though frequency selectivity of 5-pole is not better than 9-pole. However, circuit dimension can be reduced very much. The characteristic impedances of these short-circuited stubs are defined by Z_1, Z_2, Z_3, Z_4, and Z_5, and the characteristic impedances of the connecting lines are defined by $Z_{1,2}$ to $Z_{4,5}$. The terminal impedance is defined as Z_0. The electrical length was chosen to be $\theta_C = \theta_S = 35°$ for the connecting lines and the short-circuited stubs at 3.1 GHz. The characteristic impedances of these short-circuited stubs and the connecting lines, which are listed in Table 1.

Table 1: Circuit parameters for UWB band-pass filter with five short-circuited stubs when $\theta_C = \theta_S = 35°$.

Stub line impedance	Connecting line impedance
$Z_1 = Z_5 = 106.62\,\Omega$	$Z_{1,2} = Z_{4,5} = 50\,\Omega$
$Z_2 = Z_4 = 75.65\,\Omega$	$Z_{2,3} = Z_{3,4} = 50\,\Omega$
$Z_3 = 69.04\,\Omega$	$Z_0 = 50\,\Omega$

3. RESULTS AND DISCUSSIONS

Simulation design for the proposed UWB band-pass filter is carried out on the full wave simulator. To reduce the size of the conventional filter, the microstrip line was bent as a hexagon, and also these quarter-wavelength short-circuited stubs are all connected to one common via-hole, as shown in Fig. 2. There is not effect when connecting line has been bent. The five short-circuited stubs hexagon filter with a fractional bandwidth of more than 110% are presented, is shown in Fig. 3.

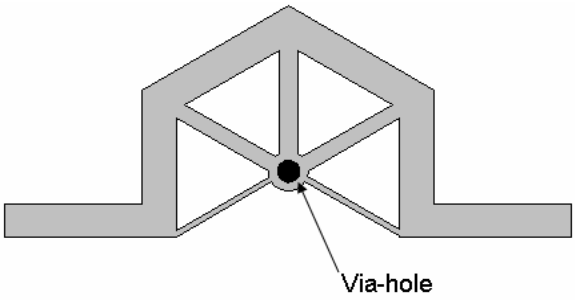


Figure 2: Structure of the novel proposed UWB filter.

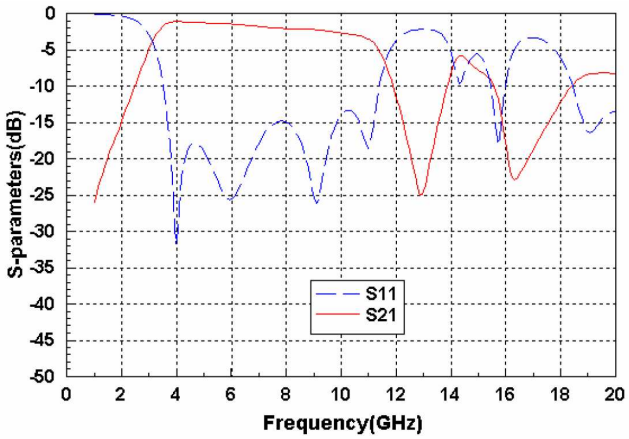


Figure 3: Simulation of the novel proposed UWB filter.

Microstrip line structure often produced periodic spurious bands, although it is an important issue when design microstrip filter. In this study, the UWB BPF added RDU-DGS to eliminate

the spurious bands, is shown in Fig. 4. There are three RDU-DGS in the ground plane, which have been set the left side of fed-in line, between the middle short-circuited stub and connecting line, and the right side of fed-in line. Those reported in [12], resistance, inductance and capacitance value can be calculated, furthermore the length of the RDU-DGS was been found. Dimension of left side of the RDU-DGS $a = 2.7\,\text{mm}$, $b = 2.1\,\text{mm}$, $c = 1.5\,\text{mm}$, dimension of middle side of the RDU-DGS $a = 2.62\,\text{mm}$, $b = 1.7\,\text{mm}$, $c = 1.42\,\text{mm}$, dimension of right side of the RDU-DGS $a = 2.7\,\text{mm}$, $b = 1.5\,\text{mm}$, $c = 1.5\,\text{mm}$, the gap between of inner and outer U-shape is 0.3 mm. Besides, the left section of $b = 2.1\,\text{mm}$, which is suppressed 13–15 GHz, the middle section of $b = 1.7\,\text{mm}$, which is suppressed 15–17 GHz, and the right section of $b = 1.5\,\text{mm}$, which is suppressed 17–20 GHz. Therefore laid aside these three RDU-DGS to be allowed out-of-band 11–20 GHz suppression below $-20\,\text{dB}$ is shown in Fig. 5.

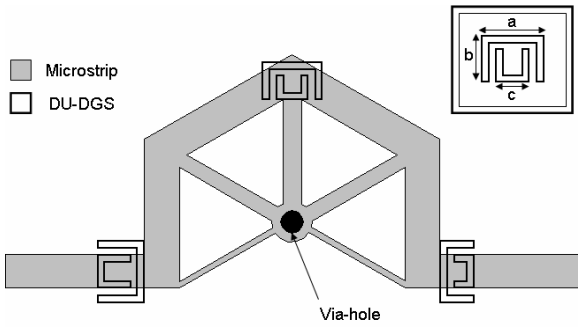


Figure 4: Structure of the novel proposed UWB filter with RDU-DGS.

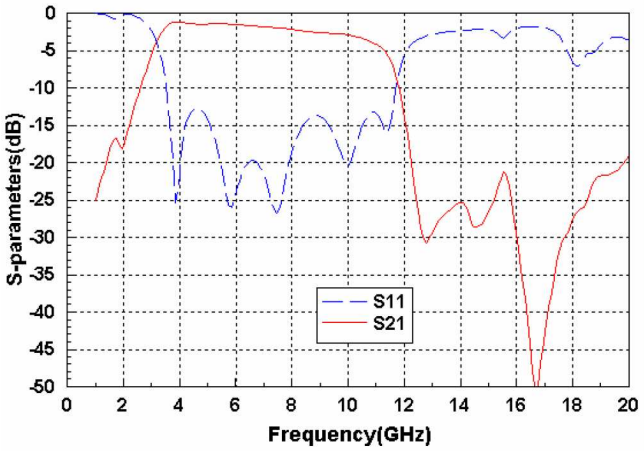


Figure 5: Simulation of the novel proposed UWB filter with RDU-DGS.

4. CONCLUSIONS

New proposed UWB band-pass filter with a FBW of about 110% has been successfully demonstrated. The filter design was based on a circuit model for a quarter-wavelength distributed band-pass filter comprised of five short-circuited stubs separated by bend redundant connecting lines. In order to improve the out-of-band performance of the filter, RDU-DGS band-stop filter was added. As such, an attractive UWB band-pass filter with five transmission poles and low insertion losses less than $-1.5\,\text{dB}$ are realized in theory and confirmed via experiment. In addition, the proposed UWB band-pass filter with RDU-DGS has a very small overall length of 16.67 mm that is approximately equal to one full wavelength at 6.85 GHz.

ACKNOWLEDGMENT

This work was sponsored by the National Science Council under the contract number of NSC 95-2221-E-024-025.

REFERENCES

1. Federal Communications Commission, "Revision of Part 15 of the commission's rules regarding ultra-wide-band transmission system," First note and order, ET-Docket 98-153, Feb. 14, 2002.
2. Saito, A., H. Harada, and A. Nishikata, "Development of band pass filter for ultra wideband (UWB) communication systems," *Proc. IEEE Conf. Ultra Wideband Systems and Technology*, 76–80, 2003.
3. Ishida, H. and K. Araki, "Design and analysis of UWB band pass filter with ring filter," *IEEE MTT-S Int. Dig.*, 1307–1310, Jun. 2004.
4. Chin, K., L. Lin, and J. Kuo, "New formulas for synthesizing microstrip bandpass filters with relatively wide bandwidths," *IEEE Microwave and Guided Wave Letters*, Vol. 14, No. 5, 231–233, Mar. 2004.
5. Li, K., D. Kurita, and T. Matsui, "An ultra-wideband bandpass filter using broadside-coupled microstrip-coplanar waveguide structure," *IEEE MTT-S Int. Dig.*, 657–678, Jun. 2005.
6. Hsu, C.-L., F.-C. Hsu, and J.-T. Kuo, "Microstrip bandpass filters for ultra-wideband (UWB) wireless communications," *IEEE MTT-S Int. Dig.*, 679–682, Jun. 2005.
7. Pozar, D. M., *Microwave Engineering*, 2nd Ed., John Wiley & Sons, New York, 1998.
8. Hong, J. S. and H. Shaman, "An optimum ultra-wideband microstrip filter," *Microw. Opt. Technol. Lett.*, Vol. 47, No. 3, 230–233, Nov. 2005.
9. Hong, J. S. and H. Shaman, "An optimum ultra-wideband bandpass filter with spurious response suppression," *IEEE WAMICON*, 1–5, Dec. 2006.
10. Hong, J. S. and H. Shaman, "A compact ulrtra-wideband (UWB) bandpass filter with transmission zero," *EuMA*, 603–605, Spt. 2006.
11. Hong, J. S. and M. J. Lancaster, *Microstrip Filters for RF/Microwave Applications*, Wiley, New York, 2001.
12. Woo, D. J., T. K. Lee, C. S. Pyo, and W. K. Choi, "Novel U-slot and V-slot DGSs for bandstop filter with improved Q factor," *IEEE Microw. Wireless Compon. Lett.*, Vol. 54, No. 6, Jun. 2006.

The Non-homogeneity of Permittivity in Microwave Dielectric Resonator

Victor N. Egorov
Eastern-Siberian branch of FSUE "VNIIFTRI", Russia

5. THE METHOD OF MEASUREMENTS AND RESULTS

Precise measurements of dielectric characteristics seem to be a necessary for exact calculating of parameters of microwave dielectric resonators, substrates and others components. Metrological application of ceramic as a dielectric standard reference material is need of permittivity homogeneity in all volume of standard sample. The high precision measurements of permittivity were performed for thermostable ceramic microwave dielectric resonators on frequency range 2–10 GHz. The "dielectric post resonator" method [1, 2] was applied for measurements by using scalar network analyzer with frequency resolution 1 Hz. In measurements the cylindrical specimen (dielectric resonator-DR) was placed closely between two flat copper mirrors with diameter 116 mm. The excitation of H_{0mp} modes was performed by semirigid coaxial cables with loop on end. Resonator transmission coefficient $\log|S_{12}|^2$ has been tuned to meaning less -30 dB. The ceramic specimens # 1, 2, 3, 4, 5, 6, 7 with permittivity ε about 9.7; 30; 39; 58; 98; 14; 78 and single crystal of sapphire Al_2O_3 has been measured by this method. Measured permittivity ε of the each ceramic specimen has the different values for operating resonance modes H_{0mp} with different radial index m and axial index p, which have different value of filling factor $K_{1E} = W_{1E}/W_{\Sigma}$, where W_{1E}, W_{Σ}-electrical energy in specimen and complete energy of resonance. The value of K_{1E} was limited as $0.92 < K_{1E} < 0.9999$ in this measurements. On the each resonance frequency of H_{0mp} modes the expanded uncertainty of ε for each ceramic specimen is less than difference between max and min measured value ε for this specimen. The analysis of experimental data shows the increasing of measured value ε when the axial index p is increasing for each radial index $m = 1, 2, 3$. Modes with different indexes m, p have a different filling factor K_{1E}, which increase with increasing of p for each m. In order to compare the dependence $\varepsilon(K_{1E})$ in different specimens the relative deviation $\delta\varepsilon = 100(\varepsilon - \bar{\varepsilon})/\bar{\varepsilon}$, (%) of measured values ε from mean value $\bar{\varepsilon}$ of specimens are shown on Fig. 1–Fig. 8 as functions of K_{1E}. The experimental data (signed as "o") were approximated by least square method as a linear function (solid line). Fig. 2 shows same chart for specimen of single crystal Al_2O_3 which has no dependence $\varepsilon(K_{1E})$. The specimens 2 (B30) and 4 (K60) have less slope of function $\delta\varepsilon(K_{1E})$ than one in specimens 3 (B40-2), 5 (D100) and in specimen 7 (B80, $\varepsilon \approx 78$).

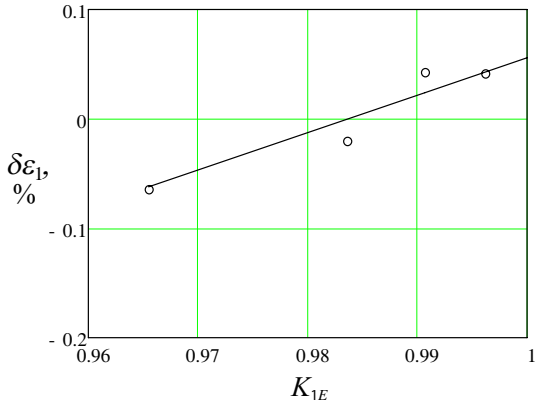


Figure 1: The deviation of measured ε from mean value $\bar{\varepsilon} \approx 9.67$ versus K_{1E} in specimen 1 (D10).

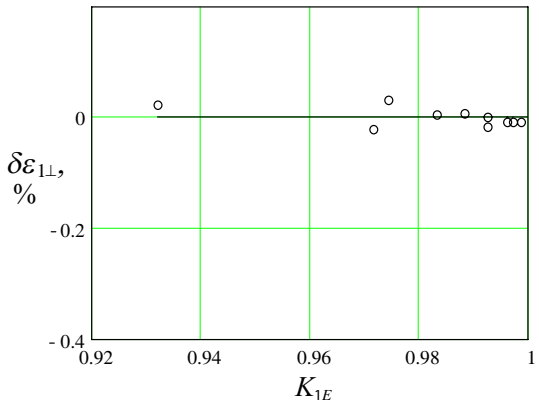


Figure 2: The deviation of measured $\varepsilon_{1\perp}$ from mean value $\bar{\varepsilon}_{1\perp} = 9.388$ versus K_{1E} in specimen of single crystal Al_2O_3.

The dependence $\varepsilon(K_{1E})$ can not explain as a frequency dispersion of specimen permittivity because the measured values ε at two closely resonance frequencies with different indexes m, p are

not equal and ε increase with frequency increasing. The residual "air-gap" in resonator between metal mirrors and top and bottom flat ends of specimen lead to small overestimate of measured value ε, which reduces when K_{1E} is grows [2]. In this reason the residual "air-gap" can not explain dependence $\varepsilon(K_{1E})$ also. The electrical field strength E_φ on cylindrical surface of specimen and in subsurface layer is reducing with K_{1E} increasing.

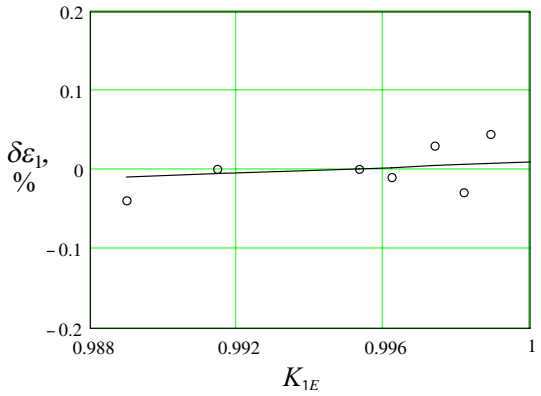


Figure 3: The deviation of measured ε from mean value $\bar{\varepsilon} \approx 30$ versus K_{1E} in specimen 2 (B30).

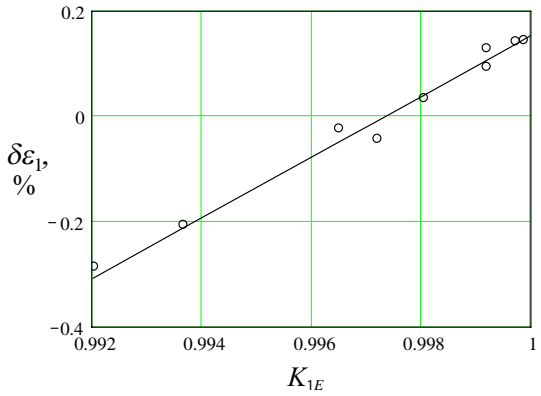


Figure 4: The deviation of measured ε from mean value $\bar{\varepsilon} \approx 39$ versus K_{1E} in specimen 3 (B40-2).

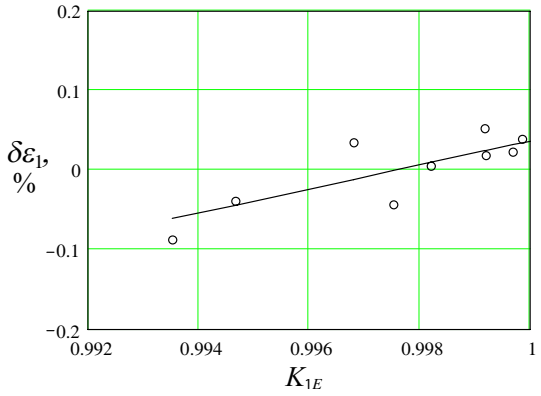


Figure 5: The deviation of measured ε from mean value $\bar{\varepsilon} \approx 58$ versus K_{1E} in specimen 4 (K60).

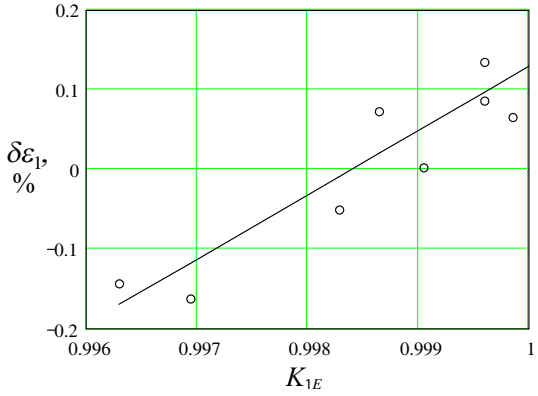


Figure 6: The deviation of measured ε from mean value $\bar{\varepsilon} \approx 78$ versus K_{1E} in specimen 7 (B80).

The dependence $\varepsilon(K_{1E})$ may be explain in assumption of radial non-homogeneity of ε with decreasing function $\varepsilon(r)$. The non-homogeneity of ε may be as result of high temperature and pressure in dielectric resonator manufacturing process (technology of ceramics) and/or mechanical treatment of DR surfaces. In this reason the axially dependence $\varepsilon(z)$ should be exist too. The zero boundary conditions $E_\varphi = 0$ on metal mirrors (specimen flat ends) surface and negligible small value of E_φ in subsurface region in specimen conceal this effect. There is other type of DR for it's observation. It is known as waveguide-dielectric resonator (WDR) [2]. The condition $E_\varphi = 0$ is satisfy on cylinder surface $r = a$ in circular metal waveguide, where specimen is placed closely and resonance frequency is below cutoff in empty parts of waveguide.

6. TWO LAYER DIELECTRC RESONATOR MODEL

The simplest model for approximation of radial non-homogeneity of permittivity is the step-constant function $\varepsilon(r) = \varepsilon_1$ for $0 < r \le r_0$ and $\varepsilon(r) = \varepsilon_3$ for $r_0 < r \le a$. The ordinary procedure of variable separation and field matching for $r = r_0$ and $r = a$ gives the equation

$$\frac{\mu_1 J_1(xt)}{(xt)\, J_0(xt)} - \frac{\mu_3 J_1(vt)}{(vt)\, J_0(vt)} \frac{\left[1 + D\dfrac{N_1(vt)}{J_1(vt)}\right]}{\left[1 + D\dfrac{N_0(vt)}{J_0(vt)}\right]} = 0, \tag{1}$$

where D is

$$D = -\frac{\frac{\mu_3}{\mu_2} J_1(v) + v J_0(v) \frac{K_1(y)}{y K_0(y)}}{\frac{\mu_3}{\mu_2} N_1(v) + v N_0(v) \frac{K_1(y)}{y K_0(y)}}, \tag{2}$$

and J_0, N_0, K_0, J_1, N_1, K_1-Bessel, Neiman and modified Bessel functions of order 0, 1; $x = \chi_1 a$, $v = \chi_3 a$, $y = \gamma_2 a$, $t = r_0/a \le 1$; $\chi_1 = \sqrt{k_1^2 - h^2}$, $\chi_3 = \sqrt{k_3^2 - h^2}$, $\gamma_2 = \sqrt{h^2 - k_2^2}$-transverse wave numbers in regions 1 (central cylinder), 3 (outer layer), 2 (surrounding media); $h = p\pi/L$-axial wave number, p =1, 2, 3. . .-axial mode index; $k_i = \omega\sqrt{\varepsilon_0\mu_0\varepsilon_i\mu_i}$, i = 1, 2, 3. For $\chi_1 = \chi_3$, i.e., for $\varepsilon_1 = \varepsilon_3$, $\mu_1 = \mu_3$ and for $t = 1$ the ordinary characteristic equation for H_{0mp} modes in homogeneity DR [3] follows from (1), (2).

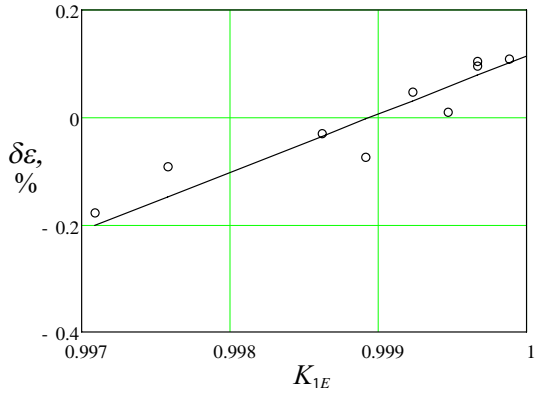


Figure 7: The deviation of measured ε from mean value $\bar{\varepsilon} \approx 98$ versus K_{1E} in in specimen 5 (D100).

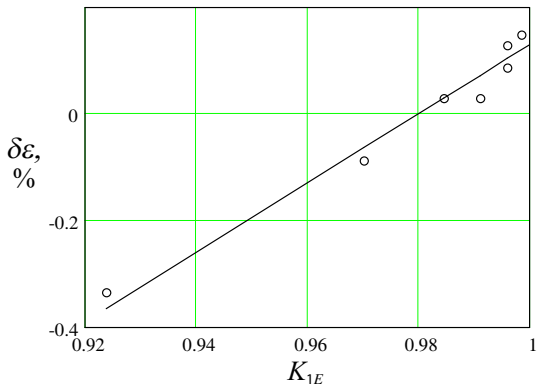


Figure 8: The deviation of measured ε from mean value $\bar{\varepsilon} \approx 14$ versus K_{1E} in in specimen 6 (D14).

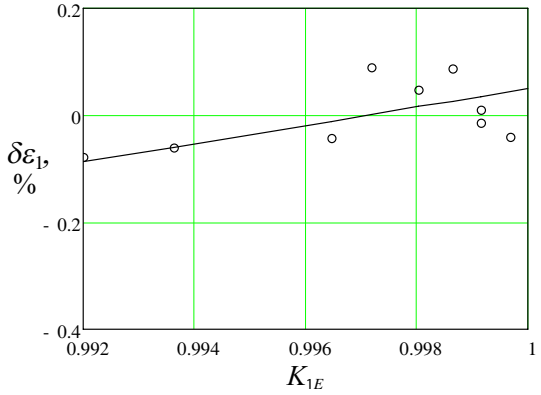


Figure 9: The deviation of measured ε_1 in specimen 3 (B40-2) for two layer DR model (see also Fig. 4).

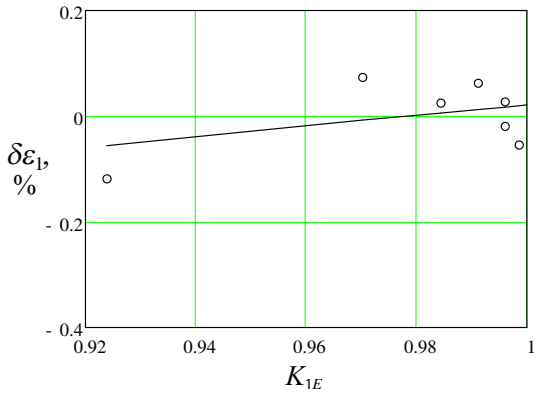


Figure 10: The deviation of measured ε_1 in specimen 6 (D14) for two layer DR model (see also Fig. 8).

The search of values r_0, ε_3 was performed on criteria of minimum rms deviation in results of calculation ε_1 in central region 1 of specimen 3 (B40-2) with sizes $2a = 29.994$ mm, $L = 15.143$ mm. The obtained values are $r_0 = 0.967a$ and $\varepsilon_3 = \bar{\varepsilon} - 3.25$, where $\bar{\varepsilon} = 39.372$ is the mean value of specimen permittivity in homogeneity approximation. The mean permittivity in central region $\bar{\varepsilon}_1 = 39.496$ is obtained from two layer DR model (1), (2) by this way. This treatment has reduced the rms deviation of results about 2.5 times. The residual slope of dependence $\delta\varepsilon_1(K_{1E})$ (Fig. 9) is less than initial slope (Fig. 4) about 2.5 times also. For specimen 6 (D14) with sizes $2a = 23.999$ mm, $L = 15.068$ mm the same treatment gives $r_0 = 0.983a$ and $\varepsilon_3 = \bar{\varepsilon} - 2.3$, where $\bar{\varepsilon} = 14.280$ is the mean value of specimen permittivity in homogeneity approximation. The obtained mean permittivity in central region is $\bar{\varepsilon}_1 = 14.324$. The function $\delta\varepsilon_1(K_{1E})$ (Fig. 10) has slope, less than initial function (Fig. 8). The approximation of function $\delta\varepsilon_1(K_{1E})$ on Figs. 9, 10 by linear function is not best approach. The decreasing of ε_1 when $K_{1E} \to 1$ on Figs. 9, 10 shows that two layer DR model is

not sufficient correct for description of non-homogeneity and fitting smooth function $\varepsilon(r)$ should be applied. The dependence $\varepsilon(z)$ should be estimated also.

7. CONCLUSION

The observation of dependence of measured mean ε versus electrical energy filling factor K_{1E} in ceramic specimens and the absent of it in single crystal of sapphire confirm the non-homogeneity of ε in radial and, may be, in axial directions in measured cylindrical ceramic specimens. The power of dependence $\varepsilon(K_{1E})$ (decreasing power of $\varepsilon(r)$) is different in different ceramic specimens and no direct correlated with value of ε. The specimens with $\varepsilon \approx 30$ and $\varepsilon \approx 39$ have different slope of functions $\delta\varepsilon(K_{1E})$. The two layer DR model with step-constant function $\varepsilon(r)$ confirm the decreasing behavior of $\varepsilon(r)$, but it is not sufficient correct in physical and mathematical sense for real smooth dependence $\varepsilon(r)$. The axial dependence $\varepsilon(z)$ exist probably in cylindrical ceramic DR also. The experiments in circular $H_{0m\delta}$ below-cutoff waveguide-dielectric resonator [4] can to detect this effect.

REFERENCES

1. IEC 61388-1-3; Waveguide Type Dielectric Resonators, parts 1–3 General Information and Test Conditions. Measuring Method of Complex Relative Permittivity for Dielectric Resonator Materials at Microwave Frequency.
2. Egorov, V. N., "Resonance methods for microwave studies of dielectrics (review)," *Instruments and Experimental Techniques*, Vol. 50, No. 2, 143–175, 2007.
3. Hakki, B. W. and P. D. Coleman, *IEEE Trans. Microwave Theory Tech.*, Vol. 4, 402, 1960.
4. Cohn, S. B. and K. C. Kelly, *IEEE Trans. Microwave Theory Tech.*, No. 9, 408, 1966.

Genetic Algorithms Applied to Microwave Filters Optimization and Design

M. F. Jiménez Nogales, J. Pascual García, J. Hinojosa, and A. Alvarez-Melcón
Universidad Politécnica de Cartagena, Spain

Abstract— This paper presents improvements in genetic algorithms for the optimization of microwave circuits. Several fitness functions are studied and implemented in the optimization process. They allow to quantify the importance of the specifications and of the solution approach with respect to the ideal circuit performance for a bandpass filter. The performance of the optimization method with different fitness functions is compared for the design of a Ku-band two-pole rectangular iris coupled waveguide bandpass filter.

1. INTRODUCTION

In [1], a group of experts estimated the state of microwave computer-aided design (CAD) in the year 2010. In a panel discussion, M. Mongiardo predicted much of emphasis will shift to an optimization and design environment that allows integrated design. In this sense, we propose an approach for the optimization of microwave filters based on Genetic Algorithms (GAs) [2], which overcome the disadvantages of the traditional optimization methods:

- They are much more robust, since they are largely independent of the initial conditions.
- They allow to find the overall best solutions with many potential local minima (or maxima).
- They can readily handle nonlinear and non-differentiable functions. The microwave circuit design problems are often one of these types.

GAs are stochastic search procedures that have been successfully applied in many electromagnetic syntheses and optimization problems [3]. It is known that a fitness function poorly defined can lead to incorrect solutions at the end of the GA operation. However, as for as we know, no work in the literature has studied the problem related with the definition of the fitness function to assure the degree of fulfillment of the circuit specifications. The authors only limit to present its fitness functions for its specific applications.

In this paper, we present the design and optimization of microwave filter using a rigorous analysis technique, i.e., the multimode network theory combined with the method of moments [4] linked to a GA optimizer. The method is illustrated with the design of a Ku-band rectangular waveguide bandpass filter for satellite applications. Several fitness functions are defined for discrete and continuous target specifications. The performance (computation time, quality of calculated best solution) of the GA with these several fitness functions are presented and compared.

2. GENETIC ALGORITHMS AND FITNESS FUNCTIONS

In this Section, basic GA is briefly described and several fitness functions, depending on the discrete and continuous target specifications of a bandpass filter, are then presented. The fitness functions must be defined to measure the fulfillment degree of the filter specifications. A proper evaluation of the specifications fulfillment is fundamental to reach the desired circuit solution, and to avoid problems in the GA operation. In order to assure an appropriate GA performance two different fitness functions have been developed for discrete target specifications, whereas a robust fitness function is defined when the continuous target specifications are known.

2.1. Basic GA

GAs are stochastic search techniques modeled on the principles and concepts of natural selection and evolution. In a preliminary stage, a set of candidate solutions called population is encoded as binary bit strings, which are called chromosomes. Then, an evaluation function, called fitness function, is used to evaluate the chromosomes. If the desired fitness is satisfied by at least one individual in the present population the optimization procedure is stopped. Otherwise successive GA operations, such as selection (the candidates from one generation are selected for the next generation), crossover (genetic material of one candidate is exchanged with another candidate) and mutation (genetic material of one candidate is altered), are repeated to the population until to obtain a chromosome with the desired fitness.

2.2. Fitness Functions for Discrete Target Specifications

Discrete target specifications are referred to those values of attenuation in the passband and in the rejection band, that must be achieved by the filter; i.e., in a bandpass filter the S_{11} response has to be lower than L_P dB for all discrete frequency points in the passband, while the S_{11} attenuation in the rejection band has to surpass a minimum value of L_R dB. A simple evaluation of the specifications fulfillment consist of assigning to each one of the N frequency points a value of 0 if the scattering parameter does not satisfy the corresponding specification and on the contrary a value of 1 if it satisfies the specification. The final fitness is the sum of the $\{0,1\}$ values for all the frequency points for either the S_{11} response or the S_{21}, divided by the number of points:

$$F(S_{11}) = \frac{\sum_{i=1}^{N} fit_{S_{11}}(i)}{N} \quad \text{with} \quad fit_{S_{11}}(i) \begin{cases} 1 & \text{if } S_{11}(i) < L_p \quad f_i \in PB \\ 0 & \text{if } S_{11}(i) > L_p \quad f_i \in PB \\ 1 & \text{if } S_{11}(i) > L_R \quad f_i \in RB \\ 0 & \text{if } S_{11}(i) < L_R \quad f_i \in RB \end{cases} \tag{1}$$

where N is the total number of frequency points, PB corresponds to the S-parameters in the passband, RB the S-parameters in the rejection band, and f_i is the frequency at the i-th point.

This fitness function has a great disadvantage. Small and large errors are treated equally, therefore filters with proper specifications fulfillment in certain parts of the bandwidth will be highly valued, even if they present serious errors in other frequency points. It is necessary to calibrate the error seriousness to avoid the error masking of the previous fitness function. Thus, we propose alternative continuous monotonic fitness functions, which allow the quantification of the fulfillment specification in a gradual fashion with the following equations:

$$F(S_{11}) = \frac{\sum_{i=1}^{N} fit_{S_{11}}(i)}{N} \quad \text{and} \quad \begin{cases} fit_{S_{11}}(i) = 1 - e^{a_{P1}(x-L_p)} & \text{if } f_i \in PB \text{ and } x < L_p \\ fit_{S_{11}}(i) = -1 + e^{-a_{P2}(x-L_p)} & \text{if } f_i \in PB \text{ and } x > L_p \\ fit_{S_{11}}(i) = 1 - e^{-a_{R1}(x-L_R)} & \text{if } f_i \in RB \text{ and } x > L_R \\ fit_{S_{11}}(i) = -1 + e^{a_{R2}(x-L_R)} & \text{if } f_i \in RB \text{ and } x < L_R \end{cases} \tag{2}$$

The preceding functions are similar to sigmoidal functions as seen in Fig. 1. It is clearly shown that these functions permit a graduation of the correctness between 0 and 1 for the frequency points that satisfy the specification. They also allow a graduation of the error between 0 and -1 for those frequency points that do not satisfy the corresponding specification. These functions saturate from a certain limit, thus large correct values will not mask serious errors in other frequencies. The terms a_{P1}, a_{P2}, a_{R1}, and a_{R2} control the slope of the fitness function for the passband and rejection band, respectively. If the slopes (a_{P2}, a_{R2}) are large the errors are highly penalized, while correct attenuation points are highly valued if (a_{p1}, a_{R1}) present large values. On the contrary, if the corresponding slope is small, large errors are little penalized and correct values are little recompensed. If the scattering parameter S_{21} is employed we must take into account that it is necessary to interchange the previous equations. This is because this response has to surpass an attenuation limit value in the passband, whereas it has produce lower values than an attenuation threshold in the rejection band.

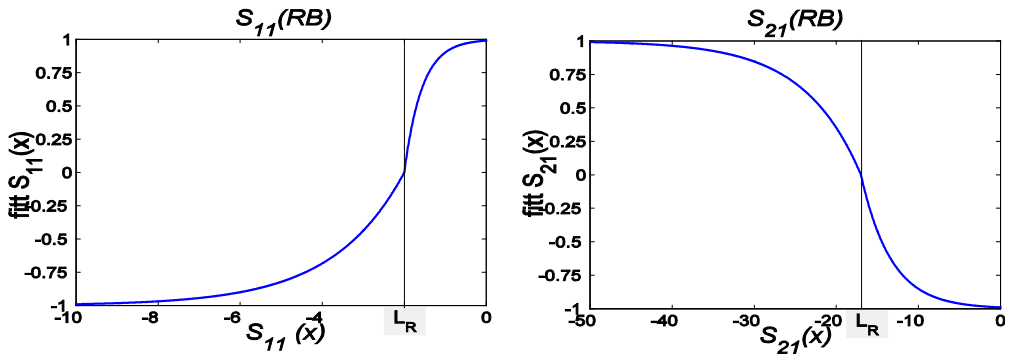


Figure 1: Illustration example of the fitness function for Equation (2) applied to the rejection band (RB).

2.3. Fitness Function for Continuous Target Specifications

We define a continuous target specification when the filter behavior is known at all frequency points. This is possible using synthesis techniques based on elliptic transfer functions, as for instance those reported in [5]. The fitness functions of Section 2.2 are not valid when the desired filter response is known. Now for each one of the frequency points we have an ideal attenuation value instead of a specification discrete limit. In the previous fitness functions the correctness of the scattering parameter at a certain frequency point depended on the attenuation value with respect to the corresponding limit; i.e., in the passband the S_{11} values must be lower than L_P dB. If the S_{11} parameter surpasses the mentioned limit then it is treated as an error, and the appropriate fitness equation is applied. If the whole circuit response is available, the difference between each ideal attenuation value and the calculated scattering parameter must be used, without taking into consideration if the scattering parameter is below or above from the ideal attenuation; i.e., if the S_{11} is -10 dB at 1 GHz and the discrete specification, used in Section 2.2, for that frequency is -11 dB, a value of -11.5 dB is more suitable for this frequency that a value of -20 dB even though -20 dB satisfies the discrete specification. For properly express the degree of closeness to the ideal known response we propose a new fitness function that has proven to be very robust. The fitness function is defined as:

$$F = a\, fit_{S_{11}}(PB) + b\, fit_{S_{11}}(RB) + c\, fit_{S_{21}}(PB) + d\, fit_{S_{21}}(RB) \tag{3}$$

where (a, b, c, d) are weighting coefficients, and $fit_{Sij}(band)$ is an average distance from the target and the calculated response, at all frequency points lying within the specific band (PB or RB) (4). For each frequency point, the distance is defined using a sigmoid (Fig. 2), which showed to exhibit the best convergence rate because of its shape. The score assigned to each distance is adjusted by means of the slope a_{Sij}. The distance is measured in logarithmic scale (4) in order to distinguish better the small differences, and to ensure in this way a good convergence.

$$fit_{S_{ij}}(B) = \frac{\sum\limits_{x\in B} \frac{1}{1+e^{a_{S_{ij}}(d(x)-b)}}}{N} \quad \text{with} \quad d(x) = \ln\left(\left\|S_{ij}(x) - S_{ij}^{T}(x)\right\| + 0.1\right) \tag{4}$$

where S_{ij}^{T} is the scattering parameter value of the target function, B is either the passband or the rejection band, and N is the number of frequency points. The weighting coefficients (a, b, c, d) are employed to control the importance during the GA operation of one of the main bands (PB or RB); i.e., if one individual and its offspring achieve a central frequency near the central frequency of the ideal response, the weighting coefficients (a, c) can be changed to overrate the results in the pass band.

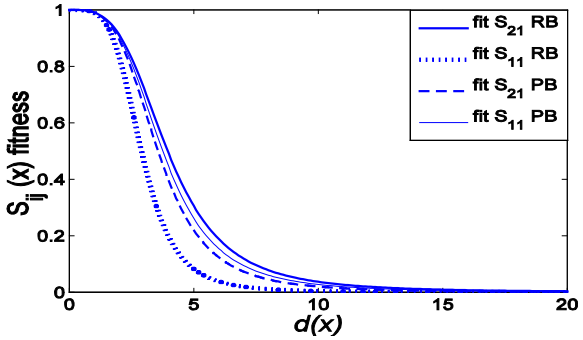


Figure 2: Illustration example of the fitness function for Equation (3).

3. EXAMPLE

The optimization technique described in Section 2 was applied to Ku-band 2-pole rectangular iris coupled waveguide bandpass filter (Fig. 3). The rectangular waveguide size is the standard: WR-75 ($A = 19.05$, $B = 9.525$ mm). The optimization parameters (W_1, W_2 and L) are the iris widths and the cavity length (S fixed at 2 mm). The design specifications are: $|S_{21}| \geq -3$ dB for $10.12 \leq f\,(\text{GHz}) \leq 10.2$ (passband), and $|S_{21}| \leq -18$ dB for $f \geq 10.3$ GHz and $f \leq 10$ GHz

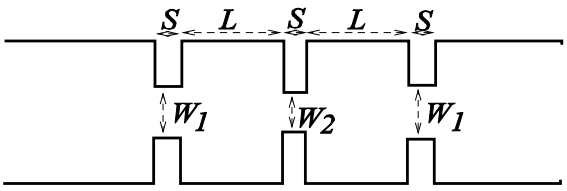
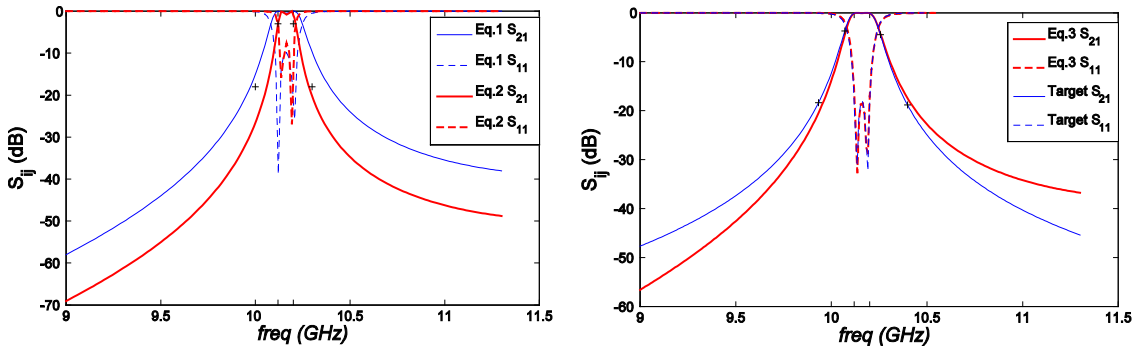


Figure 3: 2-pole rectangular iris coupled waveguide filter.

(rejection band). The analysis is performed by a rigorous technique combining the multimode network theory with the method of moments [4], for 160 points covering the frequency range 9 GHz–11.3 GHz. The set of parameters used for the GA optimizer was: A population size of 250 candidates, a binary coding of 33 bits per candidate, tournament selection, crossover and mutation probabilities fixed to 0.0303 and 0.06, respectively. Fig. 4 presents the frequency responses of the optimized filter (Fig. 3) from fitness functions using Equations (1) and (2) (Fig. 4(a)), and using Equation (3) (Fig. 4(b)) for discrete and continuous target specifications, respectively. The quality of the optimal solution depends from the number of iterations. The optimizations for the fitness functions (2) and (3) stopped when the algorithm found the desired specifications.



(a) (b)

Figure 4: 2-pole rectangular iris coupled waveguide filter.

As it was to be expected (Fig. 4(a)), the GA with the fitness function based on Equation (2) obtains an overall better solution than the fitness function based on Equation (1) after 10 iterations over 160 frequency points. The CPU time, on pentium4-based computer, was approximately 63 min. The fitness function based on Equation (2) allows a faster convergence than Equation (1) and, therefore, it requires less analysis iterations of the filter with the rigorous technique.

When the filter behavior is defined from continuous target specifications, it can be seen Fig. 4(b) the fitness function proposed in Equation (3) is very robust. The optimizer has required one additional iteration with respect to the discrete target specifications (11 iterations) to adjust the S-parameter objectives over 160 frequency points. The CPU time, on pentium4-based computer, was approximately 69 min. We can observe that the final design structure closely matches the target specifications obtained with [5].

4. CONCLUSIONS

In this work, an optimization technique based on GAs has been successfully applied to the design of Ku-band two-pole rectangular iris waveguide bandpass filter. Several fitness functions, suitable for the optimization of the filter, have been proposed. They are used to measure the degree of fulfillment depending on the discrete or continuous target specifications. Though the GAs are known to be slow to search the optimal solution, it has been showed the speed can be improved using adequate fitness functions based on sigmoidal functions, with respect to a fitness function poorly defined.

ACKNOWLEDGMENT

The authors thank Ministerio de Educación y Ciencia of Spain, which has supported with work under Grant TEC2007-67630-C03-02/TCM.

REFERENCES

1. Bootin, R. C., "Microwave CAD in the year 2010 — A panel discussion," *Int. J. of RF and Microw.*, Vol. 9, No. 6, 439–448, 1999.
2. Goldberg, J., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, 1989.
3. Rahmat-Samii, Y. and E. Michielssen, *Electromagnetic Optimization by Genetic Algorithms*, Wiley and Sons, New York, 1999.
4. Guglielmi, M., et al., "Rigorous multimode network numerical representation of inductive step," *IEEE Trans. Microw. Theory and Tech.*, Vol. 42, No. 2, 317–326, 1994.
5. Cameron, R. J., "Advanced coupling matrix synthesis technique for microwave filters," *IEEE Trans. Microw. Theory and Tech.*, Vol. 51, No. 1, 1–10, 2003.

Novel Fractal Electromagnetic Bandgap Structures to Suppress Simultaneous Switching Noise in High Speed Circuits

Kuo-Chiang Hung[1], **Ding-Bing Lin**[2], **Chin-Sheng Chang**[3]
Chun-Te Wu[4], **and I-Tseng Tang**[5]

[1]Institute of Computer and Communication Engineering
National Taipei University of Technology, Taipei, Taiwan
[2]Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan
[3]Institute of Microelectronics, Department of Electrical Engineering
Advanced Optoelectronic Technology Center, National Cheng-Kung University, Tainan, Taiwan
[4]Department of Electrical Engineering, Da-Yeh University, Changhua, Taiwan
[5]Department of Environment and Energy, Nation University of Tainan, Tainan, Taiwan

Abstract— A fractal low-period coplanar electromagnetic bandgap power/ground planes to suppress simultaneous switching noise is presented in this study. The effective suppressing bandwidth of the fractal low-period coplanar electromagnetic bandgap and meander-bridge structure is from 264 MHz to 20 GHz. The results show the presented structure is superior to other structure for suppressing simultaneous switching noise.

1. INTRODUCTION

The integration of mixed-signal systems in practical computers is forcing the integration of high-speed digital circuits with analog and radio frequency (RF) circuits. When the signal drivers or the logic circuits of a microprocessor switch simultaneously, the simultaneously switching noise (SSN) generated from the noisy circuits that can influence power integrity (PI) or the other sensitive RF/analog circuits. Therefore, an effective structure to isolate noise is required for improving the PI in mixed-signal systems.

In recent year, the exceptional configurations to mitigate the SSN in printed circuit board (PCB) were achieved by using electromagnetic bandgap (EBG) structures to provide a high impedance surface (HIS) on the power/ground planes. Those HIS structures used embedded structure between power and ground planes and connected with ground plane by vias [1, 2], which were increased the cost. Therefore, the planer EBG structures were researched for isolating wide-band SSN on the power/ground planes, as in [3, 4]. The planer EBG structures comprise two-layer power/ground planes. One layer consists of EBG structures; the other one is a continuous metal plane.

In this paper, the exceptional configurations to mitigate the simultaneous switching noise (SSN) in printed circuit board (PCB) were discussed by etching low-period coplanar EBG (LPC-EBG) structures on the power/ground planes to provide a wider bandgap bandwidth [5]. Fractal structures had been widely used to create the multiple transmission zeros at higher frequency band [6]. However, it does not have a good isolation in the lower frequency band. Therefore, this paper proposed a novel fractal LPC-EBG (FLPC-EBG) to improve the low frequency characteristics. In addition to have a lower cut-frequency, the higher inductance of meander-bridge is used to connect each FLPC-EBG cells. This paper is organized as follows: Section 2 discusses the design concept and the measurement results of FLPC-EBG structure. In Section 3, the FLPC-EBG can improved signal integrity quality of single-ended is described.

2. DESIGN CONCEPT AND IMPROVEMENT

2.1. Design Concept

Figure 1(a) shows the dimensions of FLPC-EBG and meander-bridge structure, which is applied on a two-layer FR-4 PCB substrate. The thickness of substrate is 0.4 mm and the dielectric constant is 4.4. The dimensions of the PCB are $9 \times 9\,\mathrm{cm}^2$ and the area of each unit cell is $28.5 \times 28.5\,\mathrm{mm}^2$. The ground plane is a continuous metal plane and the power plane has nine cells of FLPC-EBG etched on it. In this proposed structure, the power plane has a periodic pattern that is a two-dimensional (2-D) fractal square and meander-bridges connecting between those squares as shown in Fig. 1(c). The receiving port 1 is at (15 mm, 75 mm) and the excitation ports 2, 3, 4 are at (45 mm, 75 mm), (45 mm, 45 mm), (75 mm, 15 mm), respectively. These two structures are compared with the low-period coplanar EBG (LPC-EBG) structure proposed in [5]. The present structure is verified by a

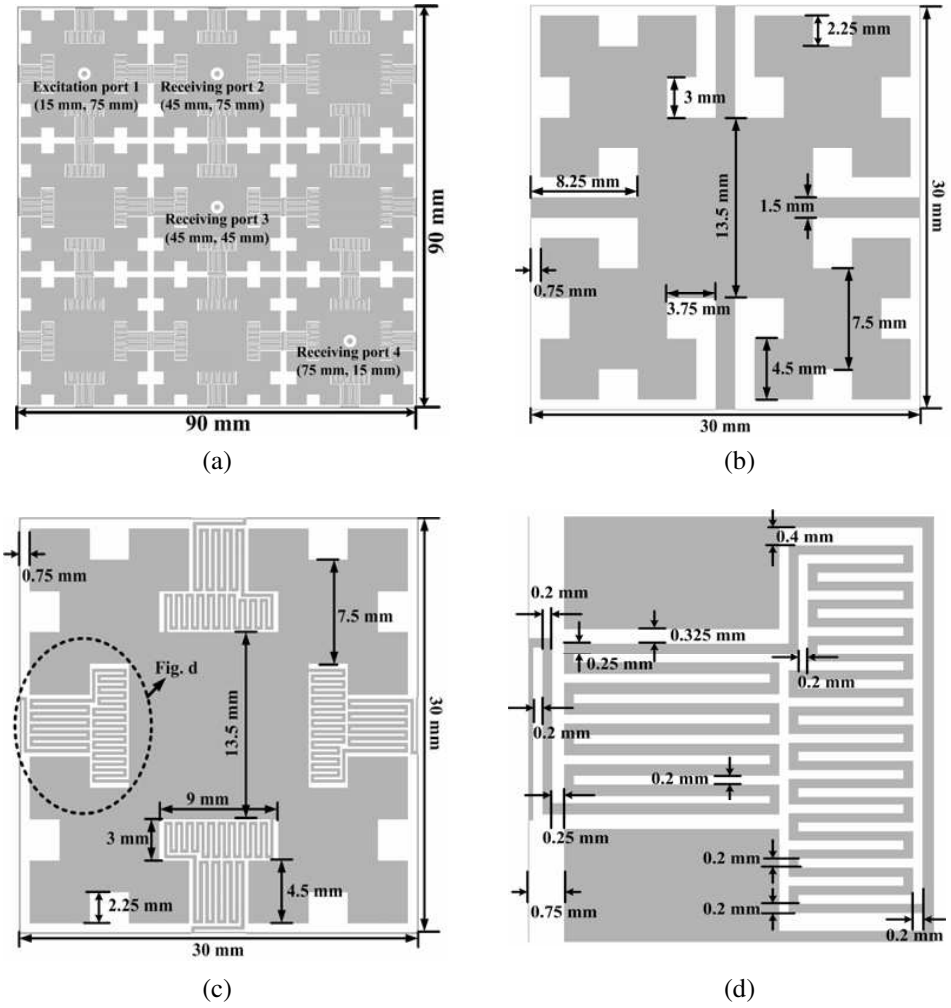


Figure 1: The proposed testing board. (a) FLPC-EBG and meander-bridge structure. (b) Unit cell of FLPC-EBG structure. (c) Unit cell of FLPC-EBG and meander-bridge structure. (d) Detail dimension of meander-bridge.

vector network analyzer (VNA) and by a 3-D simulator (HFSS). Both measurement and simulation results demonstrate the superiority of this novel structure.

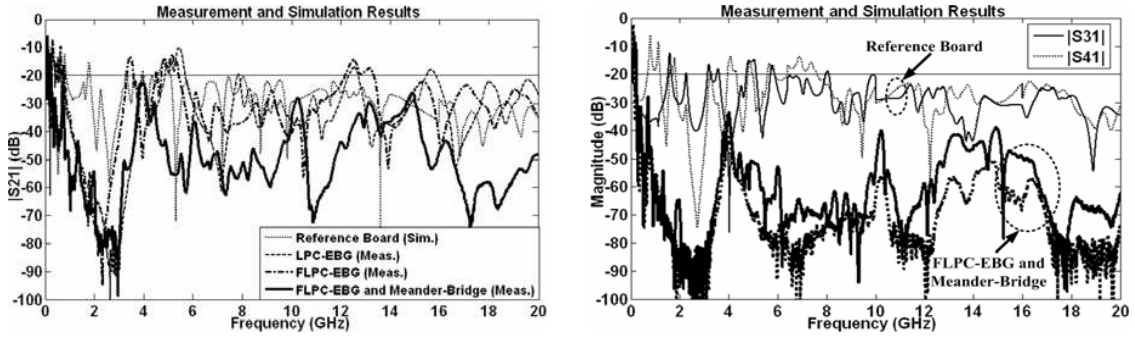


Figure 2: Measurement and simulation results of reference board, and FLPC-EBG and meander-bridge structures. (a) $|S_{21}|$, (b) $|S_{31}|$ and $|S_{41}|$.

2.2. Improve Simultaneously Switching Noise

Figure 2(a) shows the measured $|S_{21}|$ for the discussed FLPC-EBG and meander-bridge power/ground planes. The bandwidth is defined by $|S_{21}|$ of which magnitude is lower than -20 dB. The simulation

result of the reference board with both integral power and ground planes is also presented in this figure for comparison. Compared with the LPC-EBG structure, the FLPC-EBG and meander-bridge structure behaves highly efficient SSN elimination form about 264 MHz to 20 GHz (19.736 GHz bandwidth) as shown in Table 1. Fig. 2(b) shows the measurement results of $|S_{31}|$ and $|S_{41}|$ of FLPC-EBG and meander-bridge structure. From this figure, the suppressed SSN range of FLPC-EBG and meander-bridge structure is from about 221 MHz to 20 GHz (19.779 GHz bandwidth). The key factor of FLPC-EBG and meander-bridge structure is the meander bridge which can increase the inductance between two integrated squares greatly so that they can suppress the SSN at lower frequencies.

Table 1: Bandwidth and suppressed frequency ranges of three cases.

Cases	Measured Parameters	Frequency (GHz)	
		Ranges	Bandwidths
LPC-EBG [5]	$\|S_{21}\|$	$0.746 \sim 3.918$	3.172
FLPC-EBG	$\|S_{21}\|$	$0.808 \sim 3.309$	2.501
FLPC-EBG and Meander-Bridge	$\|S_{21}\|$	$0.264 \sim 20$	19.736
FLPC-EBG and Meander-Bridge	$\|S_{31}\|$	$0.235 \sim 20$	19.765
FLPC-EBG and Meander-Bridge	$\|S_{41}\|$	$0.221 \sim 20$	19.779

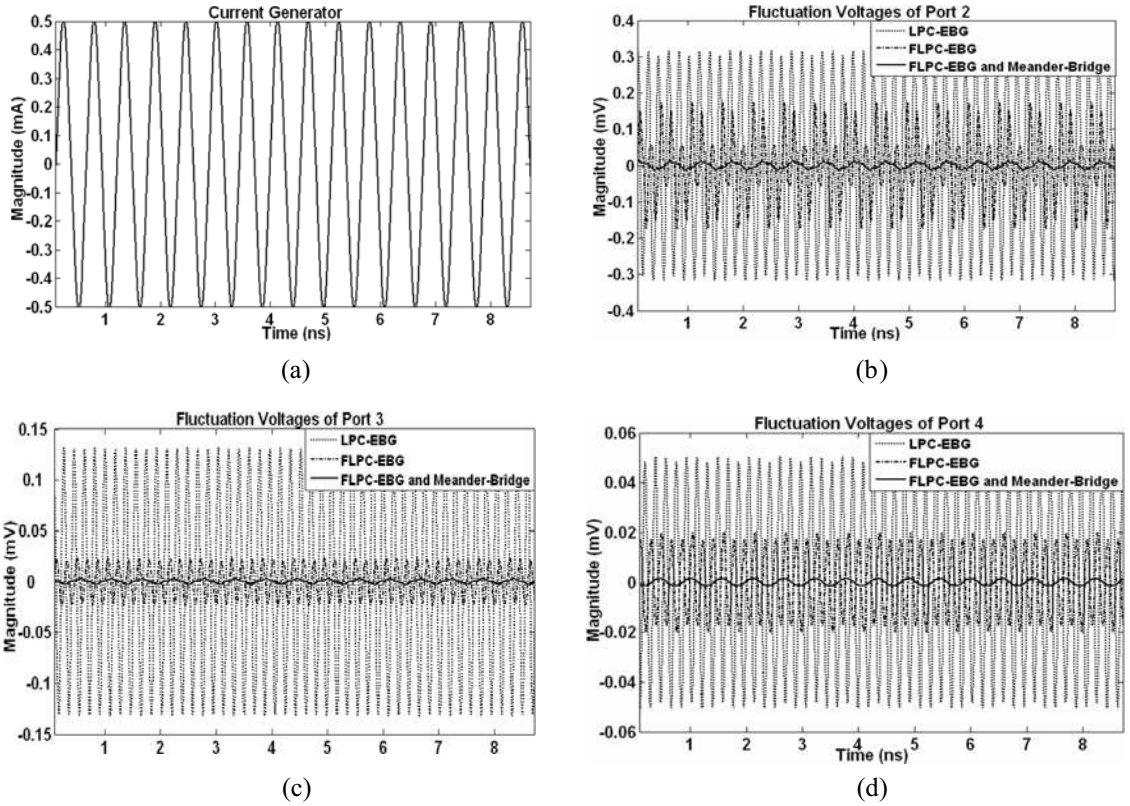


Figure 3: Fluctuation voltages at receiving port when the exciting port is located at port1. (a) Current generator at port 1, (b) Port 2, (c) Port 3, and (d) Port 4.

3. FLUCTUATION OF POWER VOLTAGE

First, a time-domain digital current source (Fig. 3(a)) is generated and then transformed into a frequency domain response by applying Fourier transform. The frequency domain response for current is then multiplied by the S-parameter measured by VNA; the time domain response is then

obtained by IFT. The data rate is 3.6 Gbps, and current swing is 1 mA (Fig. 3(a)). Table 2 presents the noise levels for the different methods.

Table 2: Voltage fluctuation of three cases.

Cases	Excited Ports	Voltage Fluctuations ($V_{p\text{-}p}$)
LPC-EBG structure	Port 2	0.6326 mV
	Port 3	0.2635 mV
	Port 4	0.1006 mV
FLPC-EBG structure	Port 2	0.3472 mV
	Port 3	0.0468 mV
	Port 4	0.04 mV
FLPC-EBG and Meander-Bridge structure	Port 2	0.0266 mV
	Port 3	0.0052 mV
	Port 4	0.00282 mV

The proposed FLPC-EBG and meander-bridge is clearly better than LPC-EBG structure. The noise level for the proposed structure is smaller than 10% of the level of the LPC-EBG structure and FLPC-EBG structure at port 2, the 4.2% decreased for the LPC-EBG structure, and the 7.66% decreased for the FLPC-EBG structure.

4. CONCLUSIONS

The FLPC-EBG and meander-bridge structure for improving the SSN between sensitive RF/analog circuits has been proposed. Compared with the LPC-EBG and FLPC-EBG structures, our physical structures can provide the higher inductance of meander-bridge for lower cut-frequency. In additional, the FLPC-EBG and meander-bridge structure can decrease the voltages fluctuations on the power/ground planes. Finally, wideband noise suppression characteristics are also achieved using proposed FLPC-EBG and meander-bridge structure.

ACKNOWLEDGMENT

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financially supporting this research under Contract No. NSC 96-2218-E-027-007.

REFERENCES

1. Chen, G. and K. L. Melde, "Cavity resonance suppression in power delivery systems using electromagnetic band gap structures," *IEEE Transactions on Advanced Packaging*, Vol. 29, 21–30, Feb. 2006.
2. Park, J., A. C. W. Lu, K. M. Chua, L. L. Wai, J. Lee, and J. Kim, "Double-stacked EBG structure for wideband suppression of simultaneous switching noise in LTCC-based SiP applications," *IEEE Microwave and Wireless Components Letters*, Vol. 16, 481–483, Sept. 2006.
3. Wu, T. L. and T. K. Wang, "Embedded power plane with ultra-wide stop-band for simultaneously switching noise on high-speed circuits," *Electronics Letters*, Vol. 42, 213–214, Feb. 2006.
4. Toyota, Y., A. E. Engin, H. K. Tae, M. Swaminathan, and K. Uriu, "Stopband prediction with dispersion diagram for electromagnetic bandgap structures in printed circuit boards," *IEEE International Symposium on Electromagnetic Compatibility*, Vol. 3, 807–811, Aug. 2006.
5. Wu, T.-L., Y.-H. Lin, T.-K. Wang, C.-C. Wang, and S.-T. Chen, "Electromagnetic bandgap power/ground planes for wideband suppression of ground bounce noise and radiated emission in high-speed circuits," *IEEE Microwave Theory and Techniques*, Vol. 53, 2935–2942, Sept. 2005.
6. Kern, D. J., D. H. Werner, A. Monorchio, L. Lanuzza, and M. J. Wilhelm, "The design synthesis of multiband artificial magnetic conductors using high impedance frequency selective surfaces," *IEEE Transactions on Antennas and Propagation*, Vol. 53, 8–17, Jan. 2005.

Free-space Optical Data Link Using Quantum Cascade Laser

Z. Bielecki[1], **W. Kolosowski**[2], **and J. Mikolajczyk**[1]

[1]Military University of Technology, 2 Kaliskiego Str., 00-908 Warsaw, Poland
[2]Telecommunications Research Institute, 20 Poligonowa Str., 04-051 Warsaw, Poland

Abstract— The paper presents construction of a broadband optical system devoted to free space optical communication link. The main elements of the system are a quantum cascade laser and an HgCdTe heterostructural photodetector operating at the wavelength of 10 µm. The described analyses showed that the system is characterized by lower sensitivity to adverse meteorological conditions when compared with the systems operating in near infrared waveband.

1. INTRODUCTION

Free-Space Optics (FSO) products are deployed in a line-of sight point-to-point configuration. Free space optical systems offer a flexible networking solution that delivers on the promise of broadband communications. Only FSO provides the essential combination of qualities required for modern networking. Since FSO transceivers can transmit and receive through windows, it is possible to mount FSO systems inside buildings, reducing the need to compete for roof space, simplifying wiring and cabling, and permitting the equipment to operate in a very favourable environment. The only essential for FSO is a line of sight between the two ends of the link. Free Space Optics is far more secure than RF technologies for several reasons: requires no RF spectrum licensing, FSO laser beams cannot be detected with RF meters or spectrum analyzers, the laser beams generated by FSO systems are narrow and invisible, making them harder to find and even harder to intercept and crack, FSO laser transmissions are optical and travel along a line of sight path that cannot be intercepted easily, data can be transmitted over an encrypted connection adding to the degree of security available in Free Space Optics network transmissions. Another advantage of FSO, when compared to RF, is significant reduction in end-to-end delay. Most FSO products are plug-and-play units independent of the transmitted protocol and data rate. Quantum cascade lasers are very suitable for such applications because their emission wavelength can be chosen in the so-called atmospheric window regions, i.e., around 3–5 µm and 8–14 µm. In addition, the fast internal lifetimes of the devices should allow for reasonable modulation frequencies of up to 5–10 GHz [1]. In this paper, we analyze FSO-10 µm system compared to shorter wavelengths, i.e., 0.8 µm and 1.5 µm.

2. ATMOSPHERIC INFLUENCES

The atmosphere is a mixture of dry air and water vapour. Carrier-class Free Space Optics systems must be designed to accommodate heavy atmospheric attenuation, particularly by fog. Longer wavelengths are favoured in haze and light fog, but under the conditions of very low visibility this long-wavelength advantage does not apply. Additionally, the fact that 10-µm-based systems are allowed to transmit hundreds times more eye-safe power makes them very attractive for modern FSO links.

The atmospheric transmission of optical signals, is given by the Beer's law equation

$$\tau_a = e^{-(\beta_a+\beta_s)L}, \tag{1}$$

where, β_a and β_s are the absorption and scattering coefficients, respectively, L is the optical distance between transmitter and receiver.

Absorption by atmospheric gases defines the atmospheric windows where absorption is less severe. Scattering is due to small particulates in the transmission path of optical waves [2]. Many scientists have tried to relate the scattering coefficient to the incident wavelength and visibility [3]

$$\beta_s(\lambda) = \frac{3.91}{V}\left(\frac{\lambda}{0.55}\right)^{-\delta}. \tag{2}$$

The exponent δ depends on the visibility V. The problem with Eq. (2) is that it fails to estimate attenuations for visibilities below 1.5 km [4]. Therefore, Maha Achour relied on the original Mie

equation to estimate attenuations due to fog and rain. So-called, total scattering coefficient is given by the following equation

$$\beta_s(\lambda) = \sum_r \pi r^2 n_r \Theta_s(x), \quad x = \frac{2\pi r}{\lambda}, \tag{3}$$

where, r is the aerosol particle radius, n_r is the concentration of radius particles, Θ_s is the scattering efficiency, and λ is the transmitted signal wavelength.

The correlation between incident wavelength and particle sizes is given by the term $r^2\Theta_s(x)$ in Eq. (3). Fig. 1 illustrates $r^2\Theta_s(x)$ versus the droplet diameter ($2r$) for five wavelengths. For fog particles, less than 5 µm in a diameter, 10-µm light is the least affected by atmosphere compared with shorter wavelengths and become more affected by particles of larger diameter. In general, the 10-µm light of QCLs suffers less attenuation than shorter wavelength light, i.e., 1.5 µm, 3 µm, and 5 µm.

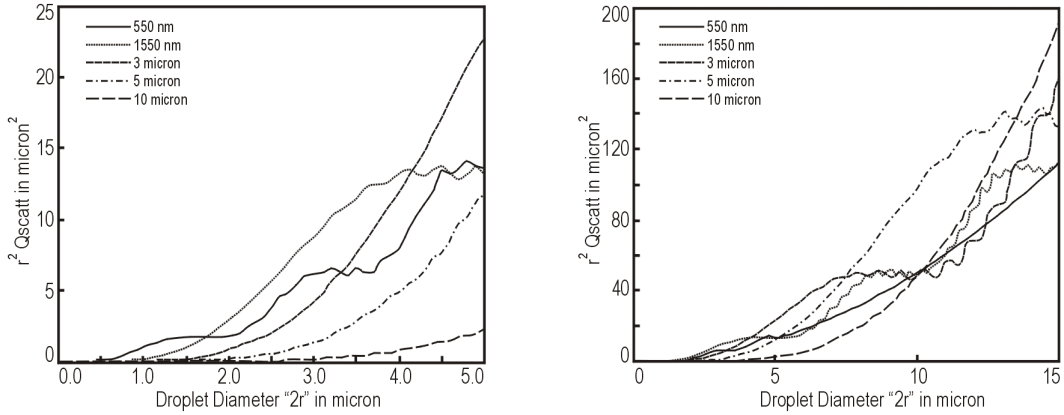


Figure 1: Attenuation versus droplet size diameter for different wavelengths [5].

Basing on Eq. (2), transmission of atmosphere versus the optical distance L for different wavelengths (0.8 µm, 1.5 µm, and 10 µm) and visibilities (900 m, 5 km, and 25 km) were calculated. For high rainfall rate, the atmosphere transmission is very low and is independent of the radiation spectrum. When the rate decreases, increase in the spectrum influence on the transmission is observed. The highest differences between transmission coefficients are noticed for low rainfall rate. In the analyses of visibilities, the atmosphere attenuation does not depend so strictly on a radiation spectrum for good weather conditions (Fig. 2). When the weather is bad, the transmission decreases more considerably for short wavelength ranges than for longer ones. In this case, the attenuation of 10-µm light is much less than shorter 0.8-µm and 1.5-µm wavelength.

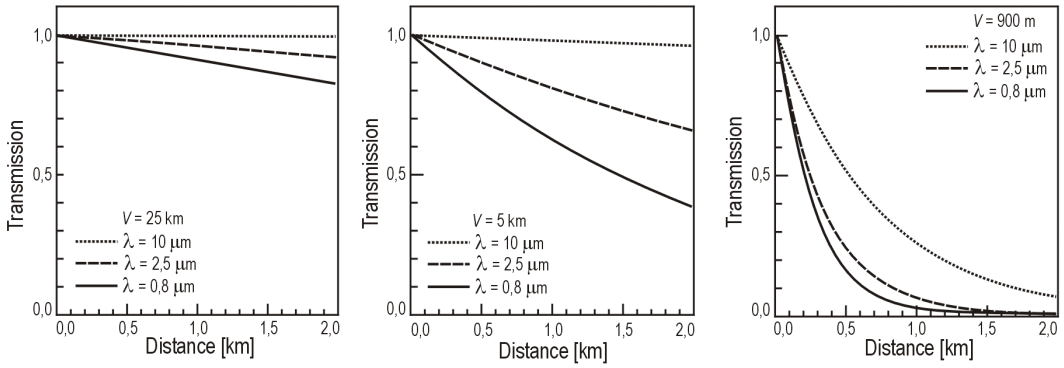


Figure 2: Atmosphere transmission versus optical distance between transmitter and receiver, for three values of wavelengths and for three visibility values.

3. TRANSMITTER AND RECEIVER OF FREE-SPACE OPTICAL COMMUNICATION SYSTEM

Progress in application of far-infrared radiation was stimulated by a development of quantum cascade lasers (QCLs). QCLs are semiconductor lasers that emit a portion of the electromagnetic spectrum in the mid- to far-infrared radiation. QCLs are unipolar and laser emission is achieved through the use of intersubband transitions in a repeated stack of semiconductor superlattices [6].

Semiconductor quantum cascade lasers achieve high power emission at room temperature and operate at various wavelengths between 3 µm to 17 µm. A unique feature of cascade lasers is ultra-fast relaxation of an excitation state of the introduced carriers as a result of dissipation of their energy on optical phonons. It decides about very short, picosecond-lifetime of these carriers, while in bipolar lasers, intrinsic lifetime is of the order of nanoseconds. Thus, cascade lasers are adapted for operation at high velocities of modulation, reaching even 100 GHz. Also, there are known the systems of laser transmitters, operating in the ranges of 5 µm and 8 µm, generating the pulses of 3–5 ps width with the repetition of about 12 GHz at the average output power of tens of milliwatts. One of the leading emergency long wavelength laser is Quantum Cascade Laser of LM-03-D Cascade Technologies firm. The laser module houses the laser. It controls the temperature of the laser as well as applying electrical pulses to the laser. The laser system generates radiation pulses with the pulse length of 0.02–10 µs, frequency of 0.01–5 10^6 Hz, power peak of 100 mW, and average power of 6 mW. The QCL laser is driven by LabVIEW system installed to our PC.

The receiver consists of a reflectance mirror of A8037-405 type, produced by Janos Technology Inc. firm, and a detection module. The mirror is used to obtain a higher optical signal on a photodiode detection area. The mirror shape is based on off-axis parabolic construction with a focal length of 3″. It provides full use of mirror surface for the received radiation. The mirror surface is covered by a thin gold layer. It provides high reflectivity in the wavelength range of 10 µm. On the receiving end, Polish Vigo Ltd. firm, $Hg_{1-x}Cd_xTe$ heterostructure detector was applied (PVI-2TE-10.6 type). The detector was optimized for detection of the wavelength radiation of about 10 µm at the working temperature of 230 K or 200 K. The temperature of the photodetector is obtained using two-stage TE-cooler. High performance and stability are achieved by using improved variable bandgap (Hg, Cd)Te semiconductors, optimized doping, improved surface processing and using hyperhemispherical lenses [8]. The main parameters are as follows: $\lambda_{op} = 10.6$ µm, detectivity 3×10^9 cmHz$^{1/2}$/W, responsivity ≥ 0.7 A/W, and response time 3 ns. Faster response may be achieved at a reverse bias and with high-frequency optimized devices. The photodiode integrated with AC coupled transimpedance preamp is characterized by transimpedance of 18000 V/A. The responsivity of the detection module is 12800 V/W. Fig. 3 shows a detection module diagram of a photoreceiver. High sensitivity of the photodiode-preamplifier circuit was achieved with individual matching of the IR detector to the preamplifier and minimizing noises of this system.

Detection module was constructed by prof. Piotrowski's group. In a common housing, there is a radiation detector, two-stage thermoelectric cooler with a temperature sensor, broadband transimpedance preamplifier, and miniature fan dissipating the heat emitted by a cooler.

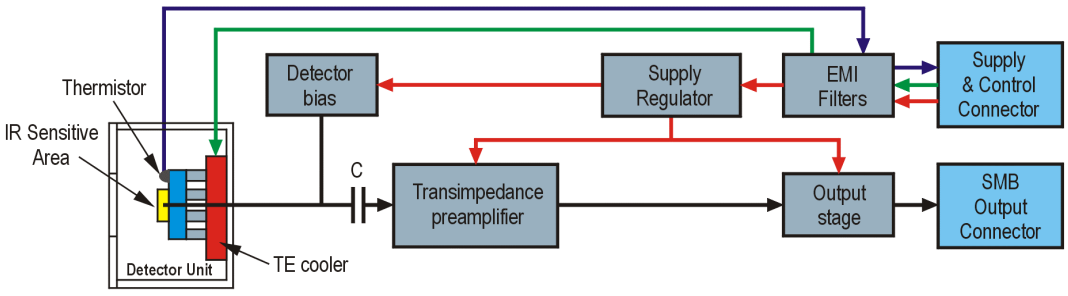


Figure 3: Diagram of a detection module [7].

Current signal from the detector is processed by a broadband transimpedance amplifier (up to 1 GHz). This amplifier forces simultaneously constant-voltage reverse supply (∼300 mV) of a detector. It creates the conditions to obtain maximum signal-to-noise ratio in a wide range of frequencies. Reverse supply causes production of low-frequency noises. However, they do not introduce large contribution into a total noise of broadband amplifier. The field of view of the detector module is about 6.6 mrad for a detector diameter of 1 mm.

The photographs of a 10-µm transmitter and photoreceiver are presented in Fig. 4.

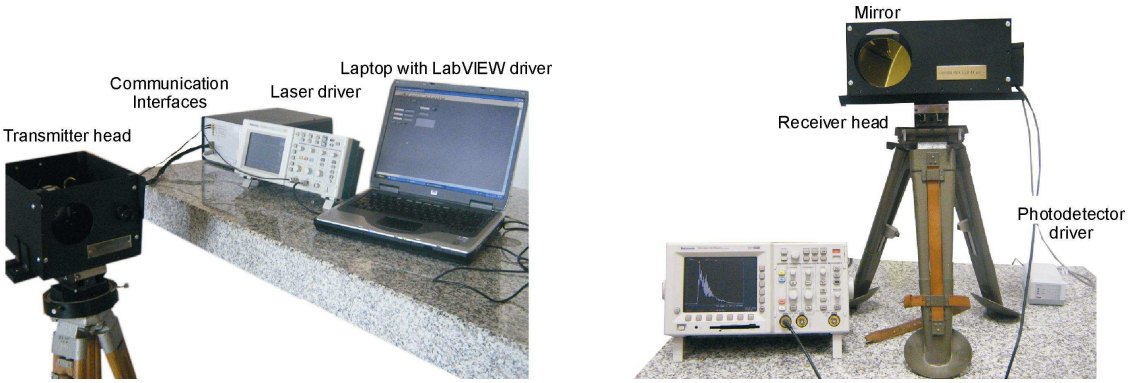


Figure 4: Photographs of FSO transmitter and photoreceiver.

4. CONCLUSIONS

The paper presents the works related to second-generation free space data link using quantum cascade lasers. Sources generating optical radiation near 10 µm and detectors sensitive to this wavelength range were analyzed. The proposed wavelength results from lower attenuation introduced by fogs (small aerosol particles) and from higher eye safety in relation to two other bands, i.e., 780–850 nm and 1520–1600 nm.

Fog is the most challenging media for FSO signals to penetrate. Fog attenuation is wavelength dependent. It can severally attenuate the signal depending on the fog type, but most common fogs attenuate 10-µm signals much less than shorter wavelengths. Therefore, 10-µm Free Space Optics systems have significantly lower attenuation in fog and give several times distance advantage over 1.5 µm systems.

Designers expect that due to application of QCL laser, that generates radiation near 10 µm, and highly sensitive detector optimized for this range of wavelengths, it will be possible to construct a second-generation optical system. The system will be characterized by larger detection range in adverse atmospheric conditions than a detection range offered by currently applied optical systems.

An optical radiation receiver should be characterized with a very high sensitivity. The Polish detector produced by Vigo System Ltd. firm was applied. High sensitivity was obtained due to a multilayer $Hg_{1-x}Cd_xTe$ heterostructure with immersion lens which was optimized for radiation detection at the wavelength of 10 µm. Detector's noises were reduced as a result of detector cooling by means of two-stage thermoelectric cooler.

ACKNOWLEDGMENT

The authors wish to thank the VIGO System S.A. staff for providing support for this work. This work was founded by the Polish Committee for Scientific Research (project No T00A 002 31/3962).

REFERENCES

1. Blaser, S., D. Hofstetter, M. Beck, and J. Faist, "Free-space optical data link using Peltier-cooled quantum cascade laser," *Electron. Letters*, Vol. 37, No. 12, 2001.
2. Maha, A., "Simulating atmospheric free-space optical propagation: Rainfall attenuation," *Proc. of SPIE*, Vol. 4635, 2002.
3. Kruse, P., et al., *Elements of Infrared Technology*, John Wiley & Sons, 1962.
4. Maha, A., "Simulating atmospheric free-space optical propagation," *Proc. of SPIE*, Vol. 4873, 2002.
5. Maha, A., "Free-space optics wavelength selection: 10 µm versus shorter wavelengths," http://www.ulmtech.com.
6. Kazarinov, R. F. and R. A. Suris, "Possibility of amplification of electromagnetic waves in a semiconductor with a superlattice," *Fizika i Tekhnika Poluprovodnikov*, Vol. 5, No. 4, 797–800, 1971.
7. User manual LS-03-D Laser System, Cascade Scientific firm. Document number: D-7010-0003.
8. www.vigo.com.pl; 2–12 µm photovoltaic detectors thermoelectrically cooled optically immersed.

Multi-band Antenna with Minimalization of Radiation towards Head

M. Wnuk[1] and R. Kubacki[2]

[1]Military University of Technology, Poland
[2]Military Institute of Hygiene and Epidemiology, Poland

Abstract— Intensive development of cellular personal communications system has been observed lately. Thus, protection of a man, and especially protection of his head against non-ionizing electromagnetic radiation generated by cellular telephones is becoming one of the most important problems. The results of elaborated microstrip antennas which have minimized radiation towards the user's head are presented in this paper.

1. INTRODUCTION

Wireless communication has been the fastest growing sector of telecommunication in recent years. Substantial contributions to this growth has been made in Europe, where many such systems have been developed: GSM (Global System for Mobile Communication), operating in the 900 MHz band, DCS-1800 (Digital Personal Communication System), using the 1800 MHz band DECT (Digital European (Enhanced) Cordless Telecommunications, using the 1900 MHz band and UMTS (*Universal Mobile Telecommunications System*) working in band circle about 2 GHz. Some experts predict that by the end of new century there will be more cordless than conventional telephones in use. In order to achieve an easy access to information, more and more modern cellular phones, which are able to work in two or three bandwidths, are being invented. This is why it is important to determine possible health effects of using such devices.

New trends require antennas to be cheap to produce and to meet tough operational requirements. The following requirements should be met by the antenna:

- a radiation pattern ensuring reliable communication regardless of the antenna's orientation;
- a wide operational frequency band, so that the spectrum of transmitted information should not be distorted;
- a minimum impact on the biological tissue of the user, to avoid health risks;
- little sensibility of the antenna parameters to the user's proximity;
- small size to facilitate mounting the antenna and using the telephone.

It is clear that some of these requirements are mutually exclusive, so a compromise has to be worked out in the designing process.

2. REQUIREMENTS CONCERNING THE ANTENNA RADIATION PATTERN

Presently, there are no formal requirements in force or informal requirements accepted by specialists concerning the recommended radiation pattern. The problem is further compounded by the fact that the user's head is in the near zone of the antenna (Table 1).

Table 1:

Class mobile station	Distance from mobile station		Limited power density
	0.5 m	1 m	
GSM 900 (0.8 W) Class 5	0.386 W/m^2	0.09 W/m^2	0.1 W/m^2
GSM 900 (2 W) Class 4	1.06 W/m^2	0.26 W/m^2	

There is concern about possible health effects of the field generated by the cellular phone, placed so close to the user's head These are not groundless fears. Research conducted so far on the effects of the PEM energy on living organisms suggests that electromagnetic energy, regardless of its amount, affects the human body. Its impact is harmless as long as it remains within the adaptation, compensation and regeneration capabilities of man, but may be harmful when it exceeds the limits of tolerance.

The point is to make sure that signals emitted by a base station are correctly received from all directions and that the user's head (especially bones, the brain and skin, which have a high level of thermal conductivity -14.6, 8.05, 4.42 $\mathrm{mW/cm^2{}^\circ C}$ respectively) is exposed to as little radiation power as possible.

A considerable part (45%) of energy, emitted by cellular phones now in use, is absorbed by the user's head, which is a health risk [L-3]. An important task is to protect users against radiation from cellular phones, which can be achieved in two ways:

- by reducing the power of electromagnetic field emissions towards the head to the necessary minimum,
- by limiting the time of exposure to such fields.

The latter condition is connected with the duration of the call, which depends largely on the user. The former condition, i.e., cutting down the amount of radiation power absorbed by the user's head, can be met — among other things — by modifying the omnidirectional radiation pattern of the antennas used in cellular phones so far. A compromise has to be reached between the requirements regarding the availability of signals received by the antenna from all directions and the protection of user's head against radiation. It is assumed that the radiation pattern is as shown in Fig. 1.

Figure 1: Requirements for antenna radiation pattern.

One of the ways of achieving this result is to replace stub antennas with an omnidirectional radiation pattern, popularly used in cellular phones, by antennas with a radiation pattern shaped in such way as to reduce radiation towards the user's head. Such antennas should be relatively small to be fit for use with cellular telephones. Microstrip antennas on a multilayer dielectric meet all these requirements.

3. MODELLING OF MICROSTRIP ANTENNAS

The throughout analysis of microstrip antennas which takes into account the structure of the layer and which is true for each frequency, is based on Green function and moment method. This method is based on solving the integral equation concerning the electric field generated by the currents flowing in the antennas element and its feeding systems. We simulate the flow of inducted current by means of distribution for base and test currents, then we test their mutual reaction by means of the functions. According to [L-5] the reaction has the form of:

$$z_{mn} = -\langle \underline{J}_m, \underline{J}_n \rangle = -\iint (\underline{E}_m \cdot \underline{J}_n)\, dxdy = -\iint (E_{mx} J_{nx} + E_{my} J_{ny})\, dxdy \tag{1}$$

The unlimited sequence of these functions is necessary for exact solution. We assume the limited number of these functions and, thus we obtain on approximate solution. The mutual reaction of the whole-analysed system can be expressed in the form of a matrix equation:

$$\begin{bmatrix} I_1 \\ I_{M+1} \\ . \\ 0 \\ I_{2M+2N+2} \end{bmatrix} = \begin{bmatrix} Y_{1,1} & . & . & . & Y_{2M+2N+2,1} \\ . & & & & . \\ . & & & & . \\ . & & & & . \\ Y_{1,2M+2N+2} & . & . & . & Y_{2M+2N+2,2M+2N+2} \end{bmatrix} \begin{bmatrix} 0 \\ V_{M+1} \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{2}$$

By solving this equation, we define the distribution of the currents flowing along the analysed structure on condition that the elements of general matrix impedance are now in our case they have the form of:

$$\begin{aligned} z_{mn} &= -(2\pi)^{-2}\cos\vartheta F^{-1}\left[F\left(G_{xx}\right)F\left(J_0\right)F'\left(J_0\right)\right]\left(x_m-x_n,y_m-y_n\right) \\ &\quad +(2\pi)^{-2}\sin\vartheta F^{-1}\left[F\left(G_{yx}\right)F\left(J_0\right)F'\left(J_0\right)\right]\left(x_m-x_n,y_m-y_n\right)=\frac{1}{Y_{m,n}} \end{aligned} \tag{3}$$

where:

- $F[G_{xx,yx}]$-Fourier transform of Green function
- $F[J_0]$-Fourier transform of base current

$\sin\vartheta$, x_n, y_n, x_m, y_m-respectively, the co-ordinates of the situated means of base and testing functions.

With the defined current distribution we can express the radiation pattern in the dipole plane by the following equation:

$$\vec{E}(\varphi)=\sum_{n=1}^{N} I_n\left[\vec{E}\left(\vec{J}_n\right)\right] \tag{4}$$

where:

I_n-coefficient of current distribution

The microstrip antennas patterns (presented in Fig. 2.) have been calculated on the basis of these relationships.

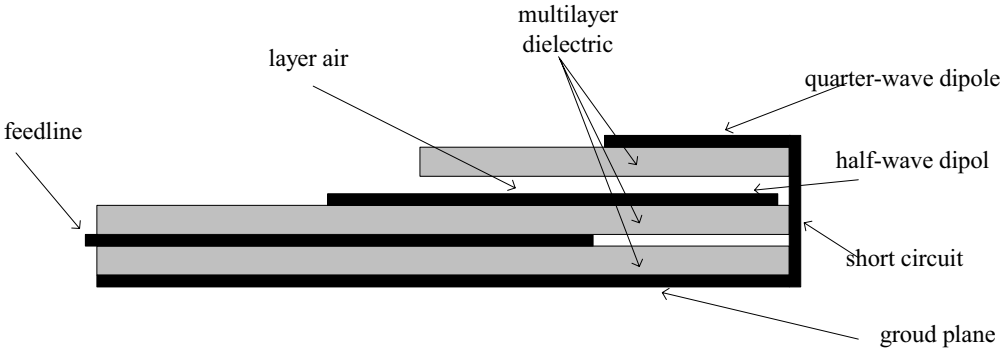


Figure 2: Dual band microstrip antenna.

4. CONSTRUCTION OF MICROSTRIP DIPOLE ANTENNAS ON A MULTILAYER DIELECTRIC

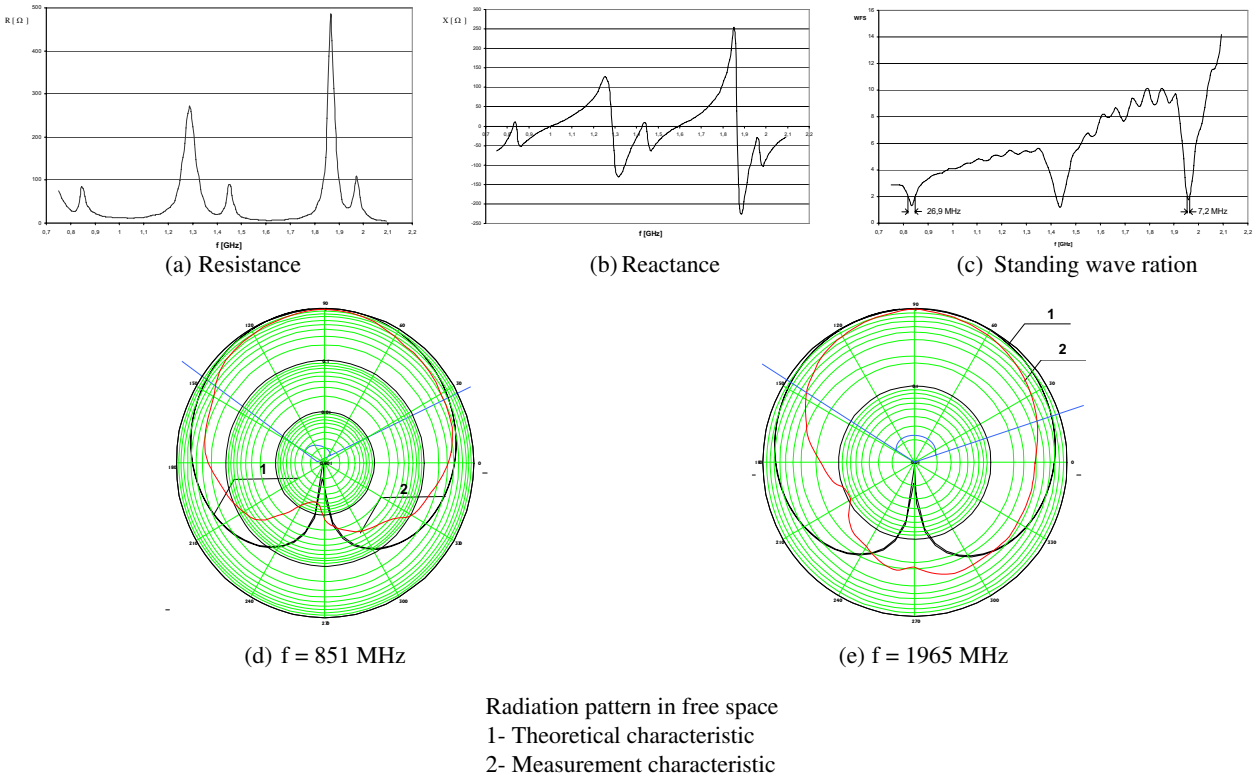
The dipole antenna is designed to operate in the 900 MHz band, but the same technology can also be used to design the antennas for other bands, e.g., 1900 MHz. It was constructed in the multilayer technology mainly to ensure an adequately wide operating band (ca 10%). Due to small size requirements (mobile communication systems) two versions of dipole antennas were constructed and tested:

- open circuit half-wave dipole
- short circuit quarter-wave dipole.

These antennas are shown in Fig. 2.

It is expected that the radiation pattern of the quarter-wave antenna in the E plane will be wide enough to achieve optimum antenna parameters. Results of measurements and calculation of antenna with feeding from an unsymmetrical strip line are shown in Fig. 3. Fig. 3(a) and (b) presents the results of measurements of impedances and Fig. 3(c) the results of measurement of standing wave ratio. In Fig. 3(d) characteristics of measured and theoretical radiation pattern for frequency band 900 MHz are presented, and in Fig. 3(e) for frequency 1900 MHz

The measurements were made in free space and in the presence of a user, in order to simulate a real life situation. The user stood on a rotating platform, holding the phone at 45° to the ground level.



(a) Resistance (b) Reactance (c) Standing wave ration

(d) f = 851 MHz (e) f = 1965 MHz

Radiation pattern in free space
1- Theoretical characteristic
2- Measurement characteristic

Figure 3: Measurement results of dipole antenna with a microstrip feedline.

5. CONCLUSIONS

The investigation confirmed that the antennas described here may be used in cellular phones operating in the 900 MHz band and in the 1900 MHz Radiation in the direction of the user's head is reduced in all the examined antennas, which was the main goal of the research work. It was assumed that a currently used antenna will be replaced by a new one, without any changes in construction of mobile phone. A new design of mobile phone will be made in close future.

REFERENCES

1. Amanowicz, M., W. Kolosowski, M. Wnuk, and A. Jeziorski, "Microstrip antennas for mobile communications," *Proc. of the Conference VTC'97*, Phoenix, USA, May 1997.
2. Bahl, I. J., "Microstrip antennas with paper-think dimensions," *Microwaves*, No. 10, 1979.
3. Guy, A. W., J. F. Lehmann, and J. B. Stonebridge, "Therapeutic applications of electromagnetic power," *Proceedings of the IEEE*, Vol. 62, No. 1, 1974.
4. Jensen, M. A. and Y. Rahmat-Samii, "Performance analysis of antennas for hand-held transceiver using FDTD," *IEEE Trans on Antennas and Propagation*, Vol. 42, No. 8, 1995.
5. Mailloux, R. J., N. P. McIlvenna, and J. F. Kernweis, "Microstrip array technology," *IEEE Transactions on Antennas and Propagation*, Vol. 42, No. 1, January 1981.
6. Bhattacharya, A., *Electromagnetic Fields in Multilayered Structures*, Artech House, London, 1994.

Electromagnetic Compatibility of the Military Handset with Hidden Authorization Function Based on MIL-STD-461D Results

Z. Piotrowski, L. Nowosielski, L. Zagoździński, and P. Gajewski
Military University of Technology, Poland

Abstract— In this paper we present algorithm and hardware architecture of the new military communication device dedicated for VHF battlefield radio stations. The results of the compatibility tests carried out for this device according to the norm MIL-STD-461D are also presented together with the figures and comments. Measurements were carried out at the Electromagnetic Compatibility Laboratory, Faculty of Electronics, Military University of Technology. The Laboratory have accreditation given by the Polish Centre for Accreditation for carrying out those military devices measurements. Described results of measurements will be a base for introduction this developed device into the Polish Armed Forces equipment.

1. INTRODUCTION

One of the most important security problems unsolved till this time in the communications links is voice authorization of the telephone or radio subscriber. Problem is more obvious when it is considered both current voice communications over digital and analog channels: VoIP, PSTN, GSM and state-of-the-art real-time, artificial, speech synthesis in the context of potential risk of impersonating, known in information technology as a spoofing. Presented algorithm and hardware architecture are the proposals for increasing security level in unprotected voice calls, as a response to a new telephone attacks named "voice spoofing".

At present, subscriber's authorization in telephone calls is performed in the most cases in subjective way using our Human Auditory System. Well known scheme of subjective authorization is based on knowledge of dedicated specific person's voice. New scheme of the authorization is based on objective way using watermarking (data hiding) technology. Additional binary signature as a inaudible watermark is embedded in the host's signal voice in the radio or telephone handset and then transmitted over communications links. At the receiving side, using dedicated hardware based on real-time processing algorithm, watermark is extracted and binary signature is compared with the set of binary signatures storage in the handset's memory. Results of the authorization process is displayed on handset's LCD. Hardware architecture of the handset is based on the fast, floating-point Digital Signal Processor TMS320C6713 and microcontroller ARM7TDMI family. Handset transmits, watermarked, analog voice, using standard voice interface in the TRC9200 battlefield VHF PR4G radio station family.

2. ALGORITHM AND HARDWARE ARCHITECTURE

Implemented numerical procedures in the Texas Instruments DSP TMS320C6713 floating point processor base on developed algorithm written in Matlab. The basic scheme of the embedder and

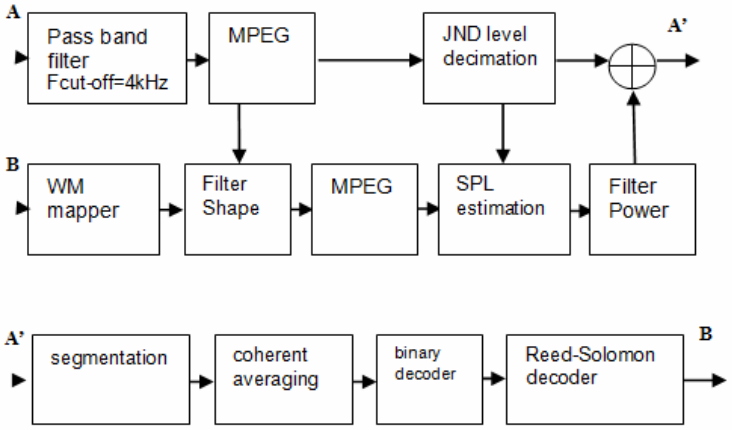


Figure 1: Base scheme of watermark embedder and decoder.

decoder are shown at Figure 1.

Embedder uses psychoacoustic procedure of computing Minimum Masking Threshold level based on ISO/IEC 11172-3:1993 standard. Host's speech signal is analyzed in frequency domain to compute all irrelevant and redundant frequency components as well as estimate the Just Noticeable Difference (JND) threshold considered as a smallest difference in a specified modality of sensory input that is detectable by a human being. Watermark (WM) is mapped into OFDM components using B binary signal pattern and then WM signal is shaped, using computed filters' characteristics, to be inaudible at presence of the host speech signal. In decoder received speech signal frames are segmented and converted to frequency domain. Spectrum is averaged and binary pattern is decoded. Reed-Solomon decoder is used for improving decoding efficiency. Basic scheme of the hardware is shown at Figure 2.

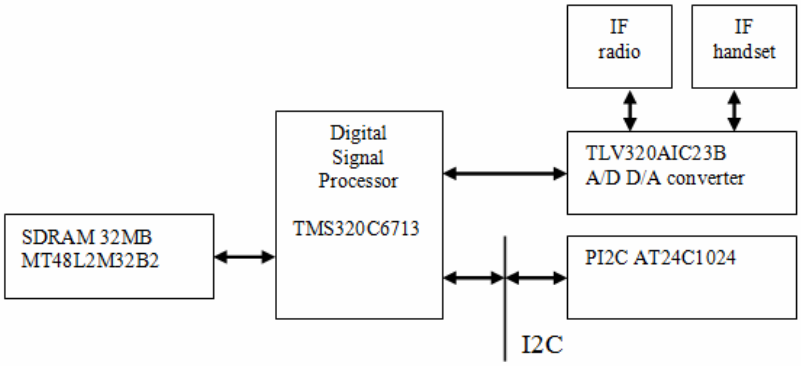


Figure 2: Basic scheme of the hardware — signal processing block.

The main processing block is based on DSP chip - Texas Instruments TMS320C6713 floating-point processor. Operational system's memory is based on SDRAM 32MB Micron chip MT48L2M32B2 and A/D and D/A converter processing the input and output data signal from following blocks: IF radio and IF handset. The PI2C AT24C1024 chip is used to store program and constant data when the power supply is cut off. The Inter Integrated Communication Interface (I2C) is connected to the control block managed by Atmel AT91SAM7S256 chip ARM7TDMI family.

3. RESEARCH METHOD PRINCIPLES

Military electronic devices should fulfill requirements described in the MIL-STD-461D standard concerning, among the others, acceptable radiating and conducting disturbances on the power supply terminals. According to the above recommendations, radiating disturbances for devices or set of devices installed on stationary or mobile objects should not exceed red line plotted values at Figure 4(a)–(c), whereas conducting disturbances on power supply terminals should not exceed the values shown at Figure 4(d). Workstation scheme for measurement of the radiating perturbations emitted by military devices is shown at Figure 3 (left plot). Tested device and receiving antennas arrangement in the unechoic chamber is shown at Figure 3 (right plot).

Measurements of the radiating disturbances emitted by military devices consisted in measurement electromagnetic field strength at the distance 1 metre from tested device using measurement antennas and receiver in the frequency range depending on maximum frequency intentionally generated by tested device. The maximum measurement range is 10 kHz ÷ 18 GHz. Measurements are taken for both vertical and horizontal polarization. Measurement process control is carried out using Rohde&Schwarz EMC32 control application installed in the control computer. This application computes, among the others parameters, electromagnetic field strength in the measurement antenna place using following formula:

$$E = U_{in} - T + K \tag{1}$$

where:

E-electromagnetic field in place of the installed measurement antenna in [dB(μV/m)],

U_{in}-input signal level in the measurement receiver in [dBμV],

T-high frequency line attenuation between measurement antenna and measurement receiver input in [dB],

K-correction coefficient of the measurement antenna allowing for antenna effective surface in [dB].

Measurements of the conducting disturbances on the power supply terminals are carried out in the same way as radiation disturbances with reservation that measurement antennas are replaced by artificial line LISN (Line Impedance Stabilization Network) which serves as the measurement sensor.

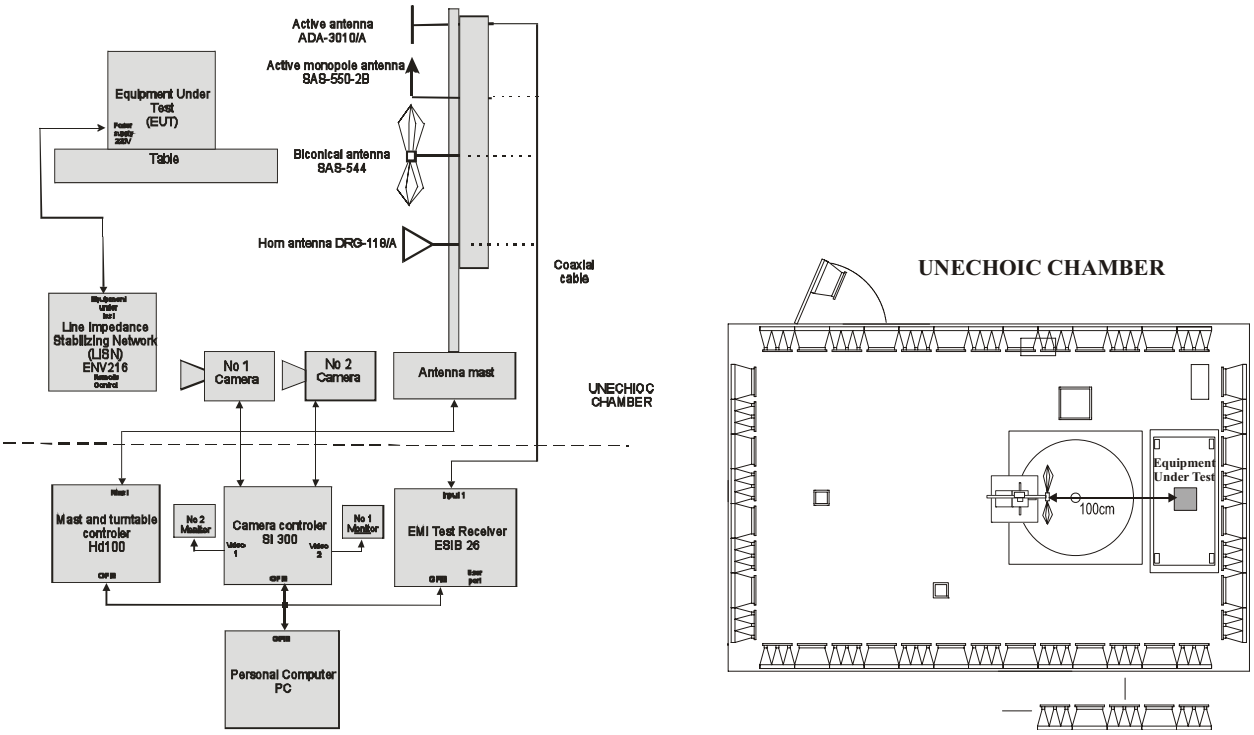


Figure 3: Workstation's block scheme for measurement of the radiating disturbances emitted by military devices (left) and tested device with receiving antenna arrangement in the unechoic chamber (right).

4. TEST BED

The measurement unit named the Equipment Under Test (EUT) consists of developed handset with hidden authorization function, working together with military VHF TRC9200 radio station PR4G family, was tested using MIL-STD-461D recommendation (RE-102, CE-102 specifications). Measurements are carried out using the receivers equipped with peak detectors. Chosing the place for the test bed it should be able to distinguish conducting perturbations inserting by tested device from the outside perturbations background. In this respect, to assess the usefulness of the place, it can be estimated the disturbances background level measured at the presence of the switched off tested device. The level of the disturbances should be at least 6 dB less than acceptable levels shown at Figures 7, 8, 9 and 10.

The recommendation MIL-STD-461D states that receiver should fulfill the requirements concerning the detector type, RBW filter waveband as well as measurement time. In order to do that the ESIB26 receiver produced by *Rohde&Schwarz* was used. This receiver, according to the producer's declaration, fulfills all recommendation's requirements. Artificial line LISN is used for isolation of the tested device from external disturbances occurring in the power supply wires and as a sensor for conducting disturbances detection in the power supply terminals. Measurement procedure is using artificial line ENV216 produced by *Rohde&Schwarz*. This device has impedance $50\,\Omega/(50\,\mu H + 5\,\Omega)$. Measured device should be connected to the artificial line from the power supply side, having uncover ending power supply lead wires. The length of the pliant power supply wire during measurements can not be longer than 2.5 m. All the power supply wires should be taken on 5 cm height above ground or dielectric base. The antennas should be used during measurements as follows:

- at the range of 10 kHz up to 30 MHz it should be used stub antenna having 103 cm height, and

placed perpendicularly to antenna counterpoise in the form of square sheet having length side not shorter than 60 cm. Antenna should be equipped in coupler to adjust antenna impedance to the measuring receivers. We used active stub antenna SAS-550-2B type

- at the range 30 ÷ 70 MHz was used biconical antenna SAS 544 type
- at the range 70 ÷ 1000 MHz was used active dipole ADA-3010/A type
- at the range form 1 GHz up to 18 GHz should be used horn antenna. The DRG-118/A antenna type can be used for this purposes. The extended measurement unreliability, estimated for procedures RE-102 and CE-102, equals 3.5 dB and 3.3 dB respectively.

5. MIL-STD-461D RESULTS

The radiating disturbances values (for vertical antennas polarization) were measured using peak detector when radio station had run in the upper frequency range. These values are shown at Figure 4(a)–(c). The results at Figure 4(a) and (b) are not exceeding, recommended by the standard, critical values in all frequency range from 2 MHz 1 GHz considering the extended unreliability of measurements. Critical radiating disturbances values are exceeded using frequency range shown at Figure 4(c). It is the result of the radio station transmission mode on the higher frequencies. These exceeded values are not considered during assessment of conformity with standard recommendation RE-102. The measurement results for conducting disturbances in the range from 10 kHz up to 10 MHz are shown at Figure 4(d). The results are not exceeding, recommended by the standard, critical values considering the extended unreliability of measurements.

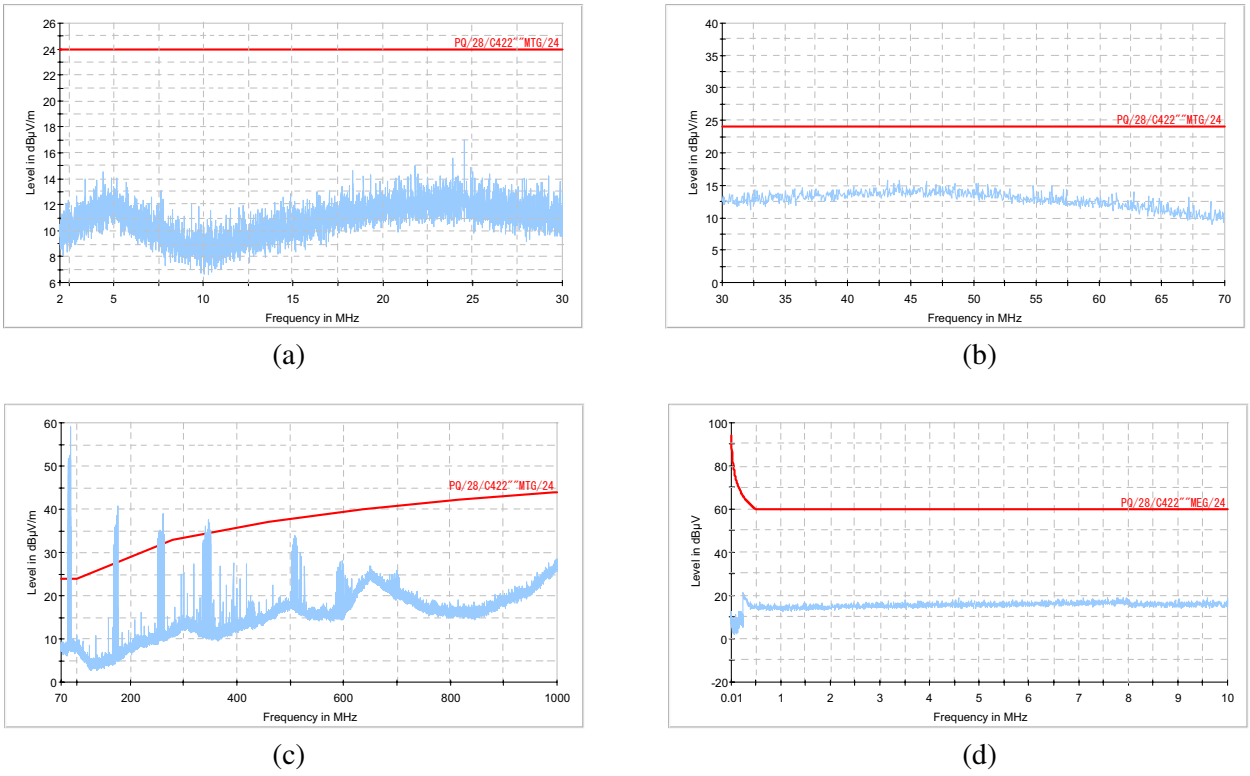


Figure 4: The radiating disturbances values, measured using peak detector, for the ranges: (a) 2 MHz–30 MHz, (b) 30 MHz–70 MHz, (c) 70 MHz–1 GHz and (d) Conducting disturbances values on the power supply terminals for the range 10 kHz–10 MHz.

6. CONCLUSIONS

The measurement test bed is dedicated for assessment of the conformity of tested military devices to the MIL-STD-461D recommendation as well as RE-102 and CE-102 procedures. Tested radio station unit with handset with hidden authorization function fulfills requirements of the RE-102 and CE-102 procedures. Handset with hidden authorization function was developed and designed in Telecommunications Institute, Electronics Faculty (EF), Military University of Technology (MUT). The measurements were carried out in the Electromagnetic Compatibility Laboratory, EF, MUT.

REFERENCES

1. Piotrowski, Z. and P. Gajewski, "Novel method for watermarking system operating on the HF and VHF radio links," *Computational Methods and Experimental Measurements XIII, CMEM XIII*, 791-800, WIT Press, Southampton, Boston, 2007, ISBN: 978-1-84564-084-2, ISSN: 1746-4064, ISSN: 1743-355X.
2. MIL STD 462D, Military standard. Measurement of Electromagnetic Interferences Characteristics, January 11, 1993.
3. MIL STD 461D, Military Standard. Requirements for the Control of Electromagnetic Interference Emissions and Susceptibility, 11 1993.
4. NO-06-A500, Procedury badań zakłóceń elektromagnetycznych i odporności na narażenia elektromagnetyczne.
5. NO-06-A200, Dopuszczalne poziomy emisji ubocznych i odporność na narażenia elektromagnetyczne.
6. Przesmycki, R., Instrukcja stanowiskowa IS-08 pomiar zaburzeń promieniowanych emitowanych przez urządzenia wojskowe. Procedura PRE-02 normy A500, LAB KEM WEL WAT, Warszawa, 2007.

Thermal Design and Electromagnetic Capability Design of the Microwave Power Module

Jinling Zhang[1], **Yinghua Lu**[2], **Biao Yang**[2], **Rongrong Li**[1], **and Jinsheng Yang**[3]

[1]School of Electronic Engineering, Beijing University of Posts and Telecommunications
Beijing 100876, China
[2]School of Telecommunication and Networks Technology
Beijing University of Posts and Telecommunications
Beijing 100876, China
[3]Beijing Vacuum Electronics Research Institute, Beijing 100016, China

Abstract— This paper has introduced the thermal design and Electromagnetic Capability (EMC) related to the Reliability of Microwave Power Module (MPM). We have given the primary result of thermal design, and have analyzed the EMC about the dependability of the MPM, and have presented the way of research. So we got the theory of the design about MPM working stability and Reliability.

1. INTRODUCTION

Microwave Power Module (MPM) is a RF power amplifier providing middle or high RF power, and it has many advantages such as small size, high efficiency, high reliability and making little noise [1–3]. MPM consists of the solid state power amplifier (SSPA) and the vacuum helix miniaturized traveling wave tube (Mini-TWT) as well as the high-voltage switch electronic power conditioner (EPC). The configuration is shown in Fig. 1. MPM is mainly used in military and satellite communications areas, such as Unmanned Aerial Vehicles (UAV), military satellite communications, interference, range, Electronic Warfare, synthetic aperture radar (SAR), the radar-power radiation sources, the weather radar, phased array radar system and traction bait etc., some country has already made MPM products series covering the frequency range of 1–40 GHz.

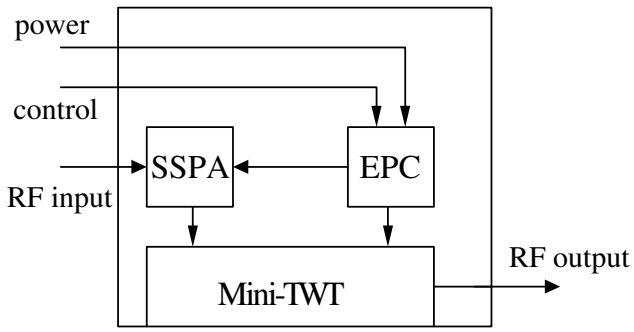


Figure 1: The configuration of MPM.

MPM is a high-density integrated as shown in Fig. 1, the working principle of MPM is that the input RF power is first amplified by the SSPA, and stimulates the Mini-TWT, then Mini-TWT amplifies the RF power (the last power amplification) and export the RF power. At the meantime the EPC with control and protection module provides electrical power to SSPA and Mini-TWT.

MPM is a high-density integrated been separated in collectivity design. There are two crucial problems in the collectivity design of MPM, one problem is the EMC inside and outside. The inside EMC has much influence in its working stability and reliability, while the outside EMC has much influence when many MPM are using in array, and it also has influence to other equipments working near the MPM. The other problem is heat transfer, every component in the MPM inside are producing heat, it is very important to transfer this heat for the steady working of the components. An integrated advisement has to be given in the collectivity design, because there is difference of the highest working condition temperature between vacuum components and solid components.

2. THERMAL DESIGN OF MPM

Heat analysis has been processed for collectivity structure using ANSYS under the foundation of the collectivity structure design of MPM. As shown in Fig. 2, only primary heating components have been considered in ANSYS modeling, such as, Mini-TWT, high frequency plane transformer and so on. The whole module's boundary has been set 25 °C for there is water-cooling device at its side, therefore we got the modeling results in Fig. 3. We can see that the highest temperature of the whole module is the collecting pole's edge of Mini-TWT which is 82 °C from the modeling results. However this modeling is only the elementary modeling results and more particular heat modeling is processing right now.

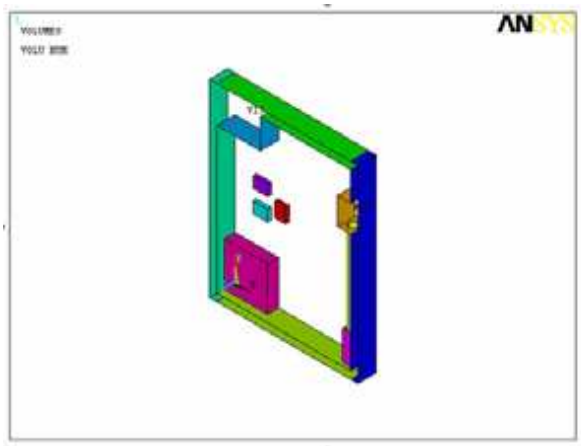


Figure 2: Simulation configuration of MPM.

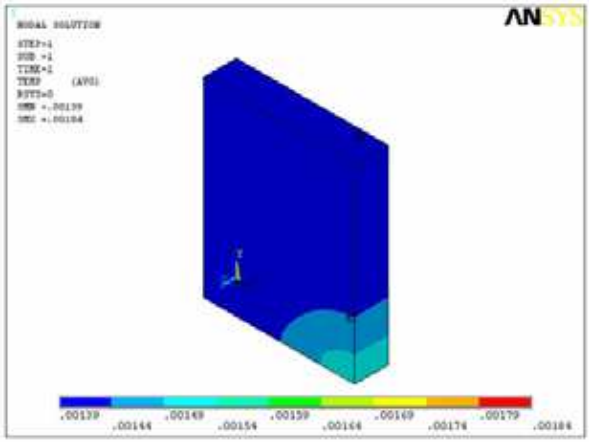


Figure 3: Simulation result.

3. ANALYSIS AND DESIGN OF EMC IN MPM

In MPM, EPC, SSPA and Mini-TWT are packaged in a box in high density, using High Voltage Wire or cables to connect. The electromagnetic interference (EMI) occurring in MPM has a relationship with the electromagnetic environment where it is, the interaction between different MPM when MPM are using in array, and the electromagnetic coupling of the internal components. As a result of that we must take measures to improve EMC [4].

Components level EMC technology is very important for achieving the EMC of MPM. If we fail on this level, it will not only affect the EMC of MPM, also directly affect the function implement, the reliability and the stability when MPM works.

3.1. Internal and External Interference Source of Switching Power Supply (EPC)

EPC is a high-voltage switching power supply. The main internal interference sources are said below: First, the large power transistor of the switching power supply works on the state of switching high-voltage and high-current. When it changes from on to off state, there will form surge voltage, and similarly from off to on, there will be surge current. Their high harmonic components become interference source by eradiating to space or transmitting through power line. Second, when it changes from off to on state, the switch transformer secondary rectify diode produce a peak of reverse current, under the restrictions of the reverse recovery characteristics. The reverse current, the diode junction capacitance as well as the down-lead inductance form a sinusoidal damping oscillation, which contains a lot of harmonic components. The sinusoidal damping oscillation constitutes interference, and may affect the work of MPM or damage some part. Moreover, when power switching transistor turns to off, the inductance leakage of high frequency transformer will pose a shutdown peak voltage, causing electromagnetic interference. In EPC, high frequency transformer has a function of isolation besides transform. However, in high-frequency circumstance, isolation effect of transformer is not complete. So high-frequency noise will transmit between the primary side and the other through distribution layer capacitance of the transformer [6]. The high-frequency noise will spread outside through the distributed capacitance between high-frequency transformer and the shell by coupling.

The external interference sources of the power switching supply are these things below: The electronic devices using in the same electromagnetism environment with the power switching supply

especially the ones using the same distribution equipment, will bring conduction and radiation EMI to the supply; Besides, the instantaneous change to the load of power switching supply will form power surge; Third, voltage disconnection may happen to the power supply because transmission transformer goes wrong. Therefore, we must solve the EMI problem of EPC.

3.2. The EMC between SSPA and Mini-TWT

SSPA is an ultra-small solid-state power amplifier. It is connected with Mini-TWT by coaxial line, and sets as Mini-TWT's former power amplifier; the output of Mini-TWT is the coaxial line with SMA joints at both ends which is close to EPC. When transmitting high frequency signal inside MPM, every part's electromagnetism radiation in use is bound to be electromagnetic interference. In particular it turns out to be weak signal transmission interference, which makes MPM performance deteriorate and stability decline. Further more, we must consider the problem of coaxial line and its joints' EMC, as a result of that there are some joints at both ends of SSPA and Mini-TWT where there will certainly be microwave leak.

3.3. The EMC between MPM

MPM internal circuitry and components will produce EMI at work which will bring electromagnetic radiation. On the one hand, the bosom of the system will form serious EMC problem due to EMI conduction and coupling; On the other hand, electromagnetic radiation will affect the devices else. Considering using MPM in array, the EMC outside MPM also has to be think better of, so as to avoid the EMI between MPM and external electromagnetic environment effects to keep the collectivity performance of the arrays.

3.4. Grounding Problem

Grounding plays a very important role in EMC. It prevents the electromagnetism field, voltage, or current which are used or formed by a point in the system from making interference to the other circuit through common grounding resistance transfer. As Mini-TWT working in high-voltage state, EPC provides Mini-TWT thousands of volt high voltage electricity and provides SSPA DC voltage. To ensure personal safety, EPC, Mini-TWT, and SSPA have to share grounding and be connected to the earth. Thus the grounding problem must be studied. At the unite-test course of developing MPM, the phenomenon that all parts don't work normally after sharing grounding appears several times.

4. MPM EMC SOLUTION METHODS

We intend to use emulation, and diagnostic tests to solve the electromagnetic compatibility problems of MPM.

4.1. Computer Emulation Analysis

The work of EMC modeling starts in two directions. One side is the restriction of EMI. And the other side is improving the capability of resisting electromagnetic interference (EMI), and accommodating to electromagnetism environment [7]. The purpose of EMC modeling is cooperating with the design and the evaluation of EMC. So usually EMC modeling and design, testing should be carried out together. We have done preliminary simulation on the performance of EPC already [8].

4.2. Near-field Test

The EMC of a system or a device is designed. And the product must pass EMC test [5]. The EMC test classifies into diagnostic testing and certification testing based on the different stages of the product. The request of certification testing is stricter; its process is more cumbersome, and costs more. But diagnostic testing can be used in every part of the development. We just need to find interference source and estimate its frequency and amplitude magnitude. Diagnostic testing is the near-field test. The interference source, radiant intensity, and radiant direction of EMI will be revealed in the near-field automatic test. Our next job intends to be test the preliminary designed MPM on EM-SCAN, so as to find the accurate position and the interference intensity of MPM interference source.

5. CONCLUSIONS

Improving the heat elimination capability and the EMC characteristic of MPM internal parts and components, as well as enhancing stability and reliability, have to be considered at the beginning of the development which is program design part. That can avoid the bad consequence that the design is congenitally deficient but can not be fixed later. Certainly, even MPM in the initial design

is very well, but in actual test parts of the guidelines still may fail to meet the requirement. For the aim of achieving the high level design of MPM and developing the MPM with high reliability, we must continue to be enriched in theory, to accumulate and explore in practice. So it can provide a powerful guarantee of consistency, stability for MPM batch production in the future.

ACKNOWLEDGMENT

The project was supported by National Natural Science Foundation of China 60671055 and 60771060.

REFERENCES

1. Armstrong, C. M., "Advances in microwave power modules," *International Conference on Microwave and Millimeter Wave Technology Proceedings*, 1–5, 2003.
2. Heinen, V. O. and K. E. Kreischer, "Vacuum electronics development at Northrop Grumman," *International Conference on Microwave and Millimeter Wave Technology Proceedings*, 293–294, 2004.
3. Hargreaves, T. A. and C. M. Armstrong, "Ku-band MPM booster helix TWT design and validation," *International Conference on Microwave and Millimeter Wave Technology Proceedings*, 5–11, 2003.
4. O, J. C., S. L. Lin, and Y. H. Lu, *Electromagnetic Compatibility Design of Electronic Equipment*, Publishing of Electronic Industry, Beijing, 2003.
5. Liu, J. H., "The measurement means of the electromagnetic compatibility and the design of the application in the practice," *Journal of Changchun Institute of Optics and Fine Mechanics*, Vol. 27, No. 2, 119, 2004.
6. Lu, H. G., "Technology for control EMI in multi-layer microwave MCM," *Safety & Emc.*, Vol. 2, 29, 2004.
7. Fu, J. H. and Q. Wu, "The summarization of microwave EDA electromagnetic simulation software," *Journal of Microwares*, Vol. 20, No. 2, 93, 2004.
8. Zhang, J. L. and Y. H. Lu, "Electromagnetic capability design of the high-voltage switch electronic power conditioner in the microwave power module," *Proceedings of Asia-Pscific Conference on Environmental Electromagnetics*, 816–819, Dalian, China, 2006.

The Method of the Calculating of Frequency Characteristics of Image Gaining and Processing Systems

Konrad Maj and Grzegorz Stępień
Remote Sensing and Geoinformatics Military University of Technology, Poland

Abstract— The basic parameter for defining the quality of the image is its resolution which can be under the influence of:

- conditions in which the image was gain (the conditions of exposure);
- parameters of image gaining system;
- parameters of processing system.

Quantitative estimation of the mentioned influences is essential in order to estimate the quality of the gaining system and image processing. That operation is possible by calculating the modulation transfer function (MTF) of the whole system. The calculation is provided via the research of the suitability of the impulse response of the system to the model signal. The model signal for the image is a suitably prepared test, which is the subject for observation and then for electronic processing into digital form.

Peak resolution of a digital image is defined by the parameters of the applied optical and detection system in the first place, and then the parameters of the processing devices such as scanners, converters and recorders. The gaining of strip tests images is the classic method of researches on influences of these devices' parameters and conditions of acquiring on the image parameters. However, in case of remote sensing imageries it is required to have done the tests of properly large sizes and covered with colours of definite contrasts. This is expensive and long-lasting procedure, but most of all subjective. The quality assessment of the gained image is carried out through the optical observation conducted by a certain group of people and then calculating the average of observation results.

In the article the authors suggest applying a new method for this kind of researches, based on defining modulation transfer function (MTF) of the gaining and image processing system on the basis of test images of much smaller sizes and easier to execute. The image is feather edge between the areas of maximum contrast. The principle advantage of the method is the objectivity of measurement, which guarantees its repeatability.

1. INTRODUCTION

The principle assumption of the authors is the fact that basic characteristics of the linear system of gaining and processing an image are: impulse response of the system and optical transfer function. The impulse response of the system is understood as a reaction on input function coming from punctual radiation source. It is also called a spot or function of broadening in the professional literature [2–4].

In case of one-dimension system its equivalent is a line broadening function, which is a response of system to input function coming from extremely narrow (infinitesimal) slit or a boundary of the two areas of high contrast of luminance.

2. THEORETICAL BASIS OF THE METHOD

From the point of view of an information theory, starting point of image processing system, e.g., luminous screen, can be consider as a multi-channelled optical line of communication. The number of channels (which have the spatial character) equals the number of discreet location in a field of screen elementary fields of local averaging of function of intensity distribution. These areas present the smallest elements of image called "pixels", but their sizes are limited by physical and technical conditions in processing stage. The number of pixels as well as a number of possible levels of greyness for a pixel, and the number of frames transferred by system in a unit of time determine the measure of informative proprieties of the system pertinent to structural, metric and temporary information represented generally by a signal.

The distortions of a signal of temporal-spatial character with the gradation of intensity transferring the mentioned three kinds of information appear in form of:

- contrast distortions connected with structural information;
- noise distortions pertinent to metric information;
- inertial distortions concerning the temporary information.

The assumptions of the presented method require further research on the first two kinds of distortions.

Real-valued function of two spatial variables defining intensity distribution in an output reference plane, which is received as a result of punctual stimulation in an input point, is called the point spread function PSF. Usually a condition of normalization is superimposed on this function:

$$\int_{-\infty}^{+\infty}\int PSF(x,y)dxdy = 1 \tag{1}$$

The PSF function is the fundamental for the picture generating process. It is, however, of the ideological character in the first place due to difficulties with its practical measuring. The response of the system to linear stimulation, so-called line spread function (LSF), strictly connected with PSF, is devoid of this defect.

This function describes the changes of intensity across the line in output plane. The condition of normalization for LSF function is:

$$\int_{-\infty}^{+\infty} LSF(u)du = 1 \tag{2}$$

where the axis u is perpendicular to the considered line. The line spread function is possible to receive by recording the signal of luminance during the movement along the spot of dispersion, a detector in a form of very narrow port. If direction of this movement marks the axis x, while the port is parallel to the axis y then it is equivalent to the relationship between PSF and LSF presented by formula:

$$LSF(x) = \int_{-\infty}^{+\infty} PSF(x,\, y)dy \tag{3}$$

There is the third function connected with already introduced spread functions, being the response of the system to the edge stimulations (discrete) so-called — Edge spread function ESF. This function describes an output image being an image projection of an input image in a form of the edge change (jump) of intensity in the input plane. The edge spread function can be received by recording the signal of luminance during the movement along the spot of dispersion, a detector in half-plane form. By keeping the direction of the movement along the axis x we can describe such a system with the following formula:

$$ESF(x) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} PSF(x,\, y)dy \tag{4}$$

This is what results from formulas (2), (3), (4):

$$LSF(x) = \frac{d}{dx}ESF(x) \tag{5}$$

These functions can be used directly to estimate the quality of processing systems or there can be calculated on their basis different profiles such as MTF. The presented method based on practical calculation MTF requires the loading of a parameter called depth of modulation of in/out signal. The parameter with large practical meaning can be defined as a ratio of M and M_0 of depth of signal modulation on entry (input) to depth of signal modulation on exit (output) and is the measure of quality of transfer of modulation by the system. The relation (ratio) depends on the frequencies of spatial f_x, f_y. Quantity M/M_0 as a function of f_x, f_y we call modulation transfer function MTF. We can describe MTF as follows:

$$MTF = \frac{I_2}{I_1} = \sqrt{F_s^2\,(f_x,\, f_y) + F_c^2\,(f_x,\, f_y)} = |F_s\,(f_x,\, f_y) - jF_c\,(f_x,\, f_y)| \tag{6}$$

where:

$$F_s(f_x, f_y) = \int_{-\infty}^{+\infty} \int PSF(x, y) \sin[2\pi(f_x x + f_y y)]\, dxdy$$
$$F_c(f_x, f_y) = \int_{-\infty}^{+\infty} \int PSF(x, y) \cos[2\pi(f_x x + f_y y)]\, dxdy \quad (7)$$

are sine and cosine of Fourier transform. It is easy to notice, that Equation (6) describes MTF function with the help of function modulus of Fourier transform of function PSF:

$$MTF(f_x, f_y) = |PSF(x, y)| \quad (8)$$

Therefore the conclusion is that the quantitative estimation of system credibility in function of spatial frequencies requires the calculation of characteristics of MTF, gaining and processing system considered as the whole.

3. THE METHOD OF REALIZATION OF EXPERIMENT AND DIGITAL CALCULATION OF MTF

The idea of the experiment carried out by the authors was based the use of the results of measurement of signal coming from monochromatic camera CCD recording the image of a black and white test to calculate the point spread function PSF. The test was in a form of white cardboard with mark-black plush material, stuck in the centre. A sharp edge with high contrast of luminance was made in this way. The edge was located perpendicularly to the direction of recording of CCD. The scheme of a system is presented on Figure 1.

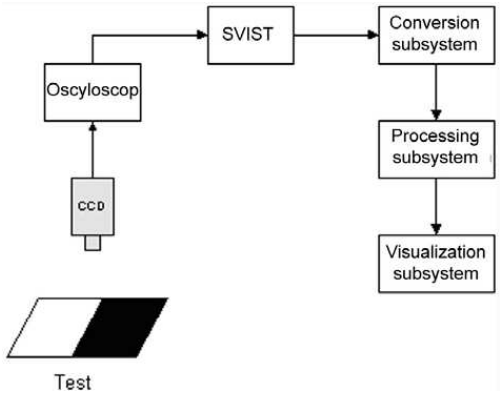


Figure 1: The scheme of testing system.

The analogue signal from a camera was sent to oscilloscope. With the help of oscilloscope the input parameters of the camera were set up, such as the diaphragm and focal length as well as specific lighting of test. The adjustment of these parameters lasted until the moment when the adjustment of a signal coming from the black fragment of the test and a level of black standard analogue signal was get on oscilloscope. It was assumed that such an arrangement equals the maximum contrast of luminance of the test (Figure 2).

Then an image signal was converted by A/C conversion with the help of SVIST card and a test image was recorded into a file form. The following activities were similar to those that are applied to all other images in the system, that means transformation of graphic formats was done, sending data to subsystem data processing, recording into a new digital file form (in the system of Erdas Imaging) and was displayed on a monitor. Next the values of lightness of pixels, lying on the line crossing the border of black and white areas, were read out (Figure 3).

Obtained series of lightness measuring tests in a function of the following number of a pixel on a line created the vector of measuring data v_j. That vector was read in to the computer programme Mathematica, where the next calculations were executed.

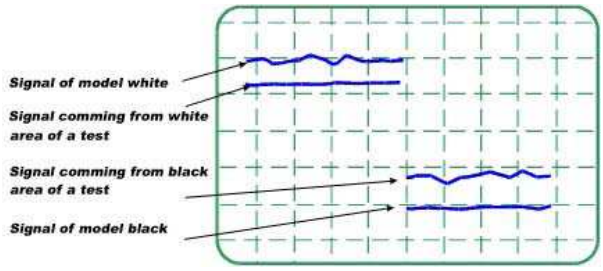


Figure 2: Display of oscilloscope with model and real signals.

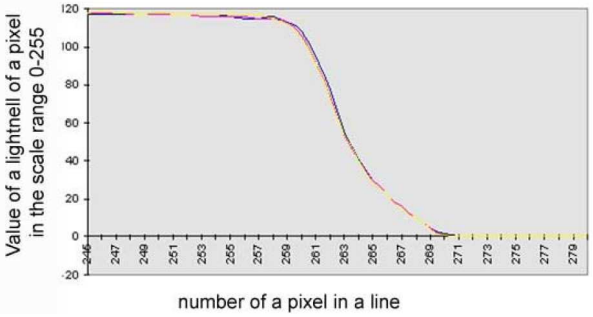


Figure 3: Values of lightness of pixels, lying on the line crossing the border of areas.

Because the information connected with the resolution of the system is contained in a shape of impulse response of an arrangement samples having numbers $j_{\min} = 245$ to $j_{\max} = 270$ were separated for further analysis. For the needs of calculation the normalised vector of data w_n was defined as:

$$n = 0, \ldots (j_{\max} - j_{\min}) \quad w_n = v_{n+j_{\min}} \quad w_n = \frac{w_n - \min(w)}{\max(w) - \min(w)} \tag{9}$$

The vector of data w_n was presented (see Figure 4.)

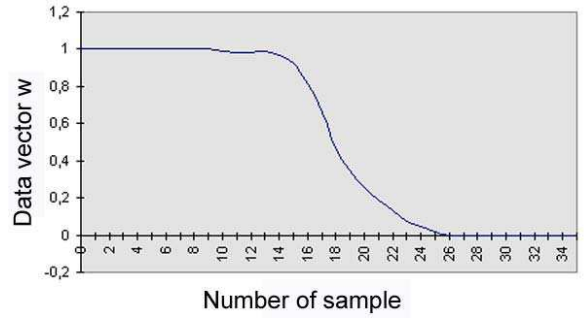


Figure 4: Vector of data w_n.

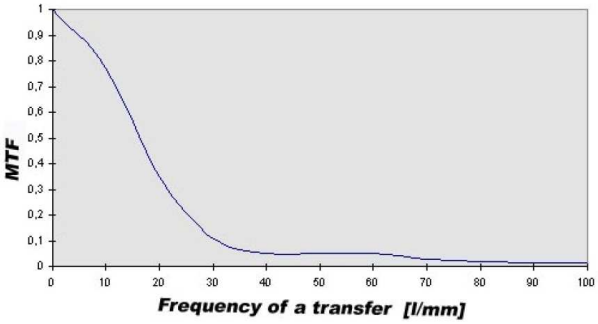


Figure 5: Modulation transfer function (MTF) of a system.

For numerical reasons calculations of Fourier transforms have to be held on the basis of sample numbers being the power of a number 2. Therefore the number of samples as a 2^8 was assumed and a discreet function was defined, describing a single impulse as a function f_k in a form:

$$f_k = if\left[k - (j_{\max} - j_{\min}),\ w_k,\ 0\right] \tag{10}$$

The domain of a function is being extended, establishing simultaneously that the function will take the zero value for points with indexes larger than $j_{\max} - j_{\min}$. To assign to a step of sampling the physical quantity in the studied system (size of a pixel) the step of sampling of recording camera was marked at $r = 0.02\,\text{mm}$. Then the Fourier transform of this function was calculated:

$$f = FFT(f)$$

The next step is the transition to physical co-ordinates assuming that the interval between samples is about $1/(r{*}n)$ where n is the number of computing points FFT. In this way we receive the optical function of transporting the OTF system, that is the dependence of amplitude of a spectrum in the function of spatial frequency presented in units 1/mm. Normalised to one unit the OTF module makes the function of transport of the MTF contrast, shown in the picture (see Figure 5.)

This profile makes the criterion of the assessment of system usefulness and its influence on image material of definite parameters. On the basis of this profile it is possible to determine what spatial frequencies are transferred in the system with a definite probability. It is possible to determine from these descriptions the contrast of objects that were recorded onto image material.

The received profile defines unambiguously that frequencies of about 15 l/mm (line per millimetre) are very well processed through the system used in the experiment. It is possible to unambiguously specify with what contrast all spatial frequencies will be transferred by reading out of the graph the coefficient of the depth of modulation.

4. FINAL CONCLUSIONS

By defining the modulation transfer function (MTF) of the digital processing system it is possible the optimal selection of the entrance parameters of the system (scanning resolution, frequency of sampling and quantisation) which will guarantee the minimization of losses caused by the applied devices. Besides that, in case of having the particular processing system it is possible to specify the required parameters of gaining the images on the basis of calculated profile MTF.

It is the universal function describing quantitatively the metric distortions which can appear on digital images. Because the way of calculating this function is clear, it guarantees objectification of the assessment of an image resolution ability transferred by any system of gaining and processing data. Applying of this function allows the considerable lowering of costs of executed tests and engagement of complicated equipment as well as large groups of people to asses the results of measurements.

REFERENCES

1. Woźnicki, J., "Podstawowe techniki przetwarzania obrazu," WKŁ, Warszawa, 1996.
2. Podręcznik, M. I., Podstawy teoretyczne, Warszawa, WKiŁ, 1988.
3. Roetling, P., E. Trabka, and E. Kinzly, "Theoretical prediction of image quality," *Journal of the Optical Society of America*, Vol. 58, No. 3, 1968.
4. Jain, A., *Fundamentals of Digital Image Processing*, Prentice-Hall, 1989.

A Comparison of Distortion Analyses Based on Volterra Series and Steady State Algorithm

J. Dobeš

Department of Radio Engineering, Czech Technical University in Prague
Technická 2, 16627 Praha 6, Czech Republic

Abstract— The most accurate way for computing the distortion coefficients is to find a steady state period of the signal, and then to determine spectrum by means of fast Fourier transform. In the paper, main features of a reliable and efficient algorithm for determining the steady state are described. However, even such accelerated method often needs a time-consuming numerical integration over many periods of the faster signal. For this reason, an efficient method for a fast estimation of the main distortion coefficients is also presented which does not contain frequent imperfections in the algorithm implementations. The algorithm uses higher-order Volterra series in a simple multistep algorithm which is compatible with a typical structure of the frequency domain part of software tools. Both methods are compared by analyses of the main intermodulation products of a low-voltage low-power CMOS RF four-quadrant multiplier.

1. INTRODUCTION

A natural and accurate method for determining intermodulation products is to find a steady state period, and then to compute the spectrum of that signal by means of fast Fourier transform. This algorithm has been implemented into our original software tool C.I.A. (Circuit Interactive Analyzer) with a possibility of the automatic determination of the unknown period of autonomous circuits. Basic theory of the steady state analysis is described in [1], some improvements especially for automating the procedure for the autonomous circuits are defined in [2]. As the implemented method for a numerical integration (which is necessary for the steady state algorithm) is accurate and very flexible (it is based on an efficient recurrent form of Newton interpolation polynomial [4, 5]), computed intermodulation products are available even for higher orders.

However, the numerical integration must still be performed over many periods of the faster signal and therefore the analysis is time-consuming. For this reason, another method for a fast estimation of the main interpolation products has also been implemented based on Volterra series. A brief introduction to using the Volterra series is shown in [6], a more comprehensive analysis can be found in [7].

A disadvantage of many of the implementations of the Volterra series consists in a creation of a new (relatively large and isolated) block of the program. In this paper, a simple modification of the method is defined, which is compatible with the frequency domain part of the program — the algorithm is built into and uses a relatively large part of the AC analysis.

2. GENERAL DESCRIPTION OF THE STEADY STATE ALGORITHM

The system of algebraic-differential equations of a circuit is generally defined in the implicit form

$$\boldsymbol{f}\left[\boldsymbol{x}(t),\, \dot{\boldsymbol{x}}(t),\, t\right] = \boldsymbol{0}, \tag{1}$$

and the problem of computing the periodic steady state response can be formulated as a solution of the nonlinear symbolic equation

$$\boldsymbol{x}_{\mathrm{ss}} = \mathcal{I}\left(\boldsymbol{x}_{\mathrm{ss}},\, t_0,\, t_0 + T_{\mathrm{ss}}\right), \tag{2}$$

where $\mathcal{I}\left(\boldsymbol{x}_{\mathrm{ic}},\, t_0,\, t_0 + T_{\mathrm{int}}\right)$ marks the solution vector after the implicit numerical integration of (1) with an initial condition $\boldsymbol{x}_{\mathrm{ic}}$ on the interval T_{int}.

Instead of usual (often very long) numerical integration of (1), much shorter one is performed. Samples of the solution are immediately collected after each of the periods (which must be determined for autonomous circuits). The samples become input for the scalar ϵ-algorithm [3] which is able to estimate a state of the system in future. The outcome of the ϵ-algorithm becomes a new initial condition for the system (1) and the process is repeated until the steady-state is detected.

A number of periods needed for an extrapolation loop depends on the number of slowly decaying transients. That number can be reduced by a low-pass filtering performed by numerical integration of a suitable length

$$\boldsymbol{x}_j^{(0)} := \boldsymbol{x}_j\left(t_0 + \Delta t_{\text{extpol}}\right) = \int_{t_0}^{t_0+\Delta t_{\text{extpol}}} \mathcal{F}\left[\boldsymbol{x}(t),\, t\right] \mathrm{d}t,\; j = 1,\, \ldots,\, j_{\max}, \tag{3}$$

where j marks an iteration of the ϵ-algorithm, and $\dot{\boldsymbol{x}}(t) = \mathcal{F}\left[\boldsymbol{x}(t),\, t\right]$ indicates a vector of functions to be integrated in a numerical way. This is only a symbolic notation because the original system (1) is always implicit and cannot be transferred to an explicit form. The parameter Δt_{extpol} is chosen so that the rapidly decaying transients are to be insignificant at $t_0 + \Delta t_{\text{extpol}}$. After that low-pass filtering, the extrapolation is assumed to be only concerned with the slowly decaying transients and the order of the extrapolation algorithm marked below by k_{extpol} can be decreased correspondingly.

The vector $\boldsymbol{x}_j^{(0)}$ represents the first sample for the extrapolation algorithm. Entire sequence of the samples is also obtained by the implicit numerical integration of the system (1)

$$\boldsymbol{x}_j^{(k)} := \boldsymbol{x}_j\left(t_0 + \Delta t_{\text{extpol}} + \textstyle\sum_{i=1}^{k} T_j^{(i)}\right) = \int_{t_0+\Delta t_{\text{extpol}}}^{t_0+\Delta t_{\text{extpol}}+\sum_{i=1}^{k} T_j^{(i)}} \mathcal{F}\left[\boldsymbol{x}(t),\, t\right] \mathrm{d}t,\; k = 1,\, \ldots,\, 2k_{\text{extpol}},\; j = 1,\, \ldots,\, j_{\max}, \tag{4}$$

where $T_j^{(i)}$ are the periods, which must be detected for autonomous circuits automatically during the integration. (They are known for the non-autonomous circuits, certainly.)

When the sampling is finished after the integration, the scalar ϵ-algorithm starts after the initialization

$$\begin{aligned} {}^{i}\epsilon_{-1}^{(k)} &:= 0, & k &= 1,\, \ldots,\, 2k_{\text{extpol}}, \\ {}^{i}\epsilon_{0}^{(k)} &:= {}^{i}x_j^{(k)}, & k &= 0,\, \ldots,\, 2k_{\text{extpol}}, \end{aligned} \quad i = 1,\, \ldots,\, n \tag{5}$$

(n marks the number of elements of the vector $\boldsymbol{x}$, i marks an element of $\boldsymbol{x}$) by the extrapolation formulae

$${}^{i}\epsilon_{m+1}^{(k)} := {}^{i}\epsilon_{m-1}^{(k+1)} + \frac{1}{{}^{i}\epsilon_{m}^{(k+1)} - {}^{i}\epsilon_{m}^{(k)}},\; m = 0,\, \ldots,\, 2k_{\text{extpol}} - 1,\; k = 0,\, \ldots,\, 2k_{\text{extpol}} - 1 - m \tag{6}$$

for all elements ${}^{i}\epsilon_{m+1}^{(k)},\; i = 1,\, \ldots,\, n$ of the vector $\boldsymbol{\epsilon}_{m+1}^{(k)}$. The outcome of the ϵ-algorithm becomes the new initial condition of the system (1), i.e.,

$$\boldsymbol{x}_{j+1}\left(t_0\right) := \boldsymbol{\epsilon}_{2k_{\text{extpol}}}^{(0)} \tag{7}$$

and the procedure is repeated until convergence is detected (ϵ_{extpol} is an allowable extrapolation error)

$$\begin{aligned} &\text{if} \quad \max_{i=1,\,\ldots,\,n} \frac{\left|{}^{i}x_j^{(k)} - {}^{i}x_j^{(k-1)}\right|}{\left|{}^{i}x_j^{(k)}\right| + {}^{i}x_{\text{null}}} \leqq \epsilon_{\text{extpol}} \quad \text{then} \quad k \in \langle 1,\, \ldots,\, 2k_{\text{extpol}} \rangle. \\ &\boldsymbol{x}_{\text{ss}} := \boldsymbol{x}_j^{(k)}, \end{aligned} \tag{8}$$

The progression of the ϵ-algorithm (5) and (6) can be represented by the following diagram (the values in the corners of the triangles create new values sequentially in the directions of the arrows)

$$\begin{array}{ccccccc} & & {}^{i}\epsilon_{0}^{(0)} & & & & \\ & \nearrow & & \searrow & & & \\ {}^{i}\epsilon_{-1}^{(1)} & & \triangleleft & & {}^{i}\epsilon_{1}^{(0)} & & \\ & \nearrow & & \searrow & & \nearrow & \\ & & {}^{i}\epsilon_{0}^{(1)} & & \triangleleft & & {}^{i}\epsilon_{2}^{(0)} \\ & \nearrow & & \searrow & & & \\ {}^{i}\epsilon_{-1}^{(2)} & & \triangleleft & & {}^{i}\epsilon_{1}^{(1)} & & \\ & \nearrow & & & & & \\ & & {}^{i}\epsilon_{0}^{(2)} & & & & \cdots \\ \cdots & & & & \cdots & & \\ & & \cdots & & & & \end{array} \tag{9}$$

The advantage of the above algorithm is a possibility of its application on the autonomous circuits. The detection of unknown periods is performed by searching the intersections of suitably chosen i_{fix}th element of the vector of circuit variables $\boldsymbol{x}_j$ with a suitably selected level

$$^{i_{\text{fix}}}x_j = x_{\text{fix}},\ i_{\text{fix}} \in \langle 1,\, n\rangle. \tag{10}$$

3. DESCRIPTION OF THE METHOD BASED ON VOLTERRA SERIES

As the resulting formulae arisen from the usage of the Volterra series are very complicated, consider for a simplicity of the explanation that a circuit system comprises only two equations, i.e., (1) can be written in the simpler form

$$f_1\left(x_1,\, x_2,\, \dot{x}_1,\, \dot{x}_2,\, t\right) = 0, \quad f_2\left(x_1,\, x_2,\, \dot{x}_1,\, \dot{x}_2,\, t\right) = 0. \tag{11}$$

The Taylor expansion of the functions f_1 and f_2 with the inclusion of second-order terms in a linearization center $^{(0)}$ is the following (of course, higher-order terms are necessary for the higher-order intermodulation products):

$$\begin{aligned}
&f_{1,2}^{(0)} + \frac{\partial f_{1,2}}{\partial x_1}^{(0)} \Delta x_1 + \frac{\partial f_{1,2}}{\partial x_2}^{(0)} \Delta x_2 + \frac{\partial f_{1,2}}{\partial \dot{x}_1}^{(0)} \Delta \dot{x}_1 + \frac{\partial f_{1,2}}{\partial \dot{x}_2}^{(0)} \Delta \dot{x}_2 \ + \frac{\partial^2 f_{1,2}}{\partial x_1 \partial x_2}^{(0)} \Delta x_1 \Delta x_2 \\
&+ \frac{\partial^2 f_{1,2}}{\partial x_1 \partial \dot{x}_1}^{(0)} \Delta x_1 \Delta \dot{x}_1 + \frac{\partial^2 f_{1,2}}{\partial x_1 \partial \dot{x}_2}^{(0)} \Delta x_1 \Delta \dot{x}_2 + \frac{\partial^2 f_{1,2}}{\partial x_2 \partial \dot{x}_1}^{(0)} \Delta x_2 \Delta \dot{x}_1 + \frac{\partial^2 f_{1,2}}{\partial x_2 \partial \dot{x}_2}^{(0)} \Delta x_2 \Delta \dot{x}_2 \\
&+ \frac{\partial^2 f_{1,2}}{\partial \dot{x}_1 \partial \dot{x}_2}^{(0)} \Delta \dot{x}_1 \Delta \dot{x}_2 + \frac{1}{2} \frac{\partial^2 f_{1,2}}{\partial x_1^2}^{(0)} \Delta x_1^2 + \frac{1}{2} \frac{\partial^2 f_{1,2}}{\partial x_2^2}^{(0)} \Delta x_2^2 + \frac{1}{2} \frac{\partial^2 f_{1,2}}{\partial \dot{x}_1^2}^{(0)} \Delta \dot{x}_1^2 + \frac{1}{2} \frac{\partial^2 f_{1,2}}{\partial \dot{x}_2^2}^{(0)} \Delta \dot{x}_2^2
\end{aligned} \tag{12}$$

A natural linearization center for this type of analysis is the operating point, i.e., the static version of (11)

$$f_1\left(x_{1_0},\, x_{2_0},\, 0,\, 0,\, 0\right) = 0, \quad f_2\left(x_{1_0},\, x_{2_0},\, 0,\, 0,\, 0\right) = 0 \tag{13}$$

must be solved in advance.

The next step is the standard frequency analysis, i.e., solving the system of the two equations

$$F_{1,2}\left(\omega\right) + \frac{\partial f_{1,2}}{\partial x_1}^{(0)} \Delta X_1 + \frac{\partial f_{1,2}}{\partial x_2}^{(0)} \Delta X_2 + j\omega \frac{\partial f_{1,2}}{\partial \dot{x}_1}^{(0)} \Delta X_1 + j\omega \frac{\partial f_{1,2}}{\partial \dot{x}_2}^{(0)} \Delta X_2 = 0, \tag{14}$$

which must be performed for the frequencies ω_1 and ω_2. In this way, we obtain the first-order products $\Delta X_1\left(\omega_1\right)$, $\Delta X_1\left(\omega_2\right)$, $\Delta X_2\left(\omega_1\right)$, and $\Delta X_2\left(\omega_2\right)$. The terms $F_1\left(\omega\right)$ and $F_2\left(\omega\right)$ represent independent signal sources of the circuit.

The second-order intermodulation products can be estimated using the second-order terms in (12) as the signal sources of the circuit (instead of the independent ones), i.e., the system

$$\begin{aligned}
&\frac{\partial f_{1,2}}{\partial x_1}^{(0)} \Delta X_1' + \frac{\partial f_{1,2}}{\partial x_2}^{(0)} \Delta X_2' + j\omega \frac{\partial f_{1,2}}{\partial \dot{x}_1}^{(0)} \Delta X_1' + j\omega \frac{\partial f_{1,2}}{\partial \dot{x}_2}^{(0)} \Delta X_2' + \frac{\partial^2 f_{1,2}}{\partial x_1 \partial x_2}^{(0)} \Delta X_1 \Delta X_2 \\
&+ j\omega \frac{\partial^2 f_{1,2}}{\partial x_1 \partial \dot{x}_1}^{(0)} \Delta X_1^2 + j\omega \frac{\partial^2 f_{1,2}}{\partial x_1 \partial \dot{x}_2}^{(0)} \Delta X_1 \Delta X_2 + j\omega \frac{\partial^2 f_{1,2}}{\partial x_2 \partial \dot{x}_1}^{(0)} \Delta X_2 \Delta X_1 + j\omega \frac{\partial^2 f_{1,2}}{\partial x_2 \partial \dot{x}_2}^{(0)} \Delta X_2^2 \\
&- \omega^2 \frac{\partial^2 f_{1,2}}{\partial \dot{x}_1 \partial \dot{x}_2}^{(0)} \Delta X_1 \Delta X_2 + \frac{1}{2} \frac{\partial^2 f_{1,2}}{\partial x_1^2}^{(0)} \Delta X_1^2 + \frac{1}{2} \frac{\partial^2 f_{1,2}}{\partial x_2^2}^{(0)} \Delta X_2^2 - \omega^2 \frac{1}{2} \frac{\partial^2 f_{1,2}}{\partial \dot{x}_1^2}^{(0)} \Delta X_1^2 \\
&- \omega^2 \frac{1}{2} \frac{\partial^2 f_{1,2}}{\partial \dot{x}_2^2}^{(0)} \Delta X_2^2 = 0
\end{aligned} \tag{15}$$

must be solved for the frequencies $\omega_1 + \omega_1$, $\omega_2 + \omega_2$, $\omega_1 + \omega_2$, and $\omega_1 - \omega_2$, which gives the second-order harmonic products $\Delta X_1'\left(\omega_1 + \omega_1\right)$, $\Delta X_1'\left(\omega_2 + \omega_2\right)$, $\Delta X_2'\left(\omega_1 + \omega_1\right)$, $\Delta X_2'\left(\omega_2 + \omega_2\right)$, and the intermodulation products $\Delta X_1'\left(\omega_1 + \omega_2\right)$, $\Delta X_1'\left(\omega_1 - \omega_2\right)$, $\Delta X_2'\left(\omega_1 + \omega_2\right)$, and $\Delta X_2'\left(\omega_1 - \omega_2\right)$. The higher-order products can be determined in the analogical way using the higher-order terms [7].

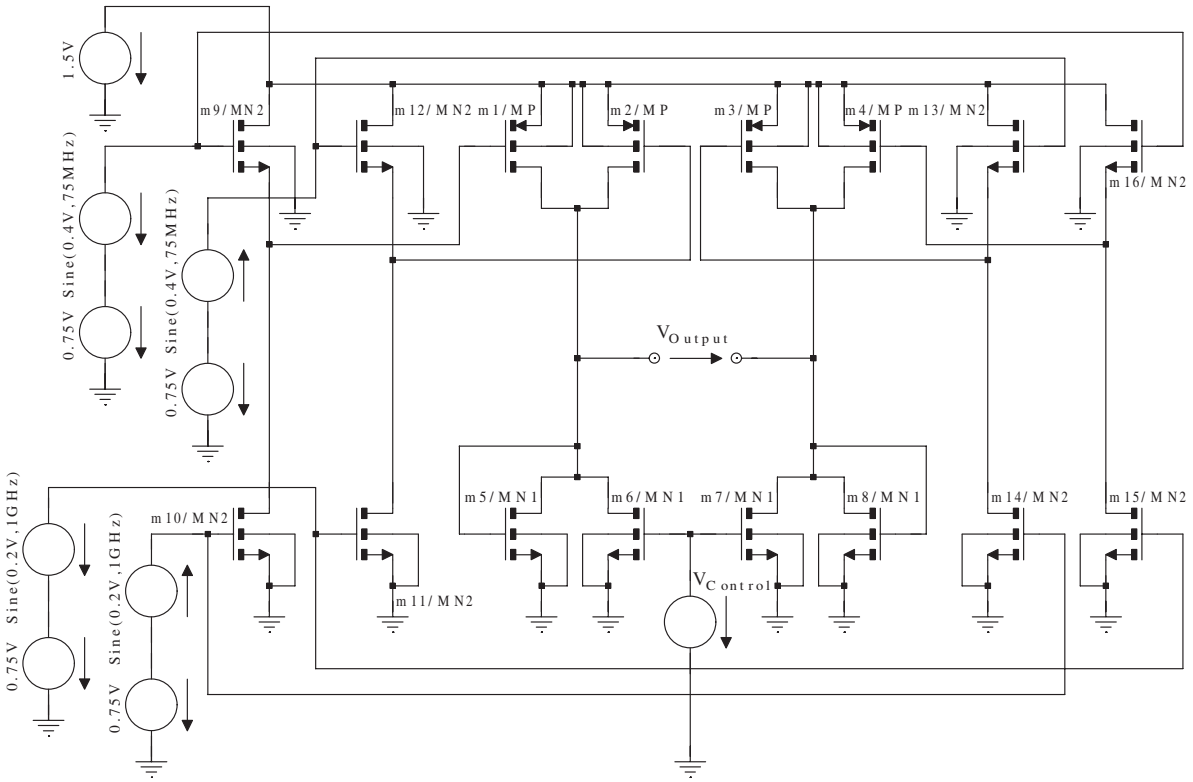


Figure 1: Low-voltage low-power CMOS RF four-quadrant multiplier with symmetrical low- (input signal) and high-frequency (local oscillator) sources.

Table 1: Intermodulation products determined accurately by the steady state algorithm and fast Fourier transform: $f_1 + f_2$, $f_1 - f_2$, $f_1 + 3f_2$, $f_1 - 3f_2$, $3f_1 + f_2$, $3f_1 - f_2$, $f_1 + 5f_2$ and $f_1 - 5f_2$ — the 3rd-order products $f_1 + 2f_2$, $f_1 - 2f_2$, $2f_1 + f_2$, and $2f_1 - f_2$ are negligible, which is very important (and expected).

V_{Control}	$V_{\mathrm{Output,1.075\,GHz}}$	$V_{\mathrm{Output,0.925\,GHz}}$	V_{Control}	$V_{\mathrm{Output,1.225\,GHz}}$	$V_{\mathrm{Output,0.775\,GHz}}$
1 V	8.97 mV	10.1 mV	1 V	0.333 mV	0.377 mV
1.1 V	10.2 mV	11.5 mV	1.1 V	0.371 mV	0.485 mV
1.2 V	11.9 mV	13.6 mV	1.2 V	0.501 mV	0.712 mV
1.3 V	14.5 mV	16.8 mV	1.3 V	0.817 mV	1.18 mV
1.4 V	18.3 mV	21.2 mV	1.4 V	1.28 mV	1.85 mV
1.5 V	22 mV	25.7 mV	1.5 V	1.22 mV	1.67 mV

V_{Control}	$V_{\mathrm{Output,3.075\,GHz}}$	$V_{\mathrm{Output,2.925\,GHz}}$	V_{Control}	$V_{\mathrm{Output,1.375\,GHz}}$	$V_{\mathrm{Output,0.625\,GHz}}$
1 V	0.217 mV	0.228 mV	1 V	49.7 µV	29.9 µV
1.1 V	0.218 mV	0.229 mV	1.1 V	59.6 µV	43.8 µV
1.2 V	0.22 mV	0.232 mV	1.2 V	79 µV	72.3 µV
1.3 V	0.227 mV	0.239 mV	1.3 V	125 µV	145 µV
1.4 V	0.238 mV	0.252 mV	1.4 V	75.9 µV	120 µV
1.5 V	0.273 mV	0.29 mV	1.5 V	387 µV	773 µV

4. PRACTICAL EXAMPLE FROM THE AREA OF CMOS RF INTEGRATED CIRCUITS DESIGN

Let us consider a four-quadrant CMOS RF multiplier in Fig. 1 [8]. All the parameters of the MOSFET model have been kindly granted by Prof. Salama. However, they have been transformed to the new ones required by the "smoothed" model with *suppressed* discontinuities [5]. The output voltage of the multiplier is strongly dependent on the controlling one connected to the gates of m6 and m7 transistors. For example, for the controlling voltages 1 and 1.5 V, the magnitudes of the output signal are about 20 and 50 mV, respectively.

First, the accurate results are determined by the steady state algorithm followed by fast Fourier transform. The significant intermodulation products are shown in Table 1 — due to the balance of the multiplier, the intermodulation products f_1+2f_2, f_1-2f_2, $2f_1+f_2$, and $2f_1-f_2$ are negligible, which is very important. On the contrary, even the intermodulation products f_1+5f_2 and f_1-5f_2 are also noticeable as also shown in Table 1.

However, the intermodulation products f_1+f_2 and f_1-f_2 can be estimated much faster using the Volterra series — the results are shown in Table 2. The error of the estimation depends on the quality of the MOSFET models (therefore the smoothing is important), and is acceptable for lesser magnitudes of the signal, of course.

Table 2: Intermodulation products estimated by the Volterra series: f_1+f_2 and f_1-f_2.

$\lvert V_{\mathrm{Output}}\rvert_{\max}$	V_{Control}	$V_{\mathrm{Output,\,1.075\,GHz}}$	$V_{\mathrm{Output,\,0.925\,GHz}}$
19.3 mV	1 V	7.02 mV	7.57 mV
22.6 mV	1.1 V	7.59 mV	8.29 mV
26.4 mV	1.2 V	8.38 mV	9.27 mV
33.1 mV	1.3 V	9.47 mV	10.6 mV
41.9 mV	1.4 V	11 mV	12.5 mV
49 mV	1.5 V	13 mV	15 mV

5. CONCLUSION

A reliable steady state algorithm has been implemented which can be used for accurate determining the shape of period. Analyzing that period by fast Fourier transform afterwards enables to calculate intermodulation products in a precise way.

An algorithm for a fast estimation of the main interpolation products has also been presented in a simple form, which could be an add-on of the standard frequency analysis, and which can efficiently reuse procedures of that analysis. The precision of the algorithm has been illustrated by an analysis of a CMOS RF four-quadrant multiplier. For a medium level of nonlinearities, the estimation can be used as a fast approximative analysis.

ACKNOWLEDGMENT

This paper has been supported by the Grant Agency of the Czech Republic, grant No. 102/08/0784, and also by the Czech Technical University in Prague Research Project No. MSM 6840770014.

REFERENCES

1. Vlach, J. and K. Singhal, *Computer Methods for Circuit Analysis and Design*, Van Nostrand Reinhold Company, New York, 1982.
2. Dobeš, J., D. Biolek, and P. Posolda, "An efficient steady-state analysis of microwave circuits," *International Journal of Microwave and Optical Technology*, Vol. 1, No. 2, 284–289, Aug. 2006.
3. Skelboe, S., "Computation of the periodic steady-state response of nonlinear networks by extrapolation methods," *IEEE Trans. Circuits Syst.*, Vol. 27, 161–175, 1980.
4. Petrenko, A. I., A. I. Vlasov, and A. P. Timtschenko, *Tabular Methods of Computer-aided Modeling*, Higher School, Kiyv, 1977.
5. Dobeš, J., "Reliable CAD analyses of CMOS radio frequency and microwave circuits using smoothed gate capacitance models," *AEÜ — International Journal of Electronics and Communications*, Vol. 57, No. 6, 372–380, 2003.
6. Massobrio, G. and P. Antognetti, *Semiconductor Device Modeling with SPICE*, McGraw-Hill, New York, 1993.
7. Wambacq, P. and W. Sansen, *Distortion Analysis of Analog Integrated Circuits*, Kluwer Academic Publishers, Boston, 1998.
8. Salama, M. K. and A. M. Soliman, "Low-voltage low-power CMOS RF four-quadrant multiplier," *AEÜ — International Journal of Electronics and Communications*, Vol. 57, No. 1, 74–78, 2003.

Modeling a Transmission Interconnect by Optimal Number of Lumped Sections

S. R. Nelatury[1], **M. N. O. Sadiku**[2], **and V. K. Devabhaktuni**[3]

[1]School of Engineering, Pennsylvania State University, Erie, PA 16563, USA
[2]College of Engineering, Prairie View A&M University, Prairie View, TX 77446, USA
[3]Department of ECE, Concordia University, Montreal, QC, H3G 1M8, Canada

Abstract— The growing demand for high-speed interconnects in VLSI circuits has turned a high impetus for their accurate modeling. The classical uniform transmission line theory based on distributed RLC circuit model has several limitations as far as CAD practices are concerned. Distributed models are being replaced with reduced-order lumped models for efficient simulation and design applications. The present literature is replete with numerous reduced-order models and macro-models. These models are developed based on an approximating criterion either in time or frequency domain with the aim of reducing computational complexity. We introduce a technique to find the optimum number of lumped sections for approximating a transmission interconnect of a given length and rise-time. This is done by first considering the simple well-known relationship between the primary constants of the line and the image parameters of each individual section viewed as a linear two-port network. Next, the sensitivity of this relationship as a function of frequency is used to develop a formula for the least number of sections needed. This method is further compared with the existing frequency domain based macro-models. Since the final formulas turn out to be simpler, they have excellent use in CAD applications.

1. INTRODUCTION

The undesirable effects such as ringing, delay, and crosstalk occurring in high-speed interconnects are of more serious concern as we push toward higher and higher frequencies [1]. For achieving appreciable signal integrity, it is important to make an accurate estimation of these effects. Accordingly, new demands are placed on the developments of computer-aided design (CAD) tools. The relations governing the voltages and currents at the terminals of interconnects are described in the frequency domain in terms of per-unit length parameters, while signal integrity assessment requires a transient analysis which is a time-domain task. One way to perform this is by translating the frequency-domain relations of the dispersive transmission line structures into the time-domain Green's functions. The problem is solved in a time-marching fashion: after each time step, relevant transmission line parameters are computed by convolving the Green's functions with previous voltage and current values. The choice of Green's functions would determine the features of the numerical method. In general, methods based on Green's functions yield acceptable results. But they consume long simulation times and large memory, and they are often not implemented in or compatible with popular general-purpose circuit simulators such as ASTAP or SPICE.

An alternative approach is the use of equivalent lumped circuits. The major advantage of the lumped element approach is that the equivalent circuits can be easily and quickly implemented in existing circuit simulators and all available features of these packages can be used. Also, it is intuitive to represent a distributed transmission line by lumped sections (e.g., T or π) or by a combination of lumped elements and ideal lossless transmission lines (hybrid cells). These sections preserve linearity, time-invariance, passivity and reciprocity. This method results in a large system of equations to be solved and demands higher CPU times. Recently, reduced-order macro-models have been suggested to overcome this problem [2–4].

The intermediate step in approximating a distributed line by a collection of discrete sections should not be overlooked. The question of what should be the number of discrete sections so as to accurately model a transmission line in a desired frequency range is not conclusively answered in the literature. This number must be large enough for accuracy reasons, but also small enough to reduce computational complexity and associated simulation times. A rule of thumb followed conventionally is to select the number of sections such that the propagation delay caused by an elementary section should be smaller than one-fifth the shortest rise-time expected. In this paper, we shall suggest a guideline for selecting the number of sections for a specified accuracy and for a specified relative bandwidth. The conventional rule of thumb is rather linear and specifies that the number of sections is proportional to the product of length of the line in wavelengths and the

relative bandwidth. However, when we attempt to approximate the line with a cascaded network of T or π sections having the same [ABCD] or image parameters, we arrive at a slightly different result.

The rest of the paper is organized as follows. In Section 2, we equate the [ABCD] parameters of the distributed and the lumped networks, and derive the sensitivity of the proportionality constant. Section 3 provides a simulation example and compares the optimal number with the conventional rule of thumb. Section 4 summarizes the idea of the paper and gives the conclusions.

2. DISTRIBUTED VERSUS LUMPED MODEL

Consider a uniform transmission interconnect of length L having the primary constants $\{R', L', G', C'\}$ with the usual meaning of each parameter (See Fig. 1). Let the distance along the line be denoted by x. Also let $\{V_R(s),\, I_R(s),\, V_S(s),\, I_S(s)\}$ stand for the voltage and current at the receiving end and those at the sending end respectively.

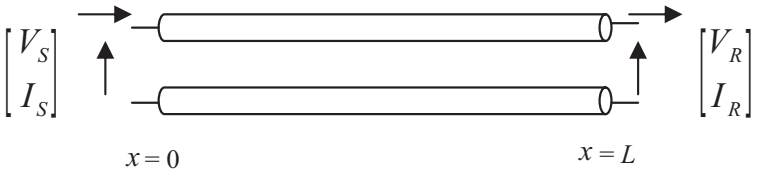


Figure 1: A uniform transmission line interconnect of length L.

From the standard equations governing a transmission line, the sending end and the receiving end voltage and currents are related by the ABCD parameter matrix as

$$\begin{bmatrix} V_R \\ I_R \end{bmatrix} = \begin{bmatrix} \cosh\gamma' L & -Z_0' \sinh\gamma' L \\ \frac{-\sinh\gamma' L}{Z_0'} & \cosh\gamma' L \end{bmatrix} \begin{bmatrix} V_S \\ I_S \end{bmatrix}. \tag{1}$$

In Equation (1), Z_0' and γ' represent the secondary constants given by

$$Z_0' = \sqrt{\frac{R' + L's}{G' + C's}} \tag{2}$$

and

$$\gamma' = \sqrt{(R' + L's)\,(G' + C's)}. \tag{3}$$

We wish to approximate the given line with N number of identical sections each of length $l = L/N$ with an equivalent circuit shown in Fig. 2. The sending end and the receiving end voltage and current variables at the nth segment of the line are denoted by $\{V_R^n(s),\, I_R^n(s),\, V_S^n(s),\, I_S^n(s)\}$. Based on the distributed model, these quantities that pertain to the nth section are in turn related by

$$\begin{bmatrix} V_R^n \\ I_R^n \end{bmatrix} = \begin{bmatrix} \cosh\gamma' l & -Z_0' \sinh\gamma' l \\ \frac{-\sinh\gamma' l}{Z_0'} & \cosh\gamma' l \end{bmatrix} \begin{bmatrix} V_S^n \\ I_S^n \end{bmatrix}. \tag{4}$$

However, using the discrete equivalent T-network shown in Fig. 2, we find that Equation (4) has an alternate form in terms of the lumped values Z and Y, i.e.,

$$\begin{bmatrix} V_R^n \\ I_R^n \end{bmatrix} = \begin{bmatrix} \left(1 + \frac{ZY}{2}\right) & -Z\left(1 + \frac{ZY}{4}\right) \\ -Y & \left(1 + \frac{ZY}{2}\right) \end{bmatrix} \begin{bmatrix} V_S^n \\ I_S^n \end{bmatrix}. \tag{5}$$

In Equation (5), Z and Y represent the series and the shunt arm impedance and admittance respectively.

Let us assume that the components bear a relation of direct proportion such that

$$Z = \kappa_1 \left(R' + L's\right) \frac{L}{N} \tag{6}$$

and

$$Y = \kappa_2 \left(G' + C's\right) \frac{L}{N}. \tag{7}$$

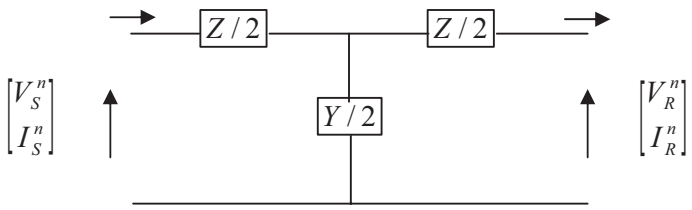


Figure 2: Equivalent circuit of the nth section of length l.

The sensitivity of the proportionality coefficients in terms of the relative bandwidth $\bar{\omega} = \frac{\delta\omega}{\omega_o}$ can be expressed as

$$\delta\kappa_2 = 2\delta\kappa_1 = \frac{4}{3}\xi^2\bar{\omega}, \tag{8}$$

where $\xi = \frac{\gamma' L}{2N}$. For an allowed tolerance in the sensitivity of $\varepsilon = \delta\kappa_2/\kappa_2$, an integer lower bound for N in terms of interconnect length in wavelengths, i.e., $\overline{L} = \frac{L}{\lambda}$, can be obtained as

$$N = \left\lceil 2\pi\overline{L}\sqrt{\frac{2\bar{\omega} - \varepsilon}{6\varepsilon}} \right\rceil, \tag{9}$$

where symbol $\lceil\ \rceil$ indicates *ceil* operation.

3. SIMULATION EXAMPLE

To illustrate how one should choose the required number of sections for a given degree of accuracy, we present a simple simulation example. The conventional rule of thumb as mentioned in [1] to have a relative bandwidth of $\bar{\omega} = 10$ is to select $N = \frac{10\tau d}{t_r}$, where $d = L$ is the length of the line, t_r is the rise-time, τ is delay per unit length. We can rewrite this relationship as

$$N = \left\lceil 1.5\pi\overline{L}\bar{\omega} \right\rceil. \tag{10}$$

Now to make a critical comparison of this result with what we have proposed in Equation (9), let us assume that the length of the line is one-fifth of a wavelength, i.e., $\overline{L} = 0.2$. In Fig. 3 we shall plot Equation (10) as a function of the relative bandwidth $\bar{\omega}$ in the range of 1–20 and superimpose values of N found from Equation (9) for $\varepsilon = 2\%$, 5%, 10%. The values of N from (10) seem to be very small for a relative bandwidth below 6. For these values, if we observe in Fig. 3, it is clear that the ε is more than 10% implying that the approximation brings about large variations in the proportionality between the lumped and distributed models. For a relative bandwidth greater than

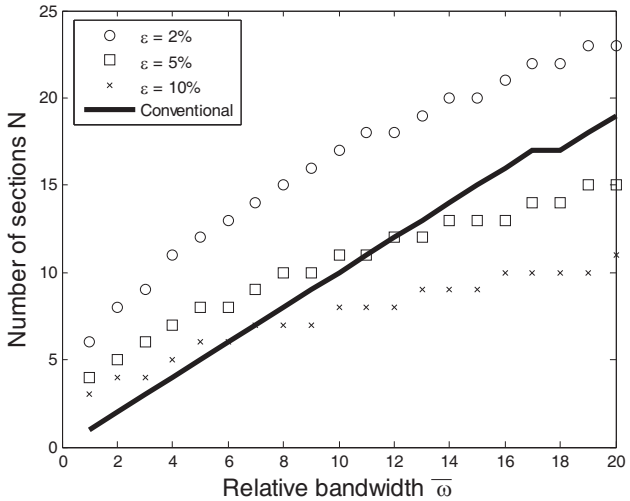


Figure 3: The required number of sections as a function of relative bandwidth for a line of one-fifth of a wavelength.

12, the conventional value for N seems to increase almost linearly and offer a tolerance of 2%–5% in ε. Although not shown in Fig. 3, for relative bandwidth greater than 35, Equation (10) gives percent sensitivity less than 2%, but the value of N would be higher. The downside of having higher N is that the circuit analysis demands higher computational time. So depending on the application at hand, if we were to make a trade-off between high accuracy and low computational cost, Equation (9) proposed in this paper would be useful.

4. CONCLUSIONS

This paper provides guidelines for approximating a distributed line with optimal number N of lumped sections. The conventional rule of thumb is almost linear and is not explicitly related to the proportionality between lumped and distributed models. We have proposed a method to select N not only based on the required bandwidth but also on the allowed amount of sensitivity in the approximation, allowing a trade-off between required accuracy and the computational complexity. This information is useful in developing reduced order macro-model for transmission interconnects.

REFERENCES

1. Achar, R. and M. Nakhla, "Simulation of high-speed interconnects," *Proc. IEEE*, Vol. 89, No. 5, 693–728, May 2001.
2. Palenius, T. and J. Roos, "Comparison of reduced-order interconnect macromodels for time-domain simulation," *IEEE Trans. Microwave Theory Tech.*, Vol. 52, No. 9, 2240–2250, Sept. 2004.
3. Gunupudi, P. K., R. Khazaka, M. S. Nakhla, T. Smy, and D. Celo, "Passive parameterized time-domain macromodels for high-speed transmission-line networks," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, No. 12, 2347–2354, Dec. 2003.
4. Grivet-Talocia, S., S. Acquadro, M. Bandinu, F. G. Canavero, I. Kelander, and M. Rouvala, "A parameterization scheme for lossy transmission line macromodels with application to high speed interconnects in mobile devices," *IEEE Trans. Electromagn. Compat.*, Vol. 49, No. 1, 18–24, Feb. 2007.

Removal of Sea Surface Related Wavefields from CSEM Data

P. M. van den Berg[1], **A. Abubakar**[2], **and T. M. Habashy**[2]

[1]Delft University of Technology, The Netherlands
[2]Schlumberger-Doll Research, Cambridge, USA

Abstract— For a controlled-source electromagnetic (CSEM) survey in a shallow water environment, the presence of the sea surface significantly hinders the interpretation of the measured data. The electromagnetic (EM) wavefields are disturbed by the sea surface. The removal of the sea-surface related wave phenomena from the data is an important step in order to robustly interpret the collected data. We propose a processing method by which the sea-surface related multiples would be removed, while a priori knowledge of the EM source wavelet becomes superfluous. The governing equations are obtained from an appropriate application of the EM reciprocity theorem that relates on one side the EM fields in the actual measurement configuration including the sea surface and on the other side the EM fields in a desired source configuration and in the absence of the sea surface, where the water layer is extended to infinity.

1. INTRODUCTION

In a controlled-source EM (CSEM) survey [1], it is necessary to interpret the measurements in such a way that a prediction of the presence of hydrocarbons in the sedimentary layers can be made. However, in a shallow water environment, the presence of the sea surface hinders this interpretation. Electromagnetic wavefields are partly reflected and partly transmitted by the sea surface. This means that the source signal is contaminated by its so-called source ghost signal and that the received signals are contaminated by the so-called receiver ghost. Further the receiver ghost can be considered as secondary source signals that are transmitted in the earth. Hence, removal of all these sea surface related electromagnetic wavefields from the data is a prerequisite step before actual interpretation of the data can take place. In this paper we show that an appropriate use of the electromagnetic reciprocity theorem leads to the mathematical equations for a consistent removal of sea-surface related wavefields.

The actual EM wavefield is denoted as $\{\hat{\mathbf{E}}(x), \hat{\mathbf{H}}(x)\}$ in the frequency domain (with frequency parameter $s = j\omega$). The position in space is denoted as $\mathbf{x} = \{x_1, x_2, x_3\}$. The coordinates x_1 and x_2 denote the horizontal directions, while x_3 denotes the vertical direction pointing into the earth. The wavefield is generated by an electromagnetic source in the sea at $\mathbf{x}^S = \{x_1, x_2, x_3^S\}$. On the sea bottom (not necessarily a horizontal plane), we measure the EM wavefield reflected by the earth geology at the sea bottom (see Figure 1, left). We assume that there exists a horizontal plane at $x_3 = x_3^{\rm obs}$ between the source level and the sea bottom. It is assumed that the sea water is homogeneous with complex permittivity $\hat{\varepsilon}$ and real constant permeability μ. In this domain, outside the sources, the electromagnetic fields satisfy Maxwell's equations

$$\nabla \times \hat{\mathbf{H}} - s\hat{\varepsilon}\mathbf{E} = 0, \quad \nabla \times \hat{\mathbf{E}} + s\mu\mathbf{H} = 0. \tag{1}$$

For very low frequencies, we may neglect the displacement currents. Then, we may write $\hat{\varepsilon} \approx \sigma/s$, where σ denotes the electrical conductivity, which is assumed to be frequency independent within the frequency band of operation.

A first step in the processing of electromagnetic CSEM data is the decomposition of the data into constituents that represent up-going and down-going wavefields. At some horizontal level above the sea bottom and below the plane $\mathbf{x}_3 > \mathbf{x}_3^S$ (Figure 1, left), the EM wavefield may be written as a superposition of up- and down-going wavefields, viz.,

$$\left\{\hat{\mathbf{E}}, \hat{\mathbf{H}}\right\} = \left\{\hat{\mathbf{E}}, \hat{\mathbf{H}}\right\}^{up} + \left\{\hat{\mathbf{E}}, \hat{\mathbf{H}}\right\}^{down}, \tag{2}$$

A proper linear combination of the electric-field signals and the magnetic field signals leads to a mechanism for selecting the up-going and down-going EM wavefields [2]. In seismics, the determination of the up-going wavefield is known as deghosting [3]. In the up-going wave field, some "reverbations" caused by the presence of the sea surface are still present. In analogy with seismics [3], we denote them as sea surface multiples.

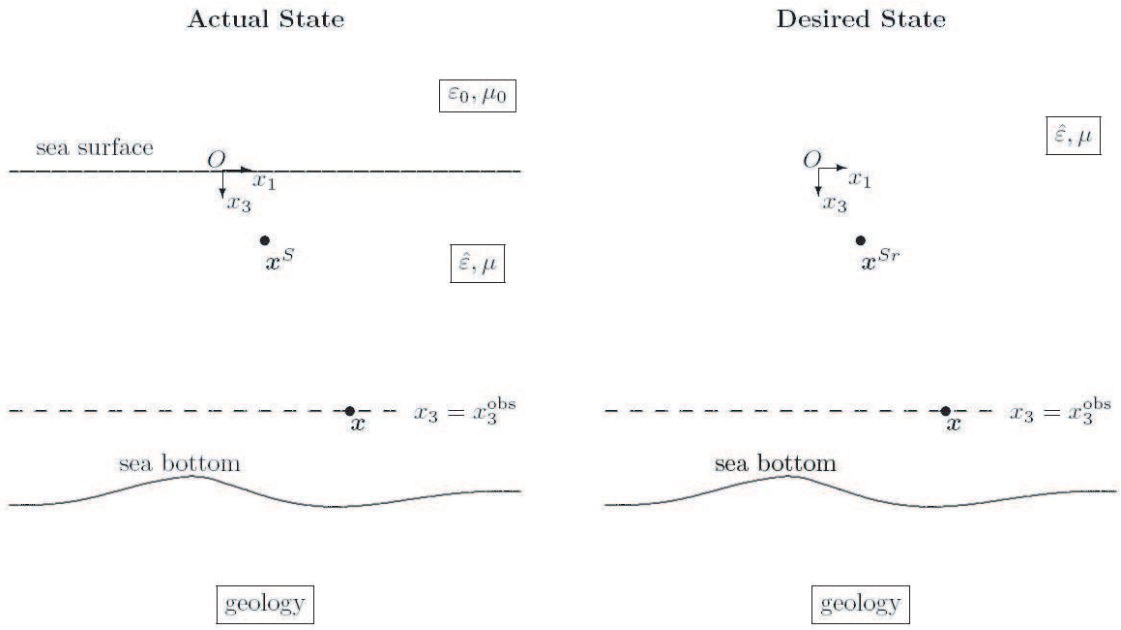


Figure 1: Actual State and Desired State.

2. REMOVAL OF SEA SURFACE MULTIPLES

We start with the propagation invariant at our observational level, $x_3 = x_3^{\text{obs}}$. Since there are no electromagnetic sources below the horizontal level $\mathbf{x}_3 > \mathbf{x}_3^S$ and assuming that the earth geology is spatially invariant, the electromagnetic field reciprocity theorem [4], Chapter 28, yields

$$\int_{(x_1,x_2)\in\mathbb{R}^2} i_3 \cdot \left[\hat{\mathbf{E}}^A \times \hat{\mathbf{H}}^B - \hat{\mathbf{E}}^B \times \hat{\mathbf{H}}^A\right] \mathrm{dA} = 0, \tag{3}$$

in which $\{\hat{\mathbf{E}}^A, \hat{\mathbf{H}}^A\}$ and $\{\hat{\mathbf{E}}^B, \hat{\mathbf{H}}^B\}$ are two admissible electromagnetic states. Writing these fields as a superposition of up- and down-going wavefields and noting that only waves travelling in opposite direction contribute, Eq. (3) becomes

$$\int_{(x_1,x_2)\in\mathbb{R}^2} i_3 \cdot \left[\hat{\mathbf{E}}^{A,up} \times \hat{\mathbf{H}}^{B,down} - \hat{\mathbf{E}}^{B,up} \times \hat{\mathbf{H}}^{A,down}\right] \mathrm{dA} = 0, \tag{4}$$

and

$$\int_{(x_1,x_2)\in\mathbb{R}^2} i_3 \cdot \left[\hat{\mathbf{E}}^{B,down} \times \hat{\mathbf{H}}^{A,up} - \hat{\mathbf{E}}^{A,down} \times \hat{\mathbf{H}}^{B,up}\right] \mathrm{dA} = 0. \tag{5}$$

Subsequently, we let State A to be a desired wavefield in a configuration where the sea surface has been removed. We denote this wavefield as the reflected wavefield $\{\hat{\mathbf{E}}^r, \hat{\mathbf{H}}^r\}$. This is the wavefield that exists in absence of the sea surface and is up-going at the observational plane (see Figure 1, right). The pertaining incident wavefield, that generates the reflected wavefield, is denoted as $\{\hat{\mathbf{E}}^i, \hat{\mathbf{H}}^i\}$. This incident field is down-going at the observational plane. State B is the actual wavefield including the sea surface multiples (Figure 1, left) and it consists of an up-going wavefield $\{\hat{\mathbf{E}}^{up}, \hat{\mathbf{H}}^{up}\}$ and a down-going wavefield $\{\hat{\mathbf{E}}^{down}, \hat{\mathbf{H}}^{down}\}$. Substituting these states into Eq. (4) we obtain

$$\int_{(x_1,x_2)\in\mathbb{R}^2} i_3 \cdot \left[\hat{\mathbf{E}}^r \times \hat{\mathbf{H}}^{down}\right] \mathrm{dA} = \int_{(x_1,x_2)\in\mathbb{R}^2} i_3 \cdot \left[\hat{\mathbf{E}}^{up} \times \hat{\mathbf{H}}^i\right] \mathrm{dA}. \tag{6}$$

This equation which is the fundament for our further discussions, holds independently of the nature of the geology. In addition, it holds for any choice of the incident wavefield $\{\hat{\mathbf{E}}^i, \hat{\mathbf{H}}^i\}$ with source position $\mathbf{x}^{Sr}$ and its corresponding reflected wavefield $\{\hat{\mathbf{E}}^r, \hat{\mathbf{H}}^r\}$. In order to indicate the different source positions $\mathbf{x}^S$ of the actual field and $\mathbf{x}^{Sr}$ of the desired sea-surface free wavefields,

we write Eq. (6) as

$$\int_{(x_1,x_2)\in\mathbb{R}^2} i_3\cdot\left[\hat{\mathbf{E}}^r\left(\mathbf{x}|\mathbf{x}^{Sr}\right)\times\hat{\mathbf{H}}^{down}\left(\mathbf{x}|\mathbf{x}^{S}\right)\right]\mathrm{dA}=\int_{(x_1,x_2)\in\mathbb{R}^2} i_3\cdot\left[\hat{\mathbf{E}}^{up}\left(\mathbf{x}|\mathbf{x}^{S}\right)\times\hat{\mathbf{H}}^{i}\left(\mathbf{x}|\mathbf{x}^{Sr}\right)\right]\mathrm{dA}, \quad (7)$$

where $\mathbf{x}=(x_1,x_2,x_3^{\mathrm{obs}})$, $\mathbf{x}^S=(x_1,x_2,x_3^S)$ and $\mathbf{x}^{Sr}=(x_1,x_2,x_3^{Sr})$. The right-hand side of this equation is known. In the left-hand side $\hat{\mathbf{H}}^{down}$ is known, while $\hat{\mathbf{E}}^r$ is unknown. For each $\mathbf{x}^{Sr}$, we have two unknown functions, viz. the two horizontal components $\hat{E}_1^r(\mathbf{x}|\mathbf{x}^{Sr})$ and $\hat{E}_2^r(\mathbf{x}|\mathbf{x}^{Sr})$. Unfortunately, for each source located at $\mathbf{x}^S$, we have one equation. We therefore need either an extra set of measurements using a different source orientation, e.g., that one data set is obtained with an electric dipole source oriented in the x_1 direction, while an other data set is obtained with an electric dipole dipole oriented in the x_2-direction. Let us consider that the field quantities in the first measurement setup are denoted by the superscript I, while the field quantities in a second measurement setup are denoted by the superscript II. Then, we end up with a system of two equations for the two unknowns $\hat{E}_1^r$ and $\hat{E}_2^r$, viz.,

$$\int_{(x_1,x_2)\in\mathbb{R}^2} i_3\cdot\left[\hat{\mathbf{E}}^r\left(\mathbf{x}|\mathbf{x}^{Sr}\right)\times\hat{\mathbf{H}}^{down,\mathrm{I}}\left(\mathbf{x}|\mathbf{x}^{S}\right)\right]\mathrm{dA}=\int_{(x_1,x_2)\in\mathbb{R}^2} i_3\cdot\left[\hat{\mathbf{E}}^{up,\mathrm{I}}\left(\mathbf{x}|\mathbf{x}^{S}\right)\times\hat{\mathbf{H}}^{i}\left(\mathbf{x}|\mathbf{x}^{Sr}\right)\right]\mathrm{dA}, \quad (8)$$

$$\int_{(x_1,x_2)\in\mathbb{R}^2} i_3\cdot\left[\hat{\mathbf{E}}^r\left(\mathbf{x}|\mathbf{x}^{Sr}\right)\times\hat{\mathbf{H}}^{down,\mathrm{II}}\left(\mathbf{x}|\mathbf{x}^{S}\right)\right]\mathrm{dA}=\int_{(x_1,x_2)\in\mathbb{R}^2} i_3\cdot\left[\hat{\mathbf{E}}^{up,\mathrm{II}}\left(\mathbf{x}|\mathbf{x}^{S}\right)\times\hat{\mathbf{H}}^{i}\left(\mathbf{x}|\mathbf{x}^{Sr}\right)\right]\mathrm{dA}. \quad (9)$$

We remark that, in practice, we only deal with a finite set of measurements points. We therefore approximate the integrals by finite summations. After choosing the source point $\mathbf{x}^{Sr}$, the $\mathbf{x}$- and $\mathbf{x}^S$-points should be chosen symmetrically located around $\mathbf{x}^{Sr}$. We further note, by using the reciprocity of Eq. (5) that in a similar way we can obtain a system of equations for $\hat{\mathbf{H}}^r$. More details can be found in [6].

We remark that, for the 2D acoustic problem, a reciprocity theorem of the type of Eq. (7) is derived [5] for the purpose of removal of sea surface multiples from marine seismic time-space data. Further, in [7], the interferometry theorem is proposed as a tool for removing the effect of the sea surface. It is equivalent to our reciprocity theorem, if we keep the actual source type and the desired source type the same. The strong advantage of the use of our reciprocity theorem is that we may choose any type of source as desired source, to arrive at the desired reflected wavefield. In this way one is able to investigate the polarization sensitivity of the geology under consideration.

3. CHOICE OF ELECTROMAGNETIC SOURCES

Although in [7] a 2D scalar example has been considered in detail with one electric line source parallel to a horizontally layered configuration, in the general 3D setup, measurements with two different types of EM sources are needed, such that, in the actual configuration both transversal 3D electric waves (TE-case, no electric field in the vertical direction) and 3D transversal magnetic waves (TM-case, with no magnetic field in the vertical direction) are present. In the desired configuration, the type of EM source can be arbitrarily chosen, with both its spatial layout and its temporal signal. This feature has a very large impact for CSEM processing.

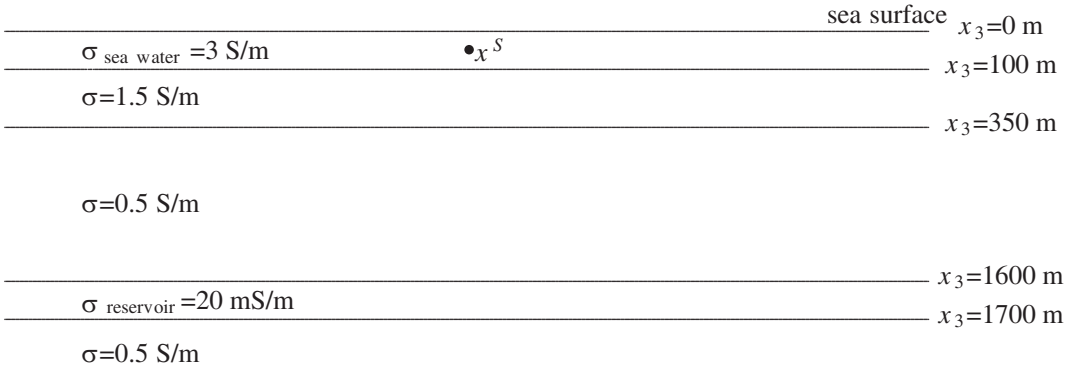


Figure 2: Horizontally layered configuration. The source at depth of 50 m operates at 0.5 Hz. The receivers are located at the sea bottom at $x_2=0$, $x_3=100$ m, with a total extent of 40 km in the x_1-direction.

In particular, a first choice of the desired configuration will be one with a vertical magnetic dipole source. In this desired configuration mainly TE waves will occur, while TM waves only arise if lateral inhomogeneities of the subsurface are present. Alternatively, a second choice of the desired configuration will be one with a vertical electric dipole source. Then, mainly TM waves will occur, while TE waves only arise if lateral changes of the subsurface are present.

In addition, in a horizontally layered configuration, the TM wave is very sensitive to differences in conductive layers, while the TE wave is rather insensitive to difference in conductive layers.

To illustrate this, we take the example of [7] (see Figure 2), and we compute the full 3D electromagnetic wavefield at the receivers for a vertical magnetic-dipole source and for a vertical electric dipole source, respectively. In the present configuration under investigation, the vertical magnetic dipole generates only TE-waves, while the vertical electric dipole generates only TM-waves. To investigate the sensitivity of the hydrocarbon reservoir layer, we also compute the responses in absence of the reservoir layer by putting $\sigma_{\text{reservoir}} = 0.5\,\text{S/m}$.

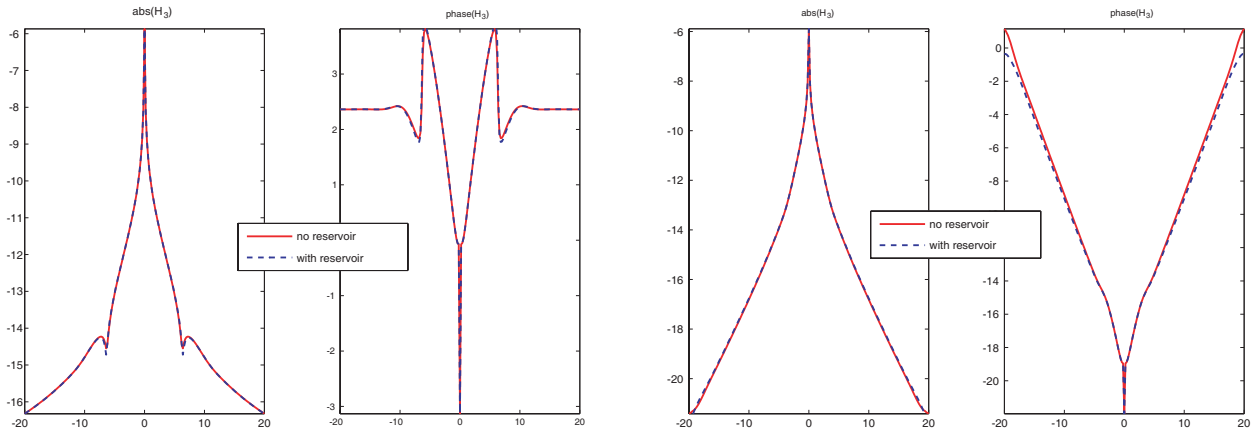


Figure 3: The logarithm of the amplitude and the phase of the magnetic field component $\hat{H}_3$ generated by a vertical magnetic dipole in a configuration with (*left*) and without (*right*) the sea surface.

In Figure 3, for a vertical magnetic dipole source, we have plotted the vertical magnetic field component, $\hat{H}_3$, both for the configuration in presence of the reservoir layer (dashed blue curve)

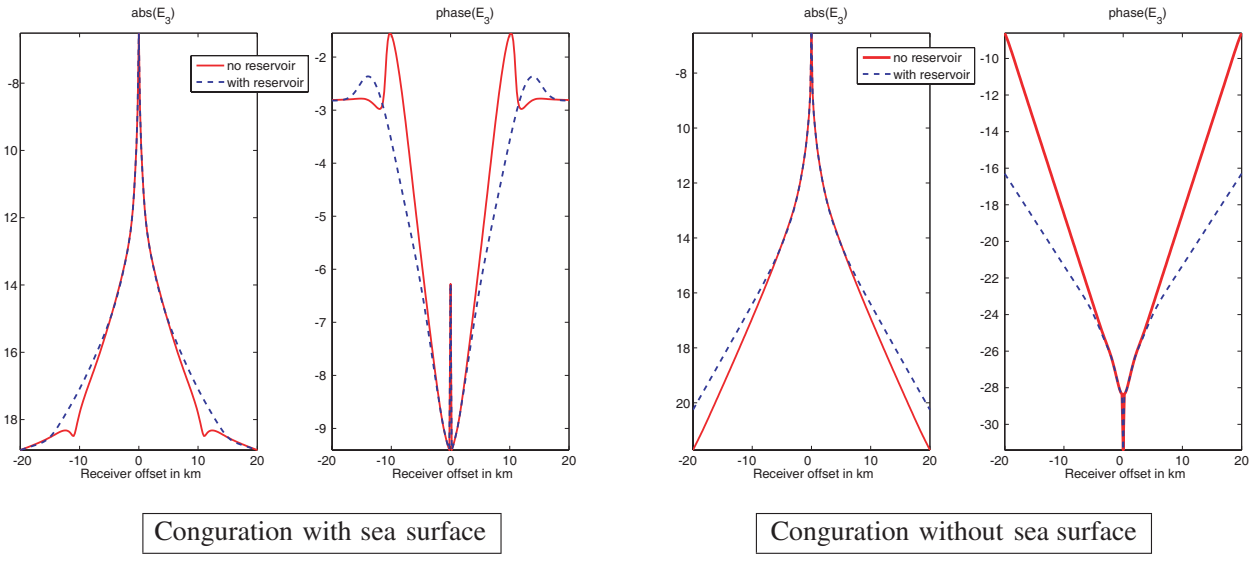


Figure 4: The logarithm of the amplitude and the phase of the vertical electric field $\hat{E}_3$ generated by a vertical electric dipole in a configuration with (*left*) and without (*right*) the sea surface.

and in absence of the reservoir layer (solid red curve). For the configuration with the sea surface, we observe from the left-hand side plots of Figure 3 that the presence of the reservoir is not visible. After removing the sea surface, we observe from the right-hand side plots of Figure 3 some very small differences due to the presence of the reservoir.

Subsequently, we have computed the vertical electric field component, $\hat{E}_3$, generated by a vertical electric dipole. These results are plotted in Figure 4. Both in the configuration with the sea surface and in the configuration without the sea surface, the presence of the reservoir is significant detectable, which shows the (well-known) sensitivity of the TM-waves. It is therefore advocated to carry out the reciprocity based sea-surface removal procedure with a vertical electric dipole source as desired source in the desired configuration where the sea surface has been removed.

4. CONCLUSIONS

We have developed a method to remove the effects of the sea surface from CSEM data measured at the sea bottom. In general, measurements of the electric and magnetic field tangential to the sea bottom are needed to decompose the EM field into up-and down going wavefields. At an observational horizontal level above the sea bottom, a reciprocity theorem has been derived to remove the sea surface effects completely. This theorem becomes operational once the CSEM surveys encompasses measurements for two independent types of EM sources. In the configuration without the sea surface, we are free to choose the source type. For an Earth with some horizontally layering, it is advocated to choose the vertical electric dipole as source, so that the TM-waves prevail in the sea-surface-free configuration. This procedure facilitates an enhanced interpretation of the difference in conductivity in the Earth and detection of a hydrocarbon reservoir.

REFERENCES

1. Eidesmo, T., S. Ellingsrud, L. M. MacGregor, S. Constable, M. C. Sinha, S. Johansen, F. N. Kong, and H. Westerdahl, "Sea bed logging (SBL), a new method for remote and direct identification of hydrocarbon filled layers in deepwater areas," *First Break*, Vol. 20, 144–152, 2002.
2. Rosten, T. and L. Amundsen, "Generalized electromagnetic seabed logging wavefield decomposition into U/D-going components," *SEG Expanded Abstracts*, Vol. 23, 600–603, 2004.
3. Fokkema, J. T. and P. M. van den Berg, *Seismic Applications of Acoustic Reciprocity*, Academic Press, London, 1993.
4. De Hoop, A. T., *Handbook of Radiation and Scattering of Waves*, Academic Press, London, 1995.
5. Vaage, S. T., J. T. Fokkema, and P. M. van den Berg, "Method for seismic exploration utilizing motion sensor and pressure sensor data," United States Patent, No. 7123543, October 17, 2006.
6. Van den Berg, P. M., A. Abubakar, and T. M. Habashy, "Removal of sea-surface related wavefields in controlled source electromagnetic survey," United States Patent Application, No. 11/554857, October 31, 2006.
7. Slob, E. C., K. Wapenaar, and R. Snieder, "Interferomtery in dissipative media: Adressing the shallow sea problem for seabed logging applications," *SEG Expanded Abstracts*, Vol. 26, 559–563, 2007.

Preprocessing of Marine CSEM Data and Model Preparation for Frequency-domain 3D Inversion

J. J. Zach[1], **F. Roth**[1], **and H. Yuan**[2]

[1]EMGS ASA, Trondheim, Norway
[2]EMGS Americas, Houston, TX, USA

Abstract— The marine controlled-source electromagnetic (MCSEM) method has been evolving into a geophysical imaging tool for increasingly complex geological settings. At the same time, 3D inversion algorithms for arbitrary survey layouts demand improved data quality compared to standard processing. Using a state-of-the-art survey acquired in the fall of 2007, we present a processing sequence starting from time-domain electromagnetic data acquired by seabed receivers to providing frequency domain data input and data weights for advanced processing. This includes determination and/or quality control of receiver orientation and time synchronization, and we show the quality of azimuthal receiver data to be adequate to be included in future inversions. Further, navigation data are adapted to a discretized grid upon determining the seafloor bathymetry. For missing or inadequate coverage from seismic surveys, the bathymetry can be extrapolated from navigation data with a spline-based algorithm, which is also described. It has proven beneficial, both in computational time and for recovering a meaningful model, to obtain a starting model for a full 3D inversion scheme by inverting reference receivers using a global, simulated annealing scheme, the result of which is imprinted onto the seafloor.

1. INTRODUCTION

We present a robust workflow for data processing and starting model preparation for frequency-domain CSEM data inversion. Increasing operational accuracy and improving equipment enables 3D inversion of complex survey layouts (e.g., [1, 2, 5, 8]). This has raised the level of requirements on procedures and supporting methods for data preprocessing and quality control compared to attribute-based processing.

2. DATA PREPROCESSING

2.1. Data Processing Workflow

The main steps of the pre-processing workflow for frequency-domain 3D-inversion are summarized in Figure 1. Subsequent to the hardware-dependent calibration of electromagnetic data, they are converted into the processing data format. To date, operational and cost reasons prohibit a sufficiently accurate direct measurement of the receiver orientation on the seafloor. Instead, the orientation is determined from the data itself, and the data for each receiver are rotated into the coordinate system of the source towline. This rotation, however, is not commutative with the estimate of the data uncertainty, or cumulative noise, which is essential for successful data inversion, and which occurs during demodulation. Hence, two demodulation steps are necessary in the processing, with the first demodulation step serving to determine the rotation angles, which are used to rotate the data in the time domain, followed by the second demodulation step in the rotated coordinate system. The most essential processing steps are outlined in detail in the following.

2.2. Data Demodulation and Noise Extraction

A short-time Fourier transform is applied to the seabed receiver data, which is recorded with a sampling frequency of typically 50 Hz. The integration window is thereby equivalent to the time it takes for the source to cross approximately one source length in about 25 periods of the base frequency, which is typically $f_0 \sim 0.25$ Hz. As an example, Figure 2 shows a receiver spectrogram resulting from a source square pulse with 0.25 Hz frequency, and Figure 3 shows a detailed power spectrum at a certain offset around the base frequency. The signal-to-noise ratio of a given frequency mode f_i is approximated by the magnitude ratio between f_i and a given frequency range $\Delta f_{int} < f_{i+1} - f_i$ in the spectral neighborhood of the mode (see the blue versus the green boxes in the right panel of Figure 3). The resulting signal and average noise amplitudes for three modes are plotted in the center panel of Figure 2. The dominant noise source is the electronics with relatively flat frequency dependence in the range from $\sim$0.1–10 Hz, and the apparent increase of noise with frequency is due to the lower source energy content in the higher modes for the square pulse used.

However, different noise characteristics for individual sensors including their electronics, as well as the superposition with natural noise sources (mainly ocean currents), renders the averaging procedure illustrated in Figure 3 a highly nonlinear operation.

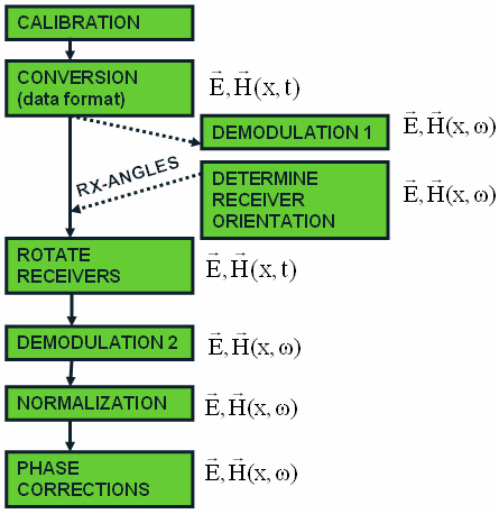


Figure 1: Main steps in preprocessing workflow for frequency-domain 3D inversion of marine CSEM data.

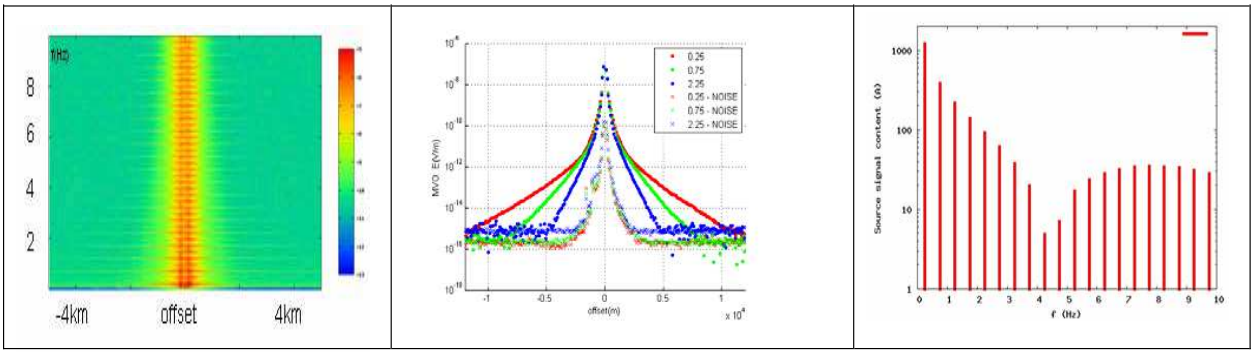


Figure 2: LHS: frequency power spectrum versus offset for one component of an electric signal upon source passage for a square-pulse CSEM source. Center panel: magnitude versus offset for the base frequency (0.25 Hz) and selected harmonics (3rd and 9th, 0.75 Hz and 2.25 Hz, respectively), along with noise (definition see Figure 3). RHS: Spectral content in electric current of a 0.25 Hz square source pulse $J_{src}(\omega) = \int dt J_{src}(t) \exp(-i\omega t)$.

2.3. Receiver Rotation

In deep water, accurate determination of the receiver orientation presently necessitates towing the source over each receiver, upon which the orientation relative to the source towline is found from symmetry considerations under assumptions of a plane-layer model for each offset [3, 4]. This relative rotation angle versus offset (see example in Figure 4) is thus used for QC and generally, the median angle is used in the further processing. The accuracy achieved is usually better than 5 degrees, however, the uncertainty can be higher for tilted bathymetries, particularly with the tilt perpendicular to the towline.

2.4. Source Normalization

The receiver data are normalized to the source signal in both magnitude and phase (see the right panel in Figure 2 for the example of a 0.25 Hz standard square pulse; see Mittet, Schaug-Pettersen 2007 for more complex source waveforms). While many earlier CSEM surveys still required corrections to the phase on close source-receiver passage, effectively rendering azimuthal data unusable in inversion schemes using both magnitude and phase, state of the art-data shows phase drifts of

typically much less than 1 degree (10 ms in time), see the right hand side in Figure 5 for the phase of multiple receivers in a recent data example.

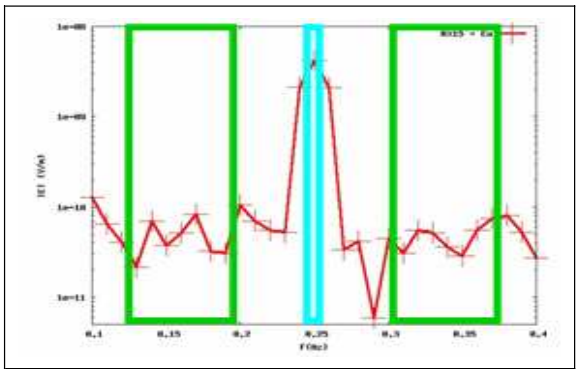


Figure 3: Receiver spectrum at offset +4 km around the base frequency (blue box). The green boxes define the windows to find a noise approximation.

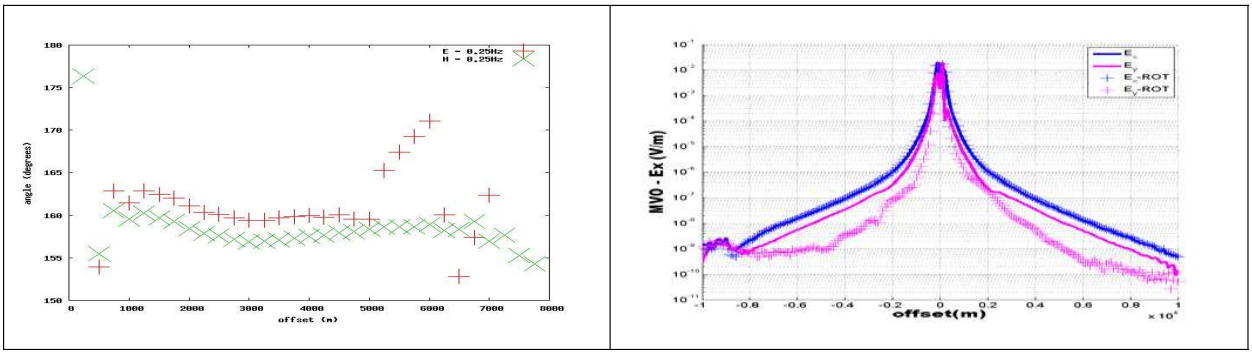


Figure 4: Inline rotation of receivers: LHS: orientation versus offset from electric (red curve) and magnetic field (green curve) at 0.25 Hz (base frequency in a square input pulse) for an example receiver. RHS: Comparison of the horizontal electric field before (Ex, Ey) and after (Ex-ROT, Ey-ROT) coordinate transformation into towline coordinates.

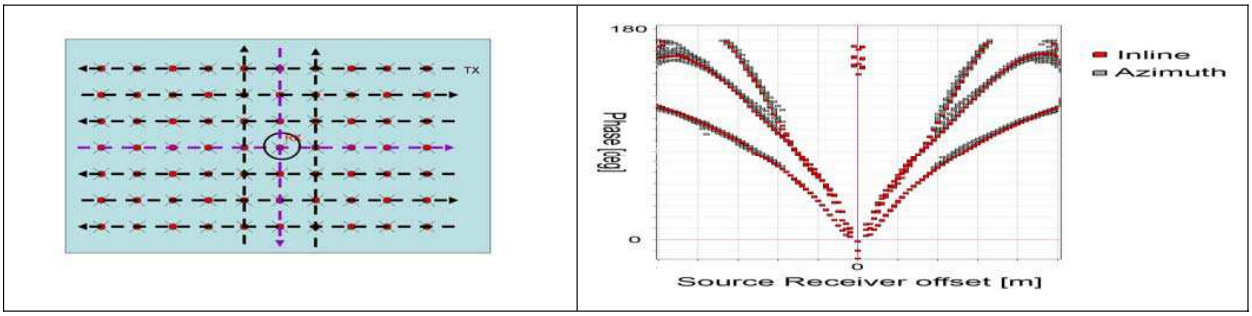


Figure 5: 3D inversion of full azimuthal data (both black and purple lines in illustration on LHS) is enabled by the consistent relative phase versus offset without any timing corrections; see RHS for phase versus offset for 1st and 3rd harmonic frequency for both inline and azimuthal data in a recent 3D-survey in the Troll oil province (2008).

3. STARTING MODEL PREPARATION

3.1. Adaptation of Navigation & Bathymetry

Since the seafloor conductivity is measured, inversion is limited to the space below mudline — in order to avoid inversion artifacts, this invariance makes it the more important to accurately represent the bathymetry and the source location relative to mudline in a finite-difference computational

grid. The source navigation data recorded during towing is comprised by a suite of continuously operating accoustic, echosounder and pressure measurements, which are recorded every 10 seconds, and from which important source parameters for accurate modeling are extracted, such as source altitude, tilt and feathering. In absence of independent bathymetry information (e.g., from 3D seismic), we construct a seafloor using water depths measured along the source towlines based on a continuous curvature spline algorithm [7].

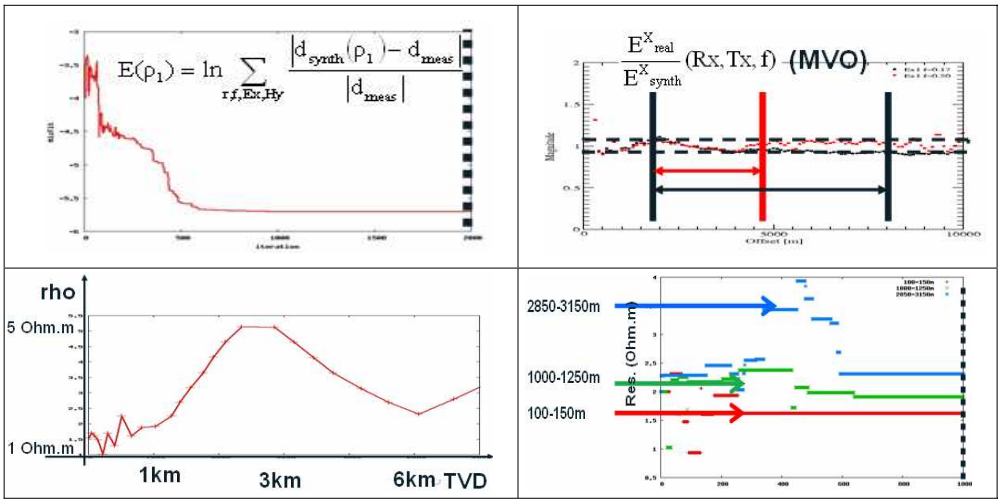


Figure 6: Example for the global inversion of two electrical frequency modes (0.17 Hz, 0.5 Hz) from a reference receiver. Upper LHS: misfit function versus iteration number. Upper RHS: ratio (in magnitude) between measured and real data after 2000 iterations, showing agreement to within 5% (black dashed lines; the agreement in the phase is approximately 5 degrees). Lower LHS: resulting resistivity-depth profile. Lower RHS: evolution of three selected resistivity bin in the simulated annealing scheme during exponential cooling.

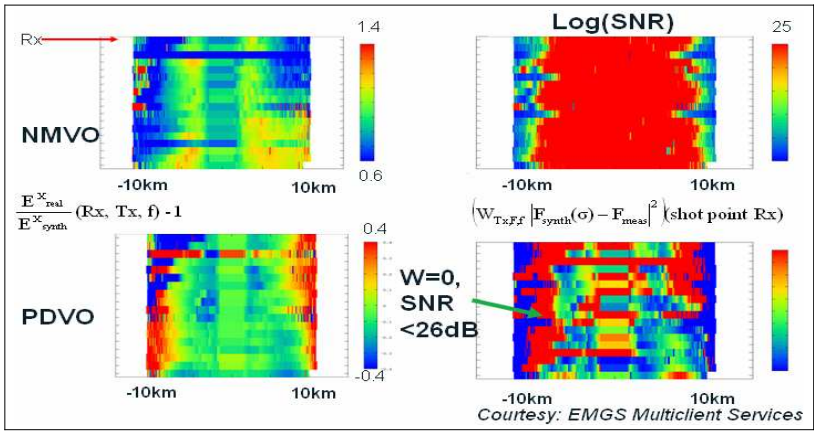


Figure 7: Convergence summary plots for iteration zero (starting model) in the inversion of a recent deep-water data example: upper/lower LHS: relative misfit in magnitude/phase versus offset in percent/radians. Upper RHS: approximate signal-to-noise ratio. Lower RHS: L2-kernel, accounting for offset ranges which are weighted down due to too low SNR, as well as offsets which an inverse scheme will act upon. The square of the inverse measured data was used to determine the weights: $W = 1/|F_{meas}|^2$. The data misfit indicates a conductive CSEM anomaly.

3.2. Determining a Starting Model with Global Plane Layer Inversion

All presently feasible algorithms for 3D inversion of CSEM data are gradient-based optimization routines in a parameter space of up to $\sim 10^7$ unknowns (e.g., [5, 8]). From the viewpoints of achieving non-unique solutions without heavy regularizations and also numerical complexity, it is therefore necessary to use a model resulting from a numerically less challenging scheme as starting model. To further prevent artifacts at discontinuities of the initial resistivity model inherent to gradient-based inversion approaches, the initial model needs to be sufficiently smooth.

We have standardized a global simulated-annealing scheme as described in Roth and Zach (2007) to invert selected reference receivers with low anomaly signature. Figure 6 shows the result from the inversion of a reference receiver in a recent marine CSEM survey, which converged to within the accuracy of the receiver data of $\sim 5\%$ within 2000 iterations (~ 2 hours on a single-processor PC). The smoothness constraint was thereby 0.5 Ohm·m per bin on a stretched grid (growth factor with depth: 1.1).

4. CONCLUSIONS

We have standardized and tested pre-processing and model preparation workflows for advanced processing of MCSEM data such as 3D inversion. Sufficient quality of all, including azimuthal, data to invert for both magnitude and phase is demonstrated. Spanning the seafloor from navigation data together with the global simulated annealing inversion of reference receivers comprise reliable and robust methods to prepare a starting model for a gradient-based inversion scheme. Paticularly, the global inversion stage has been shown to reduce the number of iterations in a full 3D scheme by up to a factor of three and has, in some cases, helped to recover resistive anomalies not found with a smooth starting model. As an example, Figure 7 shows the initial relative data misfit, approximate signal-to-noise ratio and L2-kernel of the misfit functional in an inversion scheme for a recent example line. Data preprocessing and starting model preparation were conducted using the methods described above.

ACKNOWLEDGMENT

We thank Peter van der Sman and Rene-Edouard Plessix from Shell for valuable discussions. We further acknowledge EMGS Multiclient Services for permission to publish most of the data examples shown.

REFERENCES

1. Bornatici, L., R. Mackie, and M. D. Watts, "3D inversion of marine CSEM data and its application to survey design," *EGM 2007 Conference Proceedings*, Capri, Italy, 2007.
2. Gribenko, A. and M. Zhdanov, "Rigorous 3D inversion of marine CSEM data based on the integral equation method," *Geophysics*, Vol. 72, WA73–WA84, 2007.
3. Mittet, R., O. M. Aakervik, H. R. Jensen, S. Ellingsrud, and A. Stovas, "On the orientation and absolute phase of marine CSEM receivers," *Geophysics*, Vol. 72, F145–F155, 2007.
4. Mittet, R. and T. Schaug-Pettersen, "Shaping optimal transmitter waveforms for marine CSEM surveys," *San SEG 2007 Conference Proceedings*, San Antonio, 2007.
5. Plessix, R.-E. and P. van der Sman, "3D CSEM modeling and inversion in complex geological settings," *Expanded Abstracts, SEG 2007 Annual Meeting*, San Antonio, 2007.
6. Roth, F. and J. J. Zach, "Inversion of marine CSEM data using up-down wavefield separation and simulated annealing," *SEG 2007 Conference Proceedings*, San Antonio, 2007.
7. Smith, W. H. F. and P. Wessel, "Gridding with continuous curvature splines in tension," *Geophysics*, Vol. 55, No. 3, 293–305, 1990.
8. Zach, J. J., A. K. Bjørke, T. Støren, and F. Maaø, "3D inversion of marine CSEM data using a fast finite-difference time-domain forward code and approximate Hessian-based optimization," *Accepted for SEG 2008*, Las Vegas, 2008.

Regularized and Blocky 3D Controlled Source Electromagnetic Inversion

R.-E. Plessix and P. van der Sman

Shell International E & P, The Netherlands

Abstract— Since the interpretation of controlled source electromagnetic (CSEM) data proves challenging in complex geological settings, 3D CSEM anisotropic resistivity imaging problem is formulated as an inverse problem. A least-squares misfit functional is minimized with a quasi-Newton algorithm to cope with the large number of unknowns. Furthermore, model and data weights are applied to speed up the convergence of the non-linear inversion. A-priori information obtained for instance from seismic interpretation can be included either in a blocky inversion, i.e., where the resistivity cube is parameterized with a small number of parameters or with a regularized inversion. Since earth resistivity contrasts can be high and spatially well defined, minimum norm support regularization terms are implemented. Although the blocky inversion proves quite powerful, we show that results can be misleading in complex geological settings. A combination of the blocky and regularized inversions may provide a more robust approach to interpret CSEM imaging results. We illustrate this with a deep-water example.

1. INTRODUCTION

Controlled source electromagnetic (CSEM) surveys consist of recording electromagnetic responses to an emitting electromagnetic source. In off-shore applications, a horizontal electric dipole source is towed over electromagnetic receivers [1]. This scanning technique helps detecting a thin and elongated resistive layer in a more conductive background. Because a reservoir charged with hydrocarbon can be several orders of magnitude more resistive than the same reservoir filled with brine, CSEM surveying is an important tool for oil and gas exploration. Crude interpretation of CSEM data based on anomalies detected after normalization of the data with computed or measured reference responses proves to be insufficient and even erroneous in complex geological settings in which other resistive bodies like, for example, hydrate, salt, or basalt, are present [2]. A more powerful approach consists of formulating the imaging problem as an inverse problem [3]. A least-squares misfit between observed data and computed synthetics is minimized with a quasi-Newton algorithm [4]. This local optimization approach, combined with a decoupling of the computational and inversion grids [5], allows us to deal with large 3D CSEM inverse problems containing typically between 5 and 20 millions of unkowns. Furthermore, model and data weights are added to speed up the convergence of the non-linear inversion [6].

Because of the diffusive nature of the electromagnetic waves in the conductive earth, the inverse problem is highly non-linear. To reduce the non-uniqueness and include a-priori information, we implemented two approaches: a blocky inversion scheme where the resistivity cube is described with a small number of parameters defining large homogeneous blocks and a regularized inversion. We did not use the classic square norm of the resistivity gradient as regularization term because this favours smooth results. To allow for large and spatially sharp resistivity contrasts, we prefered regularization terms based on the minimum norm support [7].

2. IMAGING APPROACH

Given the observed data, d, the imaging problem is formulated as an inverse problem:

$$\text{Find } m^{opt} \text{ such that } J(m^{opt}) \leq J(m)\ \forall m. \tag{1}$$

Here, m is the model parameter vector, m^{opt} is the optimal model parameter, and J is the least-squares misfit functional:

$$J(m) = \frac{1}{2}\|W^d(d - c(m))\|^2 + \alpha R(m) \tag{2}$$

where c are the synthetics solution of the Maxwell/Ohm equation with conductivity $\sigma(W^m m)$, W^m some model weigths, W^d some data weights and R a regularization term with α the regularization coefficient. In our implementation, the Maxwell/Ohm equations are discretized on a stretched computational grid with the finite integration technique and the associated linear system is solved

with an iterative multi-grid based solver [8]. The computational grid is automatically defined from the minimum resistivity value and the frequency and is decoupled from the inversion grid.

We don't directly invert for the conductivity on the computational grid because it would be expensive for multiple-frequency inversion. This approach however involves a mapping from the inversion grid to the computational grid that may introduce numerical inaccuracy. This decoupling of the computational and inversion grid also allows us to define a blocky and a regularized inversion.

In a blocky inversion, the domain is mainly divided in large regions/blocks and m is the vector constituded by the logarithm of the resistivity in each large region/block. A procedure to map this coarse discretization to a regular grid is then required to compute the electromagnetic fields with the finite integration technique.

In a regularized inversion, m is the vector constituted by the logarithm of the resistivity in each cell of the regular inversion grid. The inversion is stabilized with the minimum norm support functional:

$$R(m) = \frac{\|m - m^{ref}\|^2}{\beta^2 + \|m - m^{ref}\|^2}, \tag{3}$$

with m^{ref} a reference model, deduced, for instance, from a-priori information. The coefficient β means that a difference between the current model and the reference model larger than roughly 3β is hardly penalized. In our numerical tests, $\beta = 0.2$. We also tested the gradient minimum norm support

$$R(m) = \frac{\|\nabla m\|^2}{\beta^2 + \|\nabla m\|^2}. \tag{4}$$

3. DEEP-WATER REAL DATA SET

To illustrate the behavior of blocky and regularized inversions, we inverted a data set recorded in deep water offshore Malaysia. The acquisition is a fairly standard line acquisition with 24 receivers spread along 20 km. We only processed the inline electric responses to a towed inline horizontal electric dipole. The geological setting is rather complicated because of large bathymetric variations, the presence of hydrates, an increase background resistivity in the thrusted zone, where the interpreted deep targets sit, and the possibility of stacked resistive layers corresponding to a shallow accumulation and two deeper targets. A structural interpretation based on seismic data is displayed in Figure 1.

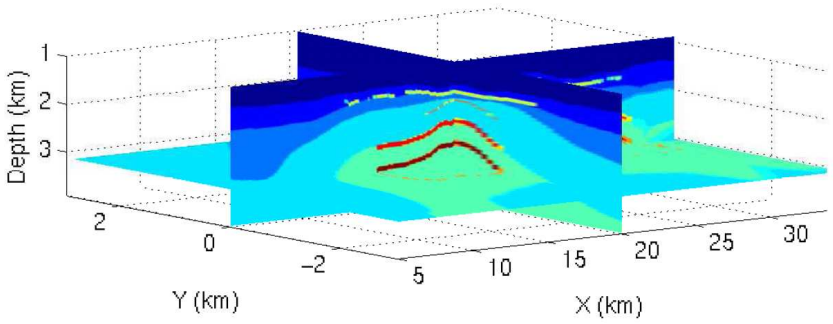


Figure 1: A 3D structural interpretation.

The amplitude of the data at 0.25 Hz normalized by the responses of a constant earth conductivity is plotted in Figure 2. Although a conventional analysis of anomalies provides some hints, the actual complexity of this problem renders this interpretation hopeless. Hence, we applied the isotropic inversion approach using four frequencies at 0.25, 0.75, 2.25, and 6.75 Hz. The initial guess is a constant earth conductivity at 2.0 Ohm·m.

4. BLOCKY INVERSION

Based on the seismic interpretation, see Figure 1, we carried out two blocky inversion. In the first blocky inversion, we did not consider the two possible deep targets. The model then consisted in 6 main blocks. In the inversion, we assumed a constant conductivity in each of the 4 large background layers. However, we allowed variable resistivity in the hydrates and shallow accumulation zones. In the second blocky inversion, the model consisted in 8 main layers. We however allowed

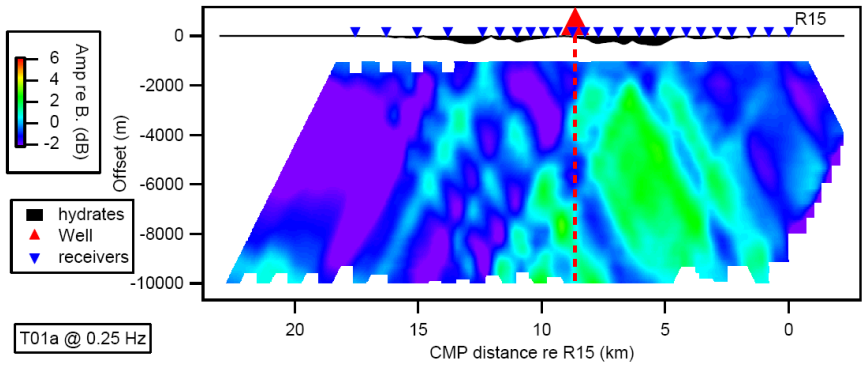


Figure 2: Normalized amplitude plotted at 0.25 Hz. The red line indicates the well location.

variable resistivity in the hydrates, shallow accumulation and deep target zones. The results after 50 iterations are displayed in Figure 3.

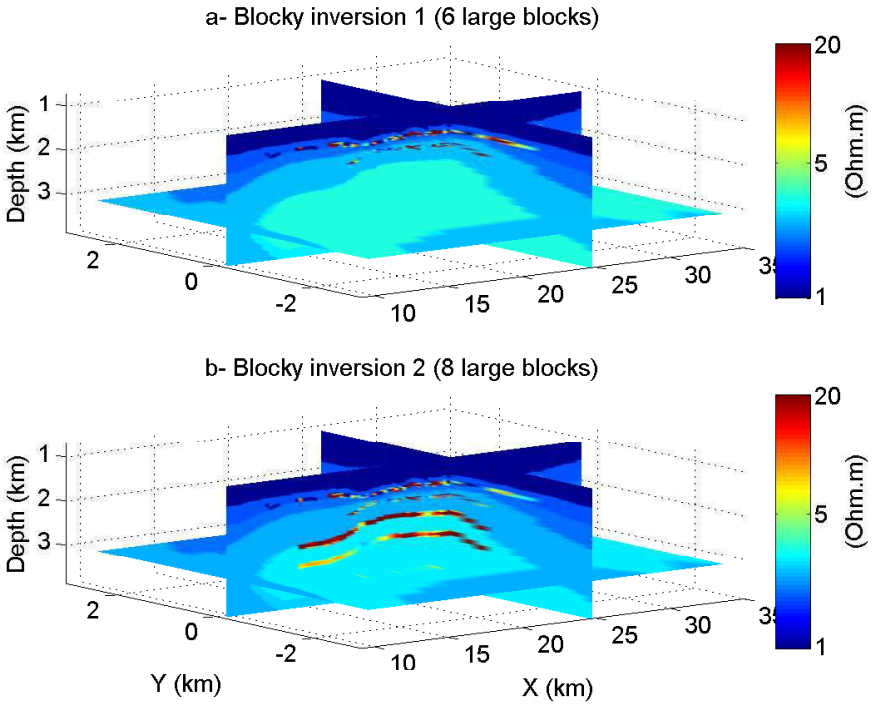


Figure 3: Optimal resistivity model in a logarithm scale of the two blocky inversions after 50 inversions.

Both resistivity models explain the data equally well, roughly up to the noise level. Unfortunately, those results give completely different images in the trust zone where the deep targets sit, which prevent us from assessing the presence of the deep targets. This illustrates the limitations of a blocky inversion approach with the diffusive electromagnetic inversion. A sensitivity analysis shows the trade-off between the shallow accumulation, the first deep target and the background resistivity of the thrust zone since the depth difference between the shallow accumulation and the first target is of the order of the skin depth at 6.25 Hz. We then have a lack of resolution in the CSEM data and forcing a high resolution results with blocky inversions leads to ambiguous results. The sensitivity analysis also shows that the data are not really sensitive to the deepest target. The blocky inversion has proven to be a powerful tool to cointerpret seismic and CSEM data sets. Unfortunately, it also has the tendancy to strongly biais the inversion. Some of the ambiguities may be resolved when more data, like multiple acquisition lines, broadside data, and multiple components, are included in the inversion.

This blocky approach tends to constrain the inversion too much, especially in the shallow part.

We already noticed this drawback as we were unable to successfully invert the data with constant resistivities in the hydrates, shallow accumulation, and target zones.

5. REGULARIZED INVERSION

To avoid the strong constraints of the blocky inversion, we carried out two regularized inversions, the first with the minimum norm support regularization and the initial guess as reference model and the second with the gradient minimum norm support regularization. We started the inversion with a strong regularization and relaxed it during the iterations. The results at iteration 75 are displayed in Figure 4.

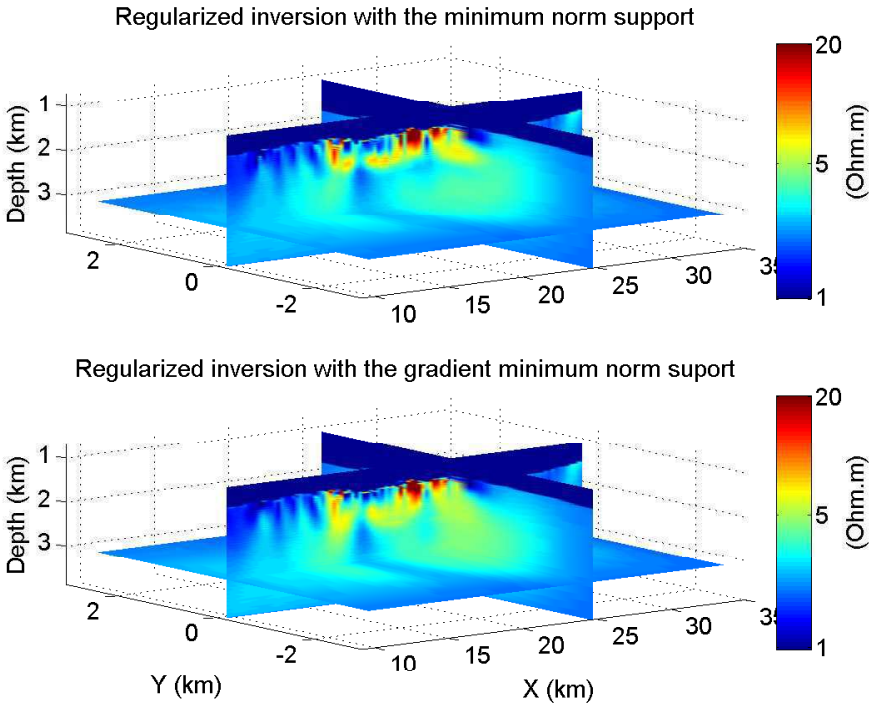


Figure 4: Optimal resistivity model in a logarithm scale of the two regularized inversions after 75 inversions.

Both regularized results explain the data equally well and are, in fact, quite similar. The use of the minimum norm support functional in the regularization term is satisfactory because, at least near the seafloor, the resistivity jumps have not been smoothed. Due to the diffusive nature of the electromagnetic waves the resolution decreases with depth. This is why we prefered the minimum norm support regularization over the classic smoothing regularization term, $\|\nabla m\|^2$, that tends to smooth results. In the final images, we can clearly detect the hydrates with relatively high resistivity and the shallow accumulation. We also notice an increase of resistivity at depth. This increase roughly correlates with the thrust zone of the structural interpretation map, Figure 1.

We also carried out a regularized anisotropy inversion to check the validity of the isotropic inversions. The optimal vertical resistivity, Figure 5, is very similar to the one obtained with the isotropic inversion. We also noticed that the data are not really sensitive to the horizontal resistivity. For a line and deep water acquisition, this result was expected because the inline response to an inline horizontal electric dipole mainly depends on the vertical resistivity. Isotropic inversion gives the vertical resistivity.

6. REGULARIZED PLUS TARGETED BLOCKY INVERSION RESULT

The previous results show that the regularized inversion is able to retrieve roughly the first 700 m below the seafloor with a satisfactory resolution, and indicates a regional increase of resistivity. The blocky inversion allows us to refine this result using a-priori information. Therefore, we decided to carry out a final partially blocky inversion. Using the optimal regularized inversion result as starting model, we inverted only the thrust region. We considered three main blocks: the background thrust

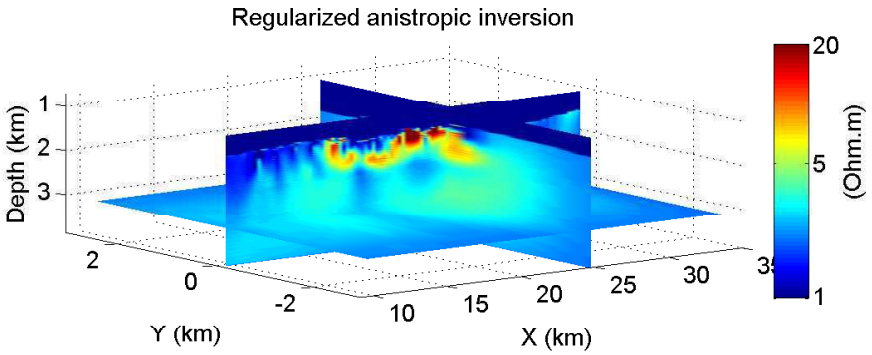


Figure 5: Optimal vertical resistivity model in a logarithm scale of the anisotropic regularized inversion after 75 inversions.

zone and the two deep targets. We impose a constant background resistivity in the thrust zone and the resistivity can freely vary in the deep targets. Outside the thrust region, the resistiviy is kept fixed and equal to the one retrieved with the regularized inversion. The final result, Figure 6, indicates an increase of resistivity of about 1 Ohm·m in the thrust zone and an uncharged first target. The result on the second target cannot be interpreted because the data are not sensitive to this target zone.

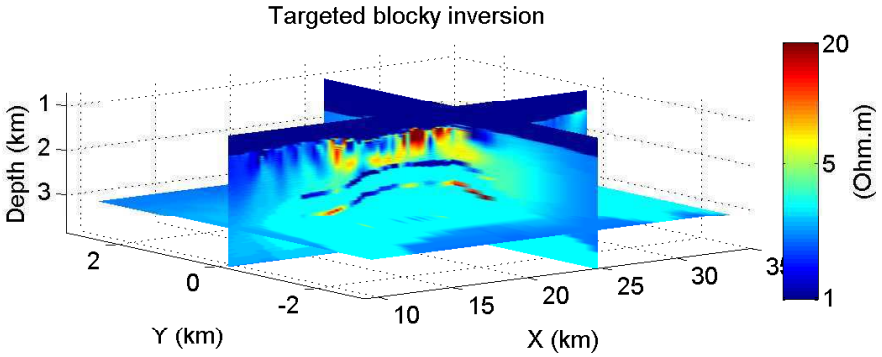


Figure 6: Optimal resistivity model in a logarithm scale obtained after a regularized inversion followed by a targeted blocky inversion.

This result is in agreement with the well log resistivity drilled before this complete analysis and that indicated the presence of hydrates, a shallow accumulation, and an increase of resistivity in the thrust zone.

7. CONCLUSIONS

We have implemented a blocky and regularized CSEM inversion based on a finite-difference solution of the Maxwell/Ohm equations and forward and backward interpolation between the computational grid and the discretization used to represent the model in the inversion. This allows us to take into account a structural interpretation in the inversion. We have illustrated the relevance of the minimum norm support regularization to stabilize the inversion and preserve the resistivity discountinuities at least in the shallow part of the image where the multiple frequency data constrain the inversion relatively well. The use of a priori information is still crucial in the interpretation of CSEMdata. However, we showed that blocky resistivity inversions based on a seismic interpretation can be misleading or at least inconclusive. A careful combination of both regularized and blocky

inversions leads to a more robust interpretation workflow as was illustrated with a real deep-water example.

ACKNOWLEDGMENT

We would like to thank PETRONAS, Sabah Shell Petroleum Co., and ConocoPhilips for the permission to show the real example and Shell International E & P for allowing the publication.

REFERENCES

1. Ellingsrud, S., T. Eidesmo, S. Johansen, M. C. Sinha, L. M. MacGregor, and S. Constable, "Remote sensing of hydrocarbon layers by seabed logging (SBL): Results from a cruise offshore Angola," *The Leading Edge*, Vol. 21, 972–982, 2002.
2. Darnet, M., M. C. K. Choo, R.-E. Plessix, M. L. Rosenquist, K. Yip-Cheong, E. Sims, and J. W. K. Voon, "Detecting hydrocarbon reservoir from controlled source electromagnetic (CSEM) data in complex settings: Application to deep water Sabah, Malaysia," *Geophysics*, Vol. 72, WA97–WA103, 2007.
3. Newman, G. A. and G. M. Hoversten, "Solution strategies for two- and three-dimensional electromagnetic inverse problems," *Inverse Problems*, Vol. 16, 1357–1375, 2000.
4. Haber, E., "Quasi-Newton methods for large-scale electromagnetic inverse problems," *Inverse Problems*, Vol. 21, 305–223, 2005.
5. Plessix, R.-E. and P. van der Sman, "3D CSEM modeling and inversion in complex geological settings," *Proceedings of the 77th SEG Annual Meeting*, 589–592, San-Antonio, USA, September 2007.
6. Plessix, R.-E. and W. A. Mulder, "Resistivity imaging with controlled-source electromagnetic data M depth and data weighting," submitted to *Inverse Problems*, 2008.
7. Zhdanov, M. S., *Geophysical Inverse Theory and Regularization Problem*, Elsevier, 2002.
8. Mulder, W. A., "A multigrid solver for 3D electromagnetic diffusion," *Geophysical Prospecting*, Vol. 54, 633–649, 2006.

Analysis of Complex SAR Raw Data Compression

Navneet Agrawal[1] and K. Venugopalan[2]

[1]Dept. of Electronics & Communication Engineering, College of Technology & Engineering
MPUAT, Udaipur 313 001, Rajasthan, India
[2]Dept. of Computer Science, University College of Science
M.L.S. University, Udaipur 313 001, Rajasthan, India

Abstract— Synthetic Aperture Radar (SAR) is active and coherent microwave radar, which produces high spatial resolution images from a moving platform — an airplane or a satellite. The radar produces 2D (range and azimuth) terrain reflectivity images by emitting a sequence of closely spaced radio frequency pulses and by sampling the echoes scattered from the ground targets. The received echoes are sampled into In-phase (I) and Quadrature (Q) components referred to as raw SAR data. In satellite systems, raw data is directly transmitted to the ground segment via a dedicated transmission link when in view with the ground segment or is stored onboard for later transmission to the ground.

For the compression of the satellite SAR raw data, the amplitude and phase are both indispensable for many applications e.g., interferometer. We propose to use two formats of the complex SAR raw data for compression: representation with real and imaginary parts and polar format [1] with magnitude and phase. In both the formats the phase information of the compressed data is preserved to a great extent. A block adaptive Max quantizer is used with 1–5 bit quantization of the components. The quality of the reconstructed data is compared in terms of compression ratio and quality parameters: signal to noise ratio (SNR), standard deviation of the phase (PSD), and mean phase error (MPE). The parameters are calculated for SAR raw data, complex data and 8-bit gray scale image. Finally, original (Fig. 4) and reconstructed gray scale images (Fig. 5) are presented.

1. INTRODUCTION

The compression of raw SAR data poses several challenges due to its noise like characteristics. The noise like characteristics arises because signals from several scatters are added incoherently with unknown phase and amplitude. Typically its adjacent samples are uncorrelated in both range and azimuth directions. Due to very poor correlation of SAR raw data it is not useful to apply redundancy reduction techniques [6] directly. Modern space borne SAR systems have an on-board hardware consisting basically of a transmitting and receiving unit and analog to digital (A/D) conversion, followed by a real time data downlink and/or storage facility. One of the main constraints in the design and operation of current space borne SAR systems is the non-availability of a downlink with a high data rate.

Usually, SAR data are represented in Cartesian format [4]. The data rate of each channel is proportional to the pulse repetition frequency (PRF), the number of sampled values in each echo, and the number of quantization bits in each sample. (PRF*N*bits). There are some possibilities to reduce the data rate, but these will deteriorate the system performance. Say reduction in PRF will increase the azimuth ambiguities unless a longer azimuth antenna is used. In this case the azimuth resolution will degrade. The system bandwidth can be reduced at the cost of a reduced range resolution. If the number of quantization bits are reduced, the digitization noise will increase which will further deteriorate the impulse response function (IRF), the image dynamic range and the radiometric accuracy. In interferometry, high phase preservation is required to prevent large errors. Therefore, complex SAR raw data compression makes greater demands on the reconstruction accuracy than in most other compression applications.

2. SAR DATA QUALITY PARAMETERS

The SAR data quality parameters are used to compare various methods of data compression. The efficacy of compression procedure is defined by compression ratio (CR) or by coding rates in bits per sample (bps). Most commonly used parameters to measure the performance are SNR (signal to noise ratio), mpe mean phase error) and psd (standard deviation of phase error). The performance of the quantizer is measured by mean square error (MSE). The MSE is defined as the expected

value E of the squared difference between the input signal s and the mapped output s'.

$$E(s-s')^2 = \frac{1}{M}\sum^{M}(s-s')^2$$

where M is the No. of input samples. MSE is usually normalized by the input signal power and the SNR is directly proportional to the normalized MSE and is a common parameter for the image quality analysis. The SNR is calculated with the help of ratio of variance of the input 8 bits raw data to the variance of the difference of the reference to the reconstructed image. Another quality parameters are mean phase error (mpe) and Phase standard deviation (Psd) & are calculated at two stages. For this we consider complex part of the image to be compared because taking only the real part will not give any angle as when imaginary part is zero the arctangent will always be zero.

3. BASIC SAR DATA COMPRESSION MODEL

The basic model used to implement the complete data compression, on the SAR image and its reconstruction is shown in Fig. 1.

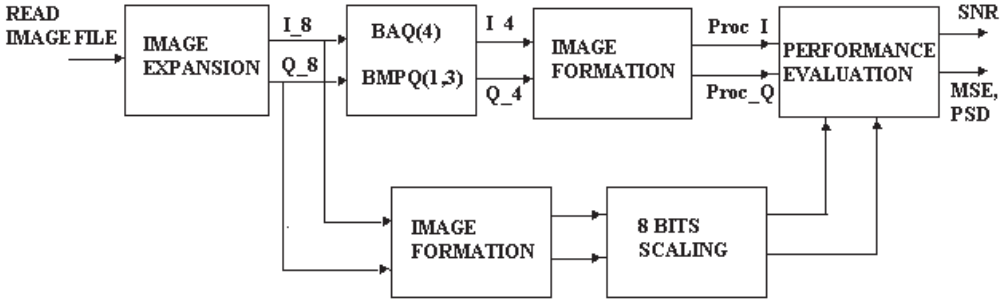


Figure 1: Basic SAR data compression model.

We first go for expansion of this image. The expansion of the image is done in azimuth and range directions. In order to expand the image in azimuth direction we generate azimuth reference chirp signal and another chirp signal is generated as a reference signal for range expansion. For azimuth expansion, we go for 1024 point fft of the input signal and 1024 point fft of the reference function as well. The two ffts are multiplied which is equivalent to convolving the two in time domain. Then we go for inverse fft of the result. This gives the expanded image in azimuth direction. The resultant image is not 256×256 after the 1024 points fft. The same process of expansion is carried out in range direction also, after giving corner turn to the azimuth-expanded image.

After expansion in azimuth and range we again go for corner turn. Now we select the image having size 256×256 out of the expanded image having size 1024×1024. The selected image of the size 256×256 is the useful image for further processing. In order to bring the 256×256 image in to 8 bits i.e., to represent the input SAR data with 8 bits confining the image gray between the values -128 to $+127$. This is raw data and we designate them as I_8 and Q_8. When we plot the histogram of real and imaginary parts, they have quite a good resemblance with the Gaussian distribution. Thus the data of the expanded image preserve the shape. Now with the data output received after image expansion, the next stage is image compression. In the compression process the fft of the azimuth reference chirp is first multiplied with the fft of the raw data (or BAQ processed data, if BAQ is used) and then the corner turn process is done by transposing the matrix. After this, inverse fft is done. Now we go for the same process in the range direction by multiplying fft of the range chirp with the fft of the azimuth-processed data. Then we again go for corner turn. Inverse fft of the resultant is done which gives the compressed image. The real and imaginary parts of the compressed image is represented as Proc_I and Proc_Q. The compressed image is also not of size 256×256 because of 1024 point fft. The desired size is extracted to 8 bits by scaling.

4. BLOCK ADAPTIVE QUANTIZATION

The most widely recognized method of raw SAR data compression is the block adaptive quantization (BAQ). BAQ uses a scalar quantizer controlled by the statistics of the raw SAR data to quantize

the raw data with fewer bits than required by the Shannon entropy. The viability of this reduction of bits is based on the observation that the entropy of the data is lower over a small time interval than the whole data set.

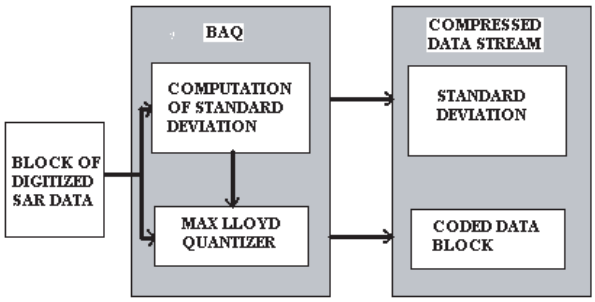


Figure 2: BAQ algorithm block diagram.

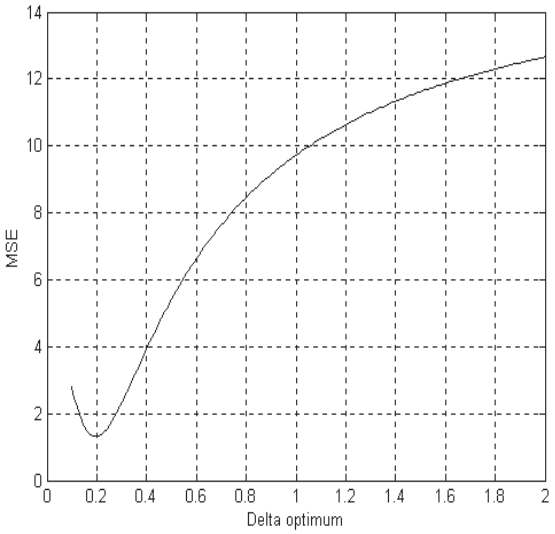


Figure 3: Scale factor curve for minimum MSE.

A block diagram of the BAQ algorithm can be seen in Fig. 1 in which the main module is a Max-Lloyd quantizer [6], which is adapted to the signal statistics block by block. The BAQ algorithm begins by dividing the raw SAR data (both the real and the imaginary parts). The block size is a trade off between the statistic of the incoming signal and the variation of the signal power. To ensure a Gaussian distribution, a minimum of 100 samples in a block is required; also a maximum block size of about the quarter of the pulse length is necessary to maintain the same average power in each of the block meaning that the signal power should be approximately constant in the block. BAQ algorithm is implemented using the following steps:

- Estimation of the variance of the block of data.
- Calculation of the optimal thresholds (Fig. 3).
- Quantization of the data using the optimal thresholds in 2/3/4/5/6/ or 8 bits.
- Transmission of the quantized data to the ground with variance.
- Calculation of the optimum threshold levels on ground.
- Reconstruction of the quantized data.

In the above process the quantization threshold values are determined by calculating the average magnitude of a block and using look-up table to provide the threshold values. Quantization is done by comparing an input sample to a set of thresholds. The outputs of the comparators are then coded to 2, 3 or 4 bits. The BAQ algorithm presents many advantages, mainly its simplicity and its low implementation complexity. It also allows an improvement of the resolution.

4.1. DATA Compression in Cartesian Format

The BAQ procedure is used for compression and decompression of complex data with coding rate from 1 bps to 4 bps for each I and Q components. The quality parameters SNR, MPE and PSD are evaluated for raw data, complex image and the 8 bits gray scale image is carried out. The visual monitoring of the gray scale image is also carried out. The results of the experiments are given in Table 1.

ASQNR performance of FBAQ and BMPQ are found to be very close to each other and to the theoretical values. As expected from theory, BAQ resulted in a small improvement (0.1 dB) over the other algorithm in I/Q domain. In the magnitude domain, BAQ and BMPQ performed similarly. For complex image SNR values are more approximately on 2 dB The rms and mean absolute errors for the algorithms show that BMPQ has significantly better performance in the signal domain than other algorithm. The phase error for the BMPQ is around the theoretical expected value of 13

degree rms, or 11 degree mean absolute while as BAQ gives phase error of around 18 degree and mean absolute phase error of 14 degree. The MPE and PSD show good results only for 4 bps quantization (CR = 2). After range and azimuth transform (For complex SAR image) phase error increases. The visual monitoring of the 256 gray scale image shows that BAQ with 4 and 3 bps have no difference from original image. The quantisation with 2 bps gives weak noise on darker sites of an image. At last quantisation with 1 bps for each channel noticeable increases the noise over all images.

Table 1: Data quality parameters: (BAQ) (gray256.bmp).

BAQ (bits per sample)	**4**		**3**		**2**		**1**	
SNR (dB) Calculated for SAR raw data	**Exp.**	**(obtained)**	**Exp.**	**(obtained)**	**Exp.**	**(obtained)**	**Exp.**	**(obtained)**
	20.5	(19.42)	14.9	(14.33)	9.5	(9.27)	4.4	(4.39)
(Mpe) & Psd calculated for SAR raw data	**Exp.**	**obtained**	**Exp.**	**obtained**	**Exp.**	**obtained**	**Exp.**	**obtained**
	(4.5°)	(5.48°)	(8.4°)	(9.19°)	(14°)	(14.39°)	(23°)	(22.30°)
	7.0	8.42	12	12.886	19	18.37	26	25.81

4.2. Data Compression in Polar Format

After transformation complex SAR data to a polar format [8] the magnitude and phase are quantized with different coding rate: magnitude with 1–4 bps and phase with 2–5 bps. The quality parameters SNR, PSD and MPE are evaluated for the SAR raw data and are given in Table 2.

Table 2: Polar format Results: (gray256.bmp) (BMPQ).

BMPQ (mag, phase) (bps)	**(3.5)**		**(2.4)**		**(1.3)**		**(0.2)**	
SNR (dB) Calculated for SAR raw data	**Exp.**	**(obtained)**	**Exp.**	**(obtained)**	**Exp.**	**(obtained)**	**Exp.**	**(obtained)**
	20.4	(19.58)	14.8	(14.09)	9.3	(8.75)	4.0	(4.98)
(Mpe) & Psd calculated for SAR raw data	**Exp.**	**obtained**	**Exp.**	**obtained**	**Exp.**	**obtained**	**Exp.**	**obtained**
	(2.8°)	(2.82°)	(5.7°)	(5.66°)	(11°)	(11.29°)	(23°)	(50.51°)
	3.3	3.26	6.5	6.53	13	13.05	26	65.11

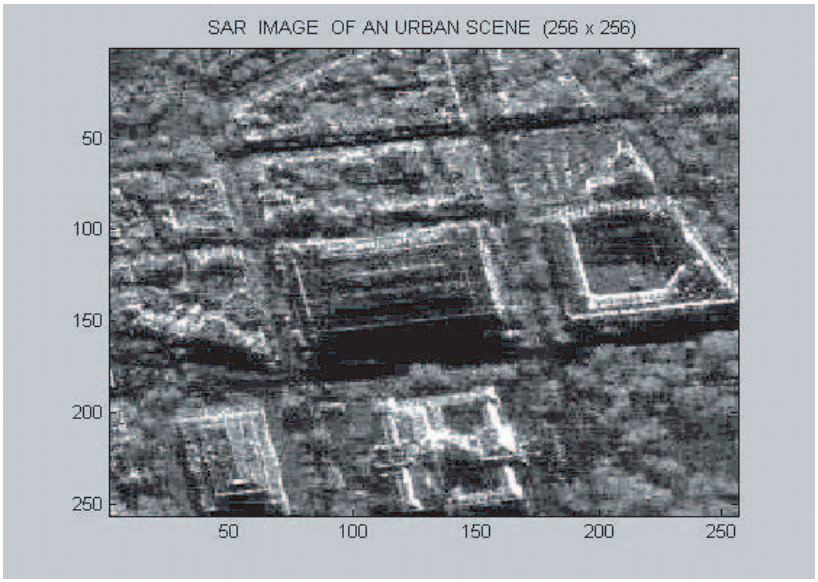


Figure 4: Input SAR image of a city.

The comparison SNR of two different formats [9] for complex SAR data: I and Q format and polar format shows that if the same coding rates are used in each channel ($4+4$ bps, $3+3$ bps etc.), the polar format gets worse values of SNR. But in a polar format there are variants of unequal quantization (magnitude + phase): $3+5$ bps, $2+4$ bps, $1+3$ bits, which for identical CR have the same SNR as I and Q format. Moreover, these variants enable phase errors twice less due to 1 bit greater for a phase coding. The histograms of I/Q data for BAQ filled up the available range poorly in 2 bits due to absence of any values at 0. BMPQ however fill the histogram space better, including filling values at zero.

The phase histogram of the original image is uniformly distributed as expected. BAQ and BMPQ have 12 possible reconstruction levels. This agrees the observed histograms of phase for the algorithms. With in the possible reconstruction levels, BMPQ retained the uniform distribution best. The phase error histograms again showed that the BMPQ has the best phase error performance, with the lowest maximum and smallest range of phase errors than BAQ.

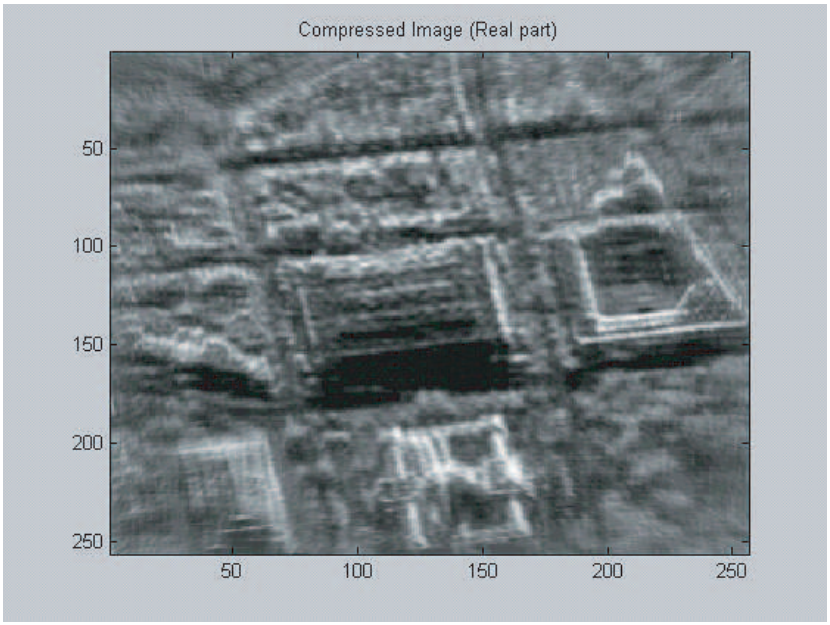


Figure 5: Reconstructed image after expansion and compression.

5. CONCLUSION

It is possible to use a polar format for complex data compression using BAQ method. This format gives large number of coding variants and permits to discover the acceptable compromises. For problems requiring better knowledge of phase more bits are selected for phase coding. All considered methods of compression are effective for SAR raw data and for range transformed data. For increasing overall data compression BAQ algorithm can be used in combination with other methods of data compression.

REFERENCES

1. Gerso, A. and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwar Academic Publishers, Boston, 1990.
2. Moreira, A. and G. Krieger, "Spaceborne synthetic aperture radar (SAR) systems: State of the art and future developments," *11th Gallium Arsenide Applications Symposium (GAAS)*, Oct. 2003, Munich.
3. DeWit, J. J. M., A. Meta, and P. Hoogeboom, "Modified range doppler processing for FM-CW synthetic apoerture radar," *IEEE Letters on Geoscience and Remote Sensing*, Jan. 2006.
4. Guan, Z.-H., D.-Y. Zhu, and Z.-D. Zhu, "Study on SAR raw data compression techniques," *Journal of Nanjing University of Aeronautics & Astronauts*, Vol. 37, No. 3, 325–329, Jun. 2005.
5. Max, J., "Quantization for minimum distortion," *IEEE Trans. Inform. Theory*, Vol. 6, 7–12, Mar. 1960.

6. Max, J., "Quantization for minimum distortion," *IEEE Trans. Inform. Theory*, Vol. 6, 16–21, Mar. 1960.
7. Kowk, R. and W. Johnson, "Block adaptive quantization of magellan SAR data," *IEEE Trans. Geosci. Remote Sensing*, Vol. 27, 375–383, Jul. 1989.
8. Peskova, S. and S. Vnotehenko, "Analysis of complex SAR data compression," 619–624, 2003.
9. Benz, U., K. Strdl, and A. Moreira, "A comparison of several algorithm for SAR raw data compression," *IEEE Transaction on Geosci. Remote Sensing*, Vol. 33, No. 5, Sept. 1995.

THz Rectangular Patch Microstrip Antenna Design Using Photonic Crystal as Substrate

Aditi Sharma, **Vivek K. Dwivedi**, **and G. Singh**

Department of Electronics and Communication Engineering
Jaypee University of Information Technology, Solan-173 215, India

Abstract— In this paper, the effects of two dimensional photonic band gap crystals (electromagnetic crystal) substrate on the performance of a rectangular microstrip patch antenna at THz (0.64–0.8 THz) frequency are simulated. Photonic crystal has been used as substrate material for high gain and highly directional resonant antenna. The bandwidth and gain of the designed antenna is 13.36% and 3.852 dB respectively which is very interesting result for this small dimension antennas. The simulation has been performed using CST Microwave Studio, which is a commercially available electromagnetic simulator based on finite difference time domain technique.

1. INTRODUCTION

The growing demand of wireless applications has presented RF engineers with continuing call for low cost, power efficient, and small size system designs. Depending on the application at hand and required system characteristics, such as data rate, environment, range, etc., the system parameters such as operating frequency transmitted power, and modulation scheme may vary widely. However, independent of the applications, compactness, wide bandwidth, high efficiency, ease of fabrication, integration and low cost are always sought in wireless systems [1]. THz region occupies a large portion of electromagnetic spectrum located between the microwave and optical frequencies and normally defined as the band from 0.1 to 10 THz. There are many forms of the communication system for sending message from one distant place to another. The basic motivation behind each new form is either to improve the transmission fidelity, to increase the data rate so that more information could be sent, or to increase the transmission distance between the relay stations. The frequencies higher than microwaves offer many advantages for technology, including wider communication bandwidths, improved spatial directivity and resolution with system compactness. At these frequencies communication with high data rates are possible, although greater bandwidths [2]. THz communication link is most likely secure communication for short-distance, point-to-point, and demanding high information data rate (multi-Mb/s to Gb/s). The effective range for free-space transmission of THz signals is also a concern. Among the practical advantages of using THz region for satellite communication system is the ability to employ smaller transmitting and receiving antennas. This allows the use of smaller satellite and a lighter launch vehicle. One of the most important components of the wireless systems is their antenna. Using carrier frequencies above 300 GHz, oscillator and amplifier sources with approximately 10% fractional bandwidth would enable very high data rate (> 10 GB/sec) wireless communications with high security protection. A photoconductive antenna is an alternative THz source because of its compactness and wide tunability at room temperature. However, the photoconductive antenna has the significant disadvantage of low output power. This is mainly due to high impedance inherent to photomixer. When an antenna with moderate input impedance is connected to a photomixer, the power transferred from photomixer to the antenna is poor due to the severe impedance mismatching. Microstrip antennas, because of their ease of fabrication/integration as well as compactness, are highly desirable for these systems. The substrate of microstrip antennas plays a very important role in achieving desirable electrical and physical characteristics. In conventional microstrip patch antenna, the antenna placed on a dielectric substrate radiates more efficiently into the dielectric substrate than the air-side. If we replace the substrate with the photonic crystal substrate, whose forbidden gap encompasses the antenna excitation frequency, in the manner that there are no surface modes, the power previously radiated into the substrate will be reflected towards the air-side. Photonic crystal is a new class of periodic dielectric structures has been developed in which electromagnetic wave propagation in any direction is completely prohibited for all frequencies within a stop band [3]. A photonic band gap material or photonic crystal is an artificial material made of periodic implants within a surrounding medium. Electromagnetic propagation through such a medium is affected by the scattering and diffraction properties of the periodic elements. Brown and Parker [4] proposed

increasing performance of planar antennas on a dielectric substrate, which suffer large radiation loss into the substrate with only 2–3% of the power radiated in the air.

In the microwave and millimeter-wave integrated circuits, the control of the radiation from a microstrip antenna is of great importance. In such circuits, the antennas are mounted on a semiconductor substrate, which enhance the performance and functionality of the circuit. But most of the power from the antenna on a dielectric substrate is radiated into the substrate. If a thin substrate is used to overcome the loss due to this trapping, another problem arises. A 180 degree phase shift comes from the reflection at the bottom conductor, causing the radiation cancel out at driving point. These problems can be solved, if the antenna is mounted on a 2-D photonic crystal, from which the radiation will be fully reflected in all directions [5–8]. By reducing or eliminating the effects of these electromagnetic inhibitors with photonic crystals, a broadband response can be obtained from inherently narrowband antennas [9–13]. It is also useful to a reduction in pattern side lobes resulting in improvements in the radiation pattern front-to-back ratio and overall antenna efficiency. Agi and Malloy [14] have experimentally and computationally studied the integration of a microstrip patch antenna with a two-dimensional photonic crystal substrate [15]. The aim of this paper is to demonstrate a gain enhancement method for the rectangular microstrip patch antennas through the use of photonic band gap materials made of planar arrays of circular blocks (circular air implants) within the dielectric layers. The high gain is due to the excitation of stronger leaky-wave fields. The organization of the paper is as follows. The Section 2 concern with geometrical configuration of the rectangular microstrip patch antenna. The Section 3 discusses the simulated results. Finally, Section 4 concludes the work.

2. ANTENNA CONFIGURATION

Figure 1 shows the geometrical configuration of the proposed rectangular microstrip patch antenna. The substrate material of this antenna is two dimensional photonic band gap crystals. The photonic band gap crystal composed of stacked layers of cylindrical air gap of 15 μm diameter at 200 μm distance between the two consecutive air gaps in the dielectric substrate. The dimension of substrate material is $1000 \times 1000\,\mu m^2$ and 200 μm thickness with dielectric permittivity 9.1. The radiating patch has dimensions $600 \times 460\,\mu m^2$. In this model of rectangular microstrip patch antenna, we have used microstrip strip line feeding [16, 17]. The dimension of the strip line is $270 \times 50\,\mu m$ with the width 36 μm.

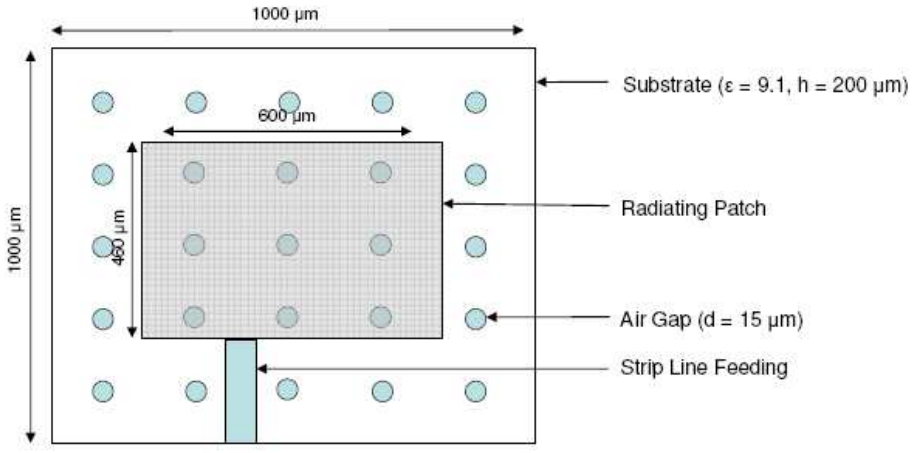


Figure 1: Geometrical configuration of the rectangular microstrip patch antenna with photonic crystal as substrate.

3. RESULTS AND DISCUSSION

Figure 2 shows the frequency versus return loss and reveals that the return loss is below −12.56 dB at frequency 720 GHz which is a center frequency of the operation range. The 10 dB impedance bandwidth as calculated from the Fig. 2 is 13.36%. The complete field expression for the microstrip structure is in terms of a continuous plane-wave spectrum. With planar material gratings, there may exist, three different propagating waves. Space waves and surface waves are similar to those in

conventional microstrip structures. Leaky waves are due to the periodic nature of the planar grating structure. The surface wave (bound wave) is a slow wave with normalized phase constant β/k_0 (k_0 — the free-space wave number) that increases with frequency. When the frequency increases to a point that the condition $\beta/k_0 \geq \lambda/q - 1$ holds, this bound wave becomes a leaky wave which is a fast wave for -1 space harmonic with a complex propagation constant. The far-zone radiated fields are due to the combination of space wave and leaky waves. The energy carried by bounded surface waves that propagate laterally is considered a loss. Antenna directivity, gain, and efficiency are determined by the energy distribution among these three propagating waves. Antenna efficiency can be maximized by partial elimination of the bound surface wave within certain directions. For high-efficiency antennas, directivity and gain are determined by the energy distribution between the space and leaky waves. Leaky-wave radiation pattern is highly directive in contrast to the low directivity of the space wave. To achieve high-antenna gain, it is necessary to excite a strong leaky wave.

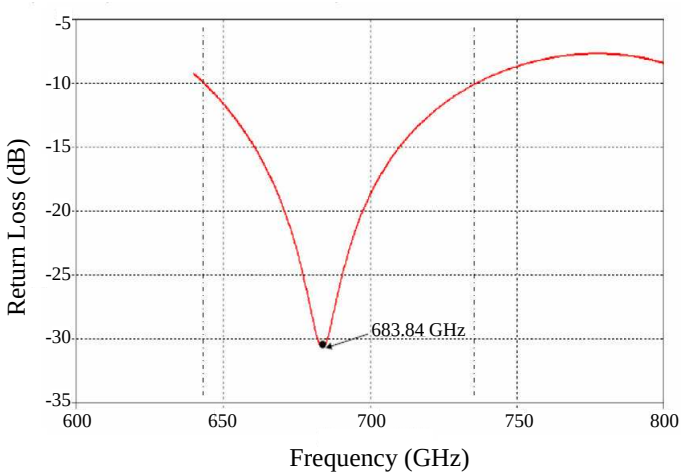


Figure 2: Frequency versus return loss of rectangular microstrip patch antenna at THz frequency using photonic crystal as substrate.

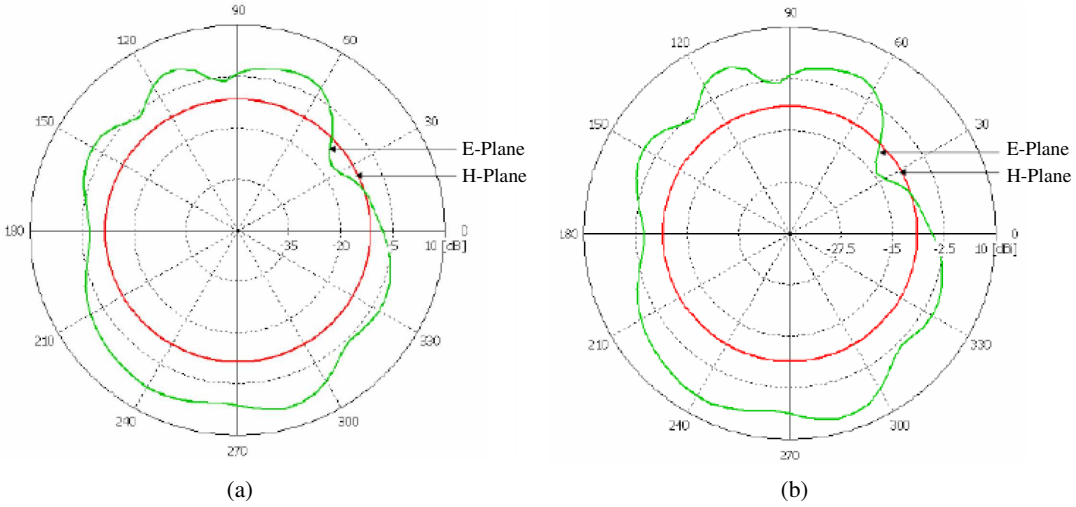


Figure 3: E and H plane far field radiation pattern of (a) gain and (b) directivity of proposed rectangular microstrip patch antenna at frequency 720 GHz.

E and H plane far field patterns for gain and directivity at frequency 720 GHz is shown in Fig. 3. The radiation efficiency, gain and directivity from Fig. 3 are 54.9%, 3.852 dB and 6.456 dBi at 36 µm strip feed line width. Significant gain enhancement is found (3.82 dB) in the E plane. It is observed that the gain decreases and the beam angle increases with increasing angle. As the

angle approaches 90°, the leaky-wave field is much weaker than the space-wave field and no gain improvement is observed (the H-plane pattern is shown in Fig. 3(a)). This observation is generally true for antennas on a thin photonic band gap substrate. The results of the directivity patterns are shown in Fig. 3(b). A particular plane with the strongest leaky-wave fields is determined by the arrangement of substrate configuration. The antenna location relative to the air blocks has a significant effect on the antenna gain. The thicker the air blocks are, the stronger the leaky wave and the higher the gain.

The effects of varying the width of micro strip feed line on radiation efficiency and gain is observed and shown in Fig. 4, which reveals that as the width of the strip line increases, the radiation efficiency and gain of the proposed microstrip patch antenna decreases continuously. At a point near to 30 µm the radiation efficiency and gain is maximum, but at this point the return loss is very high, which will affect the antenna performance. As we reached the width of 36 µm, the radiation efficiency and gain is good and the return loss decreases to approximately −30 dB at the center frequency of operation. The experimental results of the rectangular microstrip patch antenna on photonic crystal substrate at THz frequencies are still now not reported.

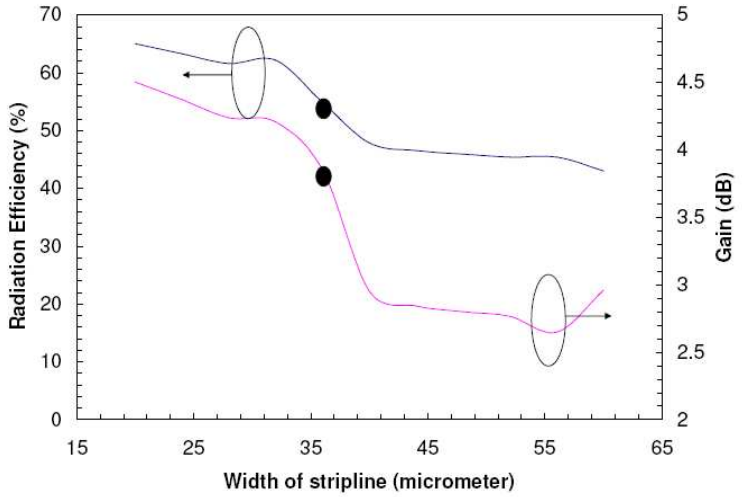


Figure 4: The effects of variation of width of microstrip strip feed line on the gain and radiation efficiency of rectangular microstrip patch antenna.

4. CONCLUSION

This paper explored a novel concept in the development of wideband rectangular microstrip patch antennas using 2-D photonic crystals as substrate at THz frequencies. We demonstrated that the gain of a printed circuit antenna can be greatly enhanced with photonic band gap materials as a substrate. It was found that significant gain enhancement is achieved by exciting strong leaky waves through proper designs of the planar periodic material structure. Photonic crystals were realized to reduce and, in some cases, eliminate surface waves, which leads to an increase in directivity, bandwidth and radiation efficiency. Using photonic crystals or we can say air gaps, promising results have been obtained by simulation through CST Microwave Studio. We will report the theoretical results as well as optimize the geometrical parameters to obtain the optimum electrical parameter of this proposed rectangular microstrip patch antenna at THz frequency in another future communication.

REFERENCES

1. Sharma, A. and G. Singh, "Rectangular microstrip patch antenna design at THz frequency for communication systems," *Int. J. Infrared and Millimeter Waves*, (under review), Dec. 2007.
2. Abdelaziz, A. A., "Bandwidth enhancement of microstrip antenna," *Progress In Electromagnetic Research*, PIER 63, 311–317, 2006.
3. Fernandes, H. C. C. and A. R. B. da Rocha, "Analysis of antennas with PBG substrate," *Int. J. Infrared and Millimeter Waves*, Vol. 24, No. 7, 1171–1176, July 2003.
4. Brown, E. R. and C. D. Parker, "Radiation properties of a planar antenna on a photonic-crystal substrate," *J. Opt. Soc. Am. B*, Vol. 10, No. 2, 404–407, 1993.

5. Ozbay, E., B. Temelkuran, and M. Bayindir, "Microwave applications of photonic crystals," *Progress In Electromagnetics Research*, PIER 41, 185–209, 2003.
6. Gonzalo, R. and B. Martinez, "The effect of dielectric permittivity on the properties of photonic band gap devices," *Microwave and Optical Technology Lett.*, Vol. 23, No. 2, 92–95, 1999.
7. Meade, R. D., A. M. Rappe, K. D. Brommer, and J. D. Joannopoulos, "Nature of the photonic band gap: Some insights from field analysis," *J. Opt. Soc. Am. B*, Vol. 10, 328–332, 1993.
8. Gonzalo, R., "Enhanced patch-antenna performance by suppressing surface waves using photonic-band gap substrates," *IEEE Trans. Microwave Theory and Tech.*, Vol. 47, No. 11, 2131–2138, Nov. 1999.
9. Zhang, Z. and S. Satpathy, "Electromagnetic wave propagation in periodic structures: Bloch wave solution of Maxwell equations," *Physical Rev. Lett.*, Vol. 65, 2650–2653, 1990.
10. Ho, K. M., C. T. Chan, and C. M. Soukoulish, "Existence of a photonic gap in periodic dielectric structures," *Phys. Rev. Lett.*, Vol. 65, 3152–3155, 1990.
11. Yang, H. Y. D., N. G. Alexopoulos, and E. Yablonovitch, "Photonic band gap materials for high gain printed circuit antennas," *IEEE Trans. Antenna and Propagation*, Vol. 45, 185–187, 1997.
12. Meade, R. D., K. D. Brommer, A. M. Rappe, and J. D. Joannopoulos, "Photonic band states in periodic dielectric materials," *Phys. Rev. B: Condensed Matter*, Vol. 44, 13772–13774, 1991.
13. Coccioli, R., W. R. Deal, and T. Itoh, "Radiation characteristics of a patch antenna on a thin PBG substrate," *IEEE Trans. Antennas Propag.*, Vol. 45, 656–659, 1998.
14. Agi, K. and J. Malloy, "Integration of a microstrip patch antenna with a two dimensional photonic crystal substrate," *Electromagnetics*, Vol. 19, 277–290, 1999.
15. Radisic, V., Y. X. Qian, R. Coccioli, and T. Itoh, "Novel 2D photonic bandgap structure for microstrip antenna," *IEEE Microwave and Guided Wave Lett.*, Vol. 8, 69–71, 1998.
16. Rumsey, I., M. P. May, and P. K. Kelly, "Photonic band gap structure for microstrip line," *IEEE Microwave and Guided Wave Lett.*, Vol. 8, 336–338, 1998.
17. Kesler, M. P., J. G. Maloney, B. L. Shirley, and G. S. Smith, "Antenna design with the use of photonic band gap materials as all dielectric planer reflectors," *Microwave and Optical Tech. Lett.*, Vol. 11, 169–174, 1996.

An Efficient BER Analysis of OFDM Systems with ICI Conjugate Cancellation Method

Vivek K. Dwivedi and G. Singh
Department of Electronics and Communication Engineering
Jaypee University of Information Technology, Solan 173 215, India

Abstract— In this paper, we have presented bit error rate (BER) analysis of the orthogonal frequency division multiplexing (OFDM) communication systems with the concept of conjugate cancellation scheme. In the first part of our work, an effort has been made to illustrate the mathematical derivation for BER of the OFDM system with frequency offset. Later, we have derived a formula for the BER analysis of OFDM system with the conjugate cancellation scheme. Bit error rate of the proposed conjugate cancellation scheme is better than the ICI self-cancellation. We also discusses carrier to interference ratio (CIR) and compared this proposed results with the others reported results.

1. INTRODUCTION

The next generation broadband multimedia communication systems will integrate various function and application in a system. This system support large data rates with sufficient robustness to radio channel impairments, requires careful choosing of modulation technique. The suitable choice seems to OFDM which is special case of multi-carrier communication system, where single data stream is transmitted over number of lower sub-carrier. OFDM communication systems can be seen as either a modulation or multiplexing technique. One main region to use OFDM is to increase the robustness against frequency selective fading or narrow band interference. This communication system is broadly considered as an effective approach for the future high speed wireless multimedia communication systems. The basic principle of OFDM is to split the high-data stream into number of lower rate data streams which are transmitted simultaneously over number of subcarriers. High spectral efficiency and multipath immunity are two major features of the OFDM technique.

However, OFDM system is very sensitive to carrier frequency offset between transmitter and a receiver, which destroys the orthogonality between sub-carriers and creates inter carrier interference (ICI). The reduction of the signal amplitude and introduction of the ICI are two destructive effects caused by carrier frequency offset in OFDM systems. So there is a need to reduce ICI. Recently several methods have been proposed and developed to estimate and adjustment of the effects of ICI. Among the all ICI cancellation schemes, the ICI-self cancellation scheme is a simple way for ICI reduction. The main idea is to modulate one data symbol onto the next sub-carrier with predefined inversed weighting coefficient "-1". By doing so the ICI signals generated within a group can be self cancelled each other. Several methods [1–6] have been developed for bit error rate (BER) analysis and ICI self cancellation of OFDM system. Kang et al. [1] and Zhao and Haggman [2] discuss the self cancellation scheme for OFDM to reduce the effects of frequency offset error. In [3] BER upper bound of OFDM system is analyzed without ICI self cancellation and [4] BER of OFDM system is analyzed using self cancellation but this method is less accurate. Yeh et al. [7] discussed ICI mitigation using conjugate cancellation but they have not provided any mathematical analysis of ICI. In [8, 9] the inter-carrier interference (ICI) by Doppler effects in time domain in orthogonal frequency division multiplexing systems is observed.

In this paper, we have presented BER analysis of the OFDM system with concept of conjugate cancellation scheme. In the first part of our work an effort has been made to illustrate the mathematical derivation for BER of OFDM system with frequency offset. Later, we have derived a formula for BER analysis of OFDM system with the conjugate cancellation scheme. The organization of the paper as follows. The Section 2 discusses the system model of the OFDM communication systems. The Section 3 is concern with the conjugate cancellation scheme for ICI. The Section 4 discusses the results and finally, Section 5 concludes the work.

2. SYSTEM MODEL

Figure 1 shows a typical discrete-time base-band equivalent model of an OFDM system. Input binary serial data stream is encoded using suitable modulation technique (M — QAM, BPSK, and QPSK). The further symbols are transferred in the serial-to-parallel converter (S/P) in this stage

duration of bits is increased. First part of parallel bit stream is subjected to IFFT block and second part is also subjected to FFT block. The modulated symbols are serialized using a parallel-to-serial converter (P/S). Now guard band addition is done because at the receiver one OFDM symbol is over lapped with the other symbol due to multipath distortion [1, 2, 7]. To eliminate the problem of inter symbol interference a guard time inserted between two symbols, duration of guard interval should be greater than maximum delay spread. Guard time consists no signal at all. Before transmission first part of data is sent as it is and conjugate is taken for second data. In next block digital signal is converted to analog via the digital-to-analog converter (D/A) before being sent down to the channel. At the receiver side, guard interval is removed and the received symbol is converted from analog to digital using the analog-to-digital converter (A/D). In next process first part of data is transferred as it is in serial to parallel converter and conjugate is taken for second part of data before transferring to the serial to parallel converter and then first and second part of data are sent in FFT block. After FFT and IFFT block data is sent for parallel to serial (P/S) conversion and then for demodulation. ICI cancellation is done after demodulation using diversity combiner.

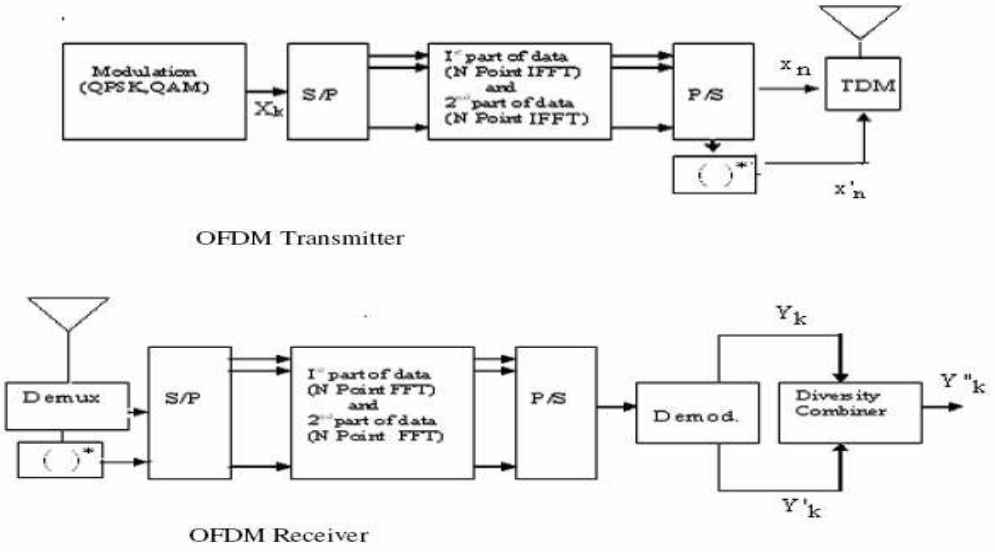


Figure 1: System model for communication systems.

3. CONJUGATE ICI CANCELLATION SCHEME

Input data bits are encoded by using suitable modulation technique like (QPSK or QAM) and output of his block is X_k. IFFT out put at the transmitter is:

$$x_n = \frac{1}{N} \sum_{k=-K}^{K} X_k e^{2\pi jnk/N}$$

where $n = 0, 1, 2, \ldots, N-1$, and $N \geq 2K+1$ where K is number of sub carries N is the period of IFFT. At received sequence after passing through the channel can be expressed as:

$$y_n = \frac{1}{N} \left[\sum_{k=-K}^{K} X_k H_k e^{2\pi jn(k+\varepsilon)/N} \right] + w_n$$

where $n = 0, 1, 2, \ldots N-1$, where H_k is channel transfer function at the frequency of kth sub carrier, ε is relative frequency offset of channel, w_n is Additive White Gaussian Noise (AWGN). Out put of DFT demodulator can be expressed as:

$$\begin{aligned} Y_k &= \sum_{n=0}^{N-1} \left\{ \frac{1}{N} \left[\sum_{n=-K}^{K} H_k X_k e^{2\pi jn(k+\varepsilon)/N} \right] + w_n \right\} e^{-2\pi jkn/N} \\ Y_k &= (X_k H_k) \left\{ (\sin \pi\varepsilon) / N \sin (\pi\varepsilon/N) \right\} e^{j\pi\varepsilon(N-1)/N} + \underbrace{I_k}_{II} + \underbrace{W_k}_{III} \end{aligned} \tag{1}$$

The first component is the modulation value X_k is modified by channel transfer function. This component experiences an amplitude reduction and phase shift due to the frequency offset. Second term is ICI term, which arises due to frequency mismatch of oscillator transmitter and receiver. After some manipulation second term ICI can be expressed as:

$$\begin{aligned} I_k &= \sum_{\substack{l=0 \\ l\neq k}}^{N-1} \frac{1}{N} X_l H_l \left(\sin \pi(l+\varepsilon-k)/\sin \pi\left((l+\varepsilon-k)/N\right)\right) \times e^{j\pi(N-1)\left(\frac{l+\varepsilon-k}{N}\right)} \\ I_k &\approx \sum_{l=0}^{N-1} X_l H_l \left(\sin \pi(l+\varepsilon-k)/\pi(l+\varepsilon-k)\right) \end{aligned} \tag{2}$$

Third is Additive White Gaussian Noise in frequency domain which can be expressed as

$$W_k = \sum_{n=0}^{N-1} w_n e^{-2\pi kn/N}$$

Now we will analyze the second part of data before transmission we will take conjugate of the original signal:

$$x'_n = \left(\frac{1}{N}\sum_{k=-K}^{K} X_k e^{2\pi jnk/N}\right)^* = \frac{1}{N}\sum_{k=-K}^{K} X_k^* e^{-2\pi jnk/N}$$

$n = 0, 1, \ldots N-1$. At the receiver a conjugate algorithm is requires a conjugate operation on received signal first and FFT is performed:

$$y'_n = \frac{1}{N}\left[\sum_{k=-K}^{K} X *_k H_k e^{2\pi jn(-k+\varepsilon)/N}\right] + w_n$$

where $n = 0, 1, 2, \ldots N-1$. Out put of DFT demodulator can be expressed as:

$$\begin{aligned} Y'_k &= \sum_{n=0}^{N-1}\left\{\frac{1}{N}\left[\sum_{k=-K}^{K} H_k X *_k e^{2\pi jn(-k+\varepsilon)/N}\right] + W_n\right\}^* e^{-2\pi jkn/N} \\ Y'_k &= (X_k H_k)\left\{\frac{(\sin \pi\varepsilon)}{N \sin\left(\frac{\pi\varepsilon}{N}\right)}\right\} e^{-j\pi\varepsilon(N-1)}/N + \underbrace{I'_k}_{II} + \underbrace{W'_k}_{III} \end{aligned} \tag{3}$$

ICI term can be expressed as:

$$\begin{aligned} I'_k &= \sum_{l=0, l\neq k}^{N-1} \frac{1}{N} X_l H_l \left(\frac{\sin \pi(l-\varepsilon-k)}{\sin \pi\left(\frac{l-\varepsilon-k}{N}\right)}\right) e^{j\pi(N-1)\left(\frac{l-\varepsilon-k}{N}\right)} \\ I'_k &\approx \sum_{l=0}^{N-1} X_l H_l \left(\frac{\sin \pi(l-\varepsilon-k)}{\pi(l-\varepsilon-k)}\right) \\ W'_k &= \sum_{n=0}^{N-1} w_n^* e^{-2\pi kn/N} \end{aligned} \tag{4}$$

Out put of receiver after using conjugate cancellation scheme is:

$$Y''_k = (Y_k + Y'_k)/2 \tag{5}$$

By putting the values in Equation (5) from Equations (1) and (3) we will get three terms Ist term is desired signal at out of the receiver is:

$$C(k) = \left[(X_k H_k)\left(\sin \pi\varepsilon\right)/\pi\varepsilon\right]$$

Second term is ICI component at output of the receiver after conjugate cancellation scheme will be:

$$I''_k = (I_k + I'_k)/2$$

Third term is Additive White Gaussian Noise at the OFDM receiver out put is:

$$W''_k = (W_k + W'_k)/2$$

In order to evaluate the statistical properties of ICI after conjugate cancellation assume $E(I''_k) = 0$ and assuming average channel gain $E\left[|H_l|^2\right] = |H|^2$ is constant and $E\left[|X_l|^2\right] = |X|^2$

$$\sigma^2_{IC} = E\left[\left|I''_k\right|^2\right] = |X|^2|H|^2(\sin\pi\varepsilon\varepsilon)^2 \times .2195$$

Bit error rate of QPSK modulated OFDM system is given in [10]:

$$\text{BER} = 1/2 * Q\left(\sqrt{E_s/N_0}\right)$$

Bit error rate of QPSK OFDM system after conjugate inter carrier interference cancellation is given by:

$$\text{BER} \leq 1/2 * Q\sqrt{|X|^2|H|^2\left\{\frac{(\sin\pi\varepsilon)}{\pi\varepsilon}\right\}^2/(N_0 + |X|^2|H|^2(\sin\pi\varepsilon\,\varepsilon)^2 \times .2195)}$$

$$= 1/2 * Q\sqrt{\frac{|X|^2|H|^2}{N_0}\left\{\frac{(\sin\pi\varepsilon)}{\pi\varepsilon}\right\}^2 / \left(1 + \frac{|X|^2|H|^2}{N_0}(\sin\pi\varepsilon\,\varepsilon)^2 \times .2195\right)}$$

$$\text{BER} = 1/2 * Q\sqrt{\frac{E_b}{N_0}\left\{\frac{(\sin\pi\varepsilon)}{\pi\varepsilon}\right\}^2 / \left(1 + \frac{E_b}{N_0}(\sin\pi\varepsilon \times \varepsilon)^2 \times .2195\right)}$$

4. RESULTS AND DISCUSSION

For simulation modulation is QPSK $N = 64$, guard interval = 7, Fig. 2 shows compression of BER between self-cancellation method and conjugate cancellation method for frequency offset 0.1 and 0.2. In this frequency offset 0.2 by using self cancellation method BER greater than 10^{-1} at SNR = 0 dB and BER less than 10^{-4} at SNR = 10 dB and for proposed method BER is 10^{-1} for 0 dB SNR and BER is just less than 10^{-5} at 10 dB SNR. But for normalized frequency offset 0.1

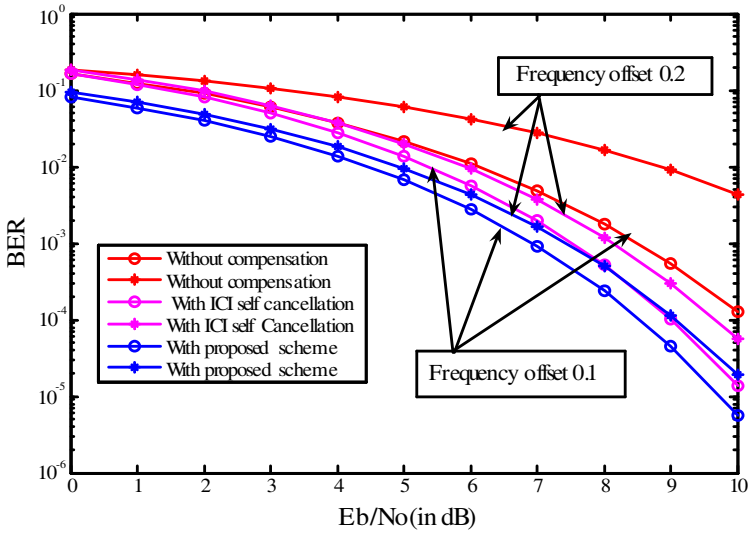


Figure 2: Comparison between proposed conjugate ICI cancellation schemes with self-cancellation scheme.

for self cancellation method BER at 0 dB SNR is greater than 10^{-1} and BER is less than 10^{-5} at 10 dB SNR but for conjugate cancellation method BER 10^{-1} at 0 dB SNR and greater than 10^{-5} at 10 dB SNR. So proposed conjugate method is better than the ICI self-cancellation method as discussed in [2].

Figure 3 shows compression of carrier to interference ratio among different methods like [2], standard OFDM system and proposed scheme for normalized frequency offset 0.25 and our result is comparable with [11]. Carrier to interference ratio (CIR) of proposed scheme is close to that of the conventional OFDM systems. It means that the effect of ICI distortion in the proposed scheme is close to the one in the conventional OFDM systems.

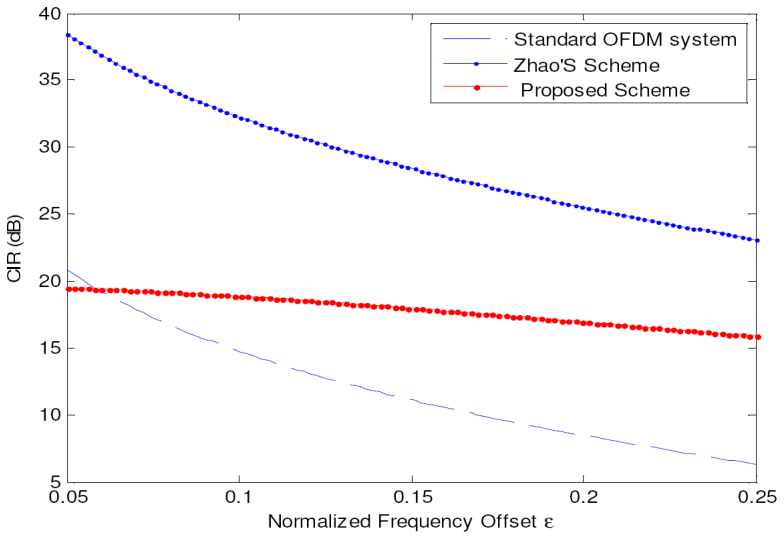


Figure 3: Carrier to interference ratio (CIR) versus normalized frequency offset.

5. CONCLUSION

In this paper, we have suggested a simple ICI cancellation scheme to reduce the frequency offset sensitivity of the OFDM system which is based upon conjugate cancellation scheme. In this scheme two sequences are transmitted in each data symbol. First sequence is original received sequence and another sequence is conjugate of the original sequence. Thus the transmitted two sequences are conjugate of each other rather than adjacent sub-carriers with opposite polarities in order to cancel ICI. It gives better bit error rate than [2], [4] and [11].

REFERENCES

1. Kang, H., Y. Kim, W. Hwang, and K. Kim, "BER performance of OFDM system with the polynomial self cancellation scheme for OFDM mobile communication systems," *Proc. IEEE APCC/OECC*, Vol. 1, 701–704, Beijing, China, Oct. 1999.
2. Zhao, Y. and S. G. Haggman, "Inter carrier interference self cancellation scheme for OFDM mobile communication systems," *IEEE Trans. Communication*, Vol. 49, No. 7, 1185–91, 2001.
3. Sathananantan, K. and R. M. A. P. Rajatheva, "Analysis of OFDM in presence of frequency offset and method to reduce performance degradation," *Proc. IEEE Globcom.*, Vol. 1, 72–76, San Fransisco, CA, Nov. 2000.
4. Fu, Y. and C. C. Ko, "Theoretical BER analysis of OFDM system with ICI self cancellation," *5th Int. Symposium on Wireless Personal Multimedia Communication*, Vol. 3, 991–994, Oct. 2002.
5. Armstrong, J., "Analysis of new and existing methods of reducing inter carrier interference due to carrier frequency offset in OFDM," *IEEE Trans. Communication*, Vol. 47, 365–369, Mar. 1999.
6. Dwivedi, V. K. and G. Singh, "Inter-carrier interference cancellation in OFDM systems," *Proc. National Conference on Wireless and Optical Communication (WOC-2007)*, 245–248, India.
7. Yeh, H. G. and Y.-K. Chang, "A conjugate operation for mitigating intercarrier interference of OFDM systems," *Vehicular Technology Conference*, Vol. 6, 3965–3969, Sept. 26–29, 2004.

8. Chang, K., Y. Han, J. Ha, and Y. Kim, "Cancellation of ICI by Doppler effect in OFDM systems," *Vehicular Technology Conference (VTC 2006-Spring)*, Vol. 3, 7–10, May 2006.
9. Chang, K., Y. Han, J. Ha, and Y. Kim, "Cancellation of ICI by Doppler effect in OFDM systems," *Vehicular Technology Conference (VTC 2006-Spring)*, Vol. 3, 1411–1415, May 2006.
10. Proakis, J. G., *Digital Communications*, Fourth Edition, McGraw-Hill Series, 2001.
11. Peng, Y.-H., Y.-C. Kuo, G.-R. Lee, and J.-H. Wen, "Performance analysis of a new ICI-self-cancellation-scheme in OFDM systems," *IEEE Trans. Consumer Electronics*, Vol. 53, 1333–1338, Nov. 2007.

A Circularly Polarized Microstrip Ferrite Phase-shifter with Uneven Excitation

Sheikh Sharif I. Mitu and M. M. Dawoud

EE Department, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia

Abstract— A low-cost and externally controlled planar microstrip ferrite phase shifter is designed, where the differential phase shift is considerably improved by introducing circularly polarized electromagnetic waves in the structure. Simulated phase response of the designed phase shifter is presented to demonstrate the improvement in the differential phase shift compared to an equivalent traditional microstrip ferrite phase shifter.

1. INTORDUCTION

Wireless communications depend on phased array antennas, where a large number of phase shifters are used to electronically steer the antenna beams or nulls in the desired direction [1]. Thus, the cost, size and integration method of the phase shifters are important factors in the design process of the array antenna. Driven and latched ferrite phase shifters are widely used in array antennas, where external biasing field is used to control the gyrotropic property of the ferrite material, which in turn controls the phase velocity and, hence, the insertion phase of the propagating electromagnetic signal. Recently, microstrip ferrite phase shifters received renewed interested due to its low cost and compact size, although they are require large magnetizing field [2]. In this paper, a novel planar phase shifter is designed, where circularly polarized (CP) electromagnetic (EM) waves are excited to strongly interact with the ferrite substrate and yield required differential phase shift without needing a large external magnetizing field. Although various other techniques are available in the literature to excite CP waves in a planar structure, they often involve complicated geometry and higher insertion loss [3, 4]

2. METHOD OF ANALYSIS

The proposed microstrip phase shifter, shown in Figure 1(a), consisted of a three way uneven power divider, three microstrip lines and reciprocal power combiner. The purpose of the power divider was to excite the microstrip lines, L_1, L_2 and L_3 with EM signal having phases of 90°, 0° and 270° and strength of 35%, 30% and 35% of the inputted signal. This particular signal distribution is needed to generate two set of EM waves, which when combined generated a circularly polarized field distribution under the center microstrip conductor. Since equal three ways power division may result in an elliptically polarized wave with stronger horizontal component of the electrical field compared to its vertical component, as shown in Figure 1(b), an uneven power divider is needed to ensure the generation of circularly polarized waves. The Ferrite with saturation magnetization (M_S) of 800 Guess, relative permittivity (ε_r) of 14 and ΔH of 10 Oe is selected as the substrate material and biased (using H_0) in the direction of propagation to tune the phase response.

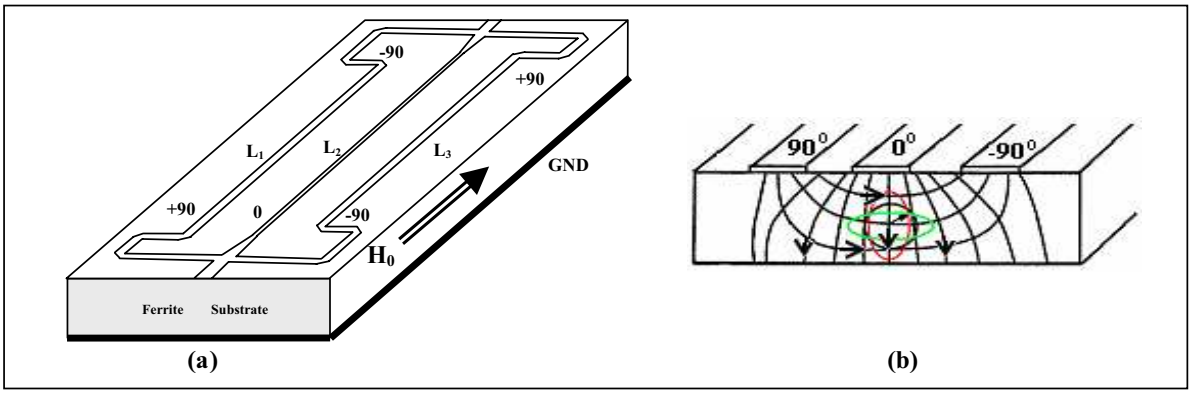


Figure 1: (a) Externally magnetized microstrip ferrite phase shifter, (b) Excited circularly polarized wave under the central conductor.

3. RESULTS

Professional software is used to simulate the designed phase shifter to observe its improved phase response compared to a linearly polarized single line microstrip ferrite phase shifter. Figure 2 shows the frequency responses (S_{11} and S_{21}) of the circularly polarized phase shifter at $H_0 = 70\,\mathrm{KA/m}$.

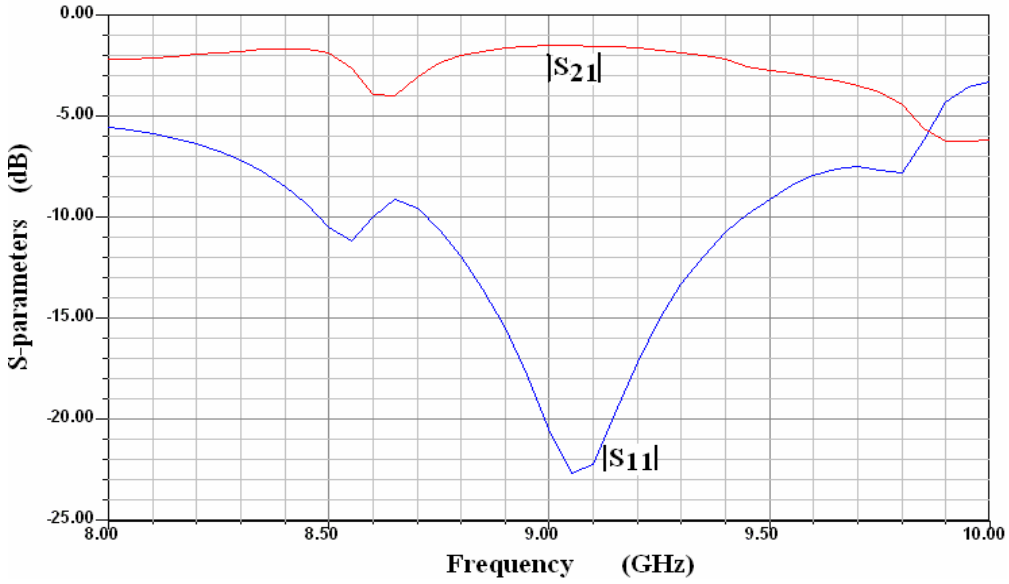


Figure 2: Frequency response (magnitude) of the designed CP microstrip ferrite phase shifter, externally biased at $H_0 = 70\,\mathrm{KA/m}$.

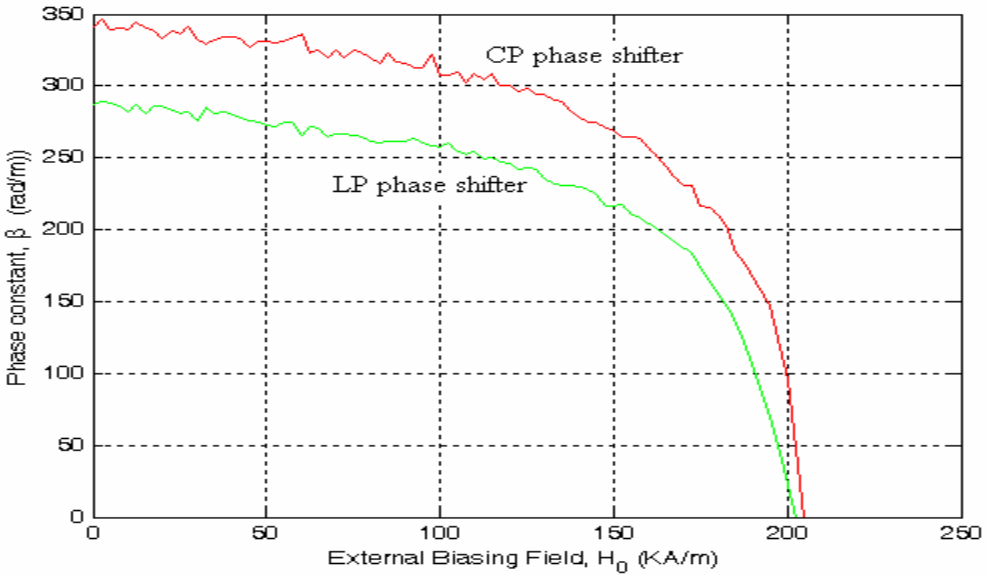


Figure 3: Comparison of phase response with changing biasing field (H_0), for the designed CP and traditional LP phase shifters.

4. CONCLUSIONS

Due to strong interaction between the propagating circularly polarized electromagnetic signal and the magnetized ferrite substrate, an increase in differential phase shift is observed for a circularly polarized microstrip ferrite phase shifter compared to its traditional linearly polarized counterpart. For the designed phase shifter, approximately 60° increase in the differential phase shift is demonstrated.

ACKNOWLEDGMENT

The authors are thankful for the support of King Fahd University of Petroleum & Minerals (KFUPM), Dhahran, Saudi Arabia for funding the research project FT-060029.

REFERENCES

1. Balanis, C. A., *Antenna Theory Analysis and Design*, 2nd ed., Wiley, USA, 1997.
2. Batchelor, J. C., et al., "Scanned microstrip arrays using simple integrated ferrite phase shifters," *IEE Proc. Microwave Antennas Propagation*, Vol. 147, No. 3, 237–241, June 2000.
3. Oates, D. E., et al., "Superconductor Ferrite phase shifter and circulators," *IEEE Trans. on Applied Superconductivity*, Vol. 7, No. 2, 2347–2350, June 1997.
4. Sorensen, R. K., et al., "Low cost nonplanar microstrip-line ferrite phase shifter utilizing circular polarization," *IEEE MWCL*, Vol. 14, No. 1, 25–27, Jan. 2004.

The Measurement of Angle-of-arrival of Microwave in a Task of Precision Landing of Aircraft

I. B. Shirokov, **A. Ponyatenko**, **and O. Kulish**
Sevastopol National Technical University, Ukraine

Abstract— A new method of aircraft's positioning based on measurement of angle-of-arrival of microwave is presented. Interferometer is placed onboard of the aircraft; several transponders are placed on the runway. Several receiving units of interferometer are not set hard and the phase stable feeder is not used. In each of interferometer's unit there is placed its own microwave oscillator. All of microwave oscillators are not synchronized.

1. INTRODUCTION

Methods of positioning of the aircraft with the amplitude methods on large distances are known. This is usual radar task and it is solved and it is implemented almost in all of airports. However accuracy of positioning is not high in this case. It serves for far homing of aircrafts and does not let to make the landing.

The measurements of angle-of-arrival of microwave give us the good opportunity to determine the mutual positioning of aircraft and runway if the transmitting and receiving microwave parts of measuring device are placed onboard of aircraft and runway respectively. The resolution of these measurements is very high and one depends on wave length/physical length ratio, so the aircraft's positioning can be carrying out with high accuracy with phase measurements on microwave.

Facility of measurement angle-of-arrival of microwaves are created out on the basis of interferometer method, where the heterodyne signal from an oscillator feeds all receiving cells which have inclusive at least a receiving antenna and the first mixer. The mixer's output signals are amplified up to limitation and then the phase difference of these signals is measured.

The angle-of-arrival of microwave can be calculated as

$$\alpha = \arcsin(\Delta\varphi/kb), \tag{1}$$

where

$\Delta\varphi$ — measured phase difference between signals from receiving cells,
k — propagation constant,
b — interferometer base.

So, making the measurements of phase difference, we can make the positioning the aircraft's axis with respect to the antenna of runway with high accuracy. In this case it is required the rather high phase stability of the feeder of heterodyne's signal. Any way, in a present case, the receiving antennas location is hard set and it is not possible to change the interferometer base during the measurement process. At the same time the antennas placement onboard of aircraft is not convenient in a most cases. The demand to the phase stability of feeders involve in complicated design solution.

2. APPROACH TO A PROBLEM

In present paper the method of automatic overlapping of a longitudinal axis of the aircraft with an axis of a runway is offered. This method is based on the definition of the relative bearing of the aircraft with respect to the axis of runway. So, if the relative bearing is known, there is a possibility to adjust the longitudinal axis of the aircraft with an axis of a runway with the steering gear of the aircraft. The present method is actual, since it allows carrying out this process automatically, with high accuracy.

The method, described in a paper, is based on homodyne method of determining of angle-of-arrival of microwave. The interferometer is placed onboard of aircraft. It is possible to abandon the principle of the feeding of interferometer's receiving units by the in-phase signals with microwave feeders with high phase stability [1–3]. The interferometer unit is represented as a several receiver-transmitter devices operating with own antennas. Here the transmitter signal is the heterodyne

signal for the receiver. Such approach represents the homodyne method of measurements. The transponders, which are located on runway, receive signals from these transmitters by theirs antennas and shift the frequencies of signals by the certain frequencies Ω_i and then they radiate these signals back in the direction of the aircraft.

The block diagram of measurement which presents mentioned approach is shown on Fig. 1.

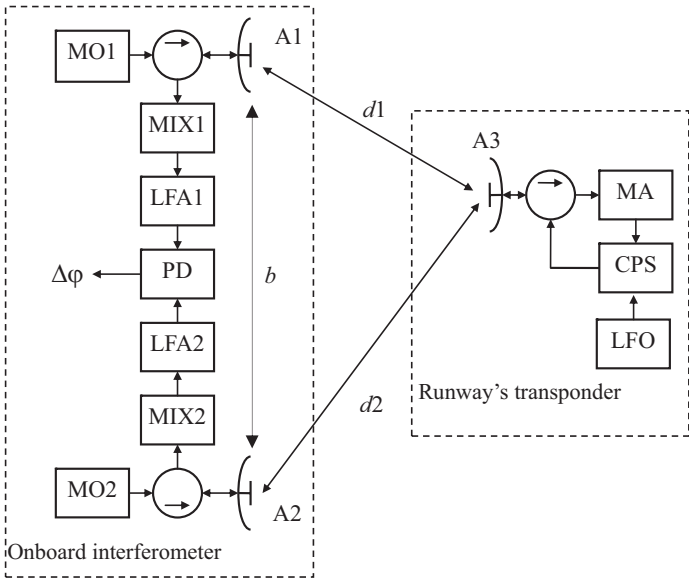


Figure 1: The block diagram of measurement of angle-of-arrival.

The equipment consists of microwave oscillators (MO1, MO2), mixers (MIX1, MIX2), low frequency amplifier-limiters (LFA1, LFA2), phase detector (PD), antennas (A1, A2, A3), Y-circulators, microwave amplifier (MA), controlled phase shifter (CPS) and low frequency oscillator (LFO).

All of microwave oscillators are characterized by the closely spaced but different frequencies and ones are not synchronized.

The microwave oscillations, which are generated with the oscillators MO1 and MO2, are radiated with antennas A1 and A2 in direction of antenna A3. These signals are presented as

$$\begin{aligned} u_1(t) &= u_{01} \cdot \sin(\omega_1 t + \varphi_{01}), \\ u_2(t) &= u_{02} \cdot \sin(\omega_2 t + \varphi_{02}), \end{aligned} \tag{2}$$

where

u_{01}, u_{02} — amplitudes;
ω_1, ω_2 — frequencies;
φ_{01}, φ_{02} initial phases of microwave oscillations.

The transponder, which is located on the runway, receives the signals from these transmitters by its antenna A3. Received signals are

$$\begin{aligned} u_{11}(t) &= u_{01} \cdot L \cdot \sin(\omega_1 t + \varphi_{01} + k_1 d_1), \\ u_{21}(t) &= u_{02} \cdot L \cdot \sin(\omega_2 t + \varphi_{02} + k_2 d_2), \end{aligned} \tag{3}$$

where

L — free space and atmospheric attenuation (taking into consideration antenna's gain);
d_1, d_2 — distances;
k_1, k_2 — propagation constants for each microwave oscillations.

Then received signals are amplified with the microwave amplifier MA. The gain of amplifier is chosen taking into account the Y-circulator's decoupling and the antenna's mismatch. The amplifying is necessary for satisfying of weak energy of microwave link. On short distances the amplifier can be excluded.

The retransmit unit shifts all of received signals by the same frequency and then radiates these signals back in the direction of interferometer antennas. The shift of microwave oscillations can be realized simply by the controlled phase shifter which is set in the transponder [4]. The controlled phase shifter creates a monotone changing of the phase shift in the microwave oscillations. The change of the microwave oscillations phase over the period T of the low frequency oscillations by 2π is tantamount to the frequency shift of microwave oscillations by the frequency $\Omega = 2\pi/T$ and phase shift, determined with the initial phase of the low frequency oscillations [4, 5]. The low frequency oscillations are generated with LFO. In a certain assumption this technical solution is equivalent to Doppler frequency shifting. The magnitude of frequency shift is chosen small. Really, F is equal 1 kHz or closely ($F = \Omega/2\pi$).

In the interferometer the frequency transformed oscillations are received by antennas of each receiving unit. The received signals are

$$\begin{aligned} u_{111}(t) &= u_{01} \cdot L^2 \cdot \sin\left[(\omega_1 + \Omega)\, t + \varphi_{01} + k_1 d_1 + k_1^1 d_1 + \varphi_{03}\right], \\ u_{211}(t) &= u_{02} \cdot L^2 \cdot \sin\left[(\omega_2 + \Omega)\, t + \varphi_{02} + k_2 d_2 + k_2^1 d_2 + \varphi_{03}\right], \end{aligned} \tag{4}$$

where

k_1^1, k_2^1 — propagation constants (taking into consideration the frequency shift Ω);
φ_{03} — initial phase of low frequency oscillations.

As far as the frequency shift Ω is much more lower than ω_1 and ω_2, we can assume that $k_1^1 \approx k_1$, $k_2^1 \approx k_2$ and Equations (4) become

$$\begin{aligned} u_{111}(t) &\approx u_{01} \cdot L^2 \cdot \sin\left[(\omega_1 + \Omega)\, t + \varphi_{01} + 2k_1 d_1 + \varphi_{03}\right], \\ u_{211}(t) &\approx u_{02} \cdot L^2 \cdot \sin\left[(\omega_2 + \Omega)\, t + \varphi_{02} + 2k_2 d_2 + \varphi_{03}\right], \end{aligned} \tag{5}$$

As well as the transmitter signal of each of units is the heterodyne signal for the receiver one, in the output of each of mixers there are formed oscillations of the same frequency Ω which is equal to the difference between the initial oscillations and the frequency transformed oscillations and it is characterizing by the initial phase which is determined by the angle-of-arrival of microwaves or difference between distances d_1 and d_2. The low frequency amplifier-limiters pick up these difference signals. So, the signals on the output of amplifiers will be

$$\begin{aligned} u_{112}(t) &= u_0 \sin\left(\Omega t + 2k_1 d_1 + \varphi_{03}\right), \\ u_{212}(t) &= u_0 \sin\left(\Omega t + 2k_2 d_2 + \varphi_{03}\right), \end{aligned} \tag{6}$$

where u_0 — amplitude of limited signals.

From (6) and it is shown that the frequencies ω_1, ω_2 and the initial phases φ_{01}, φ_{02} of initial microwave oscillations are missing. The only frequency Ω and initial phase φ_{03} of low frequency oscillation is present in both equations. In each equation there is term $2k_i d_i$ which represents the twice phase progression of microwave from antenna of each unit of interferometer to antenna of transponder, which is placed on runway.

The signals are delivered from all units through low-frequency feeder with arbitrary phase stability to the processing device, where the phase difference is measured. The output phase detector signal is

$$u_{pd} = 2 \cdot P \cdot (k_1 d_1 - k_2 d_2), \tag{7}$$

where P — phase detector constant.

Taking into account that the frequencies ω_1 and ω_2 are chosen close (the difference must be at least some greater than the frequency shift Ω), it can assume that $k_1 \approx k_2 \approx k$ and Equation (12) becomes

$$u_{pd} = P \cdot 2k\,(d_1 - d_2). \tag{8}$$

Taking into account the amount of $f_i = \omega_i/2\pi$ is, for example, 10 GHz but frequency shift is 1 kHz and difference of frequencies f_1 and f_2 is 10 kHz we can affirm the mentioned above assumption is true. In any case the difference between k_1 and k_2 is constant and one can be eliminated with the calibration procedure.

So, the presence of difference of distances d_1 and d_2 results in forming of the signal of phase detector with rating coefficient k, and phase difference of the low frequency oscillations, which is measured, is characterized by the angle-of-arrival of microwaves.

Further, these measurements with calibration procedure will be more correctly if the number of receiving cells will be more than two. The quantity of the units can be chosen arbitrary and one is characterized by the measurement conditions. In addition, their mutual placing is not hard and can be changed even during the measurement process, for example, for increasing of the interferometer base at the analysis of the small angle-of-arrival of microwaves.

Furthermore, it can transmit the initial phase of the low frequency oscillations from the first object to another with low frequency channel as described in [3]. In this case the phase difference will be measured as $2kd_i$ in each high frequency channel.

It can place several interferometer units on different axis in the same plane, or different planes [6]. Each pair of antennas represents the phase difference of receiving signals. So, in processing device there is the phase difference measuring data file, which adequately represents the mutual position of the interferometer antennas and the transponder's antenna.

If the resolution of low frequency phase meter is high (real one degree or less) and phase difference kb has large magnitude (thousands of degree), where k is the microwave propagation constant and b — interferometer base, the mutual position of objects can be calculated with very high accuracy. The interferometer antennas can be placed in arbitrary convenient points of aircraft and there is no problem to feed these units with phase stability feeders.

3. THE OVERLAPPING OF AXIS OF AIRCRAFT AND RUNWAY

The described above scheme represents the positioning of interferometer base (or aircraft axis) in relation to transponder antenna. That is true because the interferometer is placed on the aircraft. Such approach lets us to carry out the homing of aircraft. The runway is assumed as point in this case.

The certain interest represents the landing of aircraft up to its full stopping. In this case there must be carried out the overlapping of axis of runway with the axis of aircraft. Obviously, we need in two points of runway at least, or two transponders on runway. From the other hand, the patterns of transponder antennas assume the presence of the minimum in a plane of runway, or the maximum in a direction of aircraft arriving. If the aircraft will pass the antenna of first transponder during the landing, the signal from this transponder will decreased below the needed level.

In this paper it is proposed to place several transponders on a runway. The mutual placing of transponders is known. Each of transponders forms it's own frequency shift Ω_i, and in an interferometer of aircraft we can select each low frequency (Ω_i) signal, then we can measure several phase differences and then we can determine, thus, the positioning of runway in relation to the aircraft [7, 8].

The scheme of placing of transponders on runway is presented on Fig. 2, where the desirable antennas patterns are shown.

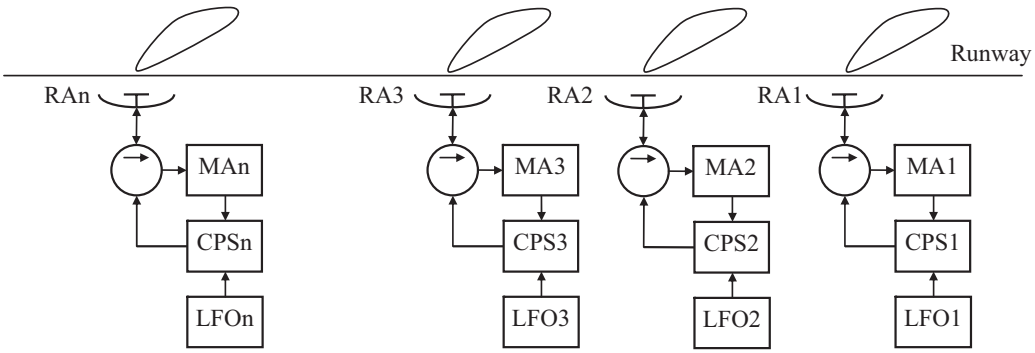


Figure 2: The scheme of placement of transponders on runway.

The scheme of interferometer of aircraft is shown on Fig. 3.

The equipment of each transponder consists of mentioned above components. The transponders distinguish one from another the only with the magnitude of low frequency oscillations.

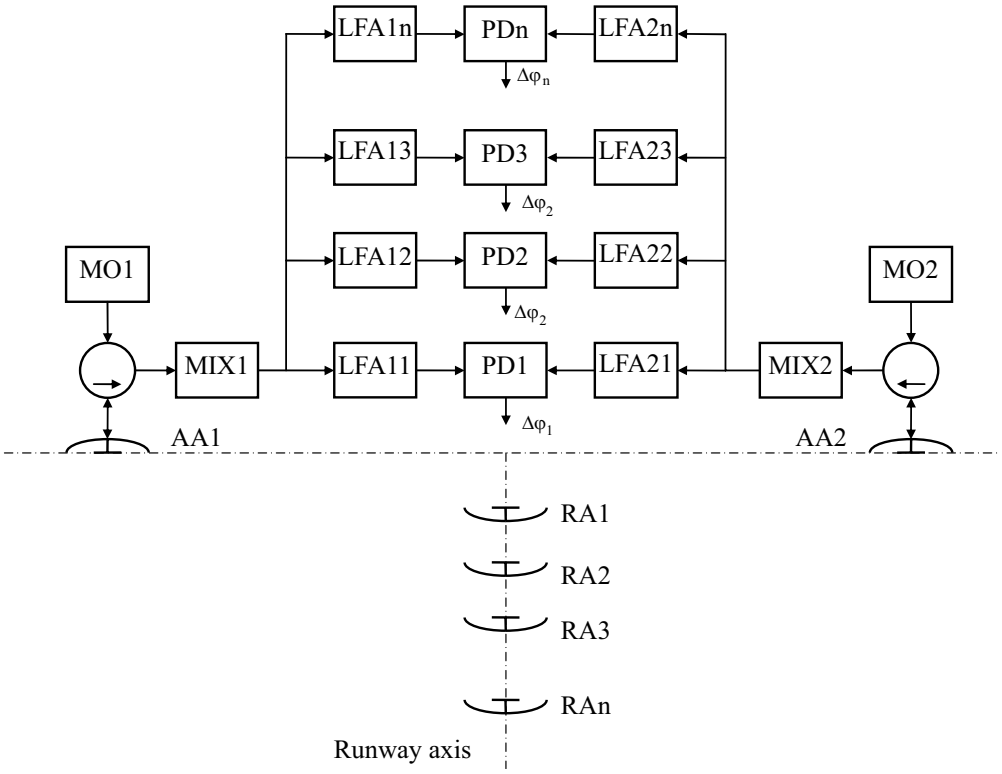


Figure 3: The scheme of interferometer of aircraft.

In a scheme of interferometer there were added low frequency amplifier-limiters (LFAji) and phase detectors (PDi).

The interferometer of aircraft radiates the same signals, as described in (2).

Each transponder receives both signals. The signals received with first transponder are the same as in (3), however the signals received with second transponder will be

$$\begin{aligned} u_{12}(t) &= u_{01} \cdot L \cdot \sin(\omega_1 t + \varphi_{01} + k_1 d_3), \\ u_{22}(t) &= u_{02} \cdot L \cdot \sin(\omega_2 t + \varphi_{02} + k_2 d_4), \end{aligned} \tag{9}$$

where d_3, d_4 — distances from antennas of interferometer to antenna of second transponder.

Similar, we can write down the signals received with other transponders. Further, we will consider signals from only two transponders.

Each transponder shifts the frequencies of all of received signals by the own frequency and then radiates these signals back in the direction of interferometer antennas. The generated with low-frequency oscillator LFO2 frequency Ω_2 has the same order that frequency Ω_1 of oscillator LFO1, but they some differs. For example, frequency F_1 is equal to 1 kHz, but F_2 is equal to 1.3 kHz.

The received with antennas of interferometer signals will be

$$\begin{aligned} u_{111}(t) &\approx u_{01} \cdot L^2 \cdot \sin[(\omega_1 + \Omega_1)t + \varphi_{01} + 2k_1 d_1 + \varphi_{03}], \\ u_{211}(t) &\approx u_{02} \cdot L^2 \cdot \sin[(\omega_2 + \Omega_1)t + \varphi_{02} + 2k_2 d_2 + \varphi_{03}], \\ u_{121}(t) &\approx u_{01} \cdot L^2 \cdot \sin[(\omega_1 + \Omega_2)t + \varphi_{01} + 2k_1 d_3 + \varphi_{04}], \\ u_{221}(t) &\approx u_{02} \cdot L^2 \cdot \sin[(\omega_2 + \Omega_2)t + \varphi_{02} + 2k_2 d_4 + \varphi_{04}], \end{aligned} \tag{10}$$

where φ_{04} — initial phase of the second low-frequency oscillations.

After mixing and limiting these signals will be

$$\begin{aligned} u_{112}(t) &= u_0 \sin(\Omega_1 t + 2k_1 d_1 + \varphi_{03}), \\ u_{212}(t) &= u_0 \sin(\Omega_1 t + 2k_2 d_2 + \varphi_{03}), \\ u_{122}(t) &= u_0 \sin(\Omega_2 t + 2k_1 d_3 + \varphi_{04}), \\ u_{222}(t) &= u_0 \sin(\Omega_2 t + 2k_2 d_4 + \varphi_{04}), \end{aligned} \tag{11}$$

By the measuring of phase difference of pairs of signals with frequency Ω_1 (one pair) and Ω_2 (another pair), and taking into account assumption $k_1 \approx k_2 \approx k$ we will obtain two signals of phase difference

$$\begin{aligned} u_{pd1} &= P \cdot 2k\,(d_1 - d_2)\,, \\ u_{pd2} &= P \cdot 2k\,(d_3 - d_4)\,, \end{aligned} \tag{12}$$

Similar, we will obtain other signals from other channels of interferometer.

If the placing of antennas of transponders is known, we can determine not only the angle-of-arrival of microwave from single antenna, but we can determine angle-of-arrival of microwave from another antenna. So, we can locate the axis of runway.

Furthermore, we can place several interferometers onboard of aircraft, so we can process more signals and we can obtain the more detail positioning of aircraft.

The interferometer units and transponders we can place in every convenient place of each object because all of these units are not connected with phase stable feeders. The frequency in all of feeders is low. These units we can place in a different planes and the quantity of these units can be chosen arbitrary. These units it can place on moving parts of aircraft. Certainly, the track of moving of these parts must be known.

If the placing of all of antennas is known in each point of time, we can control the mutual positioning of the aircraft (and (or) its moving parts) and the runway.

The only we must to do to process the output signals of phase detectors, taking into account all of mentioned above antenna placements and then we must to control the steering gear of the aircraft. Here by the operating of the position of the aircraft, the signals on the outputs of all of phase detectors are supported on the zero level, providing with that the automatic overlapping of a longitudinal axis of the aircraft with the axis of the runway.

4. CONCLUSIONS

So, in this paper it was shown there is a good opportunity to carry out mutual positioning of aircraft and runway with measurements of angle-of-arrival of microwaves. Taking into account new approach to these measurements, based on absolute independence of receiving units of interferometer, it can place each interferometer antenna and each transponder antenna in any convenient point of aircraft and runway. There are no problems to feed all of interferometer units with phase stable feeder in this case, as far as we use homodyne method where the transmitting signal is the heterodyne one for interferometer unit.

As well as phase measurements it can carry out with high resolution, we can realize the mutual positioning of aircraft and runway with high accuracy.

First aspect of increasing of efficiency from the use of the prospective method is related with the appearance of an opportunity to automate the process of homing of the aircraft at the landing beginning.

Second aspect is related with the opportunity to automate not only the process of homing, but also to automate all process of landing. Here antennas of transponders of a runway can be made not prominent, but for example slot, and they can be situated as on all length of a runway as beyond it along its axis. In this case the antennas don't disturb the process of landing itself of the aircraft on a runway; however there is an opportunity to control the position of the aircraft at its movement along the entire runway up to its complete stop.

And the last, no one transponders do not emit any microwave itself. All of them only retransmit the received signals from aircraft on homing and landing with certain amplifying. By the choosing of enough dynamic range of the microwave amplifier, these transponders can retransmit signals from several aircraft simultaneously, e.g. from one aircraft on landing, others on homing. There are no problems of electromagnetic compatibility in this case. The disturbing signals from another aircraft, or from another interferometer of single aircraft, are not synchronized and they can be eliminated by the correlation method of signal processing.

REFERENCES

1. Shirokov, I. B., L. M. Lobkova, and V. M. Iskiv, "Method of measurements of angle-of-arrival," *.C. (USSR)*, #17181491, pub. in BI #9 from 07.03.1992, G01R 29/08.
2. Shirokov, I. B., "New Approach to equipment for angle-of-arrival measurements on microwave," *IGARSS'02, Conf. Proc.*, Vol. V, 2626–2627, Toronto, Canada, June 24–28, 2002.

3. Shirokov, I. B. and S. Shaban, "Method of measurement of fluctuations of phase progression and angle-of-arrival of microwave," *Pat.*, Ukraine, # 58814, pub. in B. #8 from 15.08.2003 G01R29/08.
4. Jaffe, J. S. and R. C. Mackey, "Microwave frequency translator," *IEEE Trans. MTT*, Vol. 13, 371–378, 1965.
5. Shirokov, I. B., "The device for measurements of amplitude and phase difference," *A. C. (USSR)*, #1486942, pub. in BI #22 from 15.06.1989, 0113/04.
6. Shirokov, I. B., "Object positioning based on new approach to microwave angle-of-arrival measurements," *WMSCI'06, Conf. Proc.*, Vol. IV, 257–259, Orlando, Florida, USA, July 16–19, 2006.
7. Shirokov, I. B., "The Method of automatic adjustment of the longitudinal axis of aircraft with the axis of runaway," *Pat.*, Ukraine, # 79787, pub. in B. #11 from 25.07.2007 G08G 5/00.
8. Shirokov, I. B., "The positioning of space objects based on microwave angle-of-arrival measurements," *The 4th ESA Int, Workshop on Tracking, Telemetry and Command Systems for Space Appl.*, Darmstadt, Germany, September 11–14, 2007.

On Line Wire Diagnosis by Modified Spread Spectrum Time Domain Reflectometry

Adrien Lelong[1], **Marc Olivas Carrion**[1], **Virginie Degardin**[2], **and Martine Lienard**[2]

[1]CEA LIST, 91192 Gif-Sur Yvette, France
[2]IEMN/TELICE, Lille, France

Abstract— A method for on line wire diagnosis using reflectometry is presented. It consists in detecting and locating electric faults in wiring networks while the target system is running. Since the system is diagnosed on line, it is necessary to minimize interferences between the target system and the reflectometer. The Modified SSTDR presented here allows a spectral modification of the injected signal in order to adapt it to the application constraints. The influence of this spectral deformation on the reflectometry results is discussed here.

1. INTRODUCTION

Reflectometry consists in sending a wide band test signal down the wire which reflects back at impedance discontinuities such as branches and defects. The measured signal at the injection point is a sum of echoes which can be expressed as a simplified model.

$$y(t) = (s \star h)(t), \quad h(t) = \sum_{k \geq 0} \alpha_k \delta(t - \tau_k) \tag{1}$$

where $\star$ represents convolution product, $\delta(t)$ is the dirac distribution, $s(t)$ is the input signal and (α_k, τ_k) are the amplitude and delay of each echo which characterize the wire. τ_k provides an information about the position of the singularity corresponding to the echo. Coefficients α_k depend on transmission and reflexion coefficients of the wire. The objective of reflectometry is the measurement of $h(t)$ which is the *wire response*. Then, fault location is made by analysing $h(t)$.

The basic test signal in reflectometry is a rectangular pulse. Usually, the time-energy distribution is improved using *pulse compression* technique [1] so the signal is a digital sequence such as an MSequence or equivalent. The autocorrelation function of such a signal is a single peak, so the reflectometry result is provided by the cross-correlation between the received signal and the probing signal. This processing is used in the STDR (Sequence Time Domain Reflectometry) method [2]. From a spectral point of view, STDR is equivalent to standard Time Domain Reflectometry (TDR) [3] which consists in sending a pulse as a test signal. The interest of STDR is to reduce the amplitude of the test signal, Thus, impact of non linear effects (saturation, quantization) are minimized and noise imunity is improved. STDR is also used to generate orthogonal signals for distributed diagnosis [4].

The result of the STDR method is given by the following equation

$$\hat{h}(t) = \gamma_{sy}(t) = \int_0^{T_s} s(\tau) y(\tau + t) d\tau = (\gamma_{ss} \star h)(t) \tag{2}$$

where T_s is the duration of the signal $s(t)$ and $\gamma_{ss}(t)$ its autocorrelation function.

The main problem of on line diagnosis is not to disturb the target system and not to be disturbed by it. This results in various contraints on power spectral density of the test signal. Some of these constraints are also related to electromagnetic compatibility (EMC). Though STDR can be adequate for several applications of on line diagnosis, the spectrum of the injected signal is fixed so it cannot be adapted to applications with spectral constraints.

In order to improve the flexibility of the reflectometry in terms of spectral constraints, the signal can be modulated. That is the principle used in SSTDR (Spread Spectrum Time Domain Reectometry) [2]. The aim of SSTDR is to keep low frequency clean, which is achieved by sending an amplitude modulated pseudo noise. The modulation is made by mixing the digital signal with a high frequency sinusoidal carrier waveform. The received signal is computed in the same way as for STDR. The schematic of Figure 1 shows how SSTDR works. One of the drawbacks of this method is that it only works well when the carrier frequency is equal to the chip rate of the pseudo noise.

Indeed, since the correlation is applied on the modulated signal, the peaks of the reflectometry result are also modulated. As a consequence, the final signal is difficult to analyse when the carrier frequency is high. Furthermore, the correlation operation must be implemented by use of analog components because signal spectra are generally to wide for a cheap digital system.

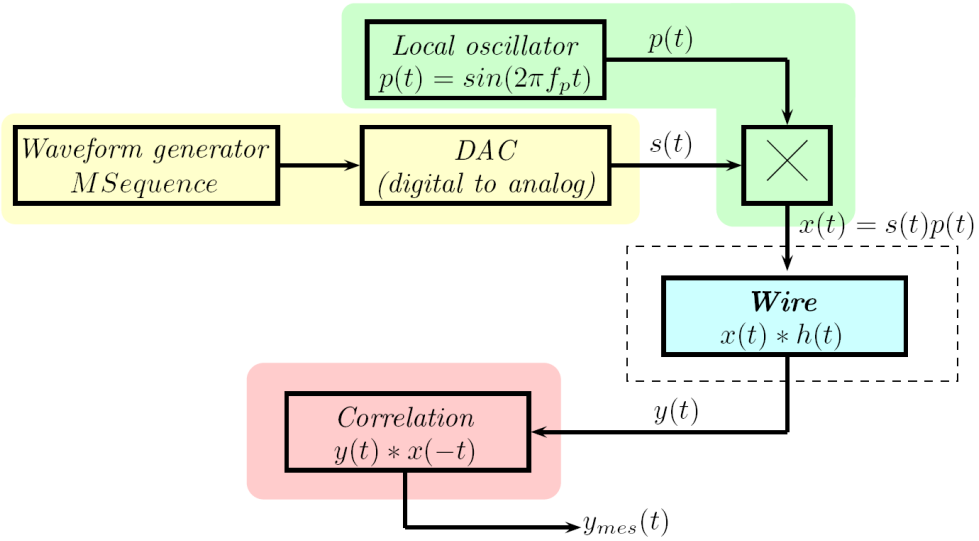


Figure 1: Schematic of the SSTDR method.

2. REFLECTOMETRY USING FREQUENCY TRANSPOSITION

In some applications, it may be necessary to inject higher frequency signals for than classical SSTDR. The method proposed in this paper is an evolution of the SSTDR. In a similar way, the test signal is a modulated M Sequence. But before passing through a matched filter, the received signal is demodulated, so the correlation can be computed in baseband and it can be easily implemented using a numeric system such as an FPGA. Only the modulation demodulation part of the system has to be made using analog components. This is illustrated by the schematic shown on Figure 2.

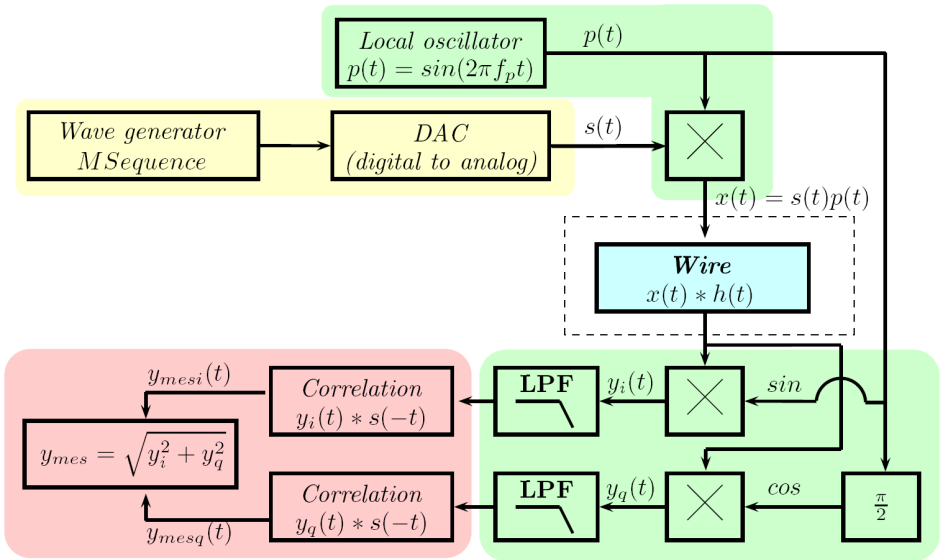


Figure 2: Schematic of the modified SSTDR method.

This process is explained by the following theory. The provided test-signal is given, in the time

domain, by

$$x(t) = s(t)p(t) = s(t)\sin(2\pi f_p t) \tag{3}$$

where f_p is the carrier frequency. Its spectrum is

$$X(f) = S(f) \star P(f) = \frac{1}{2j}\left[S(f - f_p) - S(f + f_p)\right] \tag{4}$$

The received signal, $y(t)$, is then demodulated by respectively mixing with a sine and a cosine function whose frequency is f_p. Both quadrature signals resulting from this modulation are given in the frequency domain by

$$\begin{aligned} Y_i(f) &= \frac{1}{4}\{S(f)[H(f - f_p) + H(f + f_p)] - S(f - 2f_p))H(f - f_p) - S(f + 2f_p))H(f + f_p)\} \\ Y_q(f) &= \frac{1}{4}\{S(f)[H(f + f_p) - H(f - f_p)] + S(f - 2f_p))H(f - f_p) - S(f + 2f_p))H(f + f_p)\} \end{aligned}$$

The high frequency component of these signals is cleared by a low pass filtering (LPF), only the first term of both equations, which correspond to the low frequency component, remains at the filter output. Both signals are then sampled and processed by a matched filter (equivalent to a correlation) so that $S(f)$ is compensated. Assuming that

$$H(f) = \sum_{k\geq 0} \alpha_k e^{j2\pi f \tau_k} \tag{5}$$

this leads, in time domain, to

$$y_{mesi}(t) = \frac{1}{2}\sum_{k\geq 0}^{M} \alpha_k \gamma_{ss}(t - \tau_k)\cos(2\pi f_p \tau_k) \tag{6}$$

$$y_{mesq}(t) = -\frac{1}{2}\sum_{k\geq 0}^{M} \alpha_k \gamma_{ss}(t - \tau_k)\sin(2\pi f_p \tau_k) \tag{7}$$

at the output of matched filters. As it can be seen in Equations 6 and 7, peaks of the baseband signals $y_{mesi}(t)$ and $y_{mesq}(t)$ are modulated by a cosine and sine function. This is because the phase of the carrier corresponding to each echo depends on the delay τ_k so it has to be considered has a random variable. That is why a quadrature demodulation must be done. So the result is given by

$$\begin{aligned} y_{mes} &= \sqrt{y_{mesi}^2 + y_{mesq}^2} \\ &= \left|\frac{1}{2}\sum_{k\geq 0} \alpha_k \gamma_{ss}(t - \tau_k)\right| = \frac{1}{2}\left|(h \star \gamma_{ss})(t)\right| \end{aligned}$$

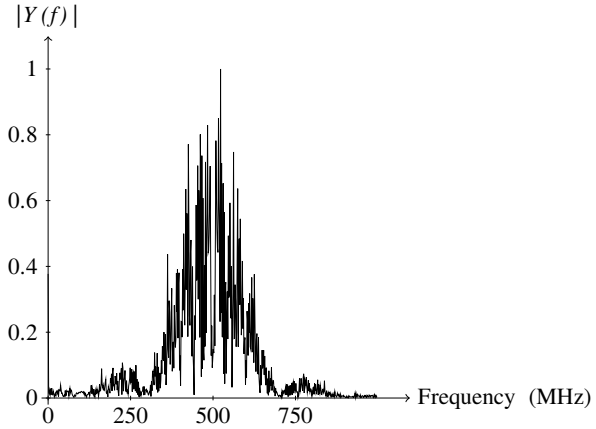


Figure 3: Spectrum of the measured signal.

This result shows that the processing is independent from the carrier frequency. In practice, this frequency is limited by the bandwidth of the wire. Thanks to this, the carrier frequency can be adjusted as desired with no influence on peaks waveform. So it is possible with this method to make a high frequency transposition of the test signal and get a correct reflectometry result. Yet, the sign information is lost because of the quadrature demodulation so the nature of the fault (short or open circuit) cannot be known, but when only the fault location is desired, the provided result is enough.

3. RESULTS OF THE MSSTDR

The Modified SSTDR was simulated using Matlab. It has been tested on a network consisting of 2 meters of coaxial cables connected to 1, 60 meters of twisted pairs. This network is illustrated by the schematic of Figure 4. The signal is injected and measured via laboratory equipment (programmable generator and digital oscilloscope). The analog part of the processing (mixers and analog filters) are also simulated by Matlab but with a high frequency sampling (10 GHz). For this experiment an MSequence with a chip rate of 200 MHz modulated by a 500 MHz sinusoidal carrier was used as a test signal. Using this signal allows to avoid perturbation of the target system below 200 MHz. Figure 3 shows the spectrum of the reflected signal.

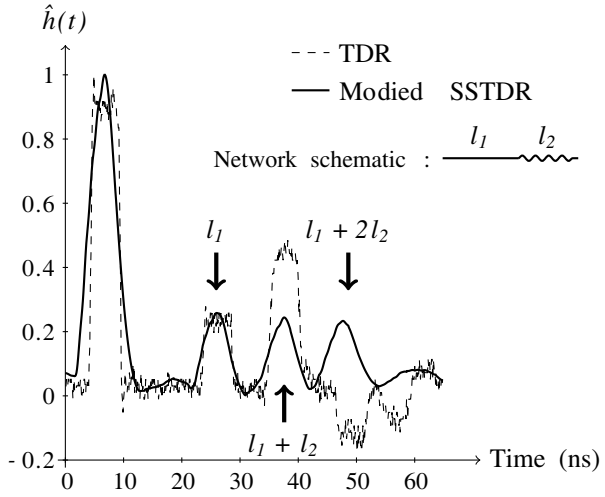


Figure 4: Result provided by a modified SSTDR and a TDR.

The received signal can be sampled at 200 MHz (or higher) after the low pass filters so that the correlation is computed on a digital signal. It is possible to achieve high sampling frequency (for

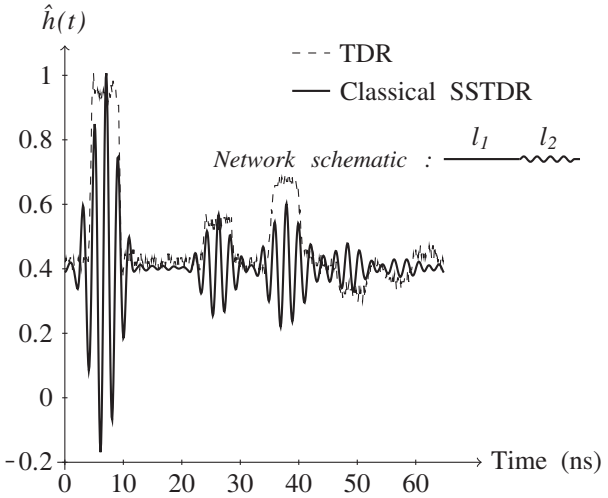


Figure 5: Result provided by a classical SSTDR and a TDR.

example 1 GHz) of the final signal y_{mes} using equivalent time sampling which means that several measurements are made and combined together. Thus, though the final result is sampled at 1 GHz the analog to digital converter of the system is still sampling at 200 MHz. In practice this can be done using vernier time delay generation [5].

Figure 4 shows the results provided by a TDR (Time Domain Reflectometry) and by the Modified SSTDR. As it can be seen on this result, the modified SSTDR can be used to efficiently locate impedance singularities in a wire. Amplitude errors between the TDR and the Modified SSTDR are caused by distorsions of the modulation and the demodulation.

The result provided by a classical SSTDR in same conditions as the previous measurement ($F_e = 200$ MHz, $f_p = 500$MHz) is shown on Figure 5. It can be compared to the modified SSTDR. On the one hand, the peak are modulated so that they are difficult to analyse. On the other hand, all post processing have to be made using analog electronic components.

4. CONCLUSION

The modified SSTDR provides a solution to avoid low frequency perturbation while a reflectometry measurement is being made. Though, the reflectometry result is distorted, the location information is available with this method for on line diagnosis. So the fault can be detected and located in real time. This location information is enough in many case for maintenance of the system which generally the objective of such methods.

REFERENCES

1. Nishiyama, E., C. Nishijima, and K. Kuwanami, "A method for fault detecting on twisted pair cable network by use of m-sequence correlation," *Industrial Electronics Society, 2004. IECON 2004. 30th Annual Conference of IEEE*, Vol. 2, 1929–1934, Nov. 2–6, 2004.
2. Smith, P., C. Furse, and J. Gunther, "Analysis of spread spectrum time domain reflectometry for wire fault location," *Sensors Journal, IEEE*, Vol. 5, No. 6, 1469–1478, Dec. 2005.
3. Carrion, M. O., F. Auzanneau, and N. Ravot, "A simple and accurate model for wire diagnosis using reflectometry," *Progress In Electromagnetics Research Symposium*, Prague, Aug. 2007.
4. Ravot, N., F. Auzanneau, Y. Bonhomme, M. O. Carrion, and F. Bouillault, "Distributed reflectometry-based diagnosis for complex wired networks," *EMC: Safety, Reliability and Security of Communication and Transportation Syst., EMC Workshop*, Paris, Jun. 2007.
5. Lee, D., J. Sung, and J. Park, "A 16ps-resolution random equivalent sampling circuit for tdr utilizing a vernier time delay generation," *Nuclear Science Symposium Conference Record, 2003 IEEE*, Vol. 2, 1219–1223, Oct. 19–25, 2003.

Coupling Theory of Asymmetric Photonic-crystal Waveguides

Chih-Hsien Huang[1], **Wen-Feng Hsieh**[1], **and Szu-Cheng Cheng**[2]

[1]Department of Photonics, National Chiao Tung University, Hsinchu, Taiwan
[2]Department of Physics, Chinese Culture University, Taipei, Taiwan

Abstract— The physical properties of asymmetric photonic-crystal directional couplers are studied by using the tight-binding theory, which considers that the field distributions of photonic-crystal waveguides (PCWs) are localized around the periodic defects. From this model, the analytic formulas which describe the dispersion of the coupler and the eigenmode patterns were derived. By considering coupling up to the third nearest-neighboring between two PCWs, analytic expressions of the dispersion relations of asymmetric photonic crystal coupler are consistent with the results from the plane wave expansion method (PWEM). Due to the broken symmetry, two dispersion curves will never cross. Nevertheless, the eigenmode patterns switch their parity as the symmetric coupler does when wavevector k passes over the decoupling point.

1. INTRODUCTION

Photonic crystal waveguide (PCW) is a structure which introduces considerable quantities of point defects in the photonic crystal (PC). The electromagnetic (EM) wave is strongly confined in the defect channel, allowing low-loss in Y-blanches and a sharp-bending. Furthermore, creating two line defects to form a directional coupler are used as add/drop filters, switches and multiplexers. To design photonic crystal devices, the simulation tools such as the PWEM and the finite difference time domain method (FDTD) have often been used. However, there is no simple analytic formula to analyze physical properties of the coupled photonic crystal waveguides (PCWs), especially to analyze coupling between PCWs [1]. Tight-binding theory (TBT) is widely used not only in condensed matter physics but also recently in the EM wave propagation in linear and nonlinear single line defects PCW analytically. They usually considered only coupling between the nearest-neighboring defects that may be good enough in the single PCW or conventional waveguides. Nevertheless, it is insufficient to only consider the nearest-neighboring defects in the directional coupler made of two identical PCWs [2]. The dispersion curves will always split or never cross if only the nearest-neighboring coupling is considered, but the dispersion do cross that was resolved by further considering coupling of the second nearest neighboring defects which causes a sinusoidal modulation to the dispersion curves.

Similar to the extended TBT [2] which considers three nearby coupling coefficients, in this paper, we firstly derive an analytic solution to describe the dispersion of asymmetric coupled PCWs. This formula provides more generalized discussion and physical insight than what is derived for the case of symmetric coupled PCWs. Especially, the coupled identical PCWs should become asymmetric due to the intensity dependent index of refraction in the nonlinear photonic crystal directional coupler. Secondly, by using the analytic formulas, the phenomena of mode switching and mode patterns in the asymmetric PCWs are extensively analyzed. Finally, the simulation results of considering PCWs of triangular lattices are consistent with our theoretic results.

2. COUPLED EQUATIONS FOR ASYMMSTRIC PHOTONIC CRYSTAL WAVEGUIDES

The asymmetric coupled PCWs in a PC with lattice constant a are formed by varying the rod sizes or refraction indices of two line defects, shown as PCW1 and PCW2 in Fig. 1. To ensure the PCWs are both single-mode waveguides, we simply reduce the radii of dielectric defects or decrease the refractive index of defects. The field distribution of a single defect in each PCW can be written as the product of time-varying and spatial-varying functions, i.e., $\mathbf{E}_{10}(\mathbf{r},\, t) = U_0 \exp(-i\omega_1 t)\mathbf{E}_{10}(\mathbf{r})$ in PCW1 and $\mathbf{E}_{20}(\mathbf{r},\, t) = V_0 \exp(-i\omega_2 t)\mathbf{E}_{20}(\mathbf{r})$ in PCW2, where U_0 and V_0 are the amplitudes of electric fields, and ω_1 and ω_2 are eigenfrequencies of the point defects in each waveguide.

Under the TBT, the propagating electric field of the PCW1 can be expanded in terms of the localized field of the individual point defects as $\mathbf{E}_1(\mathbf{r},\, t) = \sum_n u_n(\mathbf{r},\, t)\mathbf{E}_{1n}(\mathbf{r})$. Here $\mathbf{E}_{1n}(\mathbf{r}) = \mathbf{E}_{10}(\mathbf{r} - na\hat{x})$ is the localized electric field at site n, and $u_n(\mathbf{r},\, t)$ is its time-varying envelop function. Considering the coupling by Lorentz reciprocity relation [3], the coupled equations of the isolated

PCW1 can be written as

$$i\frac{\partial}{\partial t}u_n = (\omega_1 - P_0)u_n - \sum_{m=1} P_m(u_{n+m} + u_{n-m}), \tag{1}$$

where P_m is the coupling coefficient between site n and site $n+m$ and is defined as

$$P_m = \frac{\omega_1 \int dv \Delta\varepsilon(\mathbf{r})\mathbf{E}_{1n} \cdot \mathbf{E}_{1n+m}}{\int dv \left[\mu \left|\mathbf{H}_{1n}\right|^2 + \varepsilon \left|\mathbf{E}_{1n}\right|^2\right]}, \tag{2}$$

where $\Delta\varepsilon(\mathbf{r}) = \varepsilon'(\mathbf{r}) - \varepsilon(\mathbf{r})$ is the difference of perturbed and unperturbed dielectric constant, and P_0 represents for the frequency shift of the eigenmode due to the change of the dielectric constant at site n. Let k_1 and $\bar{\omega}_1$ be the wave vector and its corresponding eigenfrequency of PCW1, respectively. We substitute $u_n(\mathbf{r}, t) = U_0 \exp(ikna - i\bar{\omega}_1 t)$ into Eq. (1) and obtain the dispersion relation of PCW1:

$$\bar{\omega}_1(k) = \omega_1 - P_0 - \sum_{m=1} 2P_m \cos(mka). \tag{3}$$

Similarly, we can also derive the coupled equations and the dispersion relation of the isolated PCW2.

Now, we consider coupling between two asymetric PCWs up to the third nearest-neighboring defects with coupling coefficients α, β and γ, as shown in Fig. 1. The coupled equations of asymmetric PCWs based on the TBT can be expressed as

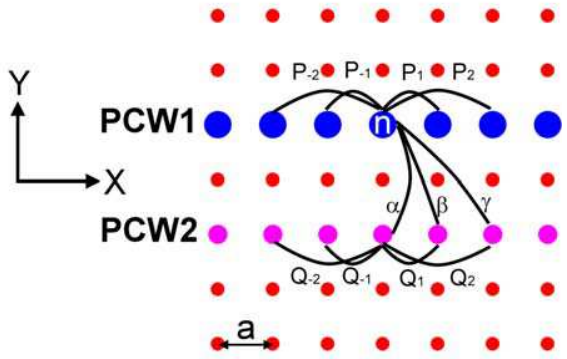


Figure 1: The structures of a photonic crystal coupler with lattice constant a. Coefficients Ps and Qs are the coupling coefficients between defects within the single waveguides. α, β and γ are the coupling coefficients between waveguides. n is the coordinate of a defect.

$$\begin{aligned} i\frac{\partial}{\partial t}u_n =& (\omega_1 - P_0)u_n - \sum_{m=1} P_m(u_{n+m} + u_{n-m}) - \alpha v_n - \beta(v_{n+1} + v_{n-1}) - \gamma(v_{n+2} + v_{n-2}), \\ i\frac{\partial}{\partial t}v_n =& (\omega_2 - Q_0)v_n - \sum_{m=1} Q_m(v_{n+m} + v_{n-m}) - \alpha u_n - \beta(u_{n+1} + u_{n-1}) - \gamma(u_{n+2} + u_{n-2}). \end{aligned} \tag{4}$$

Assume the solutions of coupled Equations (4) to be $u_n = U_0 \exp(ikna - \omega t)$ and $v_n = V_0 \exp(ikna - \omega t)$, we obtain the characteristic equations:

$$\begin{aligned} (\omega - \bar{\omega}_1)\ U_0 + g(ka)\ V_0 = 0, \\ (\omega - \bar{\omega}_2)\ V_0 + g(ka)\ U_0 = 0, \end{aligned} \tag{5}$$

where $g(ka) = \alpha + 2\beta\cos(ka) + 2\gamma\cos(2ka)$.

Solving the characteristic equations, we obtain the eigenfrequencies (dispersion relations) and eigenvectors (field amplitudes) of the coupled system as

$$\omega^{\pm}(k) = \frac{(\bar{\omega}_1 + \bar{\omega}_2)}{2} \pm \sqrt{\Delta^2 + (g(ka))^2}, \tag{6}$$

$$\chi^{\pm} = (V_0/U_0)^{\pm} = \frac{\Delta \pm \sqrt{\Delta^2 + (g(ka))^2}}{-g(ka)}, \tag{7}$$

where $\Delta = (\bar{\omega}_2 - \bar{\omega}_1)/2$ and $\chi^+\chi^- = -1$ due to the orthogonality of two eigenmodes.

From the eigenmodes of triangular and square lattices, shown in Fig. 2, we find the electric field at site n ($x = 0,\ y = 0$) of the triangular PCW has the same polarity (sign) as its first nearest-neighboring site ($x = 0,\ y = \sqrt{3}a$) and the second nearest-neighboring site ($x = a,\ y = \sqrt{3}a$) of the square PCW. Because $\Delta\varepsilon < 0$ for the air-defect PCWs, α and β are both negative values. The results of PCWS with a square-lattice structure are similar to triangular ones. Since the third nearest-neighbor coupling coefficients γ is an order of magnitude smaller than α and β, we will ignore the effect of the third nearest-neighbor coupling for simplicity and assume $\bar{\omega}_2 > \bar{\omega}_1$ in the following discussion.

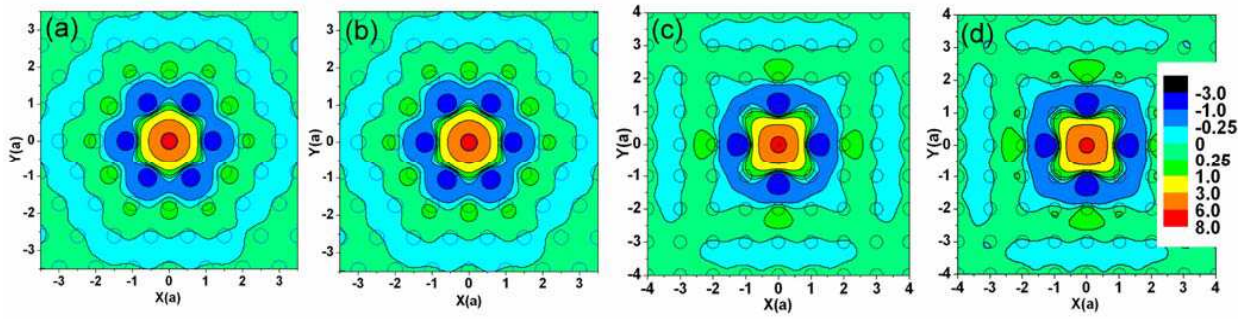


Figure 2: The electric field distribution (E_z) of a point defect. For a triangular lattice, (a) the defect eigenfrequency $f = 0.365c/a$ and refraction index $\varepsilon_r = 2.56$; (b) the defect eigenfrequency $f = 0.371c/a$ and refraction index $\varepsilon_r = 2.25$. For a square lattice, (c) the defect eigenfrequency $f = 0.364c/a$ in reduce rod ($r_d = 0.05a$); (d) the defect eigenfrequency $f = 0.376c/a$ in removed rod with square lattice. Here, c is the speed of light in vacuum.

Because $|g(ka)|$ is maximum at $k = 0$, one would expect that it has the largest splitting of dispersion curves occurring there. As mentioned above α and β are both negative, $g(ka)$ is always negative for all k if $|2\beta/\alpha| < 1$; and it is also negative for small k and may become 0 or even positive as increasing k from 0 to π, if $|2\beta/\alpha| > 1$. Under this circumstance, these two PCWs will be decoupled at $g(k_D a) = 0$ (or $k = k_D$), i.e., the field launched in PCW1 will be always confined in PCW1 without being coupled to PCW2, and vice versa. The eigenfrequencies of the coupled PCWs at $k = k_D$ are $\omega^+ = \bar{\omega}_2$ and $\omega^- = \bar{\omega}_1$ respectively. Because $g(ka) < 0$ if $|2\beta/\alpha| < 1$ or in the range $0 \le k < k_D$ if $|2\beta/\alpha| > 1$, the lower frequency mode (ω^-) has $-1 < \chi^- < 0$; namely, its eigenmode is not only superposed by the out-of-phase electric fields of PCW1 and PCW2 but also localized in the PCW1. This odd-like fundamental (low-frequency) mode is called "anti-bonding" mode which is analog to the bond mode of a two-atom molecule. On the other hand, the high frequency of an even-like mode has $\chi^+ > 1$, thus, it has in-phase electric fields in both PCWs with more field strength localized in PCW2 and is called "bonding" mode.

However, as $k > k_D$ under $|2\beta/\alpha| > 1$, $g(ka)$ becomes positive, $0 < \chi^- < 1$, the fundamental mode is a bonding mode, which has in-phase electric fields in both PCWs having more field strength localized in PCW1 with lower eigenfrequency $\bar{\omega}_1$. And the high frequency antibonding mode with $\chi^+ < -1$ which has field concentrated in PCW2. We find the fundamental mode switches from antibonding to bonding as k increased over the decoupling point k_D with no crossing point of the dispersion curves. As the previous study on coupled symmetric PCWs [2], we can simply set $\Delta = 0$ to obtain $\chi^{\pm} = \pm 1$ at all k, namely, the fundamental modes is either odd or even depending upon the sign of $g(ka)$. And the dispersion curves do cross at the decoupling point $k = k_D$ if $|2\beta/\alpha| > 1$. Furthermore, upon enlarging the separation between two PCWs into a two-rows distance, the

coupling coefficients α and β become both positive but are smaller than those of the one-row-separation PCWs. The fundamental mode becomes a bonding mode. Whether the phenomenon of the mode switching would happen or not is also determined by the criterion: $|2\beta/\alpha| > 1$.

3. SIUMLATION RESULTS

Due to the symmetry, the ratio β/α is larger for a triangular lattice than a square lattice. The criterion, $|2\beta/\alpha| > 1$, of the mode switching is easier to obtain in a triangular lattice than a square lattice [2]. Therefore, we consider a two dimensional triangular lattice PC made by dielectric rods with dielectric constant $\varepsilon_r = 12$ and radius $= 0.2a$ in the air, shown in Fig. 3(a). Two line defects as PCWs are produced by setting the dielectric constants of defect rods to be 2.56 and 2.25, respectively. The eigenfrequencies of the point defects with transverse magnetic field (TM), whose electric field is parallel to the dielectric rods, are $\omega_1 = 0.365\,(2\pi c/a)$ and $\omega_2 = 0.371\,(2\pi c/a)$, respectively. Here, c is the speed of light in vacuum. Furthermore, the dispersion curves f the isolated PCWs and coupled PCWs were calculated by the PWEM. By fitting the calculated dispersion curves to $\omega^+(k)$ and $\omega^-(k)$, we obtain the coupling coefficients between PCWs to be $\alpha = -0.00889$, $\beta = -0.00744$ and $\gamma = 0.00141$. Notice that $|\gamma| \ll |\alpha|$ and $|\beta|$.

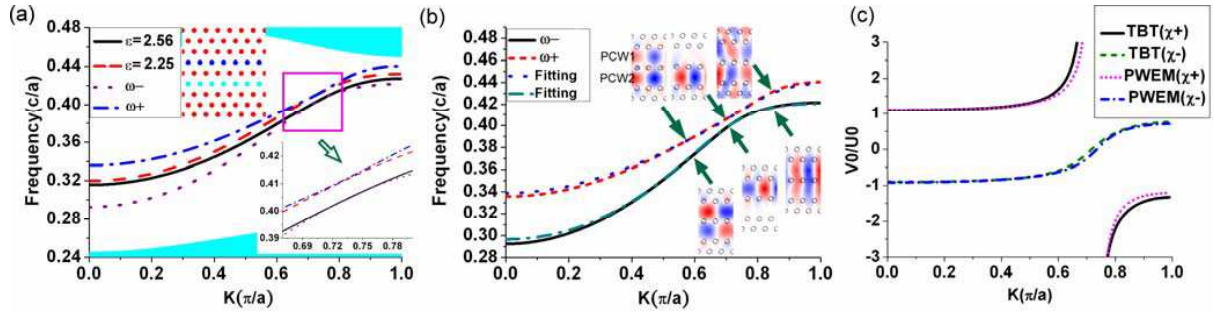


Figure 3: (a) Dispersion relations of two separate PCWs ($\varepsilon = 2.56$ and $\varepsilon = 2.25$) and the coupler in a triangular lattice. The insert is the simulation structure of the coupler. (b) The fitted dispersions from the formula derived by the TBT. The insert is the electric field distribution of each dispersion curve below, above and at the decoupling point. (c) The amplitude ratios of the modes of the coupler obtained by the PWEM and TBT.

From Eq. (6), the decupling point is located at $k_D = 0.73\pi/a$. It is noticed that the dispersion curves do not cross in the asymmetric PCWs. As shown in Fig. 3(b), the eigenmode of high (low) frequency band is bonding (anti-bonding) below the decoupling point ($k = k_D$) but mode switching occurs when $k > k_D$, namely, the eigenmode of high (low) frequency band is anti-bonding (bonding). The electric field is localized in the PCW2 for the high frequency band $\omega^+(k_D)$ and in the PCW1 for the low frequency band $\omega^-(k_D)$ at the decoupling point. The amplitude ratios ($\chi = V_0/U_0$) of the eigenmodes of the coupler obtained by the TBT are consistent with the results from the PWEN (see Fig. 3(c)). From the amplitude ratios changing their sign, we observe that the mode parity is indeed switching at k_D.

4. CONCLUSIONS

We have extended the TBT to study the asymmetric coupler. By considering coupling up to the third nearest-neighboring between two waveguides, analytic expressions of the dispersion relations of the asymmetric coupler are consistent with the results from the PWEM. Due to the broken symmetry, two dispersion curves will never cross even when the criterion $|2\beta| > |\alpha|$ is satisfied. Nevertheless, the eigenmode patterns, which are similar to bonding and anti-bonding modes of a two-atom molecular, do switch their parity as the symmetric coupler does when wavevector k passes through the decoupling point. Moreover, the amplitude ratios of the modes of the coupler obtained by the TBT are consistent with the results from the PWEM. Therefore, the extended tight-binding model is useful for providing physical phenomena of both the symmetric and asymmetric coupler.

ACKNOWLEDGMENT

We acknowledge the partially financial support from the National Science Council (NSC) of Republic of China under Contract No. NSC 96-2628-E-009-018-MY3 and NSC96-2112-M-034-002-MY3.

Chih-Hsien Huang would like to thank NSC for providing fellowship.

REFERENCES

1. Kashani, Z. G., N. Hojjat, and M. Shahabadi, "Full-wave analysis of coupled waveguides in a two-dimensional photonic crystal," *Progress In Electromagnetics Research*, PIER 49, 291–307, 2004.
2. Chien, F. S. S., J. B. Tu, W. F. Hsieh, and S. C. Cheng, "Tight-binding theory for coupled photonic crystal waveguides," *Phys. Rev. B*, Vol. 75, 125113, 2007.
3. Christodoulides, D. N. and N. K. Efremidis, "Discrete temporal solitons along a chain of nonlinear coupled microcavities embedded in photonic crystals," *Opt. Lett.*, Vol. 27, 568, 2002.

Anomalous Microwave Transmission in a Superconducting Periodic Multilayer Structure

Chien-Jang Wu[1] **and Tzong-Jer Yang**[2]

[1]Institute of Electro-Optical Science and Technology
National Taiwan Normal University, Taipei, Taiwan 106, R.O.C.
[2]Department of Electrical Engineering, Chung Hua University, Hsinchu, Taiwan 300, R.O.C.

Abstract— In this work we theoretically study microwave properties of a superconductor/dielectric periodic layered structure, in which a strongly dispersive superconductor, nearly ferroelectric superconductor (NFE SC), is taken. Microwave transmittance in the dielectric-like response has been calculated based on the transfer matrix method as well as the electrodynamics of NFE SCs. Microwave response is strongly dependent on the number of periods as well as the thickness of superconductor layer. It is found that the first anomalous transmission peak can be created when the number of periods is more than five. In addition, more anomalous peaks are generated by greatly increasing the number of periods. The presence of anomalous sharp peaks can be used to design a nicely frequency-selective filter or sampler using such a multilayer structure.

1. INTRODUCTION

Superconductor-dielectric superlattice has attracted much attention in the study of microwave and optical properties of superconducting thin films. Early in the discovery of high-temperature superconductors, Tagantsev and Traito calculated the mixed-state surface impedance of a superconductor superlattice [1, 2]. Recent research in the field of photonic crystals also leads to a work of superconducting photonic crystal [3–6]. It is shown that a one-dimensional superconductor-dielectric photonic crystal (SDPC) has a low-frequency photonic bandgap (PBG), which is not seen in the dielectric-dielectric photonic crystals. This low-frequency PBG is on the order of plasma frequency. More recently, a superlattice made of high-temperature cuprate and magnetic films is experimentally shown to have negative index of refraction, the typical features of a metamaterial [7].

To study the electromagnetic response a superconductor one will, as usual, use the two-fluid model to describe the electrodynamics. Two-fluid model and its related modified model are extensively useful in conventional superconductors and high-temperature cuprates as well. However, in addition to the familiar cuprates, there also exists a peculiar class of noncuprate superconductors called the nearly ferroelectric superconductors (NFE SCs). A NFE SC material is itself in the superconducting state and in the NFE state. The electromagnetic response of this material can be modeled by not only the two-fluid model but the lattice dynamics for the ionic host [8–10]. The theory tells that the response can be dielectriclike or Meissner-like. It has been predicted that in a single NFE SC slab in the dielectriclike response the transmission spectrum, in microwave regime, has many anomalous peaks that are sharply distributed. It is called the comblike distribution and is believed to be unique in such a material.

In this paper, we consider the superconductor-dielectric superlattice in which the superconductor is taken to be the NFE SC. Microwave transmittance will be calculated by using the transfer matrix method together with the electrodynamics of NFE SCs. The effects of number of periods on the microwave transmittance will be illustrated. The results suggest the layered structure can be designed as a nice frequency-selective filter without introducing a defect in the structure.

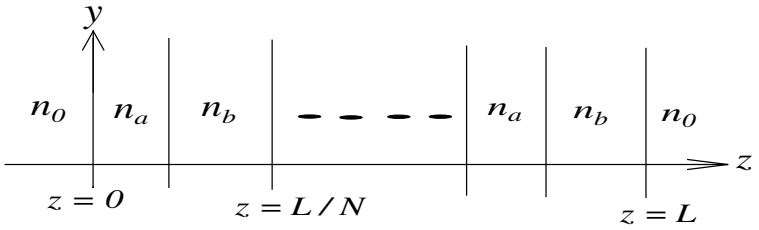


Figure 1: An N-period superconductor-dielectric multilayer structure, in which the total length of system is L, and the indices of refraction, n_a and n_b denote the superconductor and dielectric layers, respectively.

2. BASIC EQUATIONS

The superconductor-dielectric film layered structure is depicted in Fig. 1, in which an N-period layered medium is immersed in the free space characterized by refractive index, $n_0 = 1$. The thickness of superconductor slab is a while the dielectric slab has a thickness of b. The periodicity is denoted by $P = a + b$. We consider a plane wave that is normally incident on plane boundary $z = 0$ from the vacuous half space $z \leq 0$. A reflected plane wave will then also exist in the same half space, and a transmitted wave will be in the leftmost half space $z > L = NP$, where L is the total length of the system. Taking the temporal part as $\exp(-i\omega t)$, and denoting E_{i0}, E_{r0}, and E_{t0} as the amplitudes of the incident, reflected, and transmitted plane wave, respectively, one has the following matrix equation determining the transmission and reflection coefficients [11]

$$E_{t0}\begin{bmatrix}1\\ \eta_0^{-1}\end{bmatrix} = M_{cell}^{N}\begin{bmatrix}E_{i0}+E_{r0}\\ \eta_0^{-1}(E_{i0}-E_{r0})\end{bmatrix} = M_{system}\begin{bmatrix}E_{i0}+E_{r0}\\ \eta_0^{-1}(E_{i0}-E_{r0})\end{bmatrix} = \begin{bmatrix}m_{11} & m_{12}\\ m_{21} & m_{22}\end{bmatrix}\begin{bmatrix}E_{i0}+E_{r0}\\ \eta_0^{-1}(E_{i0}-E_{r0})\end{bmatrix}, \tag{1}$$

where

$$M_{cell} = \exp\left\{i\omega a\begin{bmatrix}0 & \mu_0\mu_b\\ \varepsilon_0\varepsilon_b & 0\end{bmatrix}\right\}\exp\left\{i\omega b\begin{bmatrix}0 & \mu_0\mu_a\\ \varepsilon_0\varepsilon_a & 0\end{bmatrix}\right\}. \tag{2}$$

where $\eta_0 = \sqrt{\mu_0/\varepsilon_0}$ is the free-space wave impedance.

Equation (1) then gives the transmission coefficient as follows:

$$\tau = m_{11}(1+\Gamma) + \frac{m_{12}}{\eta_0}(1-\Gamma), \tag{3}$$

where the reflection coefficient is given by

$$\Gamma = \frac{m_{11}\eta_0 + m_{12} - \eta_0^2 m_{21} - m_{22}\eta_0}{-m_{11}\eta_0 + m_{12} + \eta_0^2 m_{21} - m_{22}\eta_0}. \tag{4}$$

With $\mu_a = 1$, the relative permittivity of superconductor and its index of refraction is related by $n_a = \sqrt{\varepsilon_a}$. The expression for ε_a is given by [9]

$$\varepsilon_{NFE}(\omega) = \varepsilon(\omega) - \frac{1}{\omega^2\mu_0\lambda_L^2}, \tag{5}$$

where the dielectric in the right-hand side is

$$\varepsilon(\omega) = \varepsilon_\infty\frac{\omega_{LO}^2 - \omega^2}{\omega_{TO}^2 - \omega^2}. \tag{6}$$

A detailed description of Equation (6) can be found in Ref. [8].

3. NUMERICAL RESULTS AND DISCUSSION

In what follows we will present the numerical results for the microwave transmittance at frequencies of 15–20 GHz. The NFE SC material is electron-doped $SrTiO_3$ whose parameters are described in Ref. [9], and dielectric slab is $LaAlO_3$ with a dielectric constant of 24.

The calculated results are shown in Fig. 2. For a single period $N = 1$, the bilayer structure acts as good reflector, just as expected for a superconducting film at microwave frequency. In fact, up to $N = 5$, no substantial change can be found in the transmittance spectra. However, we find an anomalous peak with a value of unity emerges at frequency 19.7 GHz when $N = 6$. The peak then shifts to the position of lower frequency as the number of periods is increased. The appearance of anomalous peak indicates that at this frequency the electromagnetic wave can propagate through the total system. In other words, the system can be designed as a frequency-selective filter, or a sampling device that samples a specific frequency from the input signal. It is known that the maximum reflectance at certain frequency the structure can be regarded as a resonator. However if the transmittance attains a maximum, it is referred to as an antiresonator. A resonator or antiresonator can be achieved by introducing a defect in the dielectric-dielectric photonic crystal (DDPC). In addition, if the superconductor is conventional metallic superconductors or high-temperature cuprates, then the superlattice will be a nice reflector at microwave frequency, no

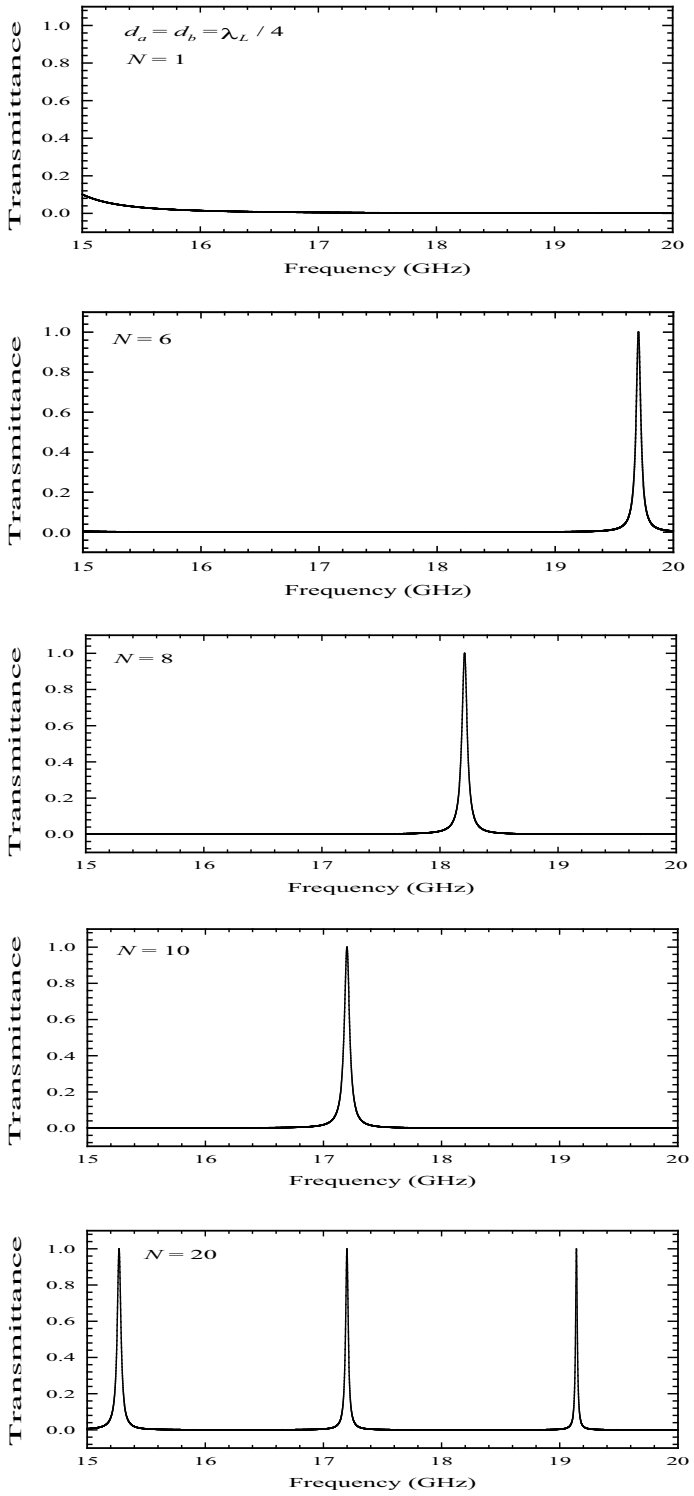


Figure 2: Calculated microwave transmittance as a function of the frequency at different periods, $N = 1, 6, 8, 10$, and 20. The thicknesses of bilayer are taken to be $d_a = d_b = \lambda_L/4$, where $\lambda_L = 18\,\mu\text{m}$.

such anomalous peak in the transmittance spectrum. The NFE SC-dielectric superlattice, however, exhibits the above combined feature altogether. It is possible achieve an antiresonator at microwave regime without introducing any defect in the structure. In addition to such properties, if the number of periods is increased greatly, say $N = 20$, then more antiresonant peaks are produced in the transmittance spectrum, making a multi-frequency sampler.

4. CONCLUSION

In summary, microwave transmittance in a NFE SC-dielectric superlattice has been theoretically investigated. For the simple case, $N = 1$, our results agree with the fact of a nice reflector. It is found the emergence of an anomalous peak at $N = 6$. At such peak frequency the incident wave is totally allowed to propagate through the structure which makes the structure antiresonant, becoming a frequency sampler or filter. The selective frequencies in filter are strongly dependent on the number of periods. The results elucidate the superior merits of such a structure, that is, it is possible to design an filter at microwave frequency without introducing an defect in the bilayer periodic structure.

ACKNOWLEDGMENT

The author, C.-J. Wu, is indebted to the financial support by from the National Science Council of the Republic of China under grant No. NSC-95-2112-M-003-027-MY2.

REFERENCES

1. Tagantsev, A. K. and K. B. Traito, *Phys. Rev. B*, Vol. 48, 10537, 1993.
2. Tagantsev, A. K. and K. B. Traito, *Phys. Rev. B*, Vol. 47, 2760, 1993.
3. Raymond Ooi, C. H., et al., *Phys. Rev. B*, Vol. 61, 5920, 2000.
4. Wu, C.-J., M.-S. Chen, and T.-J. Yang, *Physica C*, Vol. 432, 133, 2005.
5. Wu, C.-J., *Journal of Electromagnetic Waves and Applications*, Vol. 19, 1991, 2005.
6. Wu, C.-J., *Progress In Electromagnetic Research Symposium, PIERS*, Hangzhou, 2008.
7. Primenov, A., et al., *Phys. Rev. Lett.*, Vol. 95, 247009, 2005.
8. Birman, J. L. and N. A. Zimbovskaya, *Phys. Rev. B*, Vol. 64, 144506, 2001.
9. Wu, C.-J. and W.-K. Kuo, *J. Opt. Soc. Am. B*, Vol. 23, 1836, 2006.
10. Wu, C.-J., et al., *Progress In Electromagnetic Research*, PIER 39, 39, 2007.
11. Lakhtakia, A., *Optik*, Vol. 106, 45, 1997.

The Layered Metamaterial with Parabolic Dispersion

Tzong-Jer Yang[1], **Jin-Jei Wu**[1], **and Linfang Shen**[2]

[1]Department of Electrical Engineering, Chung Hua University, Hsinchu 30012, Taiwan, R.O.C.
[2]Department of Information Science and Electronic Engineering, Electromagnetic Academy
Zhejiang University, Hangzhou, Zhejiang 310027, China

Abstract— A layered metamaterial which is predicted by the effective medium theory to be an anisotropic medium with partial zero permittivities is analyzed carefully. It is shown that such a material has a parabolic dispersion relation rather than usual ones. This material may be viewed as a homogeneous medium but is spatially dispersive strongly. The structured medium may support either forward wave or backward wave, depending on the magnitude of the wave vector.

Metamaterials with special constitutive parameters that do not occur in nature may be properly synthesized by embedding suitable inclusions (whose size and separation are much smaller than operation wavelength) in a host medium. Owing to their anomalous and counterintuitive interaction with wave, these structured media have led to many new possibilities in different fields of physics and engineering [1–3]. For instance, a slab made of the so-called left-handed materials, materials possessing negative permittivity and permeability simultaneously, not only is free of usual wavelength limits but also is able to focus even evanescent waves [1]. Recently, materials with nearly zero effective permittivity and/or permeability have been demonstrated to exhibit anomalous features, allowing potential applications in directive emission [4, 5] and energy squeezing [6, 7]. In this presentation, a layered metamaterial with period much smaller than operation wavelength is studied and predicted by the effective medium theory [8] to be an anisotropic medium with partial zero permittivities and single wave vector dispersion. By solving Maxwell's equations strictly in this metamaterial, a parabolic dispersion relation is found.

A metamaterial consisting of alternating layers of metal with relative permittivity $\varepsilon_m < 0$ and dielectric with relative permittivity $\varepsilon_r > 0$ stacked along z direction is considered. The period (a) of the metamaterial is assumed to be much smaller than operation wavelength (λ). According to the effective medium theory [8], this layered material can be treated as a homogeneous medium with effective relative permittivities ε_t and ε_n in the xy plane and z direction, respectively, which are given by $\varepsilon_t = f_m\varepsilon_m + (1-f_m)\varepsilon_r$ and $\varepsilon_n = \varepsilon_r\varepsilon_m/[f_m\varepsilon_r + (1-f_m)\varepsilon_m]$, where $f_m = d/a$ is the metal filling fraction, and d is the thickness of the metal layers. For TE-polarized waves propagating in the xz plane, the dispersion relation for this structured medium is $k_x^2/\varepsilon_n + k_z^2/\varepsilon_t = k_0^2$, where $k_0 = 2\pi/\lambda$ is the wavenumber in free space. In what follows, we will focus on $\varepsilon_t \approx 0$. In such special case, the dispersion relation will be reduced to $k_x = \pm\sqrt{\varepsilon_n}k_0$ and $k_z = 0$, i.e., the structured medium seems to have a single wave vector dispersion. We will examine this by solving strictly the problem from Maxwell's equations. For this purpose, we first divide the unit cell of the layered material into three regions: region I for $-a \le z < -d/2$, region II for $-d/2 \le z \le d/2$, and region III for $d/2 < z \le a$. The TE-polarized wave with wave vector $\mathbf{k} = \hat{x}k_x + \hat{z}k_z$ has magnetic field $\mathbf{H} = \hat{y}H_y$, and the field component H_y in three regions can be written as

$$\begin{aligned} H_y^{\mathrm{I}}(x,z) &= \left[A_1\exp\left(ip_z z\right) + B_1\exp\left(-ip_z z\right)\right]\exp\left(ik_x x\right), \\ H_y^{\mathrm{II}}(x,z) &= \left[A_2\exp\left(-q_z z\right) + B_2\exp\left(q_z z\right)\right]\exp\left(ik_x x\right), \\ H_y^{\mathrm{III}}(x,z) &= H_y^{\mathrm{I}}(x, z-a)\exp(ik_z a), \end{aligned}$$

where $p_z = \sqrt{\varepsilon_r k_0^2 - k_x^2}$ and $q_z = \sqrt{k_x^2 - \varepsilon_m k_0^2}$. The nonzero components (E_x and E_z) of the electric field can be obtained straightforwardly from H_y. The dispersion relation for the layered material can be gotten by imposing matching conditions on the parallel field components (H_y and E_x) at the interface between regions I and II and the one between regions II and III. From the continuity of H_y and E_x at $z = -d/2$, we have

$$\bar{A}_1 + \bar{B}_1 = \bar{A}_2 + \bar{B}_2, \tag{1}$$

$$-i\kappa\left(\bar{A}_1 - \bar{B}_1\right) = \bar{A}_2 - \bar{B}_2, \tag{2}$$

where $\bar{A}_1 = A_1 \exp(-ip_z d/2)$, $\bar{B}_1 = B_1 \exp(ip_z d/2)$, $\bar{A}_2 = A_2 \exp(q_z d/2)$, $\bar{B}_2 = B_2 \exp(-q_z d/2)$, and $\kappa = \varepsilon_m p_z/\varepsilon_r q_z$. Then we obtain

$$\bar{A}_2 = \frac{1}{2}(1 - i\kappa)\bar{A}_1 + \frac{1}{2}(1 + i\kappa)\bar{B}_1, \tag{3}$$

$$\bar{B}_2 = \frac{1}{2}(1 + i\kappa)\bar{A}_1 + \frac{1}{2}(1 - i\kappa)\bar{B}_1. \tag{4}$$

Similarly, from the matching conditions at $z = d/2$, we can obtain

$$\bar{A}_2 \exp(-q_z d) = \frac{1}{2}\left\{(1 - i\kappa)\exp\left[-ip_z(a-d)\right]\bar{A}_1 + (1 + i\kappa)\exp\left[ip_z(a-d)\right]\bar{B}_1\right\}\exp(ik_z a), \tag{5}$$

$$\bar{B}_2 \exp(q_z d) = \frac{1}{2}\left\{(1 + i\kappa)\exp\left[-ip_z(a-d)\right]\bar{A}_1 + (1 - i\kappa)\exp\left[ip_z(a-d)\right]\bar{B}_1\right\}\exp(ik_z a). \tag{6}$$

Substitution of Eqs. (3) and (4) into Eqs. (5) and (6) yields

$$\begin{aligned}&(1 - i\kappa)\left\{\exp(-q_z d) - \exp\left[-ip_z(a-d)\right]\exp(ik_z a)\right\}\bar{A}_1 \\ =\ &-(1 + i\kappa)\left\{\exp(-q_z d) - \exp\left[ip_z(a-d)\right]\exp(ik_z a)\right\}\bar{B}_1,\end{aligned} \tag{7}$$

$$\begin{aligned}&(1 + i\kappa)\left\{\exp(q_z d) - \exp\left[-ip_z(a-d)\right]\exp(ik_z a)\right\}\bar{A}_1 \\ =\ &-(1 - i\kappa)\left\{\exp(q_z d) - \exp\left[ip_z(a-d)\right]\exp(ik_z a)\right\}\bar{B}_1.\end{aligned} \tag{8}$$

By eliminating the coefficients $\bar{A}_1$ and $\bar{B}_1$ in Eqs. (7) and (8), we can finally obtain the dispersion relation for the layered material in the form

$$\begin{aligned}\cos(k_z a) =\ &\cos\left[(1 - f_m)p_z a\right]\cosh(f_m q_z a) \\ &-\frac{1}{2}\left(\frac{\varepsilon_m p_z}{\varepsilon_r q_z} - \frac{\varepsilon_r q_z}{\varepsilon_m p_z}\right)\sin\left[(1 - f_m)p_z a\right]\sinh(f_m q_z a).\end{aligned} \tag{9}$$

This equation is the exact dispersion relation for the layered material under consideration. Note that $k_0 a \ll 1$ since $a \ll \lambda$. In the $\mathbf{k}$ region of our interest, where $|\mathbf{k}| \sim k_0$ or $|\mathbf{k}| < k_0$, the left and right sides of Eq. (9) are expanded in Taylor series to second and fourth orders, respectively, and then the equation is reduced to

$$k_z = \pm\tau\left(k_0^2 - \frac{k_x^2}{\varepsilon_n}\right)a \tag{10}$$

under the condition of $|\varepsilon_t| \ll (k_0 a)^2$, where $\tau = (1 - f_m)\varepsilon_r/2\sqrt{3}$. The dispersion relation (10) is a pair of parabolas. Evidently, the single $\mathbf{k}$ point dispersion predicted from the effective medium theory is incorrect for the layered material with $\varepsilon_t = 0$. In our previous study on photonic crystal, the parabola-like dispersion behavior was already found [9].

To describe a metamaterial with parabolic dispersion in a macroscopic view, a set of macroscopic fields via averaging the local fields over the unit cell is introduced, and then find out the relationships between them. Each averaged field component is expressed by bar on the top of each field component. From the Maxwell's equation $\partial H_y/\partial x = -i\omega\varepsilon_0\varepsilon(z)E_z$, where $\varepsilon(z) = \varepsilon_r$ in regions I and III and $\varepsilon(z) = \varepsilon_m$ in region II, we have $E_z = -k_x H_y/\omega\varepsilon_0\varepsilon(z)$. As H_y is continuous in the unit cell, we have $H_y(z) = \bar{H}_y + O(k_0 a)$, and then find $\bar{E}_z = -k_x\bar{H}_y/\omega\varepsilon_0\varepsilon_n$. Therefore, the metamaterial can be characterized by an effective relative permittivity $\varepsilon_z = \varepsilon_n$ in the z direction, as predicted by the effective medium theory. Moreover, from one of the Maxwell's equations $\partial E_x/\partial z - \partial E_z/\partial x = i\omega\mu_0 H_y$, we obtain $\partial E_x/\partial z + ik_x^2 H_y/\omega\varepsilon_0\varepsilon(z) = i\omega\mu_0 H_y$. By integrating this equation from $z = -a/2$ to $z = a/2$, we have $[E_x(a/2) - E_x(-a/2)]/a + ik_x^2\bar{H}_y/\omega\varepsilon_0\varepsilon_n = i\omega\mu_0\bar{H}_y$. Since $E_x(a/2) = E_x(-a/2)\exp(ik_z a)$, and $E_x(-a/2) = \bar{E}_x + O(k_0 a)$, we obtain $\bar{E}_x = [(k_0^2 - k_x^2/\varepsilon_n)/\omega\varepsilon_0 k_z]\bar{H}_y$. Now, an effective relative permittivity (ε_x) for the metamaterial in the x direction is introduced, we may then get

$$\varepsilon_x = \frac{k_z^2}{k_0^2 - k_x^2/\varepsilon_n} = \sigma\tau(k_z a), \tag{11}$$

where $\sigma = \pm 1$, which takes the same sign as Eq. (10). The dispersion Eq. (10) can thus be expressed as $k_x^2/\varepsilon_z + k_z^2/\varepsilon_x = k_0^2$, and then $k_z\bar{E}_x - k_x\bar{E}_z = \omega\mu_0\bar{H}_y$ is found. As Maxwell's equations are all

satisfied by the macroscopic fields, the metamaterial with parabolic dispersion can be viewed as a special homogenous medium that is spatially dispersive strongly. The macroscopic energy flow is $\bar{\mathbf{S}} = \hat{x}\bar{S}_x + \hat{z}\bar{S}_z$, where

$$\bar{S}_x = \frac{k_x}{2\omega\varepsilon_0\varepsilon_n}\left|\bar{H}_y\right|^2, \tag{12}$$

$$\bar{S}_z = \sigma\frac{1}{2\omega\varepsilon_0\tau a}\left|\bar{H}_y\right|^2. \tag{13}$$

Clearly, the sign "+" in Eq. (10) [i.e., $\sigma = 1$ in Eqs. (11) and (13)] corresponds to waves travelling in the positive z direction, and the sign "−" (i.e., $\sigma = -1$) to waves travelling in the negative z direction. Interestingly, one may get $|\bar{S}_z| \gg |\bar{S}_x|$ from Eqs. (12) and (13) for $|k_x a| \ll 1$, indicating that waves travel almost along the z direction in this medium. Furthermore, waves with $|k_x| < \sqrt{\varepsilon_n}k_0$ are forward waves with k_z parallel to $\bar{S}_z$, whereas waves with $|k_x| > \sqrt{\varepsilon_n}k_0$ are backward waves with k_z anti-parallel to $\bar{S}_z$.

In conclusion, a metamaterial predicted by the effective medium theory to be an anisotropic medium with partial zero permittivities and single wave vector dispersion has been investigated theoretically. By solving Maxwell's equations strictly in this metamaterial, a parabolic dispersion relation has been found. But the metamaterial may still be treated as a homogeneous medium with an effective permittivity being not zero but a function of wave vector. The structured medium may support either forward wave or backward wave, depending on the magnitude of the wave vector.

We would like to acknowledge the National Science Council of ROC for the financial support through the grant number NSC 96-2112-M-216-001 and NSC 96-2811-M-009-028. L.S. sincerely thanks the hospitality of NCTU when he visited there.

REFERENCES

1. Pendry, J. B., *Phys. Rev. Lett.*, Vol. 85, 3966, 2000.
2. Smith, D. R. and D. Schurig, *Phys. Rev. Lett.*, Vol. 90, 077405, 2003.
3. Hu, L. B. and S. T. Chui, *Phys. Rev. B*, Vol. 66, 085108, 2002.
4. Enoch, S., G. Tayeb, P. Sabouroux, N. Guerin, and P. A. Vincent, *Phys. Rev. Lett.*, Vol. 89, 213902, 2002.
5. Yuan, Y., L. F. Shen, L. Ran, T. Jiang, J. T. Huangfu, and J. A. Kong, *Phys. Rev. A*, Vol. 77, 053821, 2008.
6. Edwards, B., A. Alu, M. E. Young, M. Silveirinha, and N. Engheta, *Phys. Rev. Lett.*, Vol. 100, 033903, 2008.
7. Silveirinha, M. and N. Engheta, *Phys. Rev. Lett.*, Vol. 97, 157403, 2006.
8. Aspnes, D. E., *Am. J. Phys.*, Vol. 50, 704, 1982.
9. Shen, L. F., T. J. Yang, and Y. F. Chau, *Appl. Phys. Lett.*, Vol. 90, 251909, 2007.

Design of Low-cost Microstrip Antennas for Glonass Applications

Daniel C. Nascimento, Ricardo Schildberg, and J. C. da S. Lacava
Laboratório de Antenas e Propagação, Instituto Tecnológico de Aeronáutica, Brazil

Abstract— A new design procedure for low-cost circularly polarized microstrip antennas is proposed. Instead of the conventional truncated-corner square microstrip topology, a rectangular patch with four truncated corners and an equal number of stubs is utilized here. As an application, a Glonass receiver antenna is designed. Comparisons between experimental and simulated results are presented, revealing very good agreement.

1. INTRODUCTION

The design of microstrip antennas as low-cost radiators for mobile communications can be challenging. Additional complexities are introduced into the conventional microstrip design when the FR4 laminate is used as the antenna substrate, due to its high loss tangent and inaccurate relative permittivity. The high losses in the antenna laminate have a direct effect on its radiation efficiency. Consequently, for higher efficiency, thicker substrates must be used [1]. In case, for instance, a probe-fed truncated-corner square microstrip antenna (TCSMA), frequently used for achieving circular polarization (CP), is designed following the conventional procedure, an extra limitation is posed if the patch is printed on a thick substrate: its inherently inductive input impedance can not be properly matched to a 50-Ω SMA coaxial connector. In addition, the frequency where its axial ratio is best does not coincide with the frequency where its return loss is best [2].

To overcome these limitations, a new design procedure is proposed here. Instead of the conventional topology (a square patch with two truncated corners), the new procedure utilizes a rectangular patch with four truncated corners. Besides, stubs are added to each corner to compensate for the inaccuracy of the FR4 relative permittivity. This new patch topology gives the antenna designer more flexibility and the ability to properly compensate for the undesirable reactive inductance.

2. CIRCULARLY POLARIZED MICROSTRIP ANTENNAS

Antennas produce circularly polarized waves when two orthogonal field components with equal amplitudes and in-phase quadrature are radiated. Probe-fed microstrip patches, classified as resonator-type antennas, can satisfy these requirements. Various configurations capable of CP operation have been reported [3], but square and circular patches are largely utilized in practice. Square and rectangular patches for Glonass applications are discussed below.

2.1. TCSMA Design

For Glonass applications, receiver antennas have to be right-hand circularly polarized over an operating range from 1.598 to 1.609 GHz. Besides, their return loss, axial ratio and radiation efficiency have to meet specific requirements over the same frequency range [4]. TCSMA radiators that are manufactured with high quality materials can comply with these specifications. Use of low-loss RF laminates with high dielectric constant in their fabrication is currently expensive. Therefore, low-cost solutions are welcome since both market and technology are now ripe for mass production [5].

The design of a TCSMA radiator printed on FR4 substrate of $\varepsilon_r = 4.31$, thickness = 1.524 mm, and loss tangent = 0.02 for GPS applications has been recently presented [2]. However, such high loss tangent dramatically affected its axial ratio (AR) and degraded its gain, resulting in poor radiation efficiency of only 35%, which is unacceptable from a practical point of view.

By increasing the FR4 thickness to 6.2 mm, a new TCSMA radiator with the geometry shown in Figure 1 on a finite ground plane ($70 \times 70\,\text{mm}^2$) was designed for operation at 1.6035 GHz resulting in the following patch dimensions: $S = 41.1$ mm, $C = 9.35$ mm, and $P = 9.55$ mm. Corresponding input impedance, axial ratio and return loss simulated with the *HFSS* package are shown in Figures 2 and 3.

As seen from these figures, the input reactance at the operating frequency is too large ($X_{in} = j30.8\,\Omega$), so the antenna is not properly matched to the 50-Ω SMA coaxial connector. In addition, the frequency where its axial ratio is best does not coincide with the frequency where its return loss is best. The following antenna parameters were also obtained: directivity = 6.08 dB, 3-dB

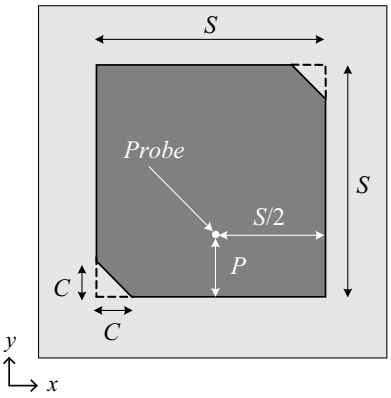


Figure 1: Typical geometry of the TCSMA radiator.

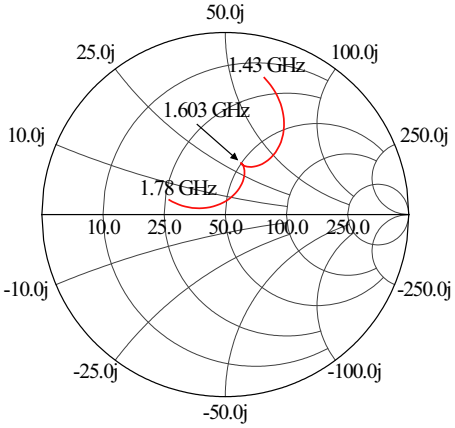


Figure 2: Input impedance of the TCSMA radiator.

AR bandwidth = 52.7 MHz (3.28 %) and radiation efficiency = 68.4%. Although these parameters comply with the Glonass requirements, return loss at 1.6035 GHz (−10.65 dB) is not acceptable. To overcome this limitation, fractal gap capacitors can be used to match the antenna to the coaxial feed line [2]. However, a different procedure is implemented next.

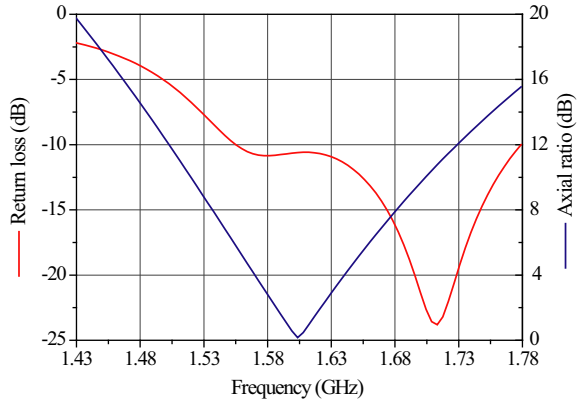


Figure 3: Axial ratio and return loss of the TCSMA radiator.

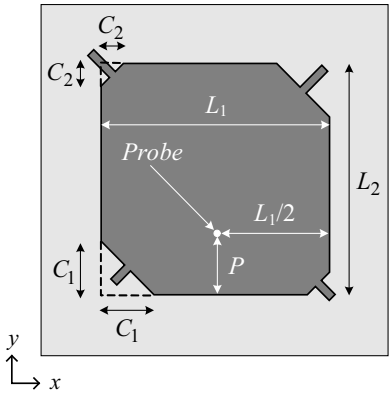


Figure 4: Proposed geometry for the TCRMA radiator.

2.2. TCRMA Design

As previously mentioned, the use of a low-cost laminate adds complexity to the antenna design due to the inaccuracy of its relative permittivity. To compensate for this effect, a topology consisting of four truncated corners with a stub added to each corner, as shown in Figure 4, is proposed. By proper trimming of each stub length, frequency deviations can be compensated for. Moreover, the new topology utilizes a rectangular patch (TCRMA) instead of the square one described in [2]. These new additional parameters (i.e., the patch sides L_1 and L_2) provide more flexibility to the antenna designer.

Keeping the FR4 thickness the same as before, i.e., 6.2 mm, a TCRMA radiator on a finite ground plane ($70 \times 70\,\text{mm}^2$) has been designed ($L_1 = 40.2$ mm, $L_2 = 42.5$ mm, $C_1 = 10$ mm, $C_2 = 4$ mm, $P = 7.25$ mm, length of the large stubs = 5.5 mm, and length of the small stubs = 3.5 mm) for operation at 1.6035 GHz, resulting in the following antenna parameters: directivity = 6.12 dB, 3-dB AR bandwidth = 52.5 MHz (3.27%), and radiation efficiency = 68%.

The corresponding input impedance, axial ratio, and return loss over the frequency range are shown in Figures 5–7. Simulated and experimental input impedance results show very good agreement. As expected, the microstrip antenna with the new geometry exhibits very good AR (0.39 dB) and return loss (−33.7 dB) characteristics at 1.6035 GHz, without any capacitive gaps. These are clearly excellent results from a practical point of view. A photo of the antenna prototype that was built is shown in Figure 8. The θ- and ϕ-components of the far electric fields radiated by

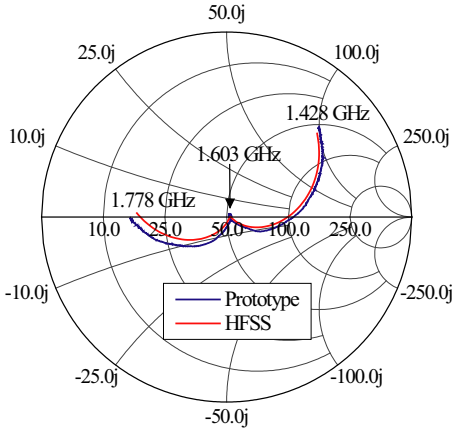


Figure 5: Input impedance of the TCRMA radiator.

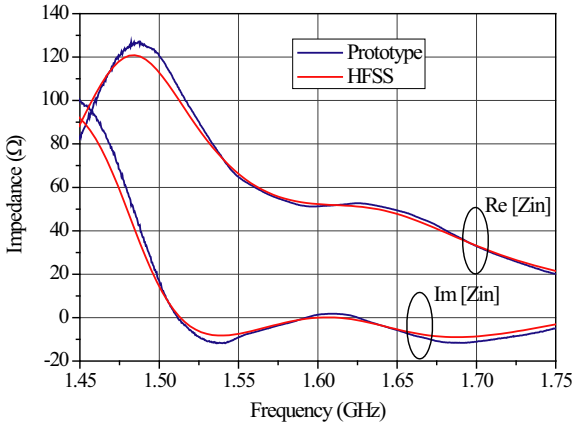


Figure 6: Input impedance of the TCRMA radiator.

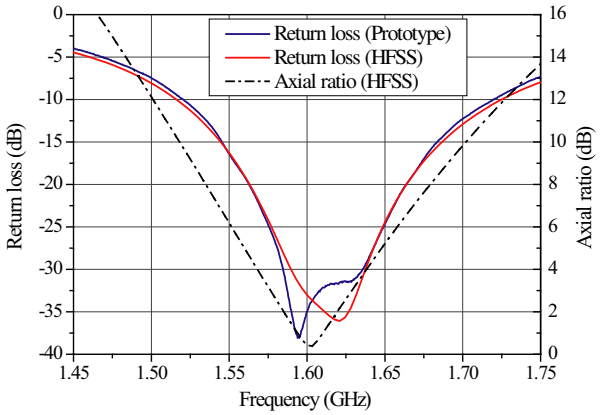


Figure 7: Axial ratio and return loss of the TCRMA radiator.

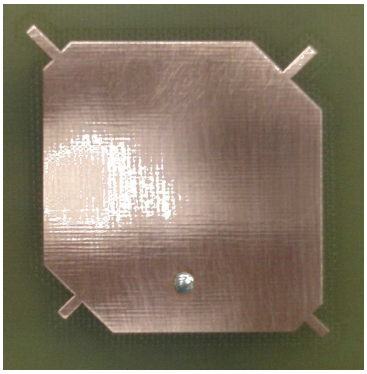


Figure 8: Photo of the Glonass antenna prototype.

the antenna are also analyzed. Experimental and simulated radiation patterns for the E_θ and E_ϕ components, plotted in the xz-plane, are shown in Figures 9 and 10, respectively.

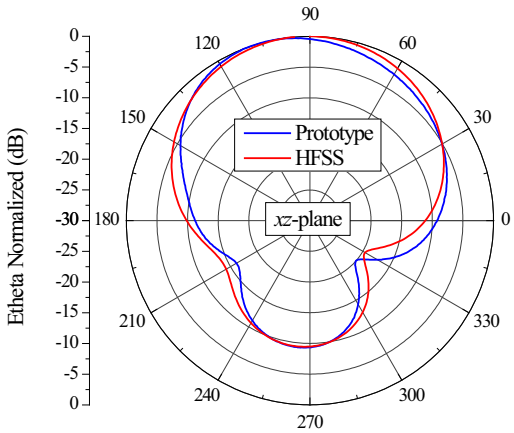


Figure 9: E_θ radiation pattern: xz-plane.

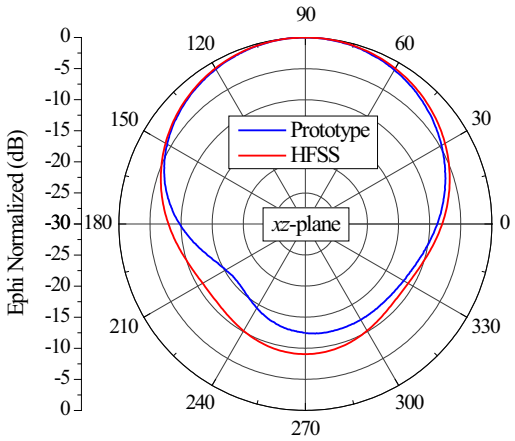


Figure 10: E_ϕ radiation pattern: xz-plane.

Results for the yz-plane are presented in Figures 11 and 12. Simulated and experimental results for this important parameter also show very good agreement.

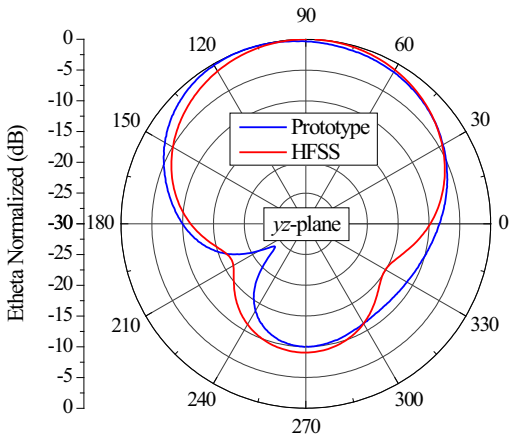


Figure 11: E_θ radiation pattern: yz-plane.

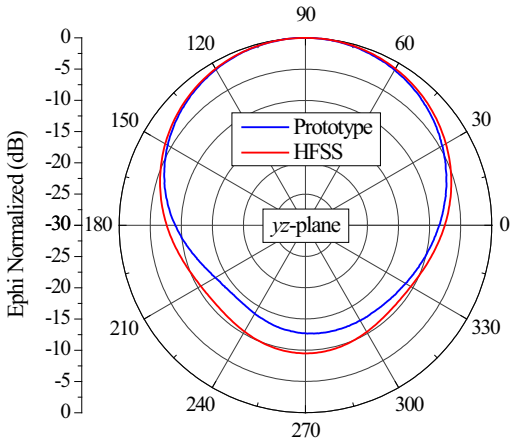


Figure 12: E_ϕ radiation pattern: yz-plane.

3. CONCLUSIONS

The design of a probe-fed truncated corner microstrip antenna on a low-cost substrate (FR4) for Glonass applications is reported. To meet design requirements, thicker substrates have to be used, but the input impedance of such antennas is inherently inductive. To overcome this problem, a new topology using a rectangular patch, rather than the usual square one, is introduced. Experimental and simulated results demonstrate it is feasible to design a thick FR4 microstrip antenna with very good AR and return loss at 1.6035 GHz without capacitor gaps or any other matching technique. This new topology can be a significant contribution to the design of low-cost circularly polarized microstrip antennas.

ACKNOWLEDGMENT

This work was partially supported by FINEP; RHODES Project.

REFERENCES

1. Nascimento, D. C., J. A. Mores Jr., R. Schildberg, and J. C. S. Lacava, "Analysis and design of low-cost microstrip antennas," *MOMAG*, CD-ROM, Belo Horizonte, Brazil, 2006 (in Portuguese).
2. Nascimento, D. C., J. A. Mores Jr., R. Schildberg, and J. C. S. Lacava, "Low-cost truncated corner microstrip antenna for GPS application," *International Symposium on Antennas and Propagation*, 1557–1560, Albuquerque, NM, USA, 2006.
3. Garg, R., P. Bhartia, I. Bahl, and A. Ittipiboon, *Microstrip Antenna Design Handbook*, Artech House Inc., Norwood, MA, 2001.
4. Nascimento, D. C., "Antennas for mobile communications," M.Sc. Thesis, ITA, 2007 (in Portuguese).
5. Gardelli, R., G. La Cono, and M. Albani, "A low-cost suspended patch antenna for WLAN access points and point-to-point links," *IEEE Antennas and Wireless Propagation Letters*, Vol. 3, 90–93, 2004.

Design of a 2.5GHz Differential CMOS LNA

Xuan Chen, **Quanyuan Feng**, **and Shiyu Li**

Institute of Microelectronics, Southwest Jiaotong University, Chengdu 610031, China

Abstract— A 2.5 GHz differential CMOS LNA which fabricated with the 0.18 µm CMOS process is proposed and the two-input and two-output architecture are designed. Complying with the aspects of noise optimization, linear gain and impedance matching, the design methodology of LNA is analyzed in detail, and the source inductive L_s, input matching capacitance and gate width W_1, W_2 is also discussed. Simulation was made by using the ADS software. Consuming 5.4 mA current at 2 V supply voltage, the proposed LNA exhibits a linear gain of 15.053 dB, noise figure of 1.910 dB, S_{11} of −50.687 dB and S_{12} of −18.132 dB.

1. INTRODUCTION

With the rapid development of the communication industry, more and more kinds of wireless communication apparatus are needed, such as small power radiation, far effective distance, big coverage range, have been the main target of the businessman and the manufacturer of the wireless communication apparatus. In this condition, swift developmental radio frequency (RF) wireless communication technology has been widely used in all fields of the world. Low noise amplifier (LNA) which is in the RF front-end circuit has the great value in this field [1–3].

The main effect of the LAN is to amplify the faint signal which is received by the antenna. It should provide enough gain, low noise figure, high linearity, great input and output matching with restraint of power consumption. There are so many merits to use differential architecture, such as abiding the noise. Millimeter wave signal can be easily leaked, and differential signal can be used to eliminate the negative effect of leaking signal. Meanwhile differential LNA can restrain common mode interference, so the noise of source voltage and underlay voltage can also be restrained.

In this paper, a modified architecture used for input matching in CMOS LNA is proposed and the noise optimization is also discussed. A two-input and two-output differential architecture is presented, in which the inductance is used as the source inductive degeneration of the C-S amplifier to enlarge the linear range.

2. DESIGN OF THE CIRCUIT ARCHITECTURE

2.1. Proposed Architecture of Input Matching

Figure 1(a) shows a single branch of the differential pair, it is the cascade architecture with source inductive degeneration, and the expression of input impedance is:

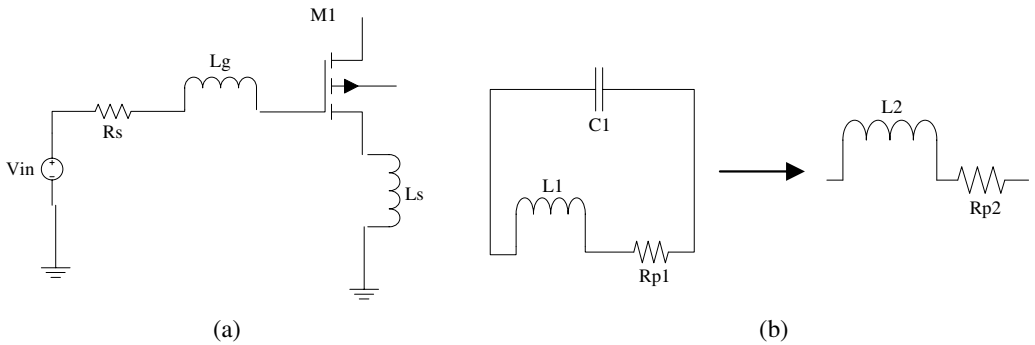


Figure 1: (a) Source inductive degeneration architecture. (b) Parallel LC network and its equivalent circuit.

$$Z_{in} = s(L_g + L_s) + \frac{1}{sC_{gs}} + \frac{g_m}{C_{gs}}L_s \approx s(L_g+L_s)+\frac{1}{sC_{gs}} + w_T L \tag{1}$$

To achieve input matching, the Z_{in} should be 50 Ω, so

$$Z_{in} = \frac{g_m}{C_{gs}}L_s = 50\,\Omega \tag{2}$$

In order to reach the input matching, however, the required gate L_g should be large enough. In practice, a large-value inductor is difficult to be implement on-chip based on a CMOS process, and it will contribute much thermal noise [3–5]. With the consideration of system integration and reducing the noise figure, a modified architecture is proposed [2], which is illustrated in Fig. 1(b). A parallel LC network is designed to replace the inductor. L_1 is an ideal inductor in series with its parasitic resistance R_{P1}. The equal impedance can be derived as $Z = jwL_2 + R_{P2}$, and ω is the operating frequency of the LNA. Where $w_{01} = \frac{1}{\sqrt{L_1 C_1}}$ is the resonant frequency of the L_1C_1 parallel network,

$$L_2 \approx \frac{L_1}{1 - w^2 L_1 C_1} = \frac{L_1}{1 - (w/w_{01})^2} \tag{3}$$

$$R_{P2} \approx \frac{R_{P1}}{(1 - w^2 L_1 C_1)^2} = \frac{R_{P1}}{(1 - (w/w_{01})^2)^2} \tag{4}$$

Here the LC parallel network is equivalent to an ideal inductor L_2 in series with a resistor R_{P2}, let us see (3), when $0 < 1 - w^2 L_1 C_1 < 1$, that $L_2 > L_1$, and as ω approaches to ω_{01}, L_2 will be greatly larger than L_1. Hence a small LC parallel network will generate a larger inductance and to replace the large L_g. So,

$$Z_{in} = \left\{ jwL_2 + jw(1 + g_m R_i)L_s - j\frac{1}{wC_{gs}} \right\} + \left(\frac{g_{m1}}{C_{gs}} L_s + R_{P2} + R_g + R_i \right) \tag{5}$$

The really definition of the parameters in (5) can be found in [2].

2.2. Noise Optimization

There are several noise source in a CMOS LNA, such as channel noise, gate induced noise, thermal noise of various parasitic resistance, etc. As [4] illustrates,

$$F = F_{\min} + \frac{R_n}{G_s} \left[(G_s - G_{opt})^2 + (B_s - B_{opt})^2 \right] \tag{6}$$

$$w_T = \frac{g_m}{C_{gs}} \tag{7}$$

And (6) can be simplified with restraint of power consumption, and F can be written as this:

$$F_{\min} \approx 1 + 2.3 \left[\frac{w}{w_T} \right] \tag{8}$$

Combining (7) and (8), and we know that, g_m is inverse ratio to F, C_{gs} is direct ratio to F. So the low noise figure properly can be implemented by reducing the C_{gs} and enhancing the transconductance g_m, and this process can be realized by modulating the width of the transistor and adjusting the bias circuit with restraint of power consumption. Meanwhile, through the introduction of L_1C_1 parallel network we know that the R_{P2} is not a fully physical resistor and it will not generate so much thermal noise as a real resistor dose, so the NF might be somehow improved [2].

2.3. The Main Circuit Architecture

Here, a 2.5 GHz LNA will be designed based on the analysis before, with the 0.18 µm CMOS process and BSIM3 model. The architecture of differential LNA was showed in Fig. 2.

There are some characteristic in this circuit:

1. A two-input and two-output differential architecture is presented.
2. Cascade topology with source inductance degeneration (L_s) is used.
3. The active mirror-biasing circuit is set up with LC parallel network in input terminal.

Cascade can reduce the Miller effect efficiently and raise S_{12}. The inductance L_s degeneration and input matching of transistor are tuned. It is the same with narrowband amplifying, and it can get low noise figure. M_3 and M_1 are current mirrors and impedance R_{bais} need to be chosen big enough so as to neglect the noise of M_3 which can be converted into an equivalent noise current source at input terminal of the LNA.

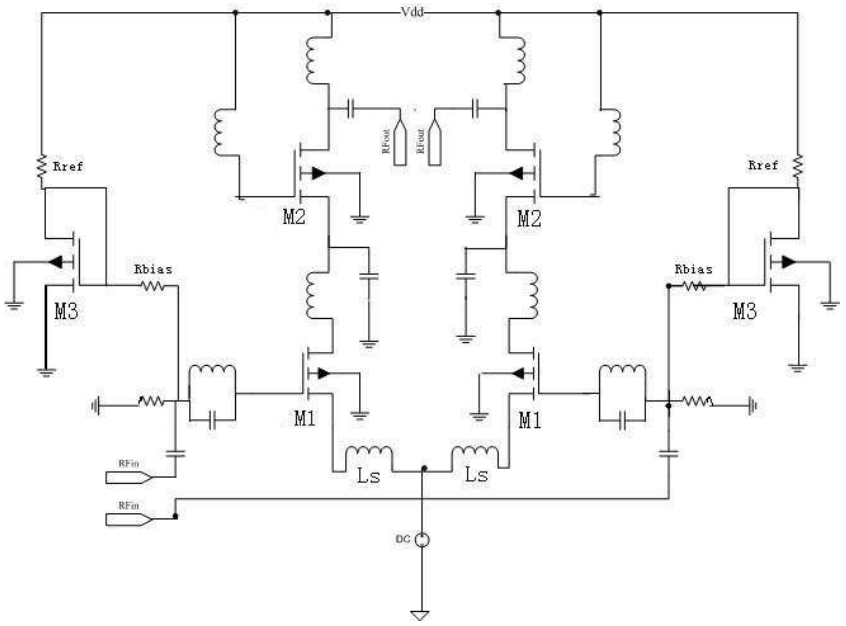


Figure 2: Circuit schematic of a differential LNA.

The active bias circuit are consist of transistor M_3, impedance R_{bais} and R_{REF}, providing transistor M_1 with gate current. M_3 and M_1 are current mirrors, they can optimize the noise and enlarge the gain, and we have to choose the best width of the transistor. Meanwhile, in order to reduce the additive power consumption of bias circuit farthest, the width of MOSFET M_3 has to be tenth as the width of the M_1, and bias impedance R_{bais} should be big enough, we choose it 6 kΩ. C_1 is a blocking capacitor. The L_1 and C_1 are set to 1.0 nH and 0.55 pF, according to (3), it can generate an inductance around 3 nH.

Because the operating frequency of the differential LNA is 2.5 GHz, each element in circuit is modulated, and the parameters of the LNA are simulated by ADS software. Then the width W_1 of

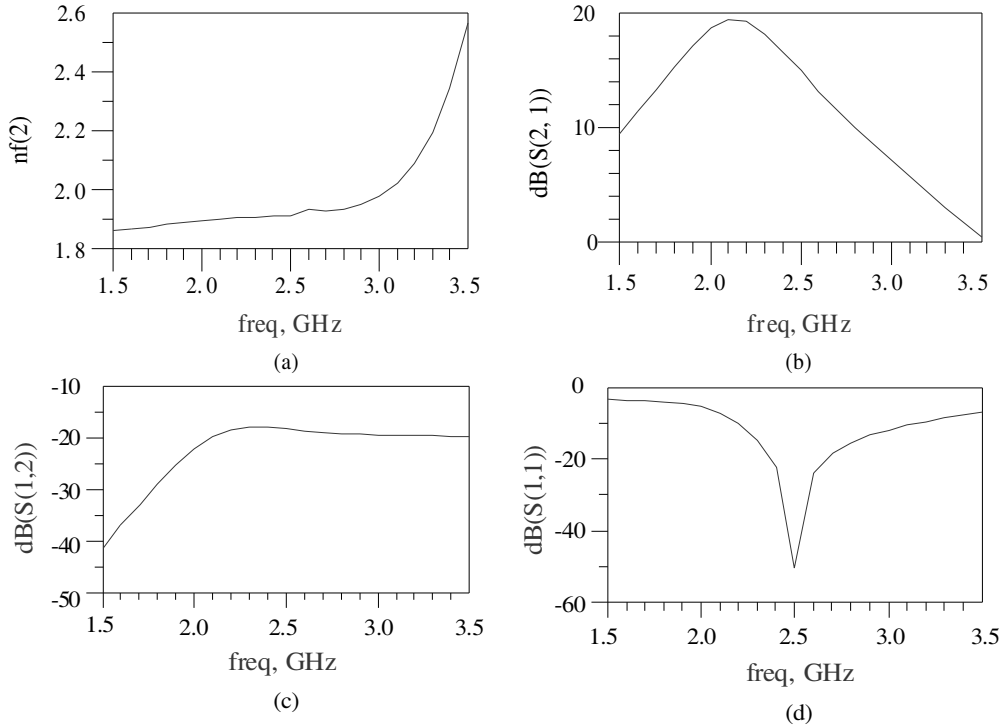


Figure 3: Simulation results. (a) Noise figure, (b) Gain, (c) S_{12}, (d) Input matching.

the transistor is scan with scope range from 100 μm to 400 μm. Considering the linear and other relative index of LNA, we set the value that: $W_1 = 250\,\mu m$, $W_2 = 195\,\mu m$, $W_3 = 25\,\mu m$.

3. RESULTS AND DISCUSSION

Based on the BSIM3 0.18 μm model, when the operating frequency is 2.5 GHz and the voltage is 2 V, the ADS software is used to simulate the CMOS LNA, and the results are show in the Fig. 3.

S-parameter is usually used to measure the performance of the LNA. S_{11} means input matching, the value of it below −10 dB is reasonable. S_{21} is the gain. NF is the noise figure. S_{12} is the reverse isolation. To reach good performance that LNA noise below 2 dB and gain beyond 10 dB are needed. As the figure shows, when the operating frequency is 2.5 GHz, S_{11} is −50.687 dB, S_{12} is −18.132 dB, S_{21} is 15.053 dB, NF is 1.910 dB and the LNA consumes 5.4 mA bias current at 2 V supply voltage. The values of the parameters are ideal and can satisfy the requirements.

According to the meaning of the stability figure **K**:

$$\mathbf{K} = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\Delta|^2}{2|S_{12}S_{21}|} \tag{9}$$

$$\Delta = |S_{11}S_{22} - S_{12}S_{21}| \tag{10}$$

When $\mathbf{K} > 1$, that $|\Delta| < 1$, the circuit is stable unconditionally. According to (9), when input and output matching are good, decreasing the absolute value of S_{12} can increase the stability figure **K**. So according to Fig. 3, LNA is stable unconditionally.

4. CONCLUSION

Through modulating and optimizing all parameters of the circuit, each index of 2.5 GHz LNA can satisfy our requirement. The two-input and two-output differential architecture can restrain common mode interference efficiently and reduce the underlay noise. In this paper, *LC* parallel network is proposed to replace gate inductance, not only the input matching is improved, but also the noise figure is degraded. The LNA can be integrated on chip more easily and compactly with small inductance. Considering the circuit application and enhancing the integration, the CMOS differential LNA is designed, and it can be utilized in wireless RF receiver which is operated at 2.5 GHz.

REFERENCES

1. Comer, D. J. and D. T. Comer, "Using the weak inversion region to optimize input stage design of CMOS op amps [J]," *IEEE Transactions on Circuits and System*, Vol. 51, No. 1, 8–14, Jan. 2004.
2. Mou, S. X., J. G. Ma, Y. K. Seng, et al., "A modified architecture used for input matching in CMOS low-noise amplifiers [J]," *IEEE Transactions on Circuits and System*, Vol. 52, No. 11, 784–788, 2005.
3. Liao, C. H. and H. R. Chuang, "A 5.7 GHz 0.18-μm CMOS gain-controlled differential LNA with current reuse for WLNA receiver [J]," *IEEE Microwave and Wireless Components Letters*, Vol. 13, No. 12, 526–528, 2003.
4. Fukui, H., "Design of microwave GaAs MESFET's for broadband low noise amplifier," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 27, No. 7, 1979.
5. Lavasani, S. H. M. and S. Kiaei, "A new method to stabilize high frequency high gain CMOS LNA [J]," *Electronics, Circuits and System*, Vol. 3, 14–17, 2003.

A Novel Anti-collision Algorithm in RFID System

Shiyu Li and Quanyuan Feng
Microelectronic Technology Lab of Southwest Jiaotong University, China

Abstract— Tag anti-collision is a significant issue in RFID system design, which determines the identification speed. Thus a novel anti-collision algorithm based on Binary Tree Searching Algorithms and BIBD $(4, 2, 1)$ is proposed to resolve this problem. A tag is divided into several sections, which each contain special sequences. Readers can efficiently get every ID by identifying each section one by one. Mathematicalanalysis and simulations results show that the algorithm presented improves the identification speed and outperforms the existing ones. It especially fits the identification condition which contains numerous tags or tags with long ID.

1. INTRODUCTION

Radio Frequency Identification (RFID) is a contactless automatic identification technology. In a RFID system, wireless RF communication techniques are used to achieve the duplex communication between the reader and tags. RFID overcomes the limitations of other automatic identification applications such as bar codes, magnetic cards and IC cards. RFID has advantages for example fast identifying speed, data encryption, longevity and is not affected by the surrounding environment. It is especially used to cooperate with other robotic equipments. In recently years RFID is one technology which developed at a rapid speed that has worldwide applications. RFID is now a third generation identification technology.

The RFID system mainly consists of three basic components: the tag, the reader and the host data processing system. Every tag has a global unique identification code (ID). When a tag enters the reader's RF range, it is powered on by induction and sends the stored ID code or other data from its inner memory to the reader. Reader can identify all tags within the interrogation zone of its antenna via wireless RF communication technology and algorithms. Through interfaces readers can be connected to a host and the stored information from the tags can be accessed with data processing systems. Additionally, when linked with a LAN, the processed data can be transmitted over the internet and shared with other network applications.

When several tags simultaneously enter into a reader's interrogation zone, they start to transmit the stored data, each tag trying to connect to the channel. As a result, the signals interfere with each other and the so-called collision happens. Resolving collisions has been a consistent research subject in wireless communication, included RFID systems. When the collision problem is solved, RFID system works effectively. In a RFID system, there are two types of collision resolution scheme: (1) probabilistic algorithm which is based on ALOHA. (2) Deterministic algorithm which detects collided bits and splits disjoint subsets of tags. In generally, probabilistic algorithm is simpler than deterministic algorithm and is adapted in practice, but its performance would decrease sharply when the tag quantity increases. Though deterministic algorithm has complicated circuits, it has obvious advantages with its faster identification speed and higher accurate rate. Especially, deterministic algorithm fits in the identification environment which contains large quantity of tags or tags with long ID. In this paper a novel anti-collision deterministic algorithm is presented.

2. THE DESIGN OF A NEW ANTI-COLLISION ALGORITHM BASED ON BIBD$(4, 2, 1)$

2.1. Manchester Coding

Manchester coding is usually adapted for collision detection in the signals that tag returns to reader in deterministic algorithm. To be brief, in Manchester coding '10' and '01' are used to instead of NRZ coding '1' and "0" respectively. For example, there are two NRZ coding sequences: '0011'and '0101' which transform to Manchester coding are '01011010' and '01100110'. The combination is '01111110'. Because in Manchester coding there is no '00' and '11' unites, the detection result is '0xx1' which has two collision bits.

2.2. Deterministic Algorithm

2.2.1. Query Tree Algorithm

There are four reader commands in query tree algorithm, as follows:

REQUEST(SNR): when reader sends this command, every tag compares its ID with SNR. If ID $\leq$ SNR, tag sends back its ID.

SELECT(SNR): the tag whose ID is equal to SNR become active, and then reader can reader-write data of this tag.

OPERATION(SNR): this command also called 'read-write' command.

UNSELECT(SNR): it is called 'QUIET' command as well. The tag (ID = SNR) would not responds commands from reader unless it reenters in the interrogation zone.

Query tree algorithm can be described as follows:

A. SNR $= \overbrace{11\ldots1}^{n}$, n is the length of tag.
B. Reader sends 'REQUEST (SNR)', the tags (ID $\leq$ SNR) respond back their IDs.
C. If collision in back sequence is found by reader, the first collision bit in back sequence is reset to '0' and SNR become this new sequence. Jump to B. If there is no collision in back sequence, jump to D.
D. 'SELECT(SNR)', 'OPERATION (SNR)' and 'UNSELECT(SNR)' commands are used by reader to get the information of the tag (ID = SNR). Jump to A.
E. The identification is end when there is no tag answer for the 'REQUEST ($\overbrace{11\ldots1}^{n}$)' command of reader.

2.2.2. Dynamic Query Tree Algorithm and Binary Tree Algorithm

A identification instance can be explained dynamic query tree algorithm and binary tree algorithm. there are three tags: '0001', '0011', '1100'.

In binary protocol, a reader broadcast 0 at t_0, two tags whose IDs 0001 and 0011 will transmit next bit whose data are all 0 and increase their counters. Next time t_1, the reader broadcast 0 (second bit data), and the two tags 0001 and 0011 also responds next bit and increase their counter. But, at this time, the reader detects collision. At t_2, the reader broadcasts 0 (third bit), and only the tag with ID 0001 transmits its data and teset its counter.

In dynamic query tree protocol, the reader requests their ID with empty prefix, and all tags can transmit their ID. As a result, received four bits are totally corrupted. Next, the reader requests with prefix 0, the tags with ID 0001 and 0011 will transmit their remaining bits $[0 \times 1]$. The reader can know the third bit is in collision, and it requests with prefix 000. Only one tag whose ID is 0001 will transmit fourth bit as one.

Dynamic query tree algorithm, which saves half repeated and useless bits, is very similar as and origin from query tree algorithm.

2.2.3. BIBD(4, 2, 1)

Balance Incomplete Block Design $(4, 2, 1)$ includes N subsets, $N = 1\times(4\times(4-1))\div(2\times(2-1)) = 6$. All the subsets of BIBD$(4, 2, 1)$ are the sequences as follows: 0011,0101,0110,1001,1010,1100.

2.2.4. A Novel Anti-collision Algorithm based on BIBD(4, 2, 1)

To improve the identification performance, a novel anti-collision algorithm based on Balance Incomplete Block Design $(4, 2, 1)$ and Binary Tree Algorithms is proposed. In this new algorithm, Tag ID is divided into several sections and each section is only equal to one of the subsets of BIBD$(4, 2, 1)$. If the length of tag is 64 bit, the maximum tag quantity in this system is 6^{16}.

Like query tree algorithm, four reader commands mentioned above are also used. According to Manchester coding principle, 00 and 11 coding can be used to act as control commands or parameters. In this algorithm 11 is used to instead of $\overbrace{11\ldots1}^{n}$ to economize communication time and bits, and n is the length of tag.

There is a counter named TC in every tag. Reader can get every UID efficiently by identifying sections form the highest one to the lowest repeatedly. The identification process is presented as follows:

A. SNR = 11.

B. Reader broadcasts 'REQUSET(SNR)' command.

IF SNR = 11: tag sends back its first section (4 bits), and its TC = 5;

IF SNR ≠ 11: If tag current section is equal to or less than the SNR, tag sends back its next section and TC = TC + 4 else tag would not answer query commands this loop.

C. According to the principle of Manchester coding and the collision of feedback sequence from tags, the reader can tell which subsets are in the next section and then evaluate the minimum value in these subsets to SNR. Jump to B step to identify the following sections.

D. When the reader receives the last section of tags, it can make sure at least one tag and can uses Select, Operation and Unselect commands to operate it or them.

E. Finally, this loop of searching finishes and then jumps to the first step.

F. The identification finishes when there is no tag answer for the 'REQUEST (11)'.

The existing subsets can be get from collision sequence because BIBD(4,2,1) only includes 6 subsets:

A. 2 collision bits in back sequence: there are 2 subsets whose collision bits are 01 and 10 from left bit to right. Ex. From 0x1x, 0011 and 0110 can be identified.

B. 3 collision bits in back sequence: there are 3 subsets in this sequence. When the no-collision bit is 0, the collision bits are 011, 101 and 110 from the highest bit to the lowest respective. When the no-collision bit is 1, the collision bits are 001, 010 and 100. Ex. From 0xxx, 0011, 0101 and 0110 can be identified.

C. 4 collision bits in back sequence: There are 2–4 subsets in this sequence. round identification method is used.

3. PERFORMANCE ANALYSIS

3.1. Mathematical Analyses

There are N tags in radio zone and spending in command time, prefix, postfix, check, etc does not take into account.

The mean probability query times of query tree algorithm is $L(N)$, $L(N) = \log_2 N + 1$. Querying a tag needs t_1 seconds, $t_1 = L(N) \times n \times 2 \div b\,(\mathrm{s})$. b is transmitted bit efficiency, and n is the length of tag.

In dynamic query tree algorithm, the length of SNR is only equals to a half of query tree algorithm's, so the identifying speed is the twice than query tree algorithm's.

The maximum query time in the new algorithm based on BIBD$(4, 2, 1)$ is fixed equal to t_2, $t_2 = (n \div 4 - 1) \times 4 \times 2 \div b\,(\mathrm{s})$.

To sum up, with the increasing of tags, the identifying time is linearly and the adding time is a constant (t_2) in the new algorithm while it is progressional and non-linear and the increasing time is a variable (t_1) in query tree algorithm. When $N = 1$, the query tree algorithm has the best performance but it still worse than the new one.

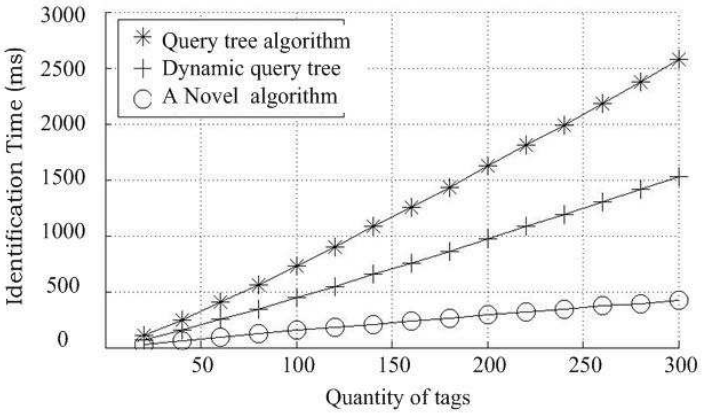


Figure 1: The identification time of different quantity of tags (ID length = 16 bits).

3.2. Simulation Analyses

Two simulation conditions which refer ISO18000-6 criterion and do not consider other spending are presented.

The first: the length of tag is fixed equal to 16 bits; the quantity of tag in the zone is from 20 to 300, and the step is 20; the bit rate is 40 kbps; simulation degree is 30. the result is shown as Figure 1.

The second: the length of tag is from 16 to 64 bits, and the step is 8; the quantity of tag is 250; the bit rate is 40 kbps; simulation degree is 30. the result is shown as Figure 2.

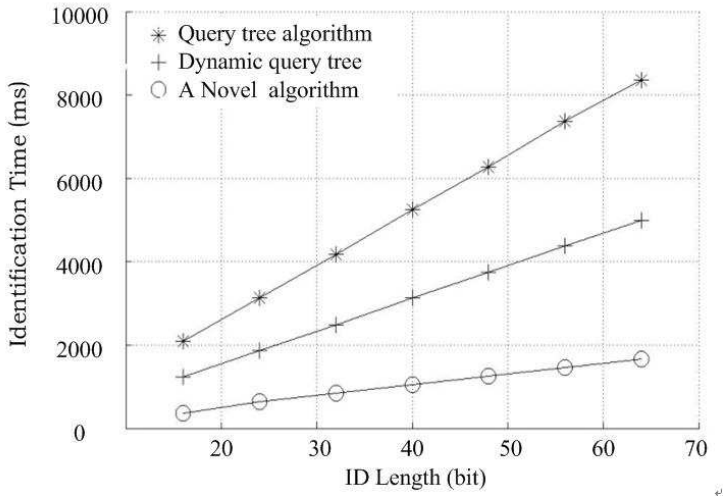


Figure 2: The identification time of different ID length (250 tags).

The simulation analyses results showed that the algorithm present improved the identification speed and outperformed existing ones. The searching speed of the new algorithm is six times that of the Query Tree Protocol. It is especially suitable for the identification conditions which contain numerous tags or tags with long UID.

4. CONCLUSIONS

Mathematics analysis and simulation results show that the algorithm presented improves the identification speed and outperforms the existing ones. It especially fits the identification condition which contains more tags or tags with long ID. the 00 and 11 coding can be used to predigest control commands or check coding. Back mechanism should be introduced ulteriorly to optimize this new algorithm.

REFERENCES

1. Liang, B., A.-Q. Hu, and Z.-Y. Qin, "Trends and brief comments on anti-collision techniques in radio frequency identification system[C]," *The 6th International Conference on ITS Telecommunications, Chengdu, China*, 241–245, IEEE Press (IEEE Catalog Number: 06EX1240, ISBN: 0-7803-9586-7), June 2006.
2. Seol, J.-M. and S.-W. Kim, "Collision-resilient multi-state query tree protocol for fast RFID tag identification[C]," *2006 International Conference on Computational Intelligence and Security, Guangzhou, China*, Vol. 2, 1159–1162, November 3–6, 2006.
3. Liu, L., Z. Xie, J. Xi, and S. Lai, "An improved anti-collision algorithm in RFID system[C]," *2005 2nd International Conference on Mobile Technology, Applications and Systems, Guangzhou, China*, 5, November 15–17, 2005.

10 GHz Two-stage Class A RF Power Amplifier in a 0.25 μm CMOS Process

Tanya Vanessa F. Abaya and Marc D. Rosales
University of the Philippines in Diliman, Philippines

Abstract— This report describes the methodology used in designing a CMOS Radio Frequency (RF) power amplifier (PA) operating at 10 GHz. The TSMC 0.25 μm CMOS process is used for design, simulation and testing. The PA is a two-stage amplifier with both driver and output stages having class-A operation supplied by a singled-ended 2.5 V source. Key specifications and corresponding characteristics are presented. Important design considerations include output power, compression point, power added efficiency and linearity. The amplifier has a 1 dB compression point of 13.18 dBm (output referred) and 1.8 dBm (input referred). For power gain, a 17.2 dB value is achieved. Linearity is quantified by an output referred harmonic IP3 of 26.17 dBm, an intermodulation IP3 of 12.2 dBm and more importantly an IM3 at compression of 31.2 dBc. Lastly, a maximum power added efficiency of 8.1% is realized.

1. INTRODUCTION

Devices of a 0.25 μm CMOS process are a relatively old technology. However, this technology is cheaper compared to HBTs or shorter channel lengths CMOS and at the same time can be more easily adapted for design in amplifiers operating in the low gigahertz range.

The power amplifier discussed is a two-stage amplifier with both stages having class A operation at a 10 GHz resonant frequency. While output power is maximized and linearity is best, the efficiency of the class A PA is least among all PA classes. Such design issues and trade-offs complicate the design and makes it difficult for the designer to achieve all specifications. In the following text, the successes and failures of the PA are described.

The discussion will start with an explanation the design process, followed by a brief description of testing. Finally, the results are presented and some conclusions and recommendations are given. An appendix is also included for further reference.

2. DESIGN

The design stage consists of an initial paper design, schematic modeling and modifications.

2.1. Paper Design

The basic PA topology shown in Figure 1 is implemented in this design. A single MOS transistor is used and loaded with a RF choke, which acts like an ideal current source. A largevalued blocking capacitor (set at 10 pF) is placed at the output to prevent DC dissipation in the load.

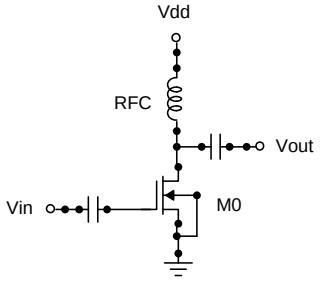


Figure 1: Class A power amplifier.

The calculations start with determining the optimal load for a given output power. Equation (1) is utilized for this purpose. Because 18 dBm (63 mW) of power must be delivered at the compression point, a maximum output power of 200 mW is assumed. Equation (1) yields an optimum load resistance of 15.625 Ω. This high a power was also set to be able to account for other losses. The optimum load would result to a required DC drain current bias of 160 mA, which should at least be

the maximum peak RF current. The transistor would then be designed to supply about 320 mA.

$$R_{\max} = \frac{V_{DD}^2}{2P_{\max}} \tag{1}$$

To properly set the DC conditions, a current mirror is used as biasing circuit. A constant gate voltage biasing may be done but this would potentially cause thermal drift problems. Here, the output transistor of the current mirror is the amplifier transistor M0 (refer to Figure 2). On the other hand, the biasing transistor M1 is sized smaller than the PA transistor to minimize current consumption [1]. A ratio of 100:1 is chosen for the widths. Consequently, the current ratio is the same.

The width of M0 is computed from the MOS saturation drain current equation. The transistor lengths are set to the minimum 0.25 µm. The width is chosen so that vgs of the mirror transistors would be about 2.5 V for maximum effective voltage [2].

As for the RFC, it should have a large enough inductance to pass through a fairly constant current [1]. Let us assume that a RFC inductance which is ten times larger than the optimum load is large enough. In this case, a value of approximately 2.5 nH will suffice.

At this point, no effort was made yet to match the optimum load to the required load resistance of 50 Ω. The input matching network was also set aside first at this stage of the design.

2.2. Schematic Simulation

The circuit schematic diagram is now set up for simulation. The inductors are first assumed to be ideal. Also, an ideal current source is used. The DC biasing and operation of the transistor is checked. Next, the output vs. input power is plotted to see if the PA has gain and see if it has enough output power. A high enough output power is noted even at input powers greater than 0 dBm in anticipation of a driver stage.

An iterative process follows consisting of changing the transistor width and finding the optimum load to achieve a high output power.

To obtain the optimum load of the amplifier, load-pull measurements were made [3]. The reflection curves are also plotted for stability purposes. Impedance matching using the Smith chart is done. After the output matching, verification is done through another set of simulations.

For maximum power transfer, input matching is applied. Again, Smith chart matching is done.

2.3. Driver Stage

Because the amplifier cannot reach the requirements set for its operation, a driver stage is designed. Again, a class A amplifier is designed to be cascaded before the previously designed output stage [4]. As with the output stage, paper design was done first. An optimal load for an output power of 20 mW is now determined as the first step of the design. Simulations and tweaking followed hand computations.

Table 1: PA key characteristics.

Specification	Obtained Value/Remark
1 dB compression point (output referred)	13.18 dBm
1 dB compression point (input referred)	1.8 dBm
Gain	17.2 dB
Power Added Efficieny	8.1% maximum
Linearity-Harmonic IP3 (output referred)	26.17 dBm
Linearity-IMD IP3 (output referred)	12.2 dBm
Linearity-IM3 at compression	31.2 dBc
Stabilty	Stable
Supply	2.5 (single-ended)
Inductor Q	14 (series resistance model)
Number of Stages	2

After a couple of simulations and changes, the input impedance of the two-stage amplifier is determined and consequently matched to the 50 Ω-source.

3. TESTING

Useful analyses applied include biasing and power related measurements. *S*-parameter analysis was also useful for impedance matching. The linearity of the PA was checked by applying a second tone of 10.001 GHz. Also, operation with a 360° conduction angle and stability was ensured [5].

As with any RF circuit, the design and implementation of inductors are problematic. The inductors were modeled and the final components used have a quality factor of 14. The final testing (and design) of the two-stage amplifier consists of inductors with a Q of 14. Also, the current sources are replaced with resistors. The iterative process described in II is repeated. The characteristics of the PA were degraded and as such needed a large amount of time for redesigning.

4. RESULTS

The final design of the power amplifier is shown in Figure 2. Table 1 presents key characteristics and considerations noted in designing and in the performance of the amplifier. The non-ideal inductors (though good) have caused the drop in output power and the consequence being the difficulty of bringing the power levels back up by adjusting matching networks and even biasing.

The gain within the range of operation is quite high and relatively linear.

Low power added efficiency is expected due to the nature of class A amplifiers, even more so for two series connected PAs of this type. The value obtained is comparable to other works [4].

The linearity of the amplifier is quantified by the IP3 measurements for harmonic and intermodulation distortion.

Lastly, the power amplifier designed is guaranteed to be stable.

An appendix is added to this paper to show plots supporting the values in Table 1.

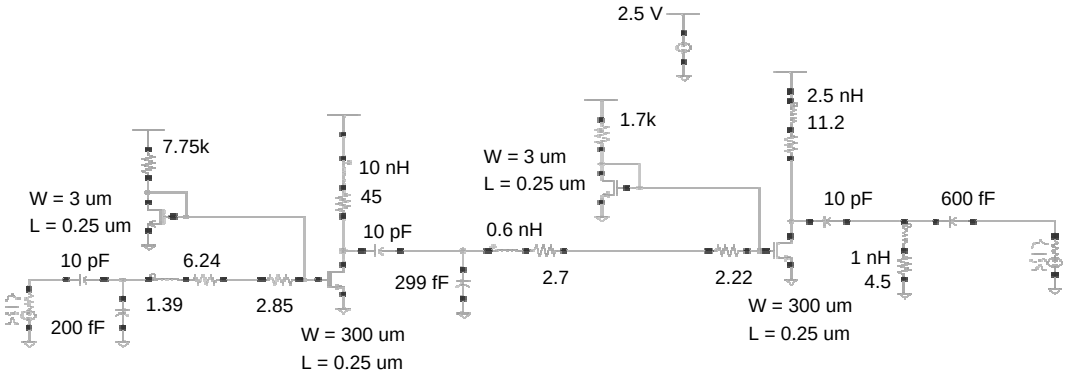
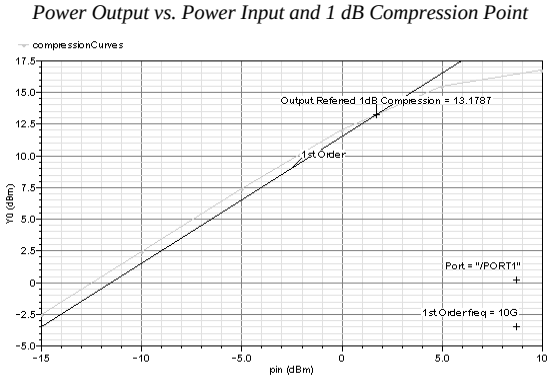
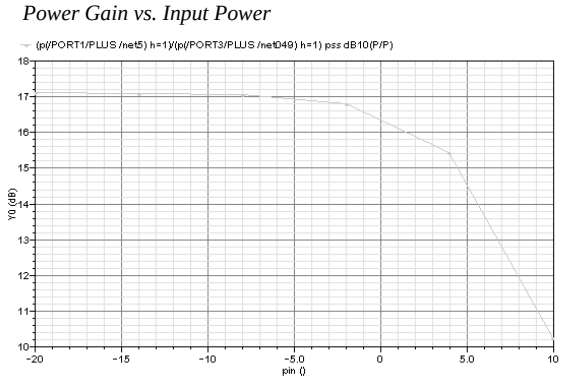
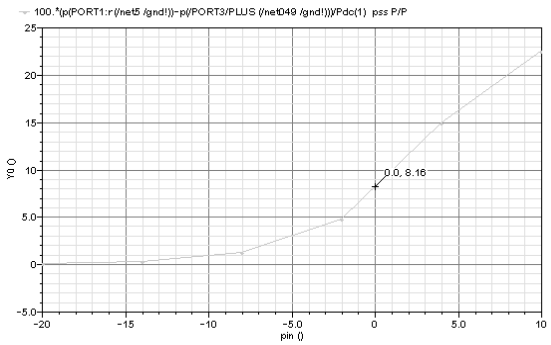


Figure 2: Final power amplifier design.

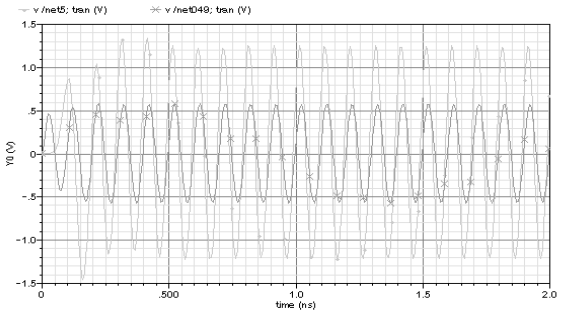




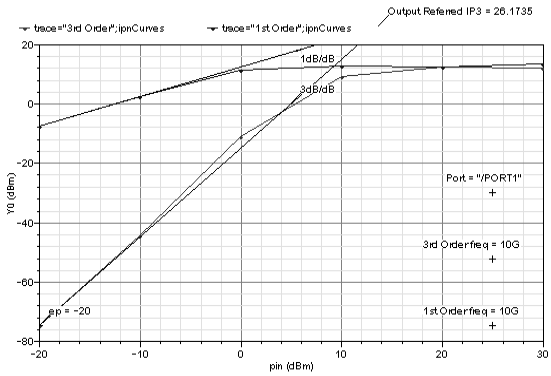
Power Added Efficiency vs. Input Power



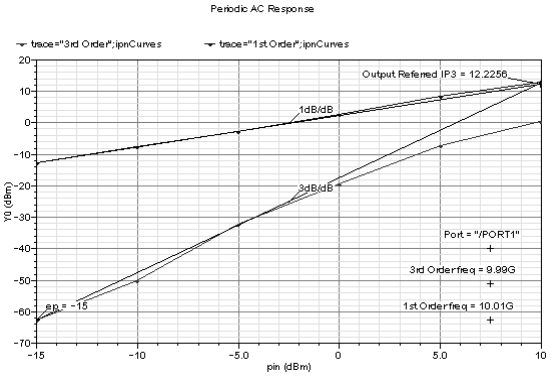
360-degrees Conduction Angle (Class A biasing)



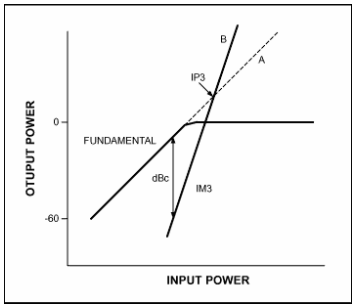
Third Order Intercept Point-Harmonic



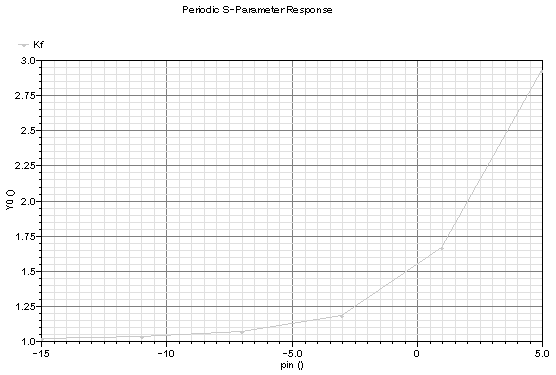
Third Order Intercept Point-Intermodulation Distortion



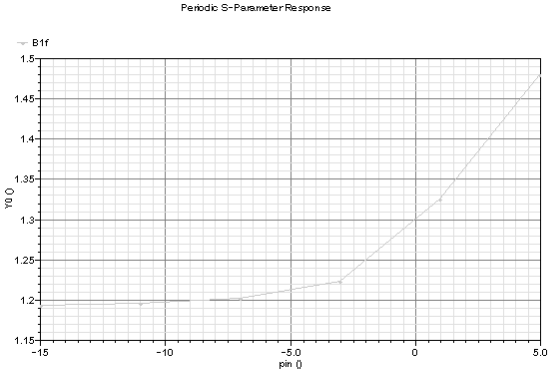
Determining IM3 Linearity [6]



Stability (Kf)



Stability (B1f)



5. CONCLUSION

Design of the PA relied heavily on hand computations for biasing and parametric analysis to ensure the optimum values of matching network elements. The PA designed was not able to meet the required output power but has other operation characteristics comparable to other designed two-stage class A amplifiers.

ACKNOWLEDGMENT

If this power amplifier were to be redesigned, more output power might be set by transforming the 50-ohm resistance to a lower output impedance. Also, a class-AB power amplifier should be explored for the same set of specifications and performance to achieve higher efficiency.

APPENDIX

This section contains plots supporting the acquired results. A subsection on determining IM3 linearity is also included and discussed.

The plots here are from power, DC, transient, linearity and stability measurements.

From the two previous plots, a comparison between the output power from the fundamental and the output power from the third IM, which is in this case 9.99 GHz (a second tone of 10.01 GHz is applied), is made. It is noted that at the 1 dB compression point (~ 1.2 dBm input referred), the third order linearity is about $13.187\,\text{dBm} - (-18\,\text{dBm}) = 31.2\,\text{dBc}$.

REFERENCES

1. Lee, T., *The Design of CMOS Radio-Frequency Integrated Circuits*, Ch. 13, Cambridge University Press, Cambridge, 2001.
2. Filanovsky, I., *EE 671 Lecture Notes CMOS Power Amplifiers in Wireless Communications Systems*, University of Calgary.
3. Razavi, B., *RF Microelectronics*, Ch. 9, Prentice-Hall, New Jersey, 1998.
4. Aniktar, H., et al., *Highly Linear Power Amplifier Design*, RISC Division, Aalborg University.
5. *PA Design Using SpectreRF Application Note*, Version 5.0 Cadence, Dec. 2003.
6. http://www.maximic. com/appnotes.cfm/appnote_number/2041/

Inductor Modeling Using 3D EM Design Tool for RF CMOS Process

Gian Paolo T. Mayuga and Marc D. Rosales
Microelectronics and Microprocessors Laboratory
Department of Electrical and Electronics Engineering
University of the Philippines Diliman, Quezon City, Philippines

Abstract— This study formulates an RF IC design flow integrating the use of 3D Electro-Magnetic (EM) design tool for inductor modeling. The 3D EM tool was used to extract the *S*-parameters that will be subsequently used as RF model for the on-chip spiral inductors in RF IC simulations. The physical layers of the CMOS process were modeled into the ASITIC 2D 1/2 tool by creating a user defined technology file (.tek), and into the 3D EM tool by CAD-supported physical geometric construction. Inductance values were chosen and geometric definitions of the inductors were determined. Comparisons were made between the gathered data from a previous on-chip implementation to the data gathered from the 3D EM tool. *S*-parameters, quality factor (Q), and the inductances versus frequency measurements were recorded. Implementation was done for a 0.25 micron CMOS process.

1. INTRODUCTION

Wireless communications has become a standard consumer commodity in everyday life. Not only are cellular phone handsets are in common persons' hands, but also Bluetooth and WLAN connectivity are as easily accessible. This 'wireless' revolution is driving consumers to be continuously connected in a mobile environment. This has also driven researchers to push the envelope to curtail to these demands. Implementation of integrated RF front-end components, digital circuitry and on-chip passive components has become the focus of these researches. RF CMOS process is a satisfactory choice for implementation of RF building blocks due to low-cost, high level integration and availability. However, introduction of on-chip passive components, such as inductors, in RF CMOS presents complications due to lossy substrate issues that limit the frequency range of operation [1].

Spiral inductors have been extensively used in modern RF CMOS ICs. Modeling for inductors has been developed in various studies in aid of inductor design. Software plays a large part in aiding ease of design. Maximizing available modeling tools help in providing accurate characterization for on-chip inductors. As improved software becomes available, new methodologies then must be adapted to successfully implement RF CMOS design in current wireless standards [2].

In line with this, it becomes imperative to investigate new methodologies and integrate them into an existing RF IC inductor design flow, as shown in Figure 1, and in general RF IC design flows as well. 3D EM design tools are industry-standard software for *S*-parameter extraction and use 3D electromagnetic field simulation. This will be implemented and compared to a previous implementation consisting of 3 inductor structures using ASITIC 2D 1/2 tool with actual on-chip results [3]. A 0.25 micron CMOS process was used for this implementation.

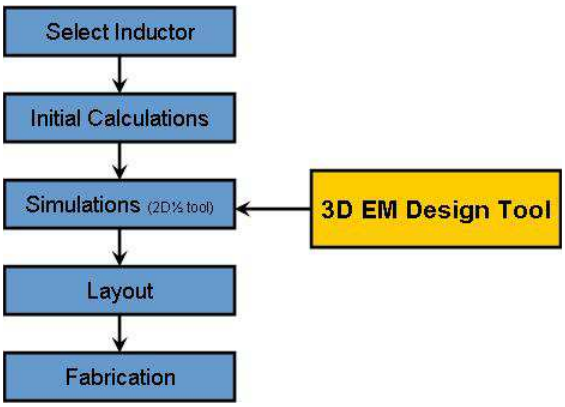


Figure 1: Design Flow for RFIC Inductor.

2. DESIGN METHODOLOGY

2.1. Spiral Inductor Geometry

Planar spiral inductors are the most widely investigated inductor structures. It is because it is much compatible with the interconnect scheme of integrated chips [4]. Integrated chips usually have 3–5 levels of interconnects. The topmost layer is the thickest with typical values of 1–2.4 μm while the lower layers are typically 0.7–1.0 μm thick. Planar spiral inductors can be implemented as square spiral, hexagonal, octagonal and circular spiral [5]. Square spirals, which have 90 degree segments, are more commonly found in fabricated circuits due to early difficulties in non-Manhattan geometries in design layout tools and fabrication foundries. Square spirals were chosen for this implementation (shown in Figure 2).

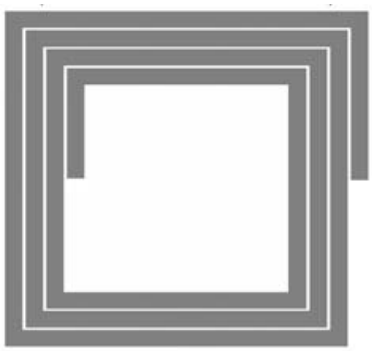


Figure 2: Square spiral structure.

The basic planar inductor structure requires at least two metal layers. The upper metal layer is used to make the spiral structure. The lower metal layer is used as underpass to connect to the inner port of the spiral. Planar spiral inductors are implemented according to geometry design parameters. The following are the geometry parameters: 1) length of the external side of the inductors (*dout*), 2) width of the metal (w), 3) metal to metal spacing (s), 4) number of turns (n), and 5) thickness of metal.

The major drawback in the design of integrated spiral inductors in most sub-micron CMOS technologies is the use of epitaxial wafers which have a heavily doped substrate [6]. Substrate losses result from ohmic losses associated with currents flowing within the semi-conducting silicon substrate below the spiral inductor.

Spiral inductance can be solved using Greenhouse's algorithm. It is the summation of the self-inductances of all the straight segments and the sum of all the positive and negative mutual inductances between all possible wire pairs [7]. Empirical formulas are also used in determining spiral inductance, which are ideal for doing initial design and could be easily integrated in the extraction part in a typical IC design process. These are as follows: 1) Modified Wheeler Formula 2) Current Sheet approximation 3) Data fitted monomial expression. The formulas presented are reported to have 1% to 3% error when compared to field solver simulations and measured data. These expressions are acceptable in circuit design and optimization because process variations can cause the parasitic capacitance in the inductor model to vary around 5% to 10%. These variations translate to an uncertainty in the impedance of the spiral that is of the same order of magnitude as the errors introduced by the expressions [5].

An optimum design of a spiral inductor for a given inductance value should aim for a maximum Q and a minimum area consumption with sufficiently high selfresonant frequency. This means that the inductors physical characteristics such as metal width, outer diameter, spacing and number of turns should be adjusted to yield an optimum inductor. Several guidelines have been recommended in aid of realizing optimal square on-chip inductors in the lower GHz range:

1) Planar spiral inductors should be hollow. At high frequencies, the innermost turns of the coil suffer from enormous increase in resistance, while their contribution to the inductance value is minimal [6].

2) Opposing sets of coupled lines must be separated at least 5 times the width of the wire segment w to allow enough magnetic flux to pass through the hollow part of the spiral [8].

3) The width of the metal inductors should not be too wide. It is reported that a 10um to 15um strip width is close to optimum for the Q factor when the frequency of operation is 1–3 GHz [6].

4) Limit the area occupied by the coil. The magnetic field of small coils penetrates less deep into the substrate. The substrate losses are less severe [6].

5) Use minimum allowable spacing between metal segments. The magnetic coupling will be maximized. Tight coupling of magnetic field maximizes the Q factor and reduces the chip area [8].

6) Use the highest metal layer to implement the spiral inductor. This makes the oxide layer as thick as possible that will reduce the substrate parasitic capacitance [8].

7) The space between the outer spiral turn and any other surrounding metal structures should be at least 5w to avoid unwanted coupling.

2.2. Inductor Implemented Using 2D 1/2 Design Tool and Actual On-chip Results

Typical values used for inductors in RF modules are in the range of 1 nH to 10 nH. Chosen inductors are in the range of 2 nH to 3 nH to be able to characterize more inductors given a limited area. Square planar spiral inductors are to be characterized. For actual on-chip inductor measurements, open de-embedding is used to cancel the effects of the test fixtures where the inductors were placed.

Given the required inductance for the square spiral structure and using the guidelines for optimum inductor design, the geometric parameters were calculated. In aid of calculation, *SpiralCalc* [9] was used. The geometry parameters are inputted, and the tool calculates the inductance of the structure. The tool solves the inductance using the Modified Wheeler, Current Sheet and Monomial Fit expressions. The inductor values are as follows:

Table 1: Summary of selected inductors and geometric parameters.

Square Spiral Inductor	Dimensions				Expected Inductance(nH)
	n	w	s	$dout$	
L1	3.5	10.88	4	160	1.8
L2	3.5	10.5	2.5	150	1.8
L3	3.5	10.1	3	180	2.7

Calculations of inductance alone are not sufficient to represent the inductors. It is also imperative to consider the associated parasitics in the inductor to fully represent them. ASITIC 2D 1/2 tool is used to simulate inductors with their associated parasitics. ASITIC uses electromagnetic simulation to analyze spiral inductors and was designed to aid in optimizing them [10]. Using a command line that accepts the geometry parameters of the inductors, the tool then generates the spiral inductor. The CMOS process parameters are modeled into a user defined technology file (.tek). In this case, the process parameters of the 0.25 micron CMOS process was described, specifically the substrate and metal layers information. ASITIC is deemed as a 2D 1/2 tool since the inductor is 'drawn' in x and y-coordinates while the '1/2' aspect is represented by indicating the thickness of the metal layers in the technology file (.tek).

On-chip results were extracted using open deembedding, details of which can be found here [11]. Open deembedding was found adequate to characterize inductors in the range of 1 nH to 3 nH for frequencies up to 6 GHz. ASITIC tool proved as a satisfactory simulation tool in aid of designing inductors.

2.3. 3D Design EM Tool

In utilizing a 3D electromagnetic tool, the S-parameters of the inductor structures will be generated from a 3D full wave electromagnetic field simulation and will be solved using finite element method. This is the main motivation in exploring the use of 3D tool. This is in comparison to the 2D 1/2 tool which uses solving from lumped circuit elements.

The parameters of the CMOS process are implemented by drawing the substrate and metals in layer by layer 3D boxes fashion (shown in Figure 3). The layers are drawn by manually inputting the 3-dimensional coordinates to form 3D boxes/layers. Each layer is defined by properties of relative permittivity and bulk conductivity. The 3D boxes are implemented from the bottom substrate up to the top most metal layer. The actual thicknesses are not mentioned since it is bound by a confidentiality agreement. The top two metal layers and via layers are then defined where the square spiral will be placed. The spiral is drawn using the same stated dimensions. The spiral is

placed in the top metal layer by defining relative coordinates so it will coincide. Vias are added as well as the underpass metal layer from the next topmost metal layer. A ring was added that connected to each port of the inductor. After solving for convergence for the 3D construction of inductor, the S-parameters, quality factor and inductance versus frequency data are extracted.

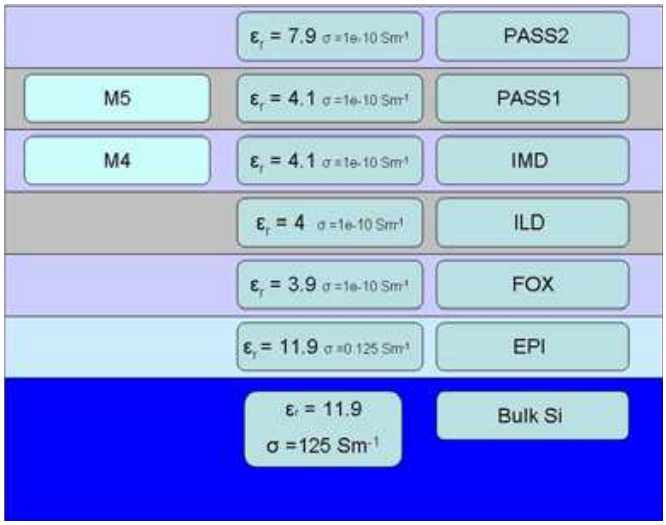


Figure 3: Process layers.

3. SIMULATION RESULTS AND ANALYSIS

The data retrieved from the 3D EM tool and ASITIC were compared. The data from the 3D tool was extracted and plots were generated against the frequency range. The data from ASITIC was also extracted, but was manually plotted for quality factor and inductance against frequency. Open deembedding was used as the method of extraction for on-chip inductors, wherein the measurements of an open test fixture is subtracted from a test fixture including the inductor. Compared to three-step de-embedding which utilizes a more complicated method of data extraction, open de-embedding provides simpler extraction with almost similar results. Measured data results shown used open de-embedding.

Inductance values were compared for ASITIC, 3D EM tool and measured results (shown in Figure 4). Minimum deviations can be seen from the results compared to the actual on-chip data, which shows both ASITIC and 3D EM tool as good simulators for inductance. This can be attributed to inductance as a function of geometry parameters.

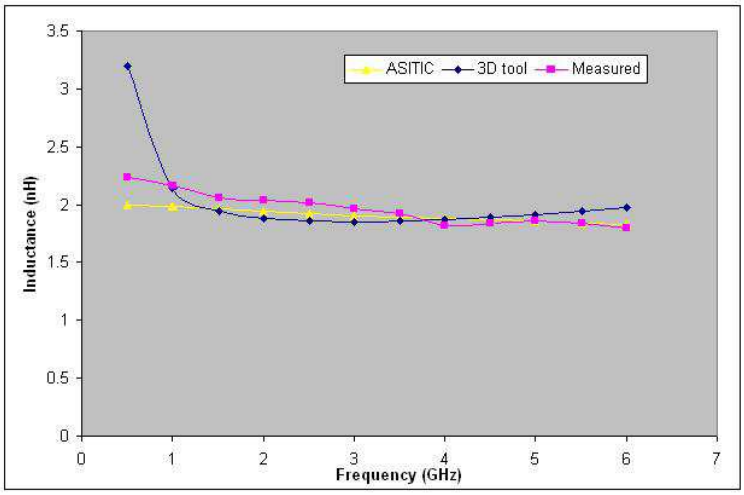


Figure 4: Inductance versus frequency.

Quality factor were compared for ASITIC, 3D EM tool and measured results (shown in Figure 5). It can be noted that the ASITIC (with eddy current off) and 3D tool results are comparable to each other but they are way off compared to the measured data. It can be noted that the eddy current engine is off for the ASITIC data where 3D tool is comparable. But when the eddy current engine is turned on, wherein ASITIC models eddy currents for conductors and bulk in "Manhattan" geometries, the Q measurements for ASITIC are found to be very close to the measured data results. It is found that the sheet resistance of metal layers gives the pronounced effect on simulated inductor Q, which can be indicated in the technology file (.tek) in ASITIC. As of this writing, the 3D EM tool is yet to be explored enough in incorporating the sheet resistance to the metal and should be further studied. Thus, ASITIC remains a valuable choice in simulating inductor quality factor (Q).

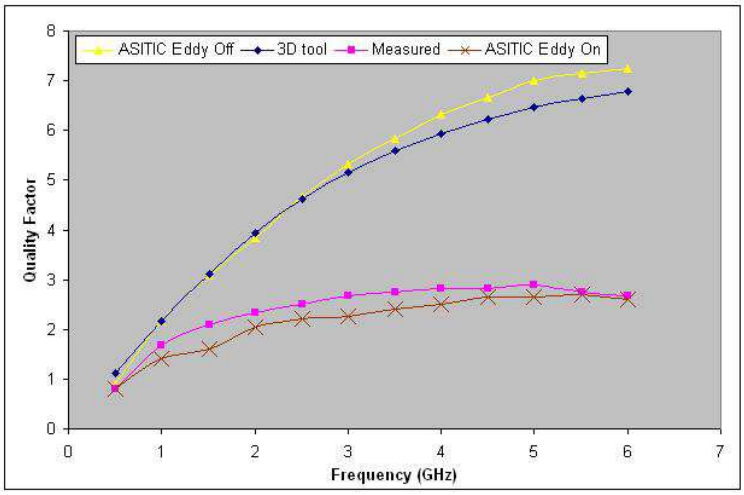


Figure 5: Quality Factor versus frequency.

The $S11$ and $S12$ parameters were compared for 3D EM tool and ASITIC (as shown in Figure 6). Cadence Schematic Design tool was used to generate the S-parameter plots. S11 plots show that an inductor is represented, as the plot is above the real axis. An ideal representation for the inductors is that they are short at DC point, thus must follow the impedance (Z0 or R) of 1 or matched to 50-ohms. It is seen that the 3D EM tool plot is closer to the impedance(R) of 1, which follows a clockwise turn in the plot, as compared to the ASITIC plot. This shows an improved capability of the 3D EM tool in terms of S-parameter extraction in characterization of the inductor. $S12$ plot shows minimum deviation between 3D EM tool and ASITIC plot.

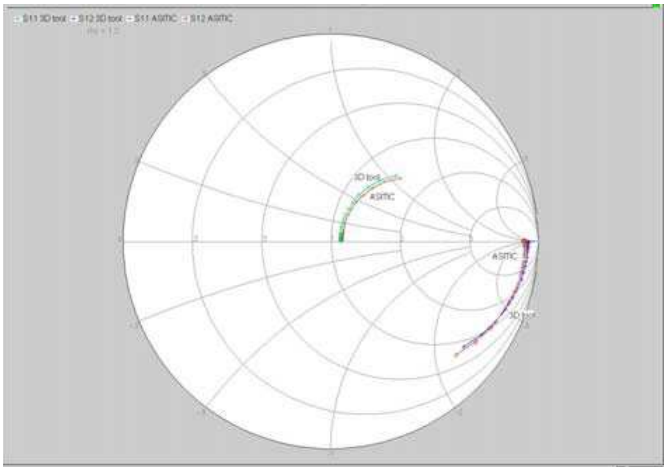


Figure 6: $S11$ and $S12$ parameters on Smith Chart for 3D EM tool and ASITIC.

Data were extracted for both L2 and L3 in practically the same manner. Similar trends can be seen for both L2 and L3 in terms of inductance, Q and the S-parameters. These trends prove

the capabilities of the 3D EM design tool when integrated into the RF IC design flow. It is to be noted that a black box approach was done in implementing the inductors in this study, wherein the inductors were pre-selected and then characterized. The 3D EM tool improves inductor selection in RFIC systems, using the S-parameter data as RF model for the inductor.

4. CONCLUSION

The 3D EM tool improves the design methodology in implementing on-chip inductors. It provides improved S-parameter generation by utilizing full-wave electromagnetic field simulation. Though the layout-driven 3D tool gives full freedom in implementation, it proves to be a slower tool due to modeling through geometric construction and its use of finite element method that requires iterations for convergence. ASITIC 2D 1/2 tool remains an important complementary tool for its reliability, speed, and ease of use.

REFERENCES

1. Lee, T. H., *The Design of CMOS Radio Frequency Integrated Circuits*, Cambridge University Press, Cambridge, UK, 1998.
2. Vye, D., et al., "New design methodologies for high performance RF CMOS," *Microwave Journal*, Vol. 49, No. 11, 148, November 2006.
3. Rosales, M. D., et al., "Monolithic spiral inductors for a 0.25 µm digital CMOS process," *Proc. IEEE Region 10 Conference*, Nov. 21–24, 2005.
4. Yue, C. P., et al., "On-Chip spiral inductors with patterned ground shields for Si-based RF IC's," *IEEE J. Solid-State Circuits*, Vol. 33, 743–752, May 1998.
5. Mohan, S., et al., "Simple accurate expressions for planar spiral inductances," *IEEE J. Solid-State Circuits*, Vol. 34, 1419–1425, 1999.
6. Craninckx, J., et al., "A 1.8-GHz low-phase-noise CMOS VCO using optimized hollow spiral inductors," *IEEE J. Solid-State Circuits*, Vol. 32, 736–744, 1997.
7. Greenhouse, H. M., et al., "Design of planar rectangular microelctronic inductors," *IEEE Transactions on Parts, Hybrids, and Packaging*, 101–110, 1974.
8. Long, J. R., et al., "The modeling, characterization, and design of monolithic inductors for silicon RF IC's," *IEEE J. Solid-State Circuits*, Vol. 32, 357–369, 1997.
9. "SpiralCalc," http://smirc.stanford.edu/spiralCalc.html.
10. Niknejad, A. M., et al., "Analysis and optimization of monolithic inductors and transformers for RF IC's," University of California, Berkeley, California, 1998.
11. Rosales, M. D., et al., "De-embedding techniques on a 0.25 µm digital CMOS process," *Proc. IEEE Region 10 Conference*, Nov. 14–17, 2006.

On the Absolute Measure of Wavelengths in Telecommunications

S. L. Vesely[1] **and A. A. Vesely**[2]

[1]I. T. B.-C. N. R., Italy
[2]via L. Anelli 13, Milano 20122, Italy

Abstract— Physics is considered the quantitative scientific discipline par excellence, as it expresses measured amounts as numbers of units of measure. Furthermore, nowadays it is possible to derive any unit of measure from a few constants, deemed fundamentals, such as the speed of light in a vacuum. Although telecommunications science proficiently uses logarithmic ratios for studying the transmission and reception of electromagnetic signals, it relies on absolute frequencies and wavelengths in theoretical matters. Here we suggest that, data transmission being the main concern, telecommunications may quit pursuing absolute measures of wave related quantities, with no attenuation whatsoever of either the precision of measurements or the physical interpretation of the phenomena.

1. INTRODUCTION

Even if nowadays it is customary to avoid according a high degree of objectivity to spare numeric data, for exact sciences the concept prevails that numbers, associated with the relevant unit of measure, express quantitative properties directly. Accordingly, any measurement is expected to be displayed as the simple *ratio* to the corresponding unit, although units of different sizes and special units of measurement can be used. Traditionally, electrical engineers leverage that opportunity by making practical use of logarithms to speed up numerical computations, to express values in wide ranges, or just because theoretical filters' amplitude responses have symmetrical skirts when plotted on bilogarithmic paper. Those practices imply that the logarithm of a measured quantity is accorded the same degree of objectivity and the same reproducibility as the "raw datum". Since today transmitted information is more tied to the signal than, if at all, to the absolute frequency of the transmission carrier, it may be interesting to consider if transforming back logarithmic values into absolute quantities is a necessary step in telecommunication sciences. If not, as an alternative to logarithmic values to express measurements, a *logarithmic measure* could be autonomously based on the cross-ratio. Historically, the lengths pertaining to Euclid's geometry are actually cross-ratios. In the era of the arithmetization of geometry (Hilbert's foundation), F. Klein also established a logarithmic measure based on the cross-ratio[1], for the purposes of non Euclidean geometries. Cross-ratios can associate a geometric model with any numeric field, possibly the complex numbers, that Cauchy characterized as numbers "$mod(x^2 + 1)$".

2. ON THE MEANING OF PHYSICAL CONSTANTS

Despite the accurate determination of numerical values has always been assessed at a remarkable significance by exact sciences, before the French revolution measurements objectivity itself was frustrated by arbitrarily choosing units of measures based on futile reasons. Thus, the degree of objectivity that the French Academy of Sciences ratified for quantitative measurements for commercial and scientific purposes, is rooted in the rationale for specifying the standard units of fundamental quantities such as mass, time, length, and their multiples. Nowadays, it is believed that any further progress consisted in the discovery of constants fixed in nature and tied to the dimensions of basic and quantized constituents of the matter, which could be determined thanks to technical advancements in measurement methods. Those advancements have been supported by electronics, often with scanty feedback by different measuring techniques, such as chemical or mechanical determinations[2]. Thus, on the one hand the values of the standard unit quantities are empirical and need to be updated when, as measuring techniques advance, experimental determinations improve. On the other hand, since high precision absolute electronic measurements are hard to make, much of the already acquired data needs to be either processed again hinging upon new constants and making use of known physical laws, or re-established resting upon dimensional analysis evidence.

[1] "Würfe" according to von Staudt.

[2] Nowadays researchers can avail themselves of optical indicators for titrations. Densitometry is optical, as well as deformation measures, etcetera.

Let's first consider the corpus of the accepted physical laws. Large parts of it fit in quantum electrodynamics, but some in other more continuum based theories, such as general relativity and electromagnetism. Before Maxwell's theory flew into telecommunications, *absolute measures* were conceived to allow a single measurement system valid for the whole of physical sciences. This understanding leaded to seek dynamical explanations of the inner working of electrical measurement devices. The need to describe electrical quantities in terms of their alleged mechanical counterparts has been upheld by conspicuous scientists, such as Lord Kelvin (absolute electrometer), W. Weber and R. Kohlrausch (speed of light). However, light emissions [1] turned out to be so different from classical constructs that they could not fit in with the dynamic approach to absolute measures sustained by the proponents of mechanical reductionism.

Now for the dimensional foundations of units of measure. Until physics and geometry are two subjects apart from each other, and the latter doctrine belongs to mathematics, it will be difficult to empirically substantiate dimensions without conceding too much to conventionalism. For example, consider the relationship between capacity and volume. As a further example, which we'll consider in some detail below, think about angular vs. linear measurements of lengths. Nevertheless, during the 1900's physical structures started to be considered isomorphic to geometry, and a progressive geometrization of physics took place as a result. In our opinion, that trend can only make sense if it also makes sense to study the physics of an empty space assuming that measurements rely upon the units of scale of it. On that subject, let's recall how E. Schrödinger associated the difficulties of microscopic physics with the possibility that length measurements that are valid in the macroscopic world may change drastically at atomic or subatomic scales.

3. LENGTH AS A PHYSICAL DIMENSION

Distance as used in land surveying does not match the requirements for being regarded as a concrete example for the linear physical dimension. Indeed, land surveyors measure objects all too different from one another, depending on whether distances are obtained on steep mountain terrains, using a theodolite rather than a graduated tape or a stick. Those techniques don't rely on the same geometrical definition of length. In particular, the Euclidean concept of distance is not equivalent to the angular difference obtained by optical telemetry, since lengths and angles are not equivalent themselves. However, it is enough to choose a projective geometry or a suitable metric on a space with constant positive curvature in order to remedy this hassle: provided that straight lines are closed curves, linear and angular distances can be defined such that they are equivalent.

Since 1983, the definition of the meter is based on a time-of-flight concept of length. The SI spells it as "the length of the path traveled by light in vacuum during a time interval of 1/299792458 of a second". A *movement* could actually bear a geometrical meaning provided that the underlying geometry were somehow built upon it, but an associated operationalist approach scarcely adds to the knowledge of the world structure. As for velocity based distance measurements, the RADAR allows to identify the positions of objects in the surrounding space w.r.t. its own position by the "echo time of flight". However, that expression is a short way of saying that the received echo contains all the data on the distance that can be retrieved from a signal. In fact the path length-data is encoded in the echoed signal, because it is possible to obtain it demodulating the carrier frequency. Since a signal depends neither on the carrier's absolute ν nor on its λ[3], the only open question is if an absolute time delay w.r.t. the carrier would allow to relate the echo modulation with metric units, irrespective of any empirical device calibration. Historically, the assumption that electromagnetic wave parameters have a physical meaning by themselves has been made by Michelson, Fabry, and Perot [2] in order to establish that the standard meter has a length equivalent to 1553163.5 complete wave lengths[4] of Cd red light in air at 14.93 °C and 760 mmHg. In their experiments, they took interferometric measures of *optical paths differences* on transparent secondary length samples. The length of the standard rod has thereafter been expressed in terms of quantized fundamental units assuming that one can recover absolute length magnitudes from *differences*. As a matter of facts, while the comparison with the prototype meter is based upon interferometry, the attribution of $\lambda = 643.84696$ nm to the Cd emission postulates an

[3]However, signals echoed from metallic objects are stronger at certain frequencies.

[4]Wavenumber expressed as 1/m. We take this occasion for recalling that the length of a stationary wave between two consecutive nodes is one half of the length exhibited by the sinusoid that represents it. "Complete" means the whole of it: In quantum mechanics, the discrete nature of the length reflects an essential characteristic of energy exchanges. Therefore, the frequency to energy conversion factor, e/h, contains *two* discrete quantities, e and h, that both express quantization as an unique fundamental property of nature.

univocal relationship between diffraction and color dispersion. Today interferential methods allow to increase the accuracy of wavelength determinations far beyond diffraction measurements, but absolute values ultimately still relay on monochromatic light's diffraction data. Besides, although grating manufacturing techniques have greatly improved since Rowland's time, the way wavelengths are assigned to spectral lines has not been altered since.

4. ON THE LENGTH OF OPTICAL WAVES

To better focus on the geometric relationship between groove spacing and optical pitch, let's consider a plane transmission grating with N grooves and a period $\ell = s + b$, where b is the transparent part of the array. Be it uniformly illuminated perpendicularly to the grating with light through a slit oriented parallel to the rulings and width δ (in a range from a few μm's to a few mm's). By placing a converging lens right behind the linear lattice, on its focal plane[5] we'll observe the image of the slit multiplied by the so-called diffraction pattern of the grating. The pattern shows a central bright spot, and further equally spaced feebler peaks arranged more or less symmetrically, the gap between one another being proportional to $1/\ell = \nu_\ell$[6]. They are the relative maxima of increasing diffraction order. Physically, only the *total* transmitted light intensity depends on the size of b. Besides, the longer the part of the grating, in our case $N\ell = L$, that contributes to the diffraction pattern, the higher the resolving power, with no increment of the number of visible diffraction orders. However, their visibility is irrelevant for geometrical reasoning. To quote some figures, by the end of the 1800's researchers managed to uniformly illuminate a grating of an overall length of an inch and three-quarters having about 30000 grooves (∼2500 grooves per inch[7]) and thereby showing less than 10 sufficiently visible monochromatic diffraction orders. In that conditions, chromatic dispersion can be observed in the spots illuminated with white light deflected from the direction of incidence. If diffraction were to be considered well described in geometric approximation, the accompanying dispersion would be encoded by a different parameter[8], or at any rate would be considered a higher order effect. Instead, at the time when absolute light wavelength measures were being taken, an *octave* consisting of exactly the whole visible spectrum, that is the numerically well ordered spectral range circa from 0.385 to 0.770 μm, used to be accommodated between the first and the second diffraction orders. Because of the constant spacing between any two adjacent diffracted monochromatic beams, the scale between, say, 400 and 800 nm, as well as that between any subsequent octave, was assumed to be *linear*[9]. That assumption allowed to determine wavelength *ratios* by Rowland's coincidence method. For this purpose, neglecting effects due to increased angles from the normal to the grating for higher orders, let the ratio λ_r/λ_s between any pair of λ be independent of the diffraction order m, and let $m(\lambda_s - \lambda_r) = m\Delta\lambda \forall m$ be true. If $n \times \lambda_r$ and $m \times \lambda_s$ coincide on the screen at order m, we have $n\lambda_r = m\lambda_s = m(\Delta\lambda + \lambda_r)$, hence $(n - m)/m = \Delta\lambda/\lambda_r$. Under the hypothesis that the color encoding is strictly linear, which is supposed true for the *normal spectrum*, the dispersion $\Delta\lambda/\lambda$ is constant, and that enables calibrating an eyepiece micrometer directly in $\Delta\lambda$. Evidently, the coincidence method can be applied if all the radiations received lie within the first octave, while spectra *overlap* only from the second diffraction order onward, thereby disrupting their sequential arrangement.

5. ROWLAND'S DIVIDING ENGINE

The absolute evaluation of one λ_r is geometrically based upon the distance ℓ between parallel adjacent lines ruled on a concave speculum metal surface[10]. Let's determine the length of the latter. As it is clumsy to mark N consecutive lines on a cylindrical surface at exactly equal distances ℓ apart by transferring the graduations from an attached meter stick without accumulating a relevant error, Rowland's method [3] was devised so as to measure angles along the circumference of a wheel with a sufficiently large diameter, and then scale them down to the desired length ℓ by means of a gear and a precision screw. That is the heart of his dividing engine. When ruling a grating in a dividing engine, a rotary motion is imparted to the shaft and the screw converts it into a uniform

[5] We are dealing with slope fringes here, not thickness ones.

[6] Nowadays grating diffraction analysis is based on Fourier expansion. The optical pitch ν_ℓ is called the spatial frequency of the diffraction pattern, while $1/\lambda$ is the wave number such that $\nu_\lambda = c/\lambda$.

[7] Nowadays grooves densities are usually expressed in g/mm.

[8] Absolute wavelengths cannot be established using a glass prism, since its dispersion is not geometrically related to the glass's structure, being glass amorphous according to diffraction studies.

[9] In this respect, λ, 2λ, 3λ, ... are regarded as different harmonics, despite they are different diffraction orders.

[10] Concave gratings with a ruled surface $6\frac{1}{4} \times 4\frac{1}{4}$ inches allow an increased resolving power $\Delta\lambda/\lambda$ and provide focusing without additional optical elements.

translation. For each given angular rotation around its axis a diamond is lowered into contact with the burnished surface and a line is ruled. The exact value of the distance ℓ_N between Nth groove and $N+1$th, is not known; however, it is known that after a complete turn $U = (180^\circ \times d)$, where d is the nominal diameter of the screw, it has moved forward by $h = U \times \tan\alpha$, where α is the thread angle. We quote a few values from Rowland: length of the screw about 9 in, $d = 1\frac{1}{8}$ in, $\alpha = \frac{1}{2}^\circ$ (720 head divisions), h about 20 threads to the inch. That screw was deployed for ruling concave surfaces of up to $5\frac{1}{2}$ in linear length with 28876 lines/in, totaling nearly 160000 grooves. Now, by definition a full rotation is *exactly* 360°, and geometrical constructions teach how to bisect angles *exactly*. Accordingly, the highest numbers of grooves Rowland has achieved are 14438, 28876, 43314, and the distance between adjacent grooves is calculated as $1/2^N$, which is a conceptually different task from the measure obtained by juxtaposing the unit of scale $u = h/720 \approx 1/14400$ for *exactly* 158818 times. Conceptually, the second measurement does not necessarily result in a $5\frac{1}{2}$ in segment. If we were to divide a concave surface of about 7 feet curvature into equal parts by a dividing engine, and afterwards found the total length slightly different than expected, or if the screw's d were a little bit smaller than initially assumed, we could only correct the "average" ℓ.

In addition, dividing engine faults, including those generated by threading and ruling, only result in deviations from a mean value, which is null on average since it is assumed to vary periodically with the rotation. Thus, the length ℓ enjoys periodical properties totally extraneous to Euclidean geometry distance measurements.

6. EUCLIDES' POINT OF VIEWS ON THE MEASURE OF GEOMETRIC MAGNITUDES

The Euclidean recipe for measuring lengths can be worded using modern terms as follows:

Given a unit of length E, compare its integer multiples with the length M to be measured so as to find an integer k such that $kE \le M < (k+1)E$ holds, then the length of M is between the given multiples. If $kE < M$, repeat the whole operation using the smaller unit of length $E' = M - kE$. In case that procedure eventually terminates[11]*, the final unit E^* is a multiple of E, hence M is itself a rational multiple of E. Otherwise, M and E are said to be incommensurable to each other, and their relationship can be expressed using ratios*[12].

That geometrical interpretation of lengths makes a peculiar use of the number system. Euclid only compares magnitudes homologous with one another. For that reason, he introduces the area A enclosed by a circle only in book X, and does not attempt to give numerical values for it. In order to get numerical approximations, the exhaustion method accepts that, given a circle, the areas of inscribed and circumscribed regular polygons with 2^N edges/angles differ by an arbitrarily small quantity provided that N is large enough, irrespective of the number of edges that can be graphically drawn and the effective numerical values of their lengths, and thus asserts that the area of the disk lies in between. It is beyond the scope of this article to investigate how the *concept* of lengths and measurements changes as number systems provide for the so-called arithmetization of geometry, thereby introducing irrational numbers and differential quantities. It suffices to contrast classical measuring methods with the procedure of ruling a grating by a dividing engine, where the transcendental number π enters as a factor in *computing* the unitary linear movement of the precision screw when it is turned by one division of the head scale, and no *single* segment of length ℓ_N is ever measured. Just to remark how far from any classically conceivable measurement we have gone, we additionally recall that $\pi = 3.141592654\ldots$ can be defined, according to Dedekind, as a cut in the rational line. The Euclidean geometry presented by Hilbert is different from the original one, and π in particular is not among the transcendental numbers that lend themselves more naturally to Diophantine approximation [5]. The old rule $A \sim p/qr^2$ (where $p/q = 3.14$, say) asserts on a graphical basis that the magnitude A is measured by the corresponding geometrical property of a square with side r multiplied by a dimensionless constant. During the 19th century, the development of geometry endeavored to provide sound analytical foundations, thus apparently didn't care much about disentangling geometrical modeling from the underlying number system.

[11] The recipe yields a rational number if and only if it terminates after finitely many operations. The practical method of juxtaposing unitary lengths, not related to number fields, is where experimental errors may occur.

[12] Book V, definition 5, commonly attributed to Eudoxus of Cnidus: "*Magnitudes are said to be in the same ratio, the first to the second, and the third to the fourth, when equal multiples of the first and the third either both exceed, are both equal to, or are both less than, equal multiples of the second and the fourth, respectively, being taken in corresponding order, according to any kind of multiplication whatever.*"

7. ALTERNATIVE λ-SCALE CALIBRATIONS

According to the current physical approach, Bragg's reflection law, $n\lambda = 2\ell^* \sin\varphi$, permits to relate the radiation to a geometrical parameter ℓ^* of a crystal lattice, taking into account the glancing angle φ, for a given diffraction order n. The hypothesis that each spot on a Laue diagram is due to some diffraction order $n\lambda$ confer periodicity to both the λ's and the ℓ^* in a "perfect and defectless", that is unbounded, crystal. Today, when the radiation sampling devices exhibit *linear* responses, as for gratings, the mathematics of space sampling is contemplated by the Nyquist-Shannon theorem[13]. Accordingly, (i) both the sampling slit of width δ and the grating length L, which generate spatial patterns at the grating and at the screen respectively, may contribute λ-ghosts; (ii) the sampling period $1/\ell$ of the grating, which is responsible for the diffraction pattern, contributes aliasing when the impinging light is not band limited; and (iii) the grating length L, which is turned into a period under Fourier transforms, adds to the aliasing. Even in the case of linear responses, the incoming light differs from the received signal. That poses a basic question: If a linear lattice contributes to the signal, or if it *is* the source whose signal is being examined, then besides the diffraction pattern all the above listed modulations characterize it. Otherwise, if the light incident on the lattice is accounted as the signal source, then all of the effects generated by sampling are spurious.

For an alternative calibration of the wavelength range, rather than deriving it from geometrical properties, we consider the musical acoustics analogy. Indeed, since the acoustic analogy is not at all new, the very mathematical expression of the Balmer series, $\lambda_n = Cn^2/(n^2 - 2^2) = CH_n$ with $C = 3645.6/10^7$ mm, and $n = 3, \ldots, 7$, when extended to $n \in \mathbb{N} \cup \{0\}$ indicates the hydrogen *frequency* spectrum was initially interpreted as an oscillating solution to the wave equation, according to the oscillation theorem. As a matter of fact, the infinite product $\Pi_n\ (1 - (2x/n))^2$ shows the proposed λ^{-1} values mapped as zeros of the function $\sin(2\pi x)$. Not all analytic functions, as defined by their zeros and poles, intuitively represent oscillations; however, the reason why the latter interpretation of the λ^{-1} was relinquished is probably that the Bohr model seemed to provide a mechanically more righteous explanation of the gas response, at the time. It is well known that the natural musical scale assumes $C0 = 16$ Hz as the fundamental note, and that all the other tones may be based upon musical chords. The unison pitch interval is $C1 = 2C0$. The intervals within a given octave, like the perfect fourth $C - F$, which is $4 : 3$, or the $C - D$ major second, $9 : 8$ are indicated by their ratios. The augmented fourth $(25 : 18)$ and the minor seventh $(15 : 8)$ are also used. In our case, as the visible colour coding seems to be quite the same for interferential and dispersion effects, both of them deserve the same mathematical indexing. Thus, the first diffraction octave between violet and red is considered a bounded interval and higher octaves exhibit integer multiples of one established λ. The optical measure is related to the partitioning of the visible wavelengths into the first octave by the assumption $2\lambda_{violet} = \lambda_{red}$. Let's use the λ_{violet}-pitch scale unit as the fundamental length. As it happens for musical chords, that's not the only available choice, since we could instead pick red, orange .... The very expression of the constant dispersion $\Delta\lambda/\lambda \to d\lambda/\lambda = d(Ln\lambda) = d(ln|\lambda|) + jd(arg\lambda) + jd(2\pi n)$, $n \in \mathbb{N}$, may then hint a base 2 logarithmic partition of the scale. Because of the mathematical boundary, the corresponding variable does not vary continuously. To give the corresponding mathematical definition of "distance", let a spectral line of the first diffraction order, for instance the violet, be represented by a real coordinate x_1. The transformation f that maps it to the corresponding line x_2 in the next diffraction order be given as $x_2 = f(x_1) = kx_1$, where k is a positive constant. Iterating the transformation p times we get: $x_p = f(f(\ldots(x_1))) = f^p(x_1) = f^{(p-1)}(kx_1) = \ldots = k^p x_1$. The distance between x_1 and x_p can be formally defined in those terms as $\rho(x_1, x_p) = p = (\log_2 k)^{-1} \log_2(x_p/x_1)$. For the sub-partition let's take the raw average grating spacing ($\ell \propto 1/k = \frac{1}{2}^N$). Thus the distance $\rho(x_1, x_p)$ multiplied by the factor $\log_2 k = N$ is equal to the number of iterates which transform $x_1 \to x_p$. It shall be of the form: $\rho(x_1, x_p) = 1/N \log_2(x_p/x_1)$, where upper octaves are reached as soon as $p \geq N$. Distance additivity is implied by $\rho(x_1, x_2) + \rho(x_2, x_3) = 1/N\{\log_2(x_2/x_1) + \log_2(x_3/x_2)\} = 1/N \log_2(x_3/x_1)$. That is one way of understanding the Balmer's [6] coefficients for the principal hydrogen lines, which are: $H\alpha = 9/5$, $H\beta = 16/12$, $H\gamma = 25/21$, $H\delta = 36/32$ and $H_{Hagenbach} = 49/45$. Although function f has been defined as a transformation, the latter ratios only stand for *relations* between elements, irrespective of their absolute wavelength values. In general, let (x_a, x_b, x_c, x_d) indicate the cross-ratio of four coordinates, for instance $[(x_a - x_c)/(x_a - x_d)] : [((x_b - x_c)/(x_b - x_d)]$. The

[13] It states, without hints at geometrical interpretations, that a correctly sampled band limited signal, for instance a dispersion signal, can be *fully* recovered.

logarithm[14] of the cross ratio can be used to define a distance function like the one we have just shown for musical chords[15]. In order to obtain numerical values from the cross-ratio, we then need to specify in addition whether the invariants of f separate the coordinates whose distance we want to measure or not. Such a definition yields a mapping that ultimately depends on the signal being examined and provides values that can be treated without the need to distinguish between angular pulsations and wavelengths.

8. PHASE MODULATION IN CONCAVE GRATINGS

A concave cylindrical grating permits to acquire the so called *normal dispersion spectrum*, used for absolute measurements of λ, because each spectral order is obtained without requiring any additional lens and shows no optical lever effect, which planar gratings do[16]. To see how it works let's first consider a cylindrical *mirror* of radius R. When it is uniformly illuminated through a slit oriented according to the cylinder's axis at a distance R from the vertex, it shapes an image of the slit, laid upon it with unitary enlargement[17]. Within geometrical approximations, the real image preserves the same features as the slit's position moves along a circumference, called Rowland circle, which passes through the vertex and the center with radius $R/2$, provided that normal light incidence conditions are preserved. It is known that a mirror behaves as a *phase plate* for coherent paraxial beams. As shown by Zernike [4], striae appear on the image of the mirror surface obtained with monochromatic light as soon as one is able to transform the phase modulation of the stationary wave into amplitude modulation[18]. When observing the image of the slit produced by a ruled mirror on a screen put on the Rowland circle, multiple images of the slit are observed, that result in the diffraction pattern. Observing a single diffraction order, rather than the whole pattern, incident and reflected beams are not overlapping. However, shape defects and deviations from perfect grating periodicity do affect the portion of the pattern shown. For example, if a concave mirror is ruled according to Rowland, the grooves' projections are at equal distances from each other on the plane of a chord orthogonal to the axis of the grating, the "plane grating" being the orthographic projection of the cylinder's grooves. Let ℓ be the spacing between two adjacent grooves at the center of either the cylinder or the plane, since there they coincide by hypothesis. The spacings on the cylinder increase toward the outer areas as $ds = \ell d(\sec\theta)$, where θ is the angle between the chord and the tangent to the circle at the chosen point. The result is a modulation of the grating period, that is, a frequency modulation, that alters the dispersion pattern on the Rowland circle depending on the angle between the incident and the reflected beams. Of course, any other *non-local* trait of the device affects both the diffraction pattern and the measured frequencies. Instead of shifting the position of a single spectral peak, the phase modulation changes their overall order in dispersion spectra, whatever the diffraction order, and even in the absence of any special effect, such as, e.g., Raman scattering.

9. CONCLUSIONS

In this article we have noted that real number values referring to absolute lengths have been assigned to dispersed light by linking waves parameters to the grating's spacing and applying Rowland's coincidence method. We have analyzed how light takes on periodic properties as a consequence of diffraction/interference. The acquired periodicity allows to take care of the passive answer of the receiving devices, and thus enables to mathematically handle the signal. However, periodicity also hampers the assignment of absolute numerical values to wavelengths. Since in mathematics periodicity is not a prominent feature of real numbers, they probably are not the best numerical labels for telecommunications coding purposes. Finally, we have recalled that Euclid introduced cross-ratios into geometry for tackling lengths and areas incommensurable with each other. Thus, measurements based upon cross-ratio don't preclude geometrical modeling at all, even if, obviously, the statement of a specific scaling problem depends on the measurement system in effect. As a conclusion, we note that the geometric representation of electromagnetism would not be hindered if measurements were expressed as cross-ratios, as long as they concern the information related to signals received regardless of bias.

[14]Natural logarithms may be taken in case partitioning the λ's into octaves is not interesting.

[15]By choosing $x_c = 0$ and letting $x_d \to \infty, [(x_a - x_c)/(x_a - x_d)] : [((x_b - x_c)/(x_b - x_d)] = [x_a/x_b] : [((x_a - x_d)/(x_b - x_d)] \to x_a/x_b$.

[16]For plane gratings, if order 0 is in focus, the outermost regions of the diffraction pattern are not.

[17]We are neglecting Airy's patterns at the borders.

[18]Armstrong showed that at radio frequencies.

REFERENCES

1. Ritz, W., "2 zur theorie der serienspektren," *Ann. Phys.*, Vol. 317, No. 10, 264–310, 1903.
2. Meggers, W. F., "The primary standard of wave-length," *Rev. Mod. Phys.*, Vol. 14, No. 2–3, 59–63, 1942.
3. Rowland, H. A., *Physical Papers*, The John Hopkins Press, Baltimore, 1902.
4. Zernike, F., "How I discovered phase contrast," *Nobel Lecture*, December 11, 1953.
5. Liouville, J., "Sur des classes très-étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques," *JMPA*, Vol. 16, Sér. I, 133–142, 1851.
6. Balmer, J. J., "Notiz über die spektrallinien des wasserstoffs," *Ann. Phys. Chem.*, Vol. 25, 80–85, 1885.

Compact Folded Dipole Antenna for DTV Signal Reception

Ding-Bing Lin[1], **Shiao-Ting Wu**[1], **and Chao-Hsiung Tseng**[2]

[1]Department of Electronic Engineering
Graduate Institute of Computer and Communication Engineering
National Taipei University of Technology, Taipei, Taiwan
[2]Department of Electronic Engineering
National Taiwan University of Science and Technology, Taipei, Taiwan

Abstract— In this paper, a compact simple folded dipole antenna for digital television (DTV) signal reception in the 470–850 MHz band is proposed. The antenna consists of two metal thin meander lines and narrow rectangular radiating patches. It is feeded by 50 Ω SMA connector and printed on a dielectric substrate FR4. By appropriate tuning geometrical parameters, the final designed antenna has only compact volume with $80 \times 28 \times 0.4\,\text{mm}^3$ .The antenna can be excited to provide a wide bandwidth (−5 dB return loss) of larger than 55%.

1. INTRODUCTION

The DTV reception has thus become very attractive for application in mobile communication devices such as the laptop computers, mobile phones, personal media players [1]. Similarly, it is also very attractive for vehicle owners to have their vehicles equipped with a DTV signal reception device. For these perspective trends, it is anticipated that a variety of mobile antenna for DTV signal reception will be increasingly required. The conventional dipole antenna are usually with a center-fed connector structure. It has narrow bandwidth, which is difficult to cover the DTV signal reception. The novel end-fed modified planar dipole antenna with additional resonant mode adjacent to the antenna fundamental (0.5-wavelengh) resonant mode can be excited. These two resonant modes can be formed into a wide operating band of larger than conventional center-fed planar dipole [2, 3]. Also note that the concept of using two adjacent resonant mode excited in the proposed dipole antenna for broadband operation are different from that of using two separate resonant mode excited in the reported dipole antenna with two U-slotted arms or two L-slotted arms for dual-band operation [4, 5].

In this paper, we present a novel side-fed modified planar dipole antenna having wide bandwidth of larger than 55% which can cover DTV receiver frequency from 470–850 MHz. The proposed antenna also has a planar configuration and is easy to fabricate at low cost by printing on a dielectric substrate. Details of the proposed antenna are described in this article and the experimental results of constructed prototype are presented.

2. ANTENNA DESIGN

Figure 1(a) shows the proposed compact folded dipole antenna. The antenna is fed by 50 Ω SMA connector and printed on a 0.4-mm thick FR4 substrate and look like business card shape (width 28-mm and length 80-mm). The antenna consists of two metal thin meander lines and narrow rectangular radiating patches. It can generate a fundamental resonant mode centered at about 520 MHz (0.5-wavelength), allowing the antenna to cover the lower-edge frequency (470 MHz) of the DTV band. However, in order to cover the upper-edge frequency (850 MHz) we used inverted U shaped to generate a high-order resonant mode. In addition, we use the rectangular radiating patches to create the slow wave effect for reducing the antenna size. Fig. 1(b) shows the concept used for broadening the impedance matching bandwidth. The metal thin meander line take on inductance effect and the patches take on capacitance effect. In this case, good impedance matching across a wide frequency range for DTV signal reception in the UHF band can be achieved.

3. MEASUREMENTS

The proposed antenna with dimensions given in Fig. 1 was fabricated and studied. Fig. 2 shows the measured return loss for proposed antenna with and without rectangular patches. It is seen that, with rectangular patches, matching bandwidth can be shifted to lower frequency band due to the influence of slow wave effect. The impedance bandwidth is enough to cover the DTV signal reception. Fig. 3 shows the measured return loss of the coupling gap d varied as 0.5, 1 and 2 mm.

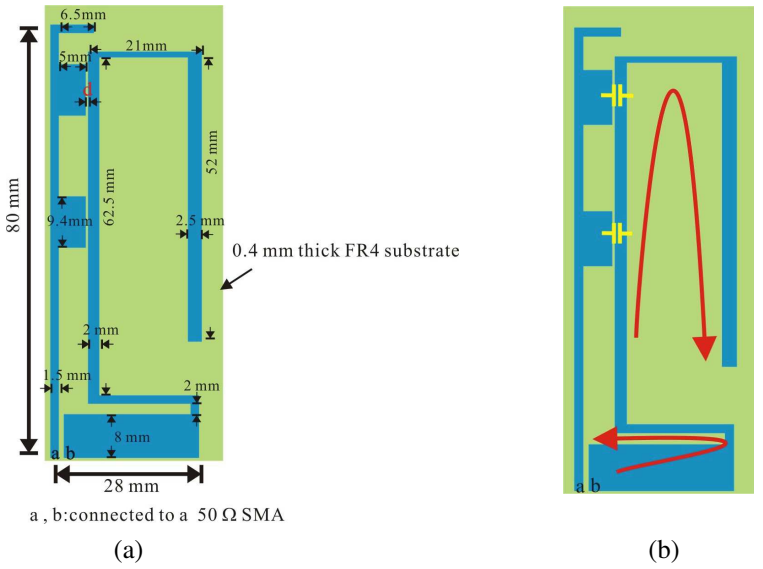


Figure 1: (a) The proposed compact folded dipole antenna. (b) To create both the inductance and capacitance effect for impedance matching bandwidth.

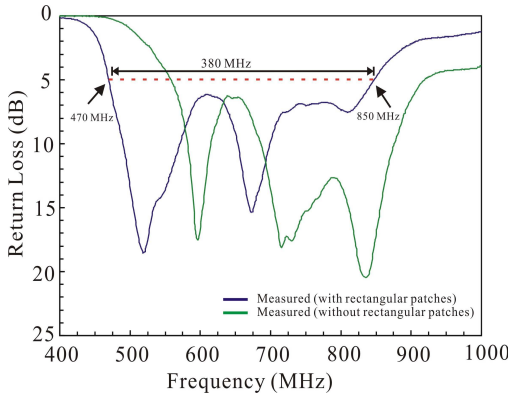


Figure 2: Measured return loss for proposed antenna with and without rectangular patches.

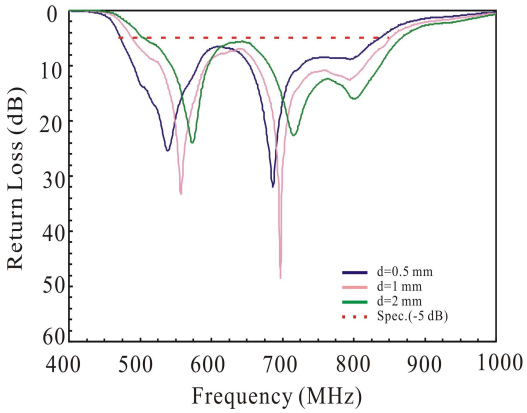


Figure 3: Measured return loss with different coupling gaps.

By changing the value of d, the mutual coupling between the metal thin meander line and narrow rectangular radiating patches effect can provide the way of the controlling the required bandwidth and impedance matching over the DTV frequency band. Fig. 4 plots the measured radiation patterns at 520, 683 and 815 MHz, respectively. The radiation pattern is almost omni-directional at x-y plane. This antenna has a good radiation pattern over the DTV frequency reception.

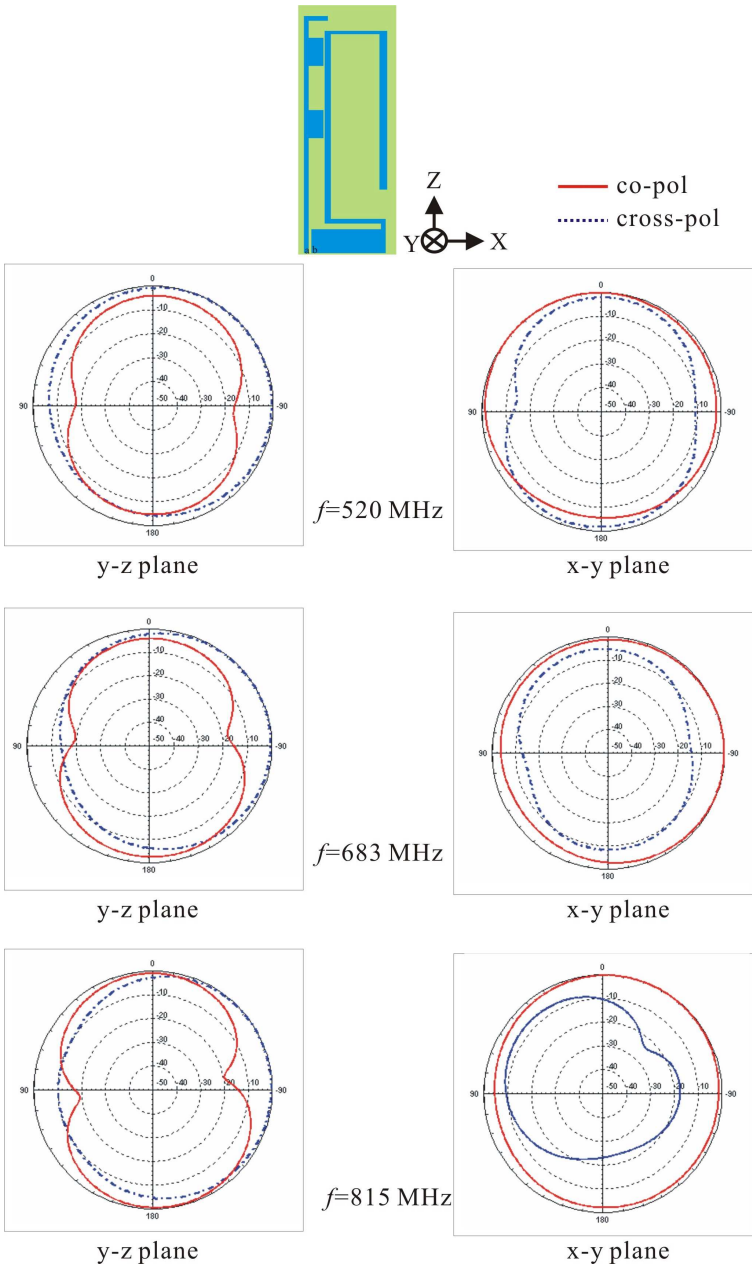


Figure 4: Measured far-field radiation pattern at 520 MHz, 683 MHz and 815 MHz for the proposed antenna.

4. CONCLUSIONS

A compact folded dipole antenna for DTV signal reception in 470–850 MHz band has been proposed and studied. The antenna is with a simple structure and is easy to fabricate with a low cost. The proposed antenna using narrow rectangular radiating patches and coupling gap d can provides the way of the controlling the required bandwidth and impedance matching over the DTV frequency band. This antenna has an omni-directional radiation pattern. The proposed antenna also shows good radiation characteristics for frequencies across the operating band.

ACKNOWLEDGMENT

The authors would like to thank the National Science Council of the Taiwan, for financially supporting this research under Contract No. NSC 96-2218-E-027-007.

REFERENCES

1. Digital Television, Major Initiatives of Federal Communications Commission, [Online]. Available: http://www.fcc.gov/dtv/.

2. Chi, Y. W., K. L. Wong, and S. W. Su, "Broadband printed dipole antenna with a step-shaped feed gap for DTV signal reception," *IEEE Transactions on Antennas and Propag.*, Vol. 55, 3353–3356, Nov. 2007.
3. Iizuka, H., T. Watanabe, K. Sakakibara, and N. Kikuma, "Stub-loaded folded dipole antenna for digital terrestrial TV reception," *IEEE Antennas and Wireless Propagation Letters*, Vol. 5, 260–261, Dec. 2006.
4. Iizuka, H., T. Watanabe, K. Sato, and K. Nishikawa, "Modified H-shaped antenna for automotive digital terrestrial reception," *IEEE Trans. Antennas Propag.*, Vol. 53, 2542–2548, Aug. 2005.
5. Su, C. M., H. T. Chen, and K. L. Wong, "Printed dual-band dipole antenna with U-slotted arms for 2.4/5.2 GHz WLAN operation," *Electron. Lett.*, Vol. 38, 1308–1309, Oct. 24, 2002.

Study on Radiation Characteristics of A Conical Conformal Phased Array

Yinsuo Song
Luoyang Optoelectro Technology Development Center, Henan, China

Abstract— This paper provides an effective method for analyzing a conical conformal phased array antenna. Firstly, a fictitious elementary antenna is supposed to be located at the origin of an array coordinate; then, a layout for the antenna to be transformed to a given position and a polarization direction can be obtained, generally by rotation thrice and translation twice.

Based on the idea above, formulas to calculate the radiation pattern of the conformal array on a frustum surface have been deduced, with the elementary number and polarization direction specified and the simulation software is programmed to validate the model. A series of calculations have been conducted on the azimuth and the pitching scan performance of the array composed of 26×95 element slots in order to study the radiation characteristics of frustum conformal phased antenna focusing on the radiation characteristic at bigger scan angle (axial direction) and measures improving the radiation characteristic.

1. INTRODUCTION

Compared with conformal antenna applied in communication field, researches on missile-borne seeker conformal phased array antenna have been making a slow progress. One of the reasons is that existing commercial software is difficult to analysis the electrically large bodies, and the normal geometry about conformal phased antenna, such as cirque or column and so on, has been studied concentratively, but there are a few researches on the interceptive cone or frustum surface which show better pneumatic performance. When the phase mode theory, which is set up during the study of cirque or column conformal arrays, is applied, the local frustum surface is often equivalent to the column. However, the pose angle of the elementary antenna on the frustum surface, which presents the maximum radiation direction and polarization angle of the elementary antenna, will be lost in this equivalence, which conflicts with the fact.

Daniel Augustin, etc analyze one-dimensional phased array antennas printed on inclined planes, and provide a method for analyzing a cone conformal antenna: Supposing the cone consists of many little planes on which there are elementary antennas, the radiation field of the elementary antenna can be approximately replaced by the result of the plane antenna.

This paper applies the method above into the analysis of a large two-dimensional cone phase scan cone conformal array and provides an effective method for the conformal antenna: a fictitious elementary antenna is supposed to be located at the origin of an array coordinate; then, a layout for the antenna to be transformed to a given position and a polarization direction can be obtained, generally by rotation thrice and translation twice. The formulas to calculate the radiation pattern of the conformal array on a frustum surface with the elementary number and polarization direction specified are presented and the relative simulation software is programmed. Based on the correctness of the model, a series of calculations are conducted on the azimuth and the pitching scan performance of the array composed of 26×95 element slot antennas focusing on the radiation characteristic at bigger scan angle (axial direction) and measures improving the radiation characteristic.

2. CALCULATION FORMULA

The elementary reference coordinate (x, y, z) is shown in Fig. 1. The global coordinate (X, Y, Z) is shown in Fig. 2, where the line segment $P_1 P_2$ is the generatrix of the frustum surface through the phase center of the ith elementary antenna, whose outward normal vector is $\hat{\boldsymbol{n}}$.

Firstly, the elementary reference antenna as shown in Fig. 1 is supposed to be located into the global coordinate to coincide the origins of the local coordinate (x, y, z) of the elementary antenna and the global coordinate (X, Y, Z) as well as the axes of the two coordinates. Therefore the reference antenna can be moved to a certain position on the generatrix of the frustum surface by translation in the plane (X, Y) and in the direction of Z axis respectively once. Then the transition coordinate (x', y', z') can be gained. In order to make sure the reference antenna in a required pose, a new transition coordinate (x'', y'', z'') can be firstly obtained by rotating the original transition

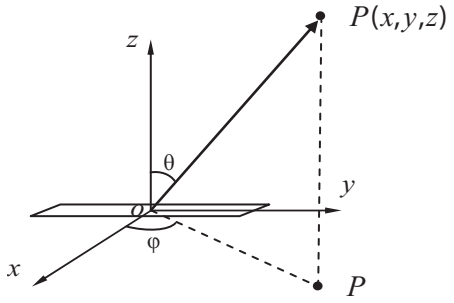


Figure 1: Elementary antenna and local coordinate.

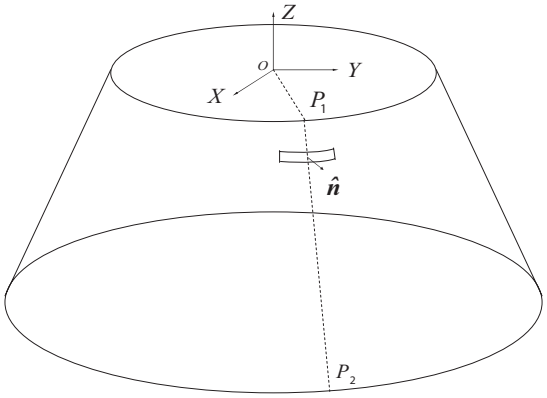


Figure 2: Frustum shape and global coordinate.

coordinate (x', y', z') around z'-axis by α angle following the right-hand rule while x''-axis is parallel with the line segment OP_1. Then we rotate the coordinate (x'', y'', z'') around y''-axis by β (the base angle) angle following the same rule to get the transition coordinate (x''', y''', z''') while z'''-axis and the vector $\hat{\boldsymbol{n}}$ coincide. Finally, the wanted coordinate (x'''', y'''', z'''') can be gained by rotating the transition coordinate (x''', y''', z''') around z'''-axis by γ angle to gain the required electric field polarization direction. So the coordinate (x'''', y'''', z'''') is the local coordinate of the ith elementary antenna named ($x^{(i)}$, $y^{(i)}$, $z^{(i)}$).

Based on the transform above, the coordinate conversion expression of field point form the global coordinate to the local coordinate of the elementary antenna.

$$\begin{bmatrix} x^{(i)} \\ y^{(i)} \\ z^{(i)} \end{bmatrix} = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}^{-1} \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} X - X_i \\ Y - Y_i \\ Z - Z_i \end{bmatrix} \tag{1}$$

Formula (1) can be substituted into the expression of the elementary antenna radiation field deduced in the coordinate in Fig. 1, and then the elementary antenna radiation field of any point $\boldsymbol{r}$ in space in the local coordinate can be obtained.

$$\boldsymbol{E}^{s,\,i}(\boldsymbol{r}) = \hat{\boldsymbol{x}}^{(i)} E^{s,\,i}_{x^{(i)}}(\boldsymbol{r}) + \hat{\boldsymbol{z}}^{(i)} E^{s,\,i}_{z^{(i)}}(\boldsymbol{r}) \tag{2}$$

To get the radiation field of the array at any point $\boldsymbol{r}$ in space, the radiation field of all the elements need to be added, which needs to transform the elementary antenna radiation field form the local coordinate to the global coordinate, which is a converse process of the above coordinate rotation conversion.

$$\begin{bmatrix} E^{s,\,i}_X(\boldsymbol{r}) \\ E^{s,\,i}_Y(\boldsymbol{r}) \\ E^{s,\,i}_Z(\boldsymbol{r}) \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} E^{s,\,i}_{x^i}(\boldsymbol{r}) \\ 0 \\ E^{s,\,i}_{z^i}(\boldsymbol{r}) \end{bmatrix} \tag{3}$$

Adding the radiation field of all the elements in the global coordinate

$$\boldsymbol{E}^{s}(\boldsymbol{r}) = \sum_{i=1}^{N} \left[\hat{\boldsymbol{X}} E^{s,\,i}_X(\boldsymbol{r}) + \hat{\boldsymbol{Y}} E^{s,\,i}_Y(\boldsymbol{r}) + \hat{\boldsymbol{Z}} E^{s,\,i}_Z(\boldsymbol{r}) \right] \tag{4}$$

where N is the total number of elementary antennas in the array, and ($\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, $\hat{\boldsymbol{z}}$) and ($\hat{\boldsymbol{X}}$, $\hat{\boldsymbol{Y}}$, $\hat{\boldsymbol{Z}}$) are respectively unit vectors of the local and global coordinates.

3. RESULTS AND VERIFICATION

Figure 3 shows the pattern of the array composed of 4×22 slot elements, which is calculated by the commercial electromagnetism simulation software CST for about 50 hours on the computer with 32 G memory and 8CPU. The same calculation can be done within 2 minutes through the MatLab

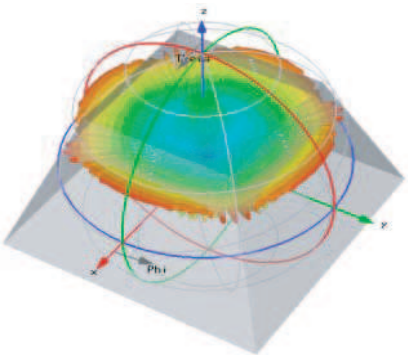


Figure 3: The pattern calculated by CST.

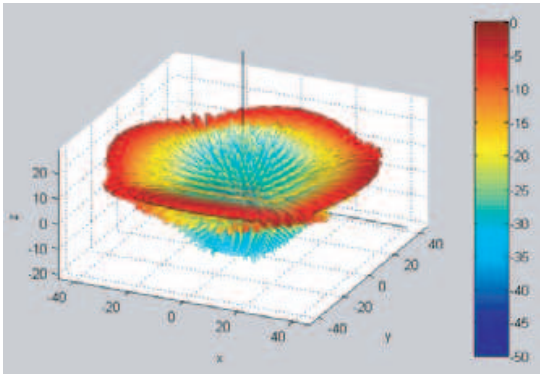


Figure 4: The pattern calculated by the MatLab.

programme compiled by the formula in this paper, as shown in Fig. 4. It is obvious that the two patterns are coincident.

The calculation of the radiation pattern of the element slot on the cone indicates that the radiation in the direction of Z-axis is strong when the electric polarization is along the generatrix, but it is weak when in the circumferential direction. Therefore the former polarization way is chosen to improve the radiation characteristic in the direction of Z-axis.

To calculate the radiation pattern of the array and its scan characteristic, three typical amplitude distributions are chosen: constant amplitude, taper and triangle (the amplitude of the middle element is the biggest and that of the two sides is descending) in the generatrix; elements in the same circumference have the same amplitude distribution. It is proved that the triangle amplitude distribution has smaller side-lobe level and better beam-shape.

The layout of the array elements is following: there are 26 elements on each generatrix of the frustum and the inclined distance between two adjacent elements is 0.5λ, and 95 elements on each circumference arranged equiangularly and the upper and bottom arc lengths of adjacent elements are respectively 0.41λ and 0.68λ.

The phase difference of adjacent elements is $\Delta\Phi$ to complete the pitching scan (in the direction of Z-axis) of the beam. The tentative calculation indicates that when $\Delta\Phi$ is changed from $-180°$ to $+180°$, the beam can be changed from $-90°$ to $+90°$. Figs. 5 and 6 present the beam scan patterns respectively for $\Delta\Phi = 0°$ (side incidence), $180°$ (axial direction) and the "active area" includes 7×26 elements. It can be seen in Fig. 6 that broadening and distortion of the beam appear in the direction of Z-axis (at bigger scan angle). Fig. 7 shows the radiation pattern with all the elements participating in and the beam-shape has been improved with obviously increasing of the back-lobe.

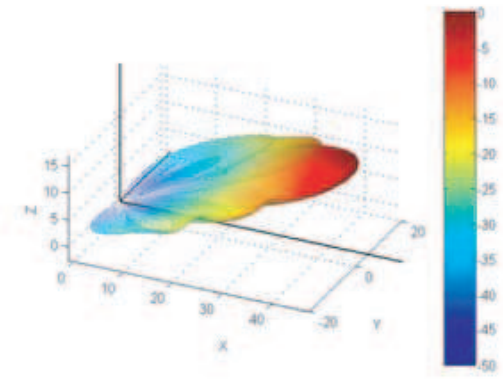


Figure 5: The pattern with $\Delta\Phi = 0°$.

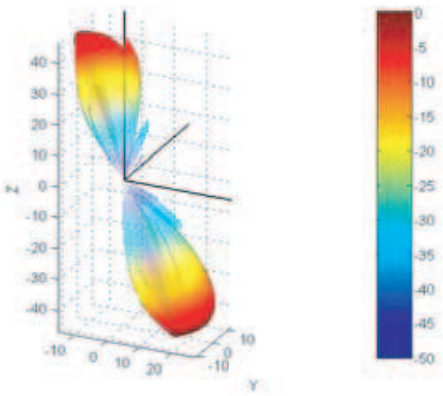


Figure 6: The pattern with $\Delta\Phi = 180°$.

The scan, whose range is $0° \sim 360$, in the azimuth plane (around Z-axis) can be achieved by changing the position of the active area. Fig. 8 shows the beam scan pattern when the angle of the azimuth scan is $90°$.

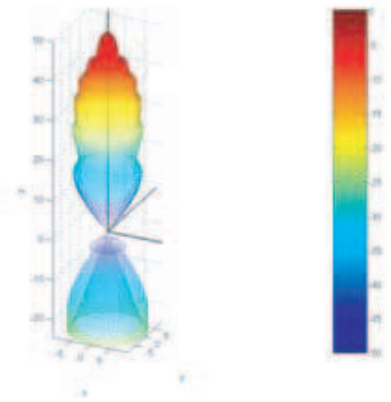


Figure 7: The pattern with all the elements.

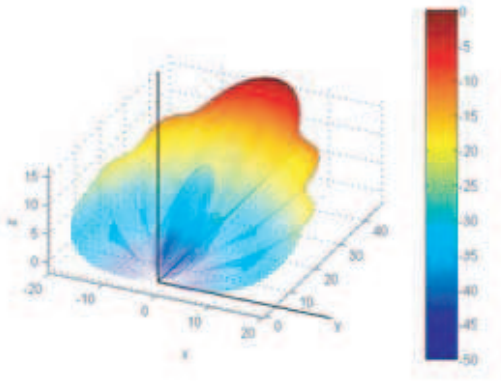


Figure 8: the pattern with azimuth scan angle 90°.

The polarization characteristic of the radiation field is also calculated and the result shows that the cross polarization value is below −25 dB at the smaller scan angle while the cross polarization value is 7 dB bigger than the main polarization value at bigger scan angle (in the direction of Z-axis).

4. CONCLUSION

The mathematical model for electromagnetic analysis of the seeker conformal phased array antenna is established. The MatLab simulation software is programmed, whose result is compared with that of the commercial electromagnetic simulation software and testify the correctness of the model. Then the MatLab simulation software can be used to study the scan characteristics of the frustum conformal antenna and some possible technical problems.

Farther studies include the amplitude/phase compensation, polarization control and synthesis of the pattern, etc.

REFERENCES

1. Josefsson, L. and P. Persson, *Conformal Array Antenna Theory and Design*, 1–108, Willey-Inter Science Publication, 2006.
2. Fondevila, J., et al., "Very fast method to synthesize conformal arrays," *Electronics Letters*, Vol. 43, No. 16, August 2, 2007.
3. Cambiaggio, E., et al., "Conformal microstrip antenna arrays: A fast analysis for conical arrays," *1997 IEEE*, 1529–1532, 1997.
4. Augustin, D., et al., "Microstrip phased array antennas printed on inclined planes," *Radioengineering*, Vol. 5, No. 2, 1–7, June 1996.
5. Harrington, R. F., *Time-harmonic Electromagnetic Fields*, addenda A, McGRAW-Hill Book Company, 1961.