Strongly Coupled Electro-optic Superlattices and the Potential as Metamaterials

Yalin Lu, **Dajani Iyad**, **and R. J. Knize**
The Physics Department, 2354 Fairchild Dr., 2A31, USAF Academy, CO 80840, USA

Abstract— Superlattices consisting of alternatively stacked rhombohedral and tetragonal phases of lead titanate doped lead magnesium niobate were demonstrated. Largely enhanced out-of-plane electro-optic (EO) coefficient was recorded, which evidences strong spontaneous polarization wave coupling with phonons (structural resonances). This coupling opens up a window to develop new metamaterials having unique permittivity and permeability properties, when using suitable combinations from both ferroelectric and ferromagnetic materials.

1. INTRODUCTION

EO effect in oxygen-octahedra ferroelectric thin films, such as (Pb,La)(Zr,Ti)O_3 (PLZT) [1], (Pb,La) TiO_3 (PLT) [2], (1-x)Pb($Mg_{1/3}Nb_{2/3}$)O_3-x$PbTiO_3$ (PMN-PT) [3], etc., has been extensively studied for exploring their potential applications as optical switch and modulator. Generally those films have been grown on common substrates, including $LaAlO_3$, sapphire, silicon, etc., with a diversity of structural qualities from amorphous to single crystalline [4]. Most recorded quadratic EO coefficients in films are normally small ($< 1.0 \times 10^{-16}(\mathrm{m/V})^2$) with only a few arguable exceptions [5], and have mainly been paid with attention in in-plane configurations.

Relaxor ferroelectric PMN-PT is a material well known for its high electromechanical coupling [6]. In our previous researches, we demonstrated good optical clarity and low waveguiding loss in a $\langle 100\rangle$ oriented PMN-PT film [7]. Large in-plane quadratic EO coefficient of $1.5 \times 10^{-16}(\mathrm{m/V})^2$ was measured. Most importantly, it was found that EO effect reaches its maximum when the film having a composition approaching to PMN-PT's morphology phase boundary (MPB) composition of $x = 0.33^3$. Meanwhile, a strong EO in-plane anisotropy was also found that the EO coefficient along in-plane $\langle 110\rangle$ is at least two orders larger than those along in-plane $\langle 001\rangle$ directions [8]. This EO anisotropy was attributed to the materials' spontaneous polarization orientations. Thus it will be reasonable to expect that it has a small out-of-plane EO coefficient since that the normal of the film is obviously one of main crystalline axes of the material.

In this paper, we describe a new type of EO superlattice, which composes of alternative PMN-PT rhombohedral (R) and tetragonal (T) phases. Such an EO superlattice can present a largely enhanced out-of-plane EO effect, because of a strong spontaneous polarization coupling via the structural resonance. The purpose of this research includes: 1) to develop an optical material having large out-of-plane EO effect, which offers an additional flexibility of using vertical designs, other than regularly used transversal designs in photonic components [9]; and 2) to verify such coupling through EO measurement, which leads to a proposal of new superlattice-type metamaterial development toward realizing unique permittivity and permeability properties.

2. FABRICATION

The superlattices were fabricated on 0.5 mm thick, r-cut Al_2O_3 substrates using a modified dip-coating process [7]. The growth temperature for both (1-x)PMN-xPTs phases [(60/40) and (74/26)] is all around 650°C. All the samples were kept with a same thickness of 0.35 μm, have a superlattice period "Γ" of 10 nm, 30 nm, 60 nm, 90 nm, and 120 nm, respectively. For comparison, single-composition films of (60/40), (74/26), and (67/33) with the same thickness were also fabricated under similar conditions. In order to measure the out-of-plane EO coefficient (requiring to apply an electrical field normal to the film), a second set of superlattices were fabricated on thin (∼50 nm) $La_{0.5}Sr_{0.5}CoO_3$ (LSCO) coated r-cut sapphires. The semi-conductive LSCO layer is required to be as thin as possible for the reason to increase its transparency in the visible range. Adding a thin layer of LSCO does not deteriorate the superlattice's crystallinity, since that PMN-PT, LSCO, and r-cut sapphire are all nearly lattice-matched [10]. On top of all these samples, a 200 nm Au coating was evaporated as a reflection layer for the out-of-plane quadratic EO measurements.

3. MEASUREMENTS

In-plane EO coefficient measurement using a pair of parallel Au electrodes with an accurately defined gap of 50μm is very straightforward. The method was discussed in our previous publications and the references therein [8]. To measure the out-of-plane EO coefficient, however, difficulty occurs when the electric field needs to be applied normally to the film [14]. In our measurement, the modulation of the polarization state of a laser beam, entering the sandwich structure from the semitransparent side (LSCO) and passing through the film when reflecting on the metal electrode (Au), will be revealed using the setup in Figure 1. We assume that the beam is passing just twice through the film (upon reflection). We neglect F-P resonance from the sandwich structure, which could simulate the EO effect and add significant error to above two-pass assumption, because that reflection from the semitransparent LSCO layer is negligible [11].

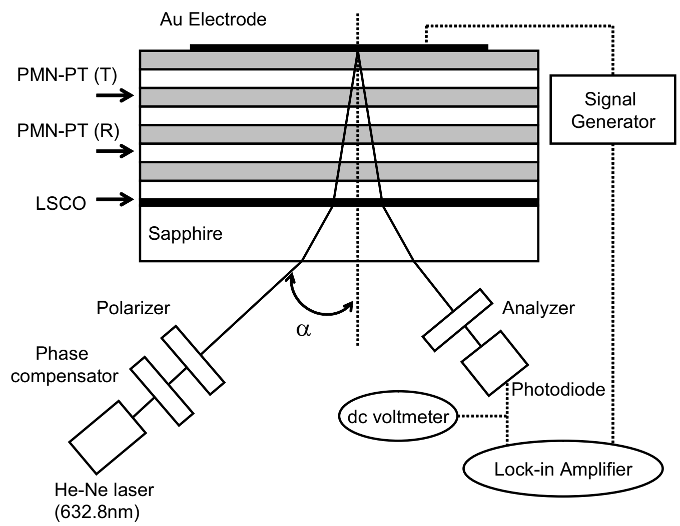


Figure 1: Set-up used to measure the out-of-plane quadratic EO coefficient.

4. RESULTS AND DISCUSSION

Figure 2 shows the XRD spectra of two superlattices with a periodicity of 10 nm and 120 nm, respectively, which are compared with those from PMN-PT (67/33), (74/26), and (60/40) films. The (002) diffraction splitting is clearly seen, which indicates the existence of the superlattice structure. X-ray rocking curve analysis of the full width at half maximum (FWHM) performed on those superlattices' (001) and (111) diffraction peaks, disclose a FWHM around 0.23° and 0.42°, respectively, which still evidences that these films on sapphires are highly out-of-plane and in-plane oriented.

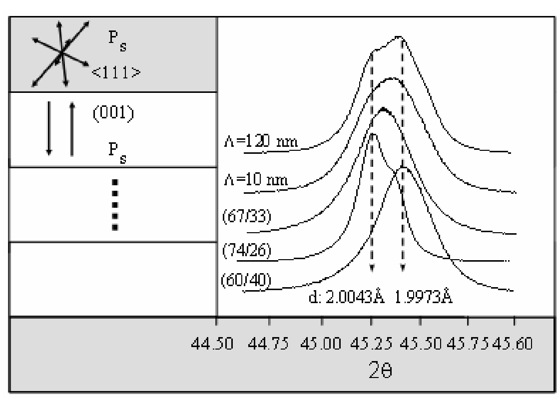


Figure 2: XRD spectra.

In Figure 3, out-of-plane quadratic EO coefficients of two superlattice samples with a periodicity of 10 nm and 30 nm, respectively, are plotted vs the incident angle α. For the two samples, the quadratic EO coefficients, which were averaged over all the values for all the incident angles, are $(1.65\pm0.25)\times10^{-16}\,(\mathrm{m/V})^2$ and $(1.47\pm0.24)\times10^{-16}\,(\mathrm{m/V})^2$, respectively, with a 16% relative error of measurement. In order to minimize errors from local heating and laser power fluctuation, we repeated the tests at least three times and improved the heat dissipation on each sample-under-measurement.

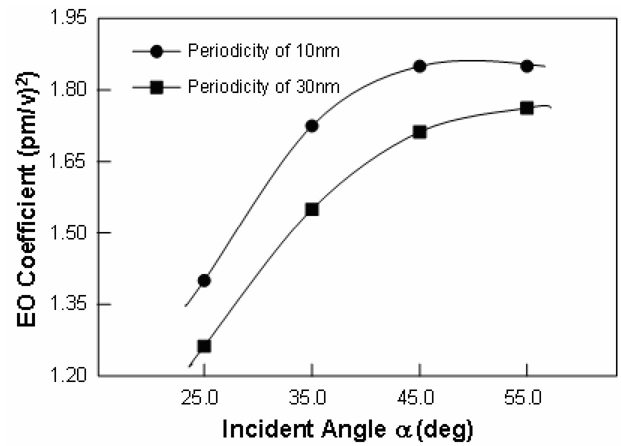


Figure 3: Out-of-plane EO vs incident angle.

Figure 4 shows both in-plane and the averaged out-of-plane quadratic EO coefficients of fiver superlattices vs their periodicities. For comparison, we also show the out-of-plane quadratic EO coefficients measured from three single-composition films: PMN-PT (60/40), (67/33), and (74/26), which are small (from 0.1 to $0.3\times10^{-16}\,(\mathrm{m/V})^2$) and inside the relative measurement error range. It is important to notice that a strong enhancement of the out-of-plane quadratic EO coefficient has been observed when comparing to single-composition films and when reducing the superlattice's periodicity. Such an enhancement coincidences very well with our dielectric and ferroelectric hysteresis measurements [10], and can be attributed to a long range coupling of spontaneous polarizations via structural phonons (resonance). Domain switching is a physical representation of such coupling, which only needs a 45° angle adjustment of dipoles between the two phases. The in-plane quadratic EO coefficient, however, only have a slight dependence on the superlattice periodicity, and maintain a large value higher than $1.5\times10^{-16}\,(\mathrm{m/V})^2$.

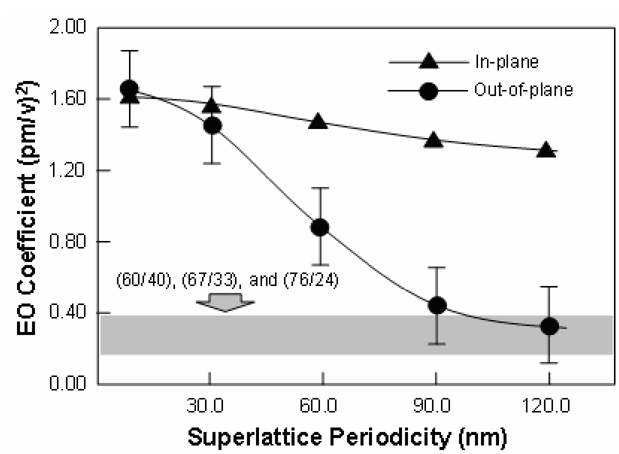


Figure 4: Dependence of both in-plane and out-of-plane EO coefficients on periodicity.

Both large in-plane and large out-of-plane EO effects in these PMN-PT superlattices make many new-conception EO devices possible. One example of such devices is a vertically aligned EO modulator having a circular aperture sandwiched by circular top- and bottom-electrodes, which can be directly integrated to a VCSEL semiconductor laser. Such integration can be exactly mode-matched and thus fully lossless. An additional advantage of such a vertically aligned modulation design is its fully polarization-independent performance, which can be very significant to fiber-optic telecommunication systems in which polarization-dependent losses (PDL) of optical components have been a very costly problem to solve.

More importantly, verification of such coupling between the spontaneous polarization wave and structural phonon proposes a new type of superlattice metamaterials having frequency-dependent and structural tailorable permittivity and permeability properties, when properly selecting the superlattice materials from ferroelectric and ferromagnetic materials. Dielectric and microwave properties of such superlattices were reported in our previous research [10], which suggests a possible resonance behavior. Our recent simulation shows a possible existence of negative permittivity or permeability or both in such superlattice structures after the resonance frequency. Both simulation and experimental verification are on-going effort and will be discussed and published separately.

ACKNOWLEDGMENT

This research is supported by the US AFOSR regular BAA, AFRL MIPR, and AFOSR Small Grant programs. Special thank will be given to Dr. Gail Brown of Wright Patterson AFRL for very useful discussion and support.

REFERENCES

1. Haertling, G. H. and C. E. Land, *J. Am. Ceram. Soc.*, Vol. 54, 1, 1971.
2. Preston, K. D. and G. H. Haertling, *Appl. Phys. Lett.*, Vol. 60, 2831, 1992.
3. Lu, Y. L., B. Gaynor, C. Hsu, G. H. Jin, and M. Cronin-Golomb, *Appl. Phys. Lett.*, Vol. 74, 3038, 1999.
4. Park, B. H., Y. Gim, Y. Fan, Q. X. Jia, and P. Lu, *Appl. Phys. Lett.*, Vol. 77, 2587, 2000.
5. Jo, W., H.-J. Cho, T. W. Noh, B. I. Kim, D.-Y. Kim, Z. G. Khim, and S.-I. Kwun, *Appl. Phys. Lett.*, Vol. 63, 2198, 1993.
6. Park, S.-E. and T. R. Shrout, *J. Appl. Phys.*, Vol. 82, 1804, 1997.
7. Lu, Y., G. H. Jin, M. Cronin-Golomb, and A. J. Drehman, *Appl. Phys. Lett.*, Vol. 72, 2927, 1998.
8. Lu, Y., J. J. Zheng, M. Cronin-Golomb, F. L. Wang, and J. Zhao, *Appl. Phys. Letts.*, Vol. 74, 3764, 1999.
9. US Patent: US6069729.
10. Lu, Y., *Appl. Phys. Lett.*, Vol. 85, No. 6, 979, 2004.
11. Michelotti, F., G. Nicolao, F. Tesi, and M. Bertolotti, *Chem. Phys.*, Vol. 245, 311, 1999.

The Classical Structure Model of Single Photon and Classical Point of View with Regard to Wave-particle Duality of Photon

Donglin Zu

IHIP, Physical School, Peking University
Beijing, 100871, China

Abstract— The enigma of the wave-particle duality of photon has remained unimpressively explained for a century since Einstein presents the concept of the photon in 1905. This article establishes a classical geometric structure model of a single photon based on field matter, educes a formula for the size of a photon; assumes that there only are two kinds of photon of right hand and left hand circular polarized, and suggests the frequency ω of photon polarization rotated to be its spin frequency. It ascribes the wavelike of photon to its spin motion and the particle-like to its translation motion. From the point of photon particle instead of wave view to re-analyze Young's double-slit interference and polarizer experiments, gives reasonable mechanism. It defines the phase velocity and the group velocity of a photon. It gives a unified and consistent understanding of quantum particle of light and classical electromagnetic waves field. Evidently, such a precisely defined conceptual model is reasonable, objective and easy to accept for classical physicists.

The puzzle [1] of the wave-particle duality of light has remained unsatisfactorily explained nearly a century since Albert Einstein first proposed the photon, the quantum unit of light $\left(E=\hbar\omega,\ \vec{p}=\hbar\vec{k}\right)$ in 1905. This concept led to the revelation that light could be particle and wave at the same time, in some unimaginable way. Light interacted with matter according to mechanical principle, it had come full circle from Newton's corpuscles to Maxwell's waves to Einstein's photons. But there remained a problem deep and stubborn, unavoidable and inexplicable, the basic paradox: How could particles of light display the inherent wave property of interference? How could they cancel each other to give darkness? How could a photon change its polarization direction when passing through a polaroid? Nothing in Einstein's hypothesis explained this. In 1917, he ever said: "For the rest of my life I will reflect on what light is" [1]. As he wrote a friend, he felt his struggle with the enigma might drive him to the madhouse. He explored the paradox in a gedanken experiment, imagining an interference experiment of an extremely weak light source that emits one photon at a time. In 1909 Taylor [2] conducted such an experiment for three months, at most only a few photons at a time. It resulted in an ordinary wavelike interference fringe. Somehow, each single photon travels through both slits to interfere with itself. It seems that appearance of wave or particle is determined by the experimental arrangement. In 1986 Grangier et al. [3] reported a modern version of Taylor's experiment based on laser, it showed that a lone photon indeed interfere with itself. So far, the nearly widespread consensus among the quantum physics community all over the world is that light is a particle and also a wave—depending on the experiment. It takes wave and particles as two different faces of reality. When you measure the properties of light, you see one aspect or the other depending on the experiment, but never both. As a cost the foundation of quantum mechanics (QM) is to sacrifice the objective process to physical phenomena. Even single- photon interference passing through a double slit is deemed [4] to be one of the defining experiments of quantum mechanics, and without other simple experiment that demonstrates the wave-particle duality so well. This led to Richard Feynman describe the double-slit interference for particles as "a phenomenon which is impossible, absolutely impossible, to explain in any classical way, and which has in it the heart of quantum mechanics. In reality it contains only mystery" [5]. In frequent discussions about quantum theory, however, the double-slit interference acts as a classic gedanken experiment for its clarity in expressing the central puzzles of quantum mechanics.

In quantum electrodynamics (QED), after quantization of the electromagnetic ($E\&M$) field, we obtained the photon field [6, 7]. The "Schrödinger's equation" for the photon is represented by Maxwell equation. Photons are bosons, so the number of photons that can be in any one state must be unrestricted. The polarization vector $\mathbf{e}$ acts for photon as the "spin part" of the wave function. The rest mass of the photon is zero, there is no rest frame, since it moves with the velocity of light in every frame of reference. There is always a distinctive direction in space, the direction of the momentum vector $\mathbf{k}$ (say the z-axis). In such a case there is clearly no symmetry with respect to whole group of rotation in three dimensions but only axial symmetry about the preferred axis, so the component of photon's angular momentum along the z-axis, is conserved. This quantity is called the helicity. For massless particle there is no rest frame and the helicity can take only two values $\wedge = \pm s$. The state of a photon having a definite momentum in fact is twofold degenerate. It is described by a "spin" wave function which is a vector $\mathbf{e}$ in x-y plane. Vector $\mathbf{e}$ with only the component $\mathbf{e}_x + i\mathbf{e}_y$ or $\mathbf{e}_x - i\mathbf{e}_y$ non-zero correspond to the components $\wedge = +1$ or -1 respectively; these are $\mathbf{e} = \mathbf{e}^{(+)}$ and $\mathbf{e} = \mathbf{e}^{(-)}$. In other words, the values $\wedge = +1$ and -1 correspond to right-hand and left-hand circular polarization of the photon. Thus the component of the photon angular momentum along the direction of its motion can have only the two values of ± 1.

So far although having some knowledge about photon mentioned above, we do not yet have unified understanding of photon and light. The wave picture is described by Maxwell's equations. Photons are invoked when light interacts with matter, as if a wave of light were breaking on the material to become a spray of quanta. But how does one become the other? How do separate photons correlate their actions to make a single wave? What is the meaning of a single, distinct photon? Until these puzzles are resolved, our comprehension of light remains incomplete.

Based on classical and quantum electrodynamics and electromagnetic engineering practices, this paper presents a distinct picture of the wave-particle duality of photon, to give a basic concept of the geometry, the size and the behavior way of photon. The aim is to establish a photon model imaginable. Using this model to re-analyze Young's double slit interference, and polarizer experiments. We can apprehend light and other frequency band of electromagnetic ($E\&M$) wave in a purely objective way, have favored sensory knowledge or spiritual awareness.

1. THE BASIC HYPOTHESES

As usual the model appears as a few hypotheses below:

(A) Since photon is the minimum unit of the classical electromagnetic field energy it is reasonable to assume that a photon consists of electric field vector matter $\mathbf{E}_p$ and magnetic field vector matter $\mathbf{H}_p$ with $\mathbf{H}_p \perp \mathbf{E}_p$, $\mathbf{E}_p$ and $\mathbf{H}_p$ overlapping at the same spherule space. We denote the orientations of the electric and magnetic field vector matter by a red and green arrow respectively as shown in Fig. 1. The wave vector k can also be ascribed to individual photon, thus the energy flow of a

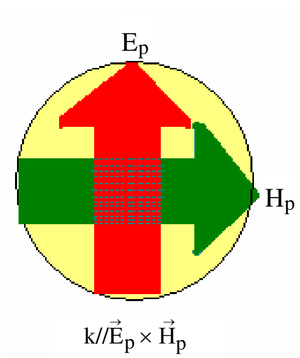


Figure 1: Classical structure of single photon.

photon $\mathbf{S}_p = \mathbf{E}_p \times \mathbf{H}_p$ along the k direction. This case is identical with the typical one of a classical E&M plane-wave. This way, a photon is very a cell of a plane E&M wave. In other words a classical plane E&M wave consists of overlapping and disposing of a lot of photons described above. Simply we can also say that a photon is very a minimum plane E&M wave. A photon has no rest mass, but motional mass. According to Einstein mass-energy relation

$$E = mc^2 = pc = \hbar\omega \tag{1}$$

the motional mass of photon is

$$m = \hbar\omega/c^2 \tag{2}$$

The momentum of the photon is $\vec{p} = \hbar\vec{k}$, here $(2\pi\hbar)$ is Plank constant. The m comes of the $\mathbf{E}_p$ and $\mathbf{H}_p$, a kind of field matter defined by Maxwell' field equations. The "field matter", a kind of "vectorial, enterable, massless matter", differs from "scalar, impenetrable, material mass" as usual. The motional mass of a photon is very made up from such field matter.

(B) Since the spin angular momentum of photon is $J = \pm\hbar$, this suggests that there are neither unpolarized photon nor linear polarized (LP) photon in our realistic world, so only exist right-hand circular polarized (RCP) and left-hand circular polarized (LCP) photons corresponding respectively to the helicity values $\wedge = +1$ and -1. The right (left)-hand CP light consists of right (left)-rotating photons, while the LP light is synchronously synthesized from coherent left-rotating and right-rotating photon pairs (Fig. 2).

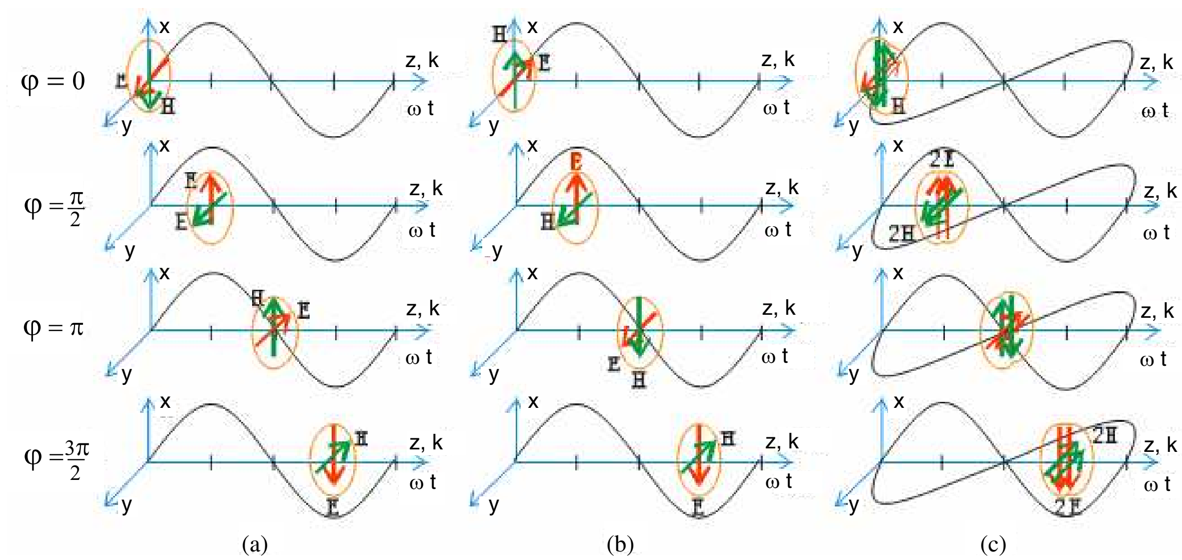


Figure 2: Classical picture of wave-particle duality of a single-photon: (a) left-rotating photon, (b) right-rotating photon, (c) linear polarized photon pair. At $\varphi = 0$, π, the photon pair is in "instant hidden state", at $\varphi = \pi/2$, $3\pi/2$ the photon pair is in "apparent state". Here only shows four phase point. Here $\varphi = kz - \omega t$.

(C) In QED [6, 7], the ω is regarded as the photon oscillation frequency, the $\pm\hbar$ is the spin angular momentum of photon. We suppose that photon never oscillates though the EM fields do, ω is the polarization-rotation angular frequency of photon, i. e., spin frequency. The wave number k is the number of its rotating periods per unit length along the motional direction. For the plane wave photons in vacuum, there is disperse relations $\vec{\omega} = c\vec{k}$ for right-hand spin, $\vec{\omega} = -c\vec{k}$ for left-hand spin. The time spent for a photon to travel a distance of a wavelength (λ) along its k direction exactly equals to that for it to rotates a full circle (2π radian) in the free space, as shown in Fig. 2. The orientations of $\mathbf{E}_p$, $\mathbf{H}_p$ define the rotating phase of photon (see Fig. 2). Here we define the translating speed of a photon as the group velocity v_g,

$$v_g = \frac{p}{m} = \frac{\hbar k}{m} \tag{3}$$

The phase speed of a photon is defined as the moving velocity of its rotating phase

$$v_p = \frac{\omega}{k} = f\lambda \tag{4}$$

The fact of $v_g = v_p = c$ in the free space means that the photon translation motion and its spin rotating one are fully synchronous. Obviously, the energy of a photon can be rewritten as

$$E = mc^2 = mv_g v_p = pv_p \tag{5}$$

In the unbounded vacuum space, a photon behaves helically, i. e., to translation at light speed c and spin at ω. At any instant moment its energy, momentum and spin angular momentum keep conservation. Its wavelike property originates from the periodicity of its helically rotating motion, instead of its wave spreading all over the place. In fact, According to our model the centroid of photon behaves beeline motion while its $\mathbf{E}_p$ and $\mathbf{H}_p$ behave helically motion. This is a necessary consequence of conservation of energy, momentum and spin angular momentum of photon.
(D) When the same frequency photons interfere each other, two typical cases occur. When the $\mathbf{E}_p(\mathbf{H}_p)$ of two coherent photons is in the same direction, the synthesized electric (magnetic) vector

equals to $2\mathbf{E}_p(2\mathbf{H}_p)$, the "apparent energy" is four times single-photon energy. This is called "apparent state" of photons. It is equivalent to four photons to present. Likewise, when the $\mathbf{E}_p(\mathbf{H}_p)$ of two coherent photons is opposite, the synthesized electric (magnetic) vector is null. It is equivalent to no photons to present. This is called "hidden state" of photons. Actually the $\mathbf{E}_p(\mathbf{H}_p)$ of two coherent photons can intersect at any angle within 0°–360°.

In a sense, a LP light is formed by the RCP light and the LCP light synchronously interfere each other. There two knots during a period. The knot points of the wave indicate the field energy being zero, however the energy flow of each photon still proceed forward (refer to Fig. 2). We call the knots as "instant hidden states". Obviously the combined energy flow or power flow does not equal to the sum of the energy flows of two photons, whereas depends on their rotating phases. Therefore for a LP light, its instant energy flow varies versus time, only the average energy flow keeps constant.

In the free space, each photon like material point, moves forward at light speed c by virtue of its inertia. The plane wave solution could describe not only a light beam, but also a single-photon.

$$\vec{E}(0,0,z,t) = E_p(\vec{e}_x + i\vec{e}_y)\exp[i(kz-\omega t)], \quad \vec{H}(0,0,z,t) = H_p(\vec{e}_x + i\vec{e}_y)\exp[i(kz-\omega t)] \tag{6}$$

$$\vec{E}(0,0,z,t) = E_p(\vec{e}_x - i\vec{e}_y)\exp[i(kz-\omega t)], \quad \vec{H}(0,0,z,t) = H_p(\vec{e}_x - i\vec{e}_y)\exp[i(kz-\omega t)] \tag{7}$$

$$\vec{E}(0,0,z,t) = E_p\vec{e}_x\exp[i(kz-\omega t)], \quad \vec{H}(0,0,z,t) = H_p\vec{e}_y\exp[i(kz-\omega t)] \tag{8}$$

here $\vec{\mathbf{e}}_x$ and $\vec{\mathbf{e}}_y$ is unit vector along x and y axis respectively. Eqs. (6), (7) and (8) describe the left-rotating, the right-rotating photon and the LP photon pair, as shown in Fig. 2(a), (b) and (c) respectively. It is easy to see that a LP photon pair can be decomposed into two CP photons, and vice versa according to equations (6), (7) and (8). An elliptic-polarized light is composed of two orthogonal LP lights with different number of photon-pairs.

(E) That light-line bends in a gravity field indicates that photon follows mechanical laws. According to equivalence principle in general relativity, the motional mass of photon has the physical significance of inertia too. We suppose that a photon is a spherical particle with radius as $\mathbf{r}_p$. According to Newton mechanics the moment of inertia of the photon is

$$I = \frac{2}{5}mr_p^2 \tag{9}$$

Thus the angular momentum of the photon

$$I\omega = \hbar \tag{10}$$

Substituting Eqs. (9) in (10), we obtain the radius of photon

$$r_p = \sqrt{\frac{5}{2}}\frac{c}{\omega} = \sqrt{\frac{5}{2}}\frac{1}{k} = \sqrt{\frac{5}{2}}\lambdabar \approx \frac{\lambda}{4} \tag{11}$$

Obviously it is reasonable that the photon size is inverse proportional to its frequency. The higher the frequency the clearer the nature of the particle-like for the photon. According to the current point of view, any elementary particle should be geometrical point without size which demanded by relativity. However, photon without rest mass has no Lorentz shrinkage in its motional direction, so its size is not restricted by special relativity. In order to construct a concept of magnitude of order, we enumerate the size of a gamma-ray photon with energy 0.511 MeV below.

$$r_p = \sqrt{\frac{5}{2}}\frac{1}{k} = \sqrt{\frac{5}{2}}\frac{\hbar c}{E} \approx 6.12 \times 10^{-13}(m)$$

In order to obtain the estimation of the field matter intensity of the photon, let

$$\left(\frac{\mu_0}{2}H_p^2 + \frac{\varepsilon_0}{2}E_p^2\right)\frac{4}{3}\pi r_p^3 = mc^2 = hv \tag{12}$$

where μ_0 is vacuum permeability, ε_0 the vacuum permittivity. Due to the electric energy equals to the magnetic energy, Eq. (12) can be substituted as

$$\varepsilon_0 E_p^2\left(\frac{4}{3}\pi r_p^3\right) = hv \tag{13}$$

Substituting Eqs. (11) in (13), we obtain

$$E_p = \left(\frac{3}{2\sqrt{10}}\right)^{3/2}\sqrt{\frac{\hbar\mu_0}{\pi^3 c}}\omega^2, \quad H_p = \left(\frac{3}{2\sqrt{10}}\right)^{3/2}\sqrt{\frac{\hbar\varepsilon_0}{\pi^3 c}}\omega^2 \tag{14}$$

For the gamma-ray photon with energy 0.511 MeV, $E_p = 9.28\times10^{16}$ (V/m); $H_p = 2.46\times10^{14}$ (A/m). Next we will use the theoretic model proposed above to explain two optical phenomena.

2. YOUNG'S INTERFERENCE EXPERIMENT

For Thomas Young's double-slit interference experiment, in spite of the probability result calculated by QM wave function theory consistent with that by classical E&M wave theory, QM by no means gives the idiographic physical mechanism occurred. From the view point of classical photon particle, the origin of this interference pattern is very difficult to understand. In order to do so Bohm introduce quantum mechanic potentials [8] in his hidden variables theory. According to the hypotheses proposed above it is easy to explain Young's experiment. The mechanism of the interference pattern formation is the same as that of classical wave. In the bright fringe of the interference pattern, two photons from two slits respectively add in phase, resulting in $2\mathbf{E}_p(2\mathbf{H}_p)$. As if the energy is four times single-photon energy, but it does not mean four photons to appear. Here we call the two as in "apparent state" of photons. In the dark fringe, two photons from two slits respectively are 180° out of phase, causing destructive interference and the synthetic field quantity $\mathbf{E}_p - \mathbf{E}_p = 0$, $\mathbf{H}_p - \mathbf{H}_p = 0$. Here we call the two in "hidden state" of photons which only do not cause light-chemical reactions or photoelectric effect. The true meaning of the interference pattern only indicates the phase message of photons in the detection plane, instead of re-distribution of photons themselves.

3. INTERFERENCE WITH POLAROID SHEET

Consider a beam of LP light incident normally on a polarizer with polarization plane and the orientation of optic axes made an angle of 45°. How do the photons cross through the polarizer? How do they change their polarization direction? Eexisting QM textbooks [9, 10] can by no means answer what the physical mechanism is. And the popular point of view deems that QM needs not answer this sort of problems. According to the hypothesis (A), (B), (C), (D) above in the model, an objective mechanism can be given.

In fact, LP light is composed of RCP and LCP-rotating photon pairs. There exists no LP single-photon. For a light beam, there always are many photon pairs impinging on the polarizer simultaneously at a time. These photon pairs possess different phase each other generally. As shown in Fig. 3 photon pairs aa' and bb' are LP photon pairs with respect to yz plane. When they

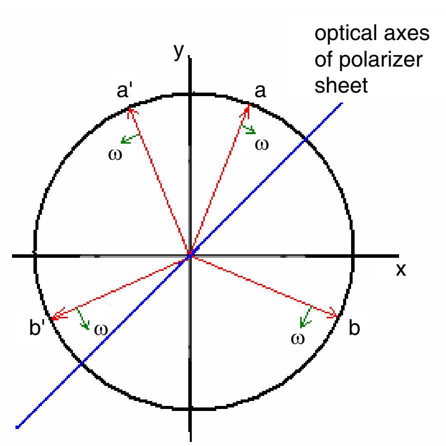


Figure 3: The polarization plane (yz) of LP light makes 45° angle with the optic axis of polarizer.

meet with the polarizer, photon a' and b combine a new LP photon pair parallel to the optical axes of polarizer, so they pass through it. While photon a and b' combine another LP photon pair perpendicular to the optical axes of polarizer, thus they is obstructed. This way the light after the polarizer, as a macroscopic effect changes its polarized plane parallel to the optic axis of polarizer. For a relative intense light beam, the phase distribution of a great number of LP photon pairs substantially is random. When they reach the polarizer, almost every photon can find a new partner to make a new pair either parallel to or perpendicular to the optic axis. Therefore the probability of the passing through to the obstructed is 50 to 50.

4. SUMMARY

A photon is characterized by its $\mathbf{E}_p$, $\mathbf{H}_p$ and $\mathbf{k}$. Because each photon consists of vector field matter, the combined field depends on their rotating phase. Like particle else the photon also follows the conservation laws of energy, momentum and angular momentum. The model presented here describes the physical mechanism of the wave-particle duality of a single photon. It can objectively illustrate the physical reality. This model gives a new and rational interpretation to Young's interference and polarizer experiments. It need not to give up the objective process to physical phenomena. According this model we may have a unified understanding between the classical electromagnetic field and the quantum particle of light.

REFERENCES

1. Perkowitz, S., *Empire of Light*, Joseph Henry Press, Washington D.C., 1998.
2. Taylor, G. I., *Proc. Cambridge Philos Society*, Vol. 15, 114, 1909.
3. Grangier, P., G. Roger, and A. Aspect, *Europhysics Letters*, Vol. 1, 173–179, 1986.
4. Wang, K., J. Xiong, and L. Gao, "From Hanbury-Brown and Twiss experiment to the second-order double-slit interference for incoherent light," *Front. Phys.*, Vol. 1, 54–66, China, 2006.
5. Feynman, R.-P., R.-B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Vol. 3, Chap. 1, Reading, Addison-Wesley, MA, 1965.
6. Berestetskii, V. B., E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics*, 2nd Edition, Pergamon Press Ltd., 1982. .
7. Gupta, S. N., Q*uantum Electrodynamics*, Gordon and Breach Science Publishers, 1977.
8. Bohm, D., *Phys. Rev.*, Vol. 85, No. 2, 166–193, 1952.
9. Fayer, M. D., *Elements of Quantum Mechanics*, Oxford University Press, 1–5, 2001.
10. Zeng, J. Y., *Quantum Mechanics*, Science Press, Beijing, 1982.

Longitudinal Power Distribution and Corresponding Temperature Distribution in a RF Waveguide CO_2 Laser

A. Rauf, Bingxin Zhang, and Jianguo Xin

Department of Physical Electronics, School of Information Science and Technology
Beijing Institute of Technology, Beijing 100081, China

Abstract— The output power of a gas laser can not be increased by increasing the input power due to the increase of the gas temperature. As the temperature of the laser gas mixture increases up to 600 K, the amplification of the amplifying medium decreases rapidly and so the output power decreases. In this paper, longitudinal power distribution and corresponding temperature distribution along the CO_2 waveguide laser are studied theoretically. Effect of voltage and current distributions along and transverse to the electrode direction respectively are accounted for the studies. It is seen that the temperature decreases exponentially along the length of the electrodes on both sides of the feed point.

Several papers [1] have already described the huge potentiality of large area diffusion cooled RF capacitive discharges for the construction of medium — to high power CO_2 lasers for materials processing. Indeed, some industrial sources based on this technology are presently appearing [2]. However, there exist some open problems that limit the simple scaling of the above-mentioned technique to higher power segments. One of the major problem is certainly the attainment of a uniform gain medium excitation because of transmission line effects [3] naturally determined by electrode dimensions comparable with the exciting RF wavelength. Typical counter measures to this difficulty are the adoption of sectioned discharges or the use of reactive elements along the electrodes. The principles for the correct design of systems based on these schemes are well established in the case of narrow channel devices [4–6], in which the equivalent-line characteristic impedance is mainly determined by the electrode structure.

In [7] it is demonstrated that the same smoothing principals could be applied to wide channel high power discharges provided that the influence of the discharge loading be taken into account. Indeed, from the analysis carried out in [7], it is possible to conclude that the main role played by the discharge loading is that of determining a new equivalent transmission line whose characteristic impedance is closer to that produced by the capacitance of the ion sheaths surrounding the electrodes rather than by the channel capacitance. This effect determines the voltage and thus power distribution closer to a lossless-line rather than lossy-line models. Moreover [8, 9], the typical impedances of sheaths and neutral plasma for the mixtures and pressure of the interests for CO_2 laser construction produce distributions hardly distinguishable from uniform transmission line models [7, 10, 11].

In this paper the temperature distribution along the longitudinal direction of the electrodes is studied theoretically by considering the longitudinal power distribution by applying the transmission line theory. Voltage distribution is considered along the longitudinal direction while the current in transverse direction of the electrodes. The voltage and current distributions along the electrodes are given by solving the simple transmission line differential equations [2, 7, 10–12]:

$$\frac{d^2I(x)}{dx^2} - \gamma^2 I(x) = 0 \tag{1}$$

$$\frac{d^2V(x)}{dx^2} - \gamma^2 V(x) = 0 \tag{2}$$

where

$$\gamma = \sqrt{ZY(x)} = \sqrt{(R + j\omega L)(G(x) + j\omega C)} \tag{3}$$

And $Z = R + j\omega L$ and $Y(x) = G(x) + j\omega C$. Where L is the longitudinal inductance per unit length, C is the electrode's structure capacitance per unit length and $G(x)$ is the local plasma conductance. We neglect the longitudinal resistance being mainly due to the metal electrodes and the values of L and C are adjusted by iteration. The general solutions of the Eqs. (1) and (2) are given by:

$$I(x) = Ae^{\gamma x} + Be^{-\gamma x} \tag{4}$$

$$V(x) = Ce^{\gamma x} + De^{-\gamma x} \tag{5}$$

where A, B, C, and D are the constants and can be determined by applying the suitable boundary conditions. Considering the feed point at the centre and applying the condition that at $x = \pm d$; $I = 0$, where "d" is length of the electrode on one side of the central feed point, we have the relation for current distribution from Eq. (4) as:

$$I(x) = (e^{rx} - e^{-r(x+2d)} + e^{-rx} - e^{r(x-2d)}) \tag{6}$$

Eq. (6) represents the current flowing along the longitudinal direction. In order to find the voltage distribution along the longitudinal direction, part of the current flowing transverse to the electrode is considered, i.e.,

$$\frac{dI(x)}{dx} = -YV(x) \tag{7}$$

From Eqs. (6) and (7) we have the relation for voltage distribution as:

$$V(x) = \frac{r}{Y}(-e^{\gamma x} - e^{-r(x+2d)} + e^{-\gamma x} + e^{r(x-2d)}) \tag{8}$$

In real facts the discharge structure is still more complicated than the uniform transmission line model considered so far. Indeed the plasma conductivity is a function of the local voltage over the electrodes and thus cannot be considered uniform. The plasma conductance is given by [8, 9]:

$$G_{pl} = \frac{1}{r_{pl} - \frac{j}{\omega c_s}} \tag{9}$$

where, r_{pl} is the plasma resistance per unit length and c_s is the sheath's capacitance per unit length. The plasma resistance is given by:

$$r_{pl} = \frac{V_\alpha}{|I_{pl}|} \tag{10}$$

where V_α is a constant voltage varying with the kind of gas mixture and pressure. From the data reported in [9] it can be taken in the interval 40–70 V. In our calculations we have taken its value as 50 V. The value of I_{pl} can be obtained from Eq. (6). The sheath's capacitance in Eq. (9) is given by [7]:

$$c_s = \frac{\varepsilon_0 A}{d_s} \tag{11}$$

where d_s is the sheath's thickness and "A" is the area of the electrode. The value of d_s can be calculated from the invariant law given in [9] as $fd_s = 42$, where, "f" is the excitation frequency. From Eqs. (9)–(11) we obtain:

$$G_{pl}(x) = jwc_s \left[1 - \frac{V_\alpha}{|V(x)|} \exp\left(j \arctan\left(\frac{\mathrm{Im}G_{pl}}{\mathrm{Re}G_{pl}}\right)\right)\right] \tag{12}$$

Considering that:

$$I_{pl}(x) = G_{pl}(x)V(x) \tag{13}$$

Since we have the relation for the power as:

$$P(x) = V(x)I(x) \tag{14}$$

The change in temperature ΔT due to change in the heat energy Q is given by:

$$Q = mC_p(T)\Delta T \tag{15}$$

where, $C_p(T)$ is the temperature dependent specific heat capacity at constant pressure for the gas mixture [13, 14] and "m" is the mass of gas mixture. We also know that $Q = P(x)\ t$, where, "t" is applied pulse width. From Eqs. (14) and (15) the relation for the temperature distribution is given as:

$$T(x) = T_0 + \frac{P(x)t}{mC_p(T)} \tag{16}$$

where, T_0 is the room temperature.

The temperature distribution along the longitudinal direction of the waveguide channel is studied theoretically while the gas is pumped through a pulse of pulse length 10 μsec [15]. The ratio of the laser gas mixture is CO_2:N_2:He:Xe=1:1:3:0.25% and the pressure of gas is considered to be as 80 torr. The values for L and C used in this analysis are 2.4 nH/cm and 40 pF/cm respectively. Since the resistance of the plasma is not a constant quantity while the discharge is in the running condition so the distribution of plasma resistance is also included.

The current distribution along the longitudinal direction of the electrode system is shown in Fig. 1. It is clear from the figure that the current at the centre is maximum and it decreases exponentially on both sides of the feed point with the increase of position. It is due to the fact that as the input energy is transmitted along the longitudinal direction, the gas molecules are heated up and so the resistance of the gas plasma increases which results in the decrease of the current.

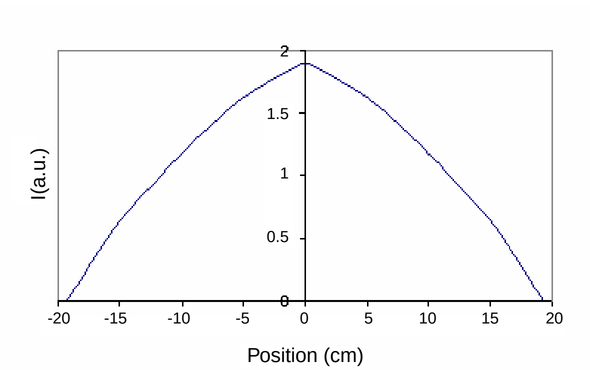


Figure 1: Current distribution along the longitudinal direction.

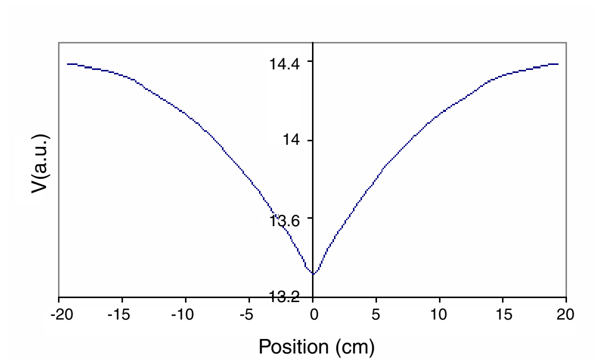


Figure 2: Voltage distribution along the longitudinal direction.

Figure 2 represents the voltage distribution along the longitudinal direction of the electrode. It can be seen that the voltage increases exponentially toward the ends of the electrodes on both sides of the feed point with the increase of position. The results are in good agreement with the results published elsewhere [2, 7].

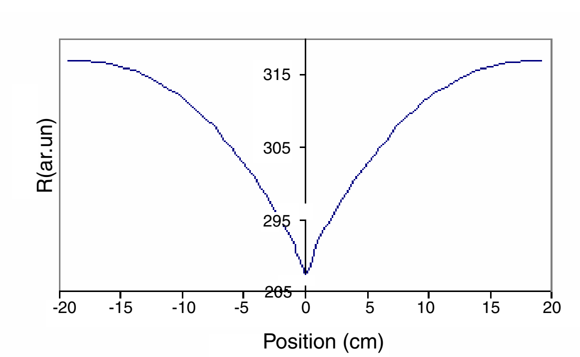


Figure 3: Resistance distribution along the longitudinal direction.

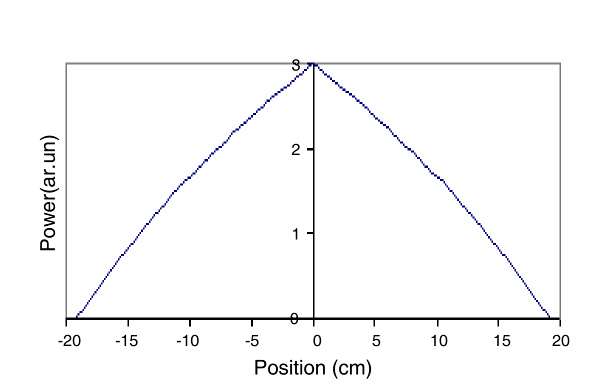


Figure 4: Power distribution along the longitudinal direction.

Figure 3 represents the resistance distribution along the longitudinal direction. It can be seen that the resistance increases towards the electrodes ends and is minimum at the central feed point. It is due to the fact that the neutral plasma resistance increases with the increase of the time in pulse mode operation. This is attributed to the different gas composition due to the discharge driven molecular dissociation and consequent formation of new species [2].

Figure 4 represents the distribution of power along the longitudinal direction. It is clear from the figure that it is maximum at the central feed point and decreases exponentially towards the ends of the electrodes. If we consider the Figs. 1, 3 and 4 it becomes clear that the current is maximum at the central feed point, the resistance of the plasma is minimum at the central feed point resulting in the maximum power at the centre. Indeed in the pulsed regime, the dynamics of power transfer to the gas has to be taken into account, producing time-varying density distributions deeply different from those of the CW operation, and often determining pressure waves during the pulse [16]. So the

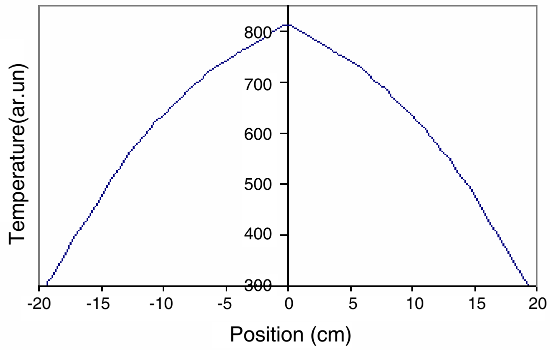


Figure 5: Temperature distribution along longitudinal direction.

inhomogeneity is to be taken into account not only caused by the locally different electron density but also to the locally different gas density and temperature. Moreover the equilibrium between dissociation and recombination of species could be a source of nonhomogeneous effects considering the relevant fraction of energy transferred to these processes.

Figure 5 represents the temperature distribution along longitudinal direction. It can be seen from the figure that the temperature decreases exponentially towards the ends of the electrode structure. This decrease of the gas temperature in the laser cavity can be explained by considering the direct relationship of power and temperature (Eq. (16)). As the power transferred to the gas molecules is maximum (Fig. 4) at the center and decreases exponentially with the increase of the position from the central feed point so the temperature is maximum at the center and decreases exponentially with the increase of the position on both sides from the central feed point.

REFERENCES

1. Hall, D. R. and H. J. Baker, "Diffusion-cooled large surface area CO_2/CO lasers," *Proc. SPIE*, Vol. 2502, 12–19, 1994.
2. Lapucci, A., S. Mascalchi, and R. Ringressi, "Parameters affecting the power distribution in the radio frequency discharge of large-area diffusion-cooled CO_2 lasers," *IEEE J. Quant. Elect.*, Vol. QE 34, No. 4, 616–621, 1998.
3. Hall, D. R. and C. A. Hill, "Radio frequency-discharge-excited CO_2 lasers," *Handbook of Molecular Lasers*, P. K. Cheo, Ed. Marcel Dekker, New York, 1987.
4. Griffith, G. A., "Improved discharge uniformity for transverse radio frequency waveguide CO_2 lasers," *Proc. SPIE*, Vol. 335, 69–71, 1982.
5. Chenausky, P. P., R. A. Hart, L. A. Newman, and N. N. Hoffman, "High-capacitance parallel-resonance RF-excited waveguide CO_2 lasers," CLEO'82, Tech. Dig., Phonix, AZ, Paper THN2, 1982.
6. Kim, Y., C. E. Youn, J. W. Ra, and Y. S. Kim, "Method for reducing the longitudinal voltage variation in transverse radio-frequency discharge waveguide lasers," *J. Appl. Phys.*, Vol. 67, 1127–1129, 1990.
7. Lapucci, A., F. Rossetti, M. Ciofini, and G. Orlando, "On the longitudinal voltage distribution in radio-frequency-discharged CO_2 lasers with large-area electrodes," *IEEE J. Quant. Elect.*, Vol. 31, 1537–1542, 1995.
8. Vitruk, P. P., H. J. Baker, and D. R. Hall, "The characteristics and stability of high power transverse radio-frequency discharges for waveguide CO_2 slab lasers excitation," *J. Phys. D.*, Vol. 25, 1767–1776, 1992.
9. Vitruk, P. P., H. J. Baker, and D. R. Hall, "Similarity and scaling in diffusion cooled RF-excited carbon dioxide lasers," *IEEE J. Quant. Elect.*, Vol. 30, 1623–1634, 1994.
10. Strohschein, J. D., W. D. Bilidia, H. J. J. Seguin, and C. E. Capijack, "Enhancing discharge uniformity in a multi-kilowatt radio frequency excited CO_2 slab laser array," *Appl. Phys. Lett.*, Vol. 68, 1043–1045, 1996.
11. Strohschein, J. D., W. D. Bilidia, H. J. J. Seguin, and C. E. Capijack, "Computational model of longitudinal discharge uniformity in RF-excited CO_2 slab lasers," *IEEE J. Quant. Elect.*, Vol. 32, 1289–1298, 1996.
12. Chopman, R. A., *Schaum'S Outline of Theory and Problems of Transmission Lines*, Mc Graw Hill Book Company, New York, 1968.

13. Assael, M. J., J. P. M. Trusler, and T. F. Tsolakis, *Thermophysical Properties of Fluids*, Imperial College Press, 1996.
14. Iynkaran, K. and D. J. Tandy, *Basic Thermodynamics, Applications and Pollution Control*, Printice Hall, Singapore, 1993.
15. Rauf, A., W. Zhou, and J.-G. Xin, "Theoretical studies of the output pulse with the variation of the pumping pulse for RF excited CO_2 pulsed waveguide laser," *J. China Ordnance*, Vol. 2, No. 1, 37–42, 2006.
16. Lapucci, A., G. Cangioli, and R. Meucci, "Frequency stability of Rf-excited CO_2 waveguide lasers in gain-switched and Q-switched regimes," *IEEE J. Quant. Elect.*, Vol. 29, 926–932, 1993.

Lasing Dynamics of a Novel Silicon Photonic Crystal Cavity

Shouyuan Shi and Dennis W. Prather
Electrical and Computer Engineering Department, University Of Delaware, Newark, DE 19716, USA

Abstract— In this paper we proposed novel silicon laser design based on the dispersion engineered photonic crystals. With the unique self-collimation property, we design an optical cavity with a Q factor of 2000. The strong mode confinement in the low index active material offers an opportunity to realize a lasing mechanism. To investigate the lasing dynamics we introduce the rate equations of atomic system into the electromagnetic polarization. With these auxiliary differential equations we solve the time evolutions of the electromagnetic waves and atomic populations by using the FDTD method.

1. INTRODUCTION

Silicon, the leading material in microelectronics during the last four decades, also promises to be the key material in the future. Recent activities have focused on the achieving active functionalities, mostly light amplification and generation. However, due to the fundamental limitation related to the indirect bandgap of the bulk Si, the development of silicon gain and laser becomes one of most challenging goal in the silicon photonics and optoelectronics. Many different approaches have been taken to achieve this goal. Recently stimulated Raman scattering effect has been used to demonstrate the light amplification and lasing in the silicon both in pulse and continuous-wave operation. [1,2]In the presented work, we attempt to design a novel silicon laser.

To this end, we will design a novel optical cavity based on the dispersion engineered photonic crystals (PhCs). To observe the lasing dynamics, we will incorporate the rate equations of a four level atomic system into the device to simulate the gain and absorption of the active material. By solving the Maxwell's equations with these auxiliary differential equations with the Finite-difference Time-domain (FDTD) method, we will track the time evolutions of the electromagnetic waves and atomic populations.

2. DISPERSION BASED PHOTONIC CRYSTAL CAVITY

The light propagation in a PhC is most appropriately interpreted through a dispersion diagram, which characterizes the relationship between the frequency of the wave, ω, and its associated wavevector, $\mathbf{k}$. Dispersion surfaces provide the spatial variation of the spectral properties of a certain band within the photonic crystal structure. Electromagnetic wave vectors propagate at directions normal to the dispersion surface as shown in Fig. 1, which stems from the relation

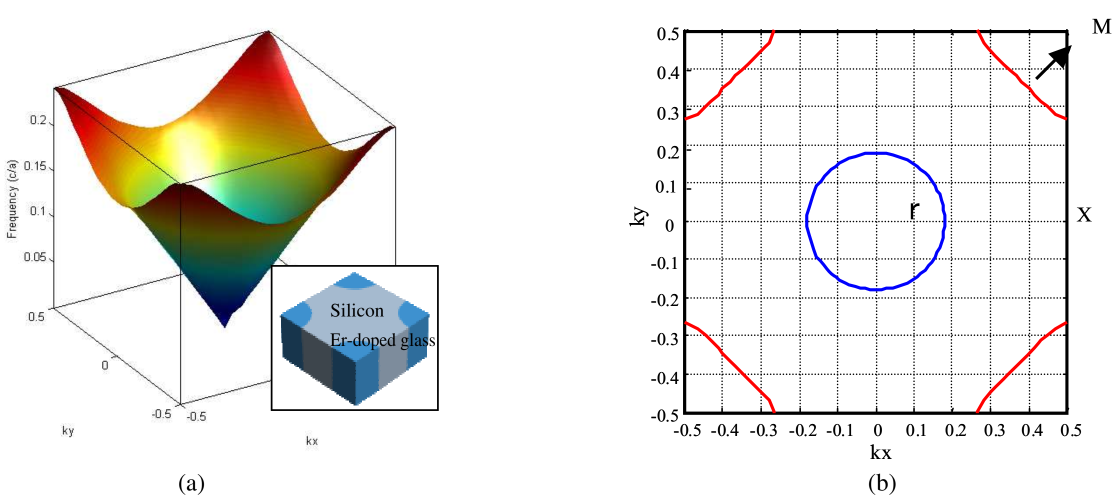
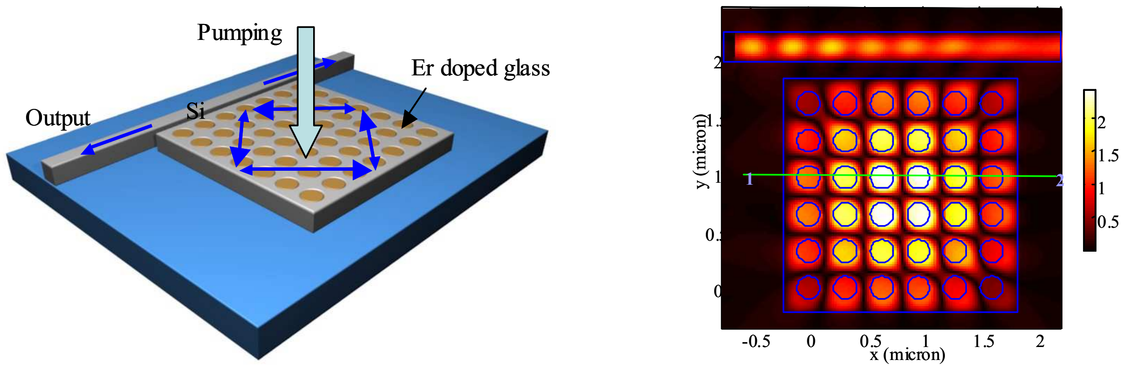


Figure 1: (a) The dispersion surface of the first band for silicon PhC with circular holes filled with Er-doped glass. (b) Flat EFCs perpendicular to ΓM direction at frequency of 0.18c/a.



(a) (b)

Figure 2: (a) Schematic chart of silicon laser based on dispersion engineered photonic crystals where the active material is introduced by backfilling the air holes of the PhC. (b) The highest cavity mode.

$\mathbf{v}_g = \nabla_k \omega(\mathbf{k})$. The equil frequency contour (EFC) can lead to beam divergence or convergence as shown in the figure. The ability to shape the EFCs, and thereby engineer the dispersion properties of the PhC, opens up a new paradigm for the design of optical devices. [3–5] For self-collimation, we desire a flat EFC, in which case the wave is only allowed to propagate along those directions normal to the sides of the straight curvatures. As such, it is possible to vary the incident wavevector over a wide range of angles and yet maintain a narrow range of propagating angles within the PhC.

Consider silicon photonic crystals perforated by a square lattice with holes back-filled by the gain medium, i.e., Er-doped glasses, as shown in the inset of Fig. 2(a). The hole has radius of 0.3a, where a is the lattice constant. The silicon and glass have refractive indices of 3.5 and 1.5, respectively. The dispersion surface of first band diagram is plotted in Fig. 1(a). By carefully selecting the frequency, one can obtain a flat curvature within certain angular range at specified frequency, i.e., 0.18c/a, as depicted in Fig. 1(b), where the blue curvature is the dispersion contour in the free space with frequency of 0.18a. Such flatness of the curvature offers self-collimation along ΓM direction within a wide incident angular range.

Self-collimation effect has been widely used in the applications of optical routing, sensors, variable beam splitters, self collimated optical emitters to enhance the emission from a light emitting diode. In the presented paper we propose a novel design of silicon laser based on the self-collimation phenomena. Fig. 2 depicts the schematic chart of our design. A photonic crystal cavity consists of silicon square slab perforated by a $N \times N$ array of square lattice. Based on the dispersion property that we discussed above the electromagnetic waves near the band edge of first band propagate along ΓM direction. With the assistance of the clear four edges of silicon slab as mirrors, a cavity can be formed along the optical path, illustrated as arrow loop. At one edge a conventional waveguide is used to coupling into or out from the cavity. The system activated by the optical pumping. The gap between the cavity and waveguide is critical design parameter.

We simulate the proposed device by using FDTD method. 7×7 array is initially simulated; the highest resonant mode is shown in Fig. 2(b). As we can observed that most optical mode are confined within the low index materials, i.e., Er-doped glass. This unique property will benefit the lasing mechanism and lower the threshold optical pumping. The quality factor is measured in this case as 2000. In addition, we found as the number of array increases the resonant frequency shift to high end, and the Q-factor will increase accordingly, which offer us another degree of freedom to tune the Q factor.

3. RATE EQUATIONS FOR A FOUR-LEVEL ATOMIC SYSTEM

To simulate a laser dynamics in an optical cavity associated with a gain or active material, a conventional rate equation model [6] can be employed for the simulations of the nonlinearity and dispersion properties, i.e., a uniform glass host containing a dispersion of erbium ions. Coupled with Maxwell's equations, we are able to model the time evolution of the atomic energy level populations as well as the optical signal propagation, amplification and absorption in the devices. In this paper we proposed a simplified realistic four-level atomic system as sketched in Fig. 3, where the principal transitions induced by the presence of the pump and signal beams are included. The absorption of pump radiation at 980 nm wavelength promotes the electrons from level 0 to level 3. After decaying to level 2, these electrons provide the signal gain at 1535 nm via the transition to level 1. Here we describe the time domain population dynamics by using a four-level rate equation formulism:

$$\left\{\begin{array}{l}\dfrac{dN_3(t)}{dt}=W_p(t)-\dfrac{N_3(t)}{\tau_3}\\ \dfrac{dN_2(t)}{dt}=\dfrac{N_3(t)}{\tau_{32}}+\dfrac{1}{\hbar\omega_a}\mathbf{E}(t)\cdot\dfrac{d\mathbf{P}(t)}{dt}-\dfrac{N_2(t)}{\tau_2}\\ \dfrac{dN_2(t)}{dt}=\dfrac{N_3(t)}{\tau_{31}}-\dfrac{1}{\hbar\omega_a}\mathbf{E}(t)\cdot\dfrac{d\mathbf{P}(t)}{dt}+\dfrac{N_2(t)}{\tau_{21}}-\dfrac{N_1(t)}{\tau_1}\end{array}\right. \tag{1}$$

where Ni is the population. The ground state population N_0 is assumed to be very large compared to the populations of the high energy levels and is basically constant with time. τ_{ij} are the lifetime associated with the transitions from energy E_i and E_j, τ_j are the lifetime from energy E_i to all low levels, and w_{ij} are the signal simulated transitions probabilities from energy E_i and E_j. ω_a is resonant frequency of the materials related to the atomic transition energy levels through $\omega_a=(E_1-E_2)/\hbar$.

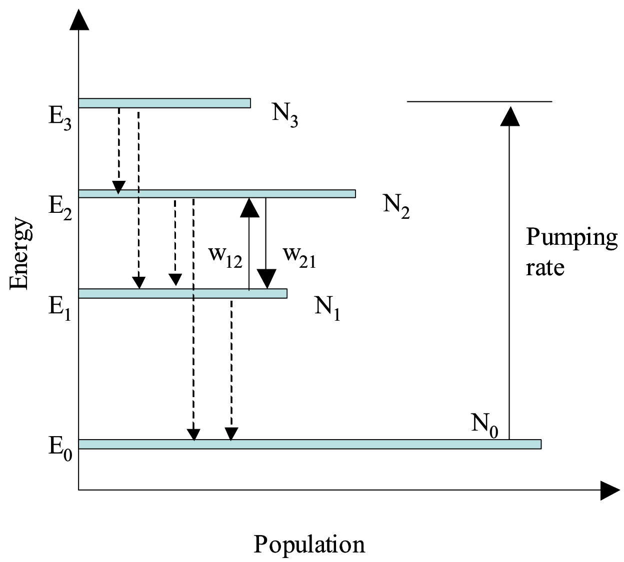


Figure 3: Simplified four energy level diagram.

Based on classical electron oscillator model, the net macroscopic polarization $\mathbf{P}(t)$ induced with the presence of applied electric field $\mathbf{E}(t)$ for an isotropic medium can be described by the following equations,

$$\frac{d^2\mathbf{P}(t)}{dt^2}+\Delta\omega_a\frac{d\mathbf{P}(t)}{dt}+\omega_a^2\mathbf{P}(t)=\kappa\Delta N(t)\,\mathbf{E}(t) \tag{2}$$

where $\Delta N(t)$ is instantaneous population difference between energy levels 1 and 2, which is given by $\Delta N(t)=N_1(t)-N_2(t)$. $\kappa=3F_{osc}e^2/m$ is a constant and $\Delta\omega_a$ is the total energy decay which describes the transition linewidth.

Then the effects of the nonlinear and active medium on the propagation of electromagnetic waves are incorporated through the polarization response of the medium. Coupled with the Maxwell's equation in our optical devices, we are able to simulate the proposed photonic crystal laser cavity.

4. DYNAMICS OF ELECTROMAGNETIC WAVE IN LASER SYSTEM

Based on the theory we developed, in this section we will introduce Finite-difference Time-domain (FDTD) to solve the coupled equations for the application of lasing dynamics.

As is well known, the FDTD method has become one of the more widely used numerical techniques for solving electromagnetic boundary value problems. In short, the FDTD method directly discretize Maxwell's equations in both time and space by using central difference technique, and iteratively updates the electrical and magnetic field components at all points in the computational space using the leapfrog fashion. In our simulation of the lasing action, we discretize the auxiliary differential equations as well, i.e., rate equations that govern the populations, in both time and space for the time marching.

As shown in the schematic view of proposed novel silicon laser in Fig. 2, the holes are back filled with an active material, i.e., highly doped Er glass. The medium is uniformly pumped by another 980 nm source. The lifetime are given by $\tau_1 = 1e - 9s$, $\tau_2 = 1.35e - 7s$, $\tau_{21} = 1.35e - 7s$, $\tau_3 = 1e - 10s$, $\tau_{31} = 1e - 6s$, and $\tau_{32} = 0.99e - 10s$. The transition frequency associated with the energy levels E_2 and E_1 was chosen as 200 THz and line width was taken to be 6 THz. The radiative decay rate for this from E_2 and E_1 transition was 8×10^8/s. The pump rate into level E_3 was chosen $2 \times 10^{29}\,\mathrm{m}^{-3}$.

In the FDTD simulation, a total time of 15 ps is simulated. The evolution of electromagnetic waves starts from an initial small noise over the photonic crystal cavity. A detector is placed in the waveguide to monitor the lasing dynamics, and the total field in the output straight waveguide is illustrated in Fig. 4(a). In the figure we can obviously lasing performance. For the clarity, we zoom the transient plot near 12ps as shown in Fig. 4(b), which indicates nearly single lasing model. In addition, a snap shot of total electric fields at time of 15 ps is plotted in Fig. 4(c), where the high intensity of electric field can be observed in the cavity area, particularly in the low index materials and steady results are outputted from both ends of straight waveguide.

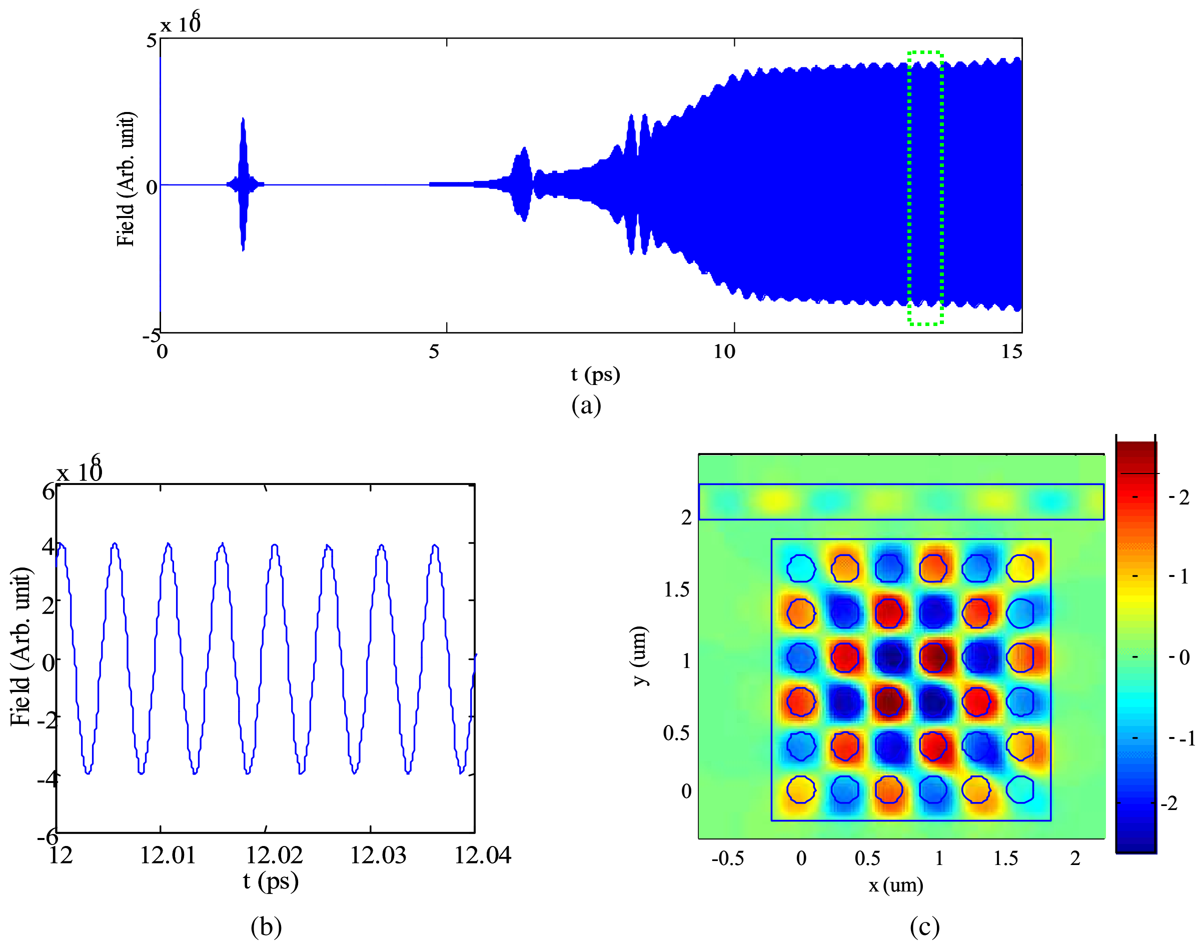


Figure 4: (a) Lasing dynamics by monitoring output in the straight dielectric waveguide, (b) steady output of single optical mode as shown in the time window in figure (a). (c) A snap shot of 2D field distribution.

5. SUMMARY

In this paper we proposed a novel silicon laser design based on the dispersion engineering of photonic crystals. By carefully designing the photonic crystals, we took advantages of the self-collimation property to suppress the optical cavity modes compared to a bulky square cavity. In addition to this, by changing the number of array of holes, we can tune the Q factor. To simulate the lasing dynamics in such optical cavity, we incorporated the rate equations of a four-level atomic system into Maxwell's with the assistance of the induced electric polarization contributed from the response of a collection of atoms. Finally, we numerically implement time evolutions of electromagnetic waves and atomic populations by using FDTD method to the application of our engineered dispersion based photonic crystal laser.

REFERENCES

1. Rong, H., R. Jones, A. Liu, O. Cohen, D. Hak, A. Fang, and M. Paniccia, "A continues-wave Raman silicon laser," *Nature*, Vol. 433, 725–728, 2005.
2. Boyraz, O. and B. Jalali, "Demonstartion of a silicon Raman laser," *Optics Express*, Vol. 12, 5269–5273, 2004.
3. Prather, D. W., S. Shi, D. Pustai, C. Chen, S. Venkataraman, A. Sharkawy, G. schneider, and J. Murakowski, "Routing optical waves without waveguides," *Optics Letters*, Vol. 29, 50, 2004.
4. Shi, S. Y., A. Sharkawy, C. H. Chen, D. M. Pustai, and D. W. Prather, "Dispersion-based beam splitter in photonic crystals," *Optics Letters*, Vol. 29, 617–619, 2004.
5. Tsia, K. K. and A. W. Poon, "Dispersion-guided resonances in two dimensional photonic-crystal embededded microcavities," *Optics Express*, Vol. 12, 5711–5722, 2004.
6. Chang, S. and A. Taflove, "Finite-difference time-domain model of lasing action in a four-level two-electron atomic system," *Optics Express*, Vol. 12, 3827–3833, 2004.

Using Intrinsic Layer to Improve the Efficiency of Organic Solar Cells

Wei Ping Chu[2], **Yu Sheng Tsai**[2], **Fuh Shyang Juang**[2]
Tao Sheng Li[1], **and Chang Han Chung**[1]

[1]Institute of Electrical-Engineering, Kun-Shan University
Yong Kang, Tainan 710, Taiwan, China
[2]Institute of Electro-Optical and Materials Science, Formosa University
Huwei, Yunlin 632, Taiwan, China

Abstract— Organic small molecule and polymer materials are being considered as solar cell material in place of silicon wafer. Organic solar cells are now a popular subject in international research. This study uses ZnPc and C_{60} as respectively the donor material and acceptor material of Organic Solar Cells (OSC) and examines the effects of different blocking layers and structures (P-N junction and P-I-N structure) on open-circuit voltage (V_{OC}), short-circuit current (I_{SC}), fill factor (FF) and efficiency of solar cell. It is found under the same thickness of 15nm, the short-circuit current generated by the BCP and $BAlq_3$ blocking layers is respectively 1.36 and 1.19 mA/cm^2. This is because the BCP has higher energy gap and Highest Occupied Molecular Orbital (HOMO) than $BAlq_3$, which enables BCP to block excitons and holes better. When the P-I-N structure is ZnPc(30 nm)/C_{60}(40 nm) and ZnPc(30 nm)/ZnPc:C_{60}(20 nm)/C_{60}(40 nm), the short-circuit current produced could be increased to 2.67 mA/cm^2 respectively. This is because the intrinsic layer increases the contact area of P-N junction, enabling more excitons to dissociate into free charges.

1. INTRODUCTION

The finite availability of natural resources such as coal, oil and natural gas as we enter into the 21st century can no longer meet the development needs of mankind. Environ-mental pollution accompanying the development is also becoming a serious issue that needs to be dealt with urgently. One of the solutions is to look for alternative energy. To this end, improving the performance of solar cell, lowering its costs, and cutting down the impact of large-scale production on the environment are the major directions for future development. The OSC fabrication involves simple process where thin film can be formed by vacuum evaporation or spin coating and flexible solar cells can be fabricated on plastic substrate (PET) [1] to turn it into portable power. In comparison with silicon wafer as solar cell material, organic materials offer the advantages of light-weight, flexible, transparent, low cost, low-temperature and large area manufacture. These promising features turn OSC into one of the products with great market potential following organic light-emitting diodes (OLED).

There have been reports on using organic molecules as solar cell material in 1970s [2]. But such devices had very low efficiency ($< 0.1\%$). No breakthrough was reported until 1986 when Tang published the first OSC [3] with more than 1% efficiency. Since then, scientists reignited their interest in the potential of OSC. The low manufacturing cost, simple process, light weight, easy processing, flexibility, and large scale production have attracted many researchers into the field.

This study uses ZnPc and C_{60} respectively as the donor material and acceptor material of OSC to explore the effects of different blocking layer materials and thickness of ZnPc and C_{60} on the open-circuit voltage (V_{oc}), short- circuit current (I_{sc}), fill factor (FF) and efficiency of solar cells. The study also develops a P-I-N structure to enhance the efficiency of photoelectric conversion.

2. EXPERIMENT

Glass substrate deposited with ITO was defined with an anode pattern using photo-lithography. The glass substrate was ultrasonically vibrated in acetone, methanol and DI water in such sequence, and then blow dried with nitrogen gas before it was put into the oven to bake for 30 minutes. Subsequently, the glass substrate was placed in microwave plasma system to rinse the ITO surface with 90 sccm oxygen under 150 W microwave power. Next, the substrate was placed in an organic evaporator and coated with in sequence a donor layer (ZnPc), acceptor layer (C_{60}), and blocking layer (BCP or $BAlq_3$), and then placed in a metal evaporator to form a cathode (Al). The evaporation rate of ZnPc and C_{60} were approximately 0.1–0.2 nm/sec. The Al deposition rate was 2–3 nm/sec.

In the P-I-N structure, the I layer was co-evaporated ZnPc-C_{60}. After the OSC was fabricated, KEITHLEY 2400 was used to measure its current and voltage.

3. RESULTS AND DISCUSSION

3.1. P-N Structure under Different Blocking Layers

Figure 1 shows voltage vs. current density of solar cells using respectively BCP (15 nm) and $BAlq_3$ (15 nm) as blocking layer. As shown, $BAlq_3$ has smaller than BCP. This is because the energy level of $BAlq_3$ Highest Occupied Molecular Orbital (HOMO) at 5.85 eV is lower than the 7.0 eV of BCP

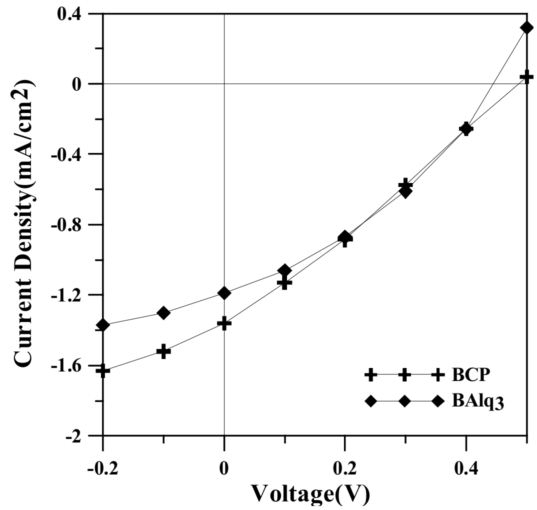


Figure 1: Optoelectrical properties of solar cell under different blocking layers.

(as shown in Fig. 2(a) and 2(b)), which weakens the ability of $BAlq_3$ to block hole diffusion as compared to BCP. Consequently, holes generated after illumination tend to diffuse to the vicinity of metal cathode and recombine with electrons, thereby reducing the number of carriers collected by the electrode and lowering the light current. In addition, the $BAlq_3$ HOMO energy level is also lower than that of C_{60} HOMO, causing the holes to migrate towards the cathode and recombine with electrons at the cathode junction, thereby decreasing the light current.

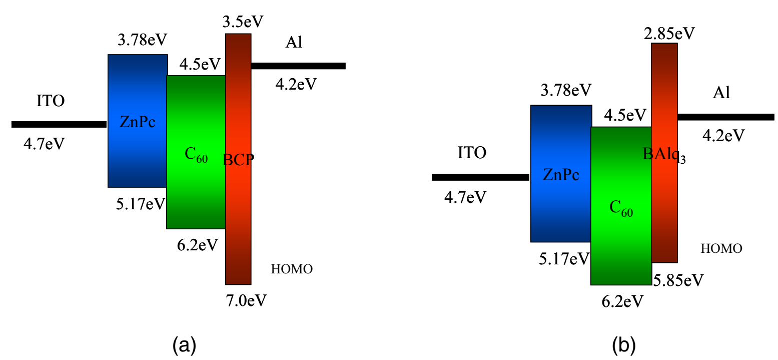


Figure 2: Band gap diagram of organic solar cells (a) using BCP as blocking layer, (b) using $BAlq_3$ as blocking layer.

3.2. P-I-N Structure with Varied Thickness

Given that excitons will only be dissociated into free charges by the built-in electric field at the P-N junction, an intrinsic layer (I) is added into the device structure. The intrinsic layer is co-evaporated ZnPc-C_{60} formed under the evaporation rate of 1:1. Its main purpose is to augment the volume of depletion region to increase the size of effective light absorption area, thereby allowing more excitons to be dissociated into free charges. The structure and band gap diagram of device with an intrinsic layer are as shown in Fig. 3 and Fig. 4. Fig. 5 shows voltage-current density of organic

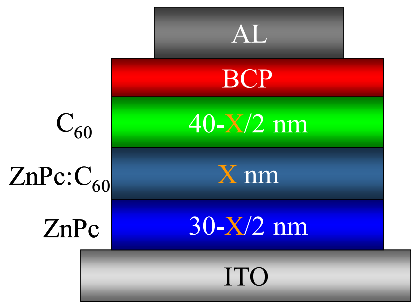


Figure 3: P-I-N structural diagram.

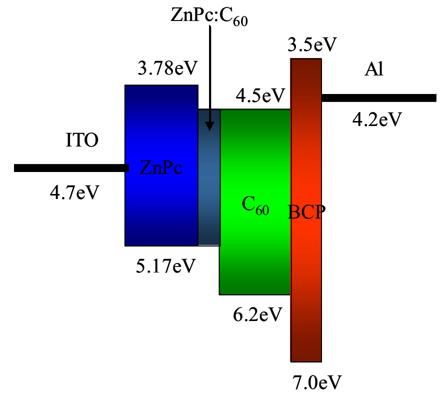


Figure 4: Band gap diagram of P-I-N device.

solar cells under varying I layer thickness. As com-pared with P-N structure without an intrinsic layer, the I_{sc} of P-I-N device increased from 1.80 to 2.67 mA/cm^2. This is because the effective light absorption area of P-I-N structure increases, enabling more excitons to be dissociated into electrons and holes, and resulting in increase in I_{sc}. Jones (2004) [4] points out that because an intrinsic layer is a polymolecular thin film formed by co-evaporation of two different materials, its conductivity is reduced, which in-creases the series resistance of device. Also as the series resistance of solar cell is inversely proportional to fill factor (FF), an increase in I layer thickness would cause increase in short-circuit current while FF becomes smaller. However when the thickness of I layer increases to 30 nm, short-circuit current becomes lower. This is because I layer is a blend of different materials, which makes it easier for carriers in I layer to be recombined and captured by the organic layer.

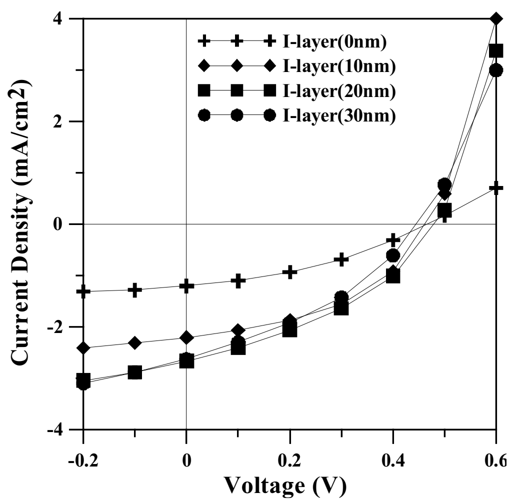


Figure 5: Optoelectrical properties of OSC with varied I layer thickness.

4. CONCLUSION

I_{sc} of 1.36 mA/cm^2 is produced when a device uses BCP as blocking layer, which is better than the Isc of 1.19 mA/cm^2 with $BAlq_3$ as blocking layer. When an intrinsic layer of ZnPc-C_{60} is added by co-evaporation under the rate of 1:1 to form a P-I-N structure, the I_{sc} is effectively increased. By shining incident light with air mass of AM1.5 100 mW/cm^2 on a solar cell with P-I-N structure (ITO/ZnPc (20 nm)/ZnPc:C_{60}(20 nm)/C_{60}(30 nm)/BCP (15 nm)/Al(100 nm)), the following measurements are ob-tained: $V_{OC} = 0.481$ V, $J_{SC} = 2.67$ mA/cm^2, FF $= 0.38$, and $\eta = 0.49\%$.

REFERENCES

1. Serdar, S. N., *Sol. State and Mater. Sci.*, Vol. 4, 373, 1999.

2. Merritt, V. Y. and H. J. Hovel, "Organic solar cells of hydroxy squarylium," *Appl. Phys. Lett.*, Vol. 29, 414, 1976.
3. Kittel, C., *Introduction to Solid State Physics*, Chap. 11, John Wiley and Sens., New York, 1996.
4. Peuman, P., V. Bulovic, and S. R. Forrest, "Efficient photon harvesting at high optical intensities in ultrathin organic double-heterostructure photovoltaic diodes," *Appl. Phys. Lett.*, Vol. 76, 2650, 2000.

Effects of Nitridation Time on Top-emission Inverted Organic Light Emitting Diode

Chien-Chang Tseng, **Liang-Wen Ji**, **Yu-Sheng Tsai**, **and Fuh-Shyang Juang**
Institute of Electrical-Engineering, Institute of Electro-Optical and Materials Science
Formosa University, Huwei, Yunlin 63208, Taiwan, China

Abstract— A top-emission inverted organic light-emitting diode (TEIOLED) was fabricated by using Al/AlNx layer as the cathode in the device structure of glass/Al/AlNx/AlQ3/NPB/MTDATA/Au/Ag, where AlNx ultra thin layer was obtained from Al layer under 90 W microwave plasma treatments in Ar and N_2 mixed gas environment. The N_2/Ar ratio and plasma treatment time were adjusted to obtain the maximum luminance and efficiency of 1206 cd/m^2 and 0.51 cd/A, respectively, both at 17 volts. The AlNx layer surface after plasma treatment was examined by atomic force microscope to study the effects of surface roughness on the electroluminescent (EL) characteristics. The AlNx layer thickness also affected the EL results apparently.

1. INTRODUCTION

One of the trends in the OLED (organic light emitting diode) display industry in recent years is to develop large-screen display. Large-size panel must be driven by thin film transistor (TFT) to enhance its brightness uniformity, resolution and life time. The conventional OLED have predominantly bottom-emission structure where emitted light is partially blocked by bottom-layer TFT and data line. To improve aperture ratio and minimize the effect of TFT on active-matrix panel, the research and development of top-emission OLED is inevitable [1, 2]. To integrate the process for OLED with other active components in the fabri-cation of active-matrix panel, including amorphous silicon (α-Si), poly-silicon (poly-Si) or complementary metal oxide semiconductor (CMOS), the whole process may be simplified if the electrode (Al pad) prepared on TFT substrate in the final step of the process is also used as the cathode of OLED. Such OLED device must employ top-emission inverted organic light emitting diodes (TEIOLEDs) structure. But when Al electrode is deposited on the substrate using thermal evaporation, spikes tend to form, which are prone to lead to organic carbonization and puncture the organic thin film after prolonged operation of OLED, causing dark spots on the emitting surface of OLED [3]. Many researchers have pro-posed inverted top-emission light-emitting diode by inserting an ultra thin insulating layer between the organic layer and metal electrode to increase electron injection from cathode, thereby enhancing the luminance efficiency. For example, Li et al., [4], Kurosaka et al., [5] and Kho et al., [6] respectively inserted a proper buffer layer of Al_2O_3/AlNx between the organic light-emitting layer and Al cathode to improve the electroluminescence of OLED. This paper utilized the plasma of Ar+N_2 gas mixture to treat Al surface to modify the surface roughness and form an ultra-thin AlNx insulating layer (or buffer layer) to improve the efficiency of cathode electron injection.

2. EXPERIMENT

Our experiment comprised the following steps: First ultra-sonically wash the glass substrate with in sequence acetone, methyl alcohol and DI water. Deposit cathode metal Al on the glass substrate by evaporation under 3×10^{-6} torr and then move the substrate to a microwave plasma chamber to un-dergo plasma treatment of Al surface with the mixture of Ar (70 sccm) and N_2 (10 sccm) under 0.35 torr and 90 W to form AlNx as buffer layer. Remove the AlNx-treated substrate from plasma chamber and place it in organic vacuum evapo-rator to deposit tris (8-quinolinolato) aluminum (AlQ_3) as green-light emitting layer, NPB as hole transport layer, and m-MTDATA as hole inject layer (HIL). Finally move the substrate to another chamber to deposit Au/Ag as metal anode by evaporation. The resulting device has a structure of glass/A1 (80 nm)/AlNx/AlQ_3 (100 nm)/NPB (60 nm) /m-MTDATA (50 nm)/Au (5 nm)/Ag (10 nm) as shown in Fig. 1. The light emitting area of the device was 36 mm^2 as defined by shadow mask. The experiment used KEITHLEY 2400 and SpectraScan PR650 to measure the L-J-V curve of device, atomic force microscope (AFM) and contact angle to observe its surface roughness, and four-point probe to measure the sheet resistance.

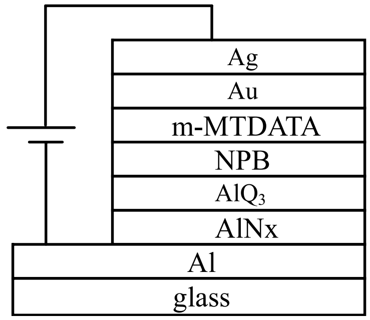


Figure 1: TEIOLED structure.

3. RESULTS AND DISCUSSION

Atomic Force Microscope (AFM) was used to observe the surface roughness of Al cathode layer after microwave plasma treatment over different durations. As shown in Fig. 2, it is found that untreated Al electrode surface has many spikes that would cause non-uniform distribution of electric field in the organic layer during the operation of OLED and render the OLED prone to damage under higher current density. Thus prior to the deposition of organic film, Al electrode should undergo surface planarization with microwave plasma, which improves surface roughness significantly. Based on the measurement of contact angle, it is also clear that there is better adhesion between the plasma-treated Al electrode and the organic layer.

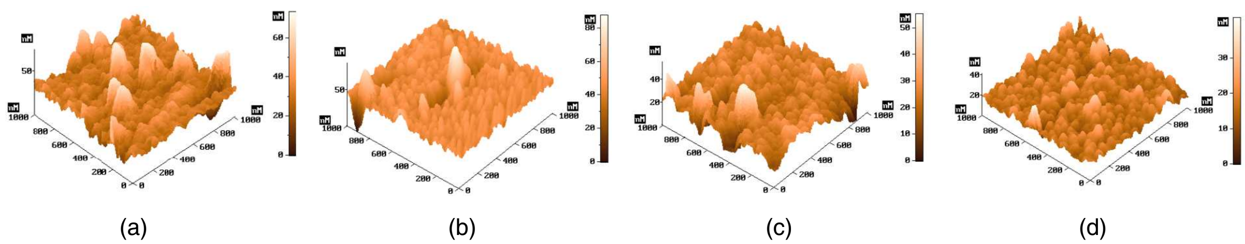


Figure 2: Plasma-treated surface of Al cathode under AFM, (a) untreated Ra = 5.249 nm, (b) 40 sec Ra = 4.112 nm, (c) 60 sec Ra = 3.747 nm, (d) 80 sec Ra = 2.551 nm.

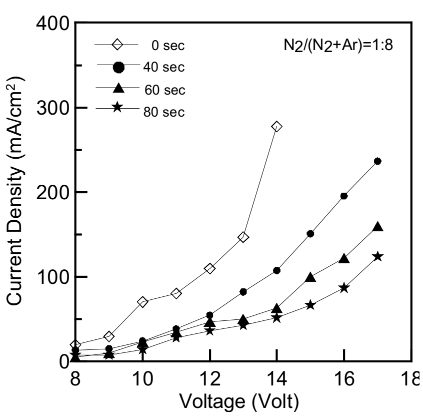


Figure 3: Current density of TEIOLED vs. voltage under different plasma treat-ment times.

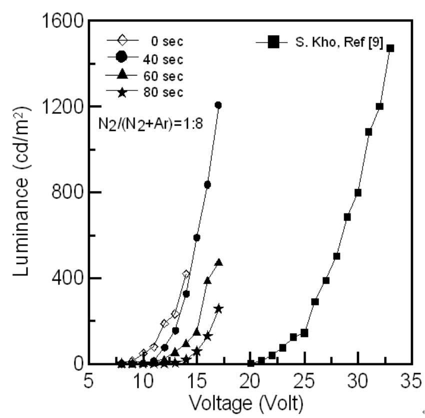


Figure 4: Luminance of TEIOLED vs. voltage under different plasma treatment times.

Figure 3, Fig. 4 anf Fig. 5 show respectively the J-V curve (current density vs. voltage), L-V curve (luminance vs. volt-age), and η-V curve (yield vs. voltage) of the same device under plasma treatment of different durations. As shown by the J-V curve, the resistance of plasma-treated Al surface increased as treatment time increased, which led to decrease in current density as shown in Table 1. Based on the L-V curve in Fig. 4, it is found increase in the time of nitrogen treatment led to increase in threshold voltage. Untreated Al cathode caused organic layer to produce joule heat under higher current and same voltage and burn more easily. As shown in Fig. 4 and Fig. 5, under optimum plasma treatment time of 40 sec, the device obtained maximum luminance and yield of 1206 cd/m^2 and 0.51 cd/A respectively under 17 V. Device made of untreated Al electrode had extremely short life-time, while treated device with an AlNx buffer layer had significantly longer life time. Our TEIOLED has lower turn-on voltage when compared with Ref. [6]

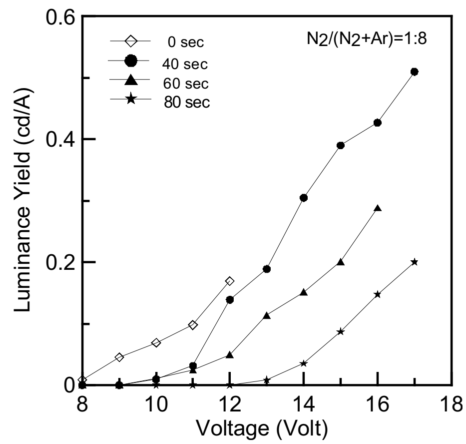


Figure 5: Luminance yield of TEIOLED vs. voltage under different plasma treatment times.

Table 1: Roughness, contact angle, resistivity of Al cathode, max. luminance and yield of TEIOLED under different plasma treatment time.

Treatment time	Roughness (Ra,nm)	Contact Angle	Resistivity (Ω-cm)	Max. Luminance (cd/m^2)	Yield (cd/A)
Untreated	5.249	22.30°	6.55×10^{-8}	418@14 V	0.17@14 V
40 s	4.112	9.88°	7.40×10^{-8}	1206@17 V	0.51@17 V
60 s	3.747	8.63°	7.51×10^{-8}	474@17 V	0.29@16 V
80 s	2.551	7.32°	7.54×10^{-8}	256@17 V	0.20@17 V

When plasma treatment time increased to more than 40 sec, the luminance and yield of device decreased gradually instead. It might be attributed to the fact that prolonged treatment time resulted in thicker AlNx and hence high resistivity, which makes injection of electron into AlQ_3

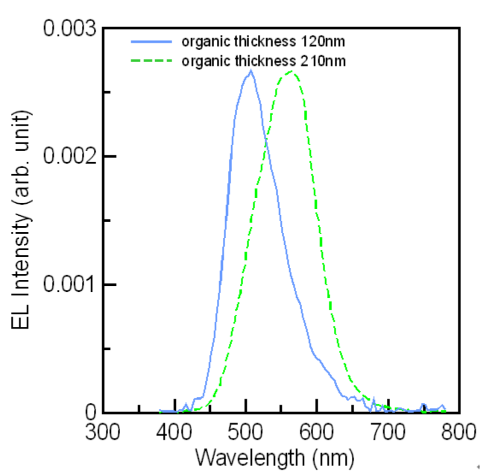


Figure 6: Shift of emission peak from 508 nm to 558 nm when thickness of organic layer increases.

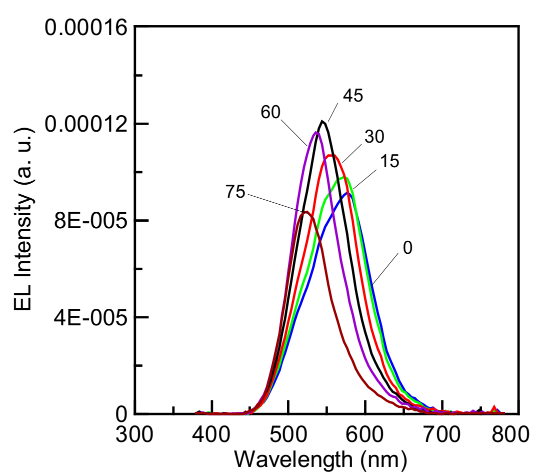


Figure 7: EL intensity of TEIOLED under different angles EL spectrum.

more difficult. The functions of AlNx may be summed up as follows: 1. Proper AlNx thickness aids the tunneling of electrons [8], for in the absence of AlNx as buffer layer, there exits a larger potential barrier at AlQ_3/Al interface to block the injection of cathode electrons into AlQ_3; 2. the mobility of surface electron in organic material is lower than that of hole, and AlNx can block the circulation of holes to prevent quenching of exciton at AlQ_3/Al interface.

In the TEIOLED structure discussed in this paper, Al metal with high reflectivity was used for cathode, and Au (5 nm)/ Ag (10 nm) was used for anode where Au was for improving hole injection from anode into organic layer and Ag with higher conductivity and transmission was for reducing sheet resistance and enhancing light output. The penetration rate of such double-layer anode was approximately 47%. In addition, change in the thickness of organic or metal electrode would cause displacement of emission peak in microcavity. In the example of changing the total thickness of organic layer as shown in Fig. 6, when the length of microcavity increased from 120 nm to 210 nm, emission peak shifted from 508 nm to 552 nm. The change in resonance wavelength is not just a result of change in microcavity length, but also due to difference in the reflection phase of top and bottom surface electrodes. Fig. 7 shows the electroluminescence (EL) spectrum obtained at different angles, which are the angles intersecting the measuring direction of PR650 relative to the normal line of light-emitting surface of device. The measurement took place at the interval of 15° between $0° \sim 75°$. It is clear that as the measuring angle increased, the peak of EL spectrum would drop from 576 nm to 524 nm to produce blue displacement phenomenon, mainly due to the intense microcavity effect [9].

4. CONCLUSION

We have demonstrated that using the microwave plasma of $Ar+N_2$ mixture to treat the surface of Al electrode over the optimum treatment time of 40 sec can form an ultra thin AlNx layer as the buffer layer for injection of electrons from cathode in TEIOLED. This surface treatment effectively improves the efficiency of electron injection, enable the number of electrons and holes injected to reach a balance and prevent the phenomenon of exciton quenching at the cathode, thereby significantly enhancing the luminance, efficiency and life-time of the device. The N_2/Ar ratio and plasma treatment time were adjusted to obtain the maximum luminance and efficiency of 1206 cd/m^2 and 0.51 cd/A, respectively, both at 17 volts.

REFERENCES

1. Kitamura, M., T. Imada, and Y. Arakawa, "Organic light-emitting diodes driven by pentacene-based thin film transistors," *Appl. Phys. Lett.*, Vol. 83, 3410–3412, 2003.
2. Pribat, D. and F. Plais, "Matrix addressing for organic electroluminescent displays," *Thin Solid Film*, Vol. 383, 25–30, 2001.
3. Elschner, A., F. Bruder, H.-W. Heuer, F. Jonas, A. Karbach, S. Kirchmeyer, S. Thurm, and R. Wehrmann, "PEDT/PSS for efficient hole-injection in hybrid organic light-emitting diodes," *Synth. Metal*, Vol. 111–112, 139–143, 2000.
4. Li, F., H. Tang, J. Anderegg, and J. Shinar, "Fabrication and electroluminescence of double-layered organic light-emitting diodes with the Al_2O_3/Al cathode," *Appl Phys Lett.*, Vol. 70, 1233–1235, 2000.
5. Kurosaka, Y., N. Tada, Y. Ohmori, and K. Yoshino, "Improvement of electrode/organic layer interfaces by the insertion of monolayer-like Alu-minum oxide film," *Jpn. J. Appl. Phys.*, Vol. 37, L872–L875, 1998.
6. Kho, S., S. Sohn, and D. Jung, "Effects of N_2 plasma treatment of the Al bottom cathode on the characteristics of top-emission-inverted organic-light-emitting diodes," *Jpn. J. Appl. Phys.*, Vol. 42, L552–L555, 2003.
7. Deng, Z. B., X. M. Ding, S. T. Lee, and W. A. Gambling, "Enhanced brightness and efficiency in organic electroluminescent devices using SiO_2 buffer layers," *Appl. Phys. Lett.*, Vol. 74, 2227–2229, 1999.
8. Dodabalapur, A., L. J. Rothberg, R. H. Jordan, T. M. Miller, R. E. Slusher, and J. M. Philips, "Physics and applications of organic microcavity light emitting diodes," *J. Appl. Phys.*, Vol. 80, 6954–6964, 1996.

Effects of Isolation-layer on Luminance Efficiency of Organic Light-emitting Diodes

Chang-Han Chung[1], **To-Sing Li**[1], **Fuh-Shyang Juang**[2], **and Day-Shan Liu**[2]

[1]Institute of Electrical-Engineering, Kun-Shan University
Yong Kang, Tainan 710, Taiwan, China
[2]Institute of Electro-Optical and Materials Science, Formosa University
Huwei, Yunlin 63208, Taiwan, China

Abstract— This study deposited SiO_2 of varying thickness by PECVD or coated photo-resist of varying thickness at the edge of patterned OLED anode and between the patterned anode and the organic layer as isolation-layer to improve the current leak of OLED. Current leak would cause the loss of charges around the emitting zone, which adversely affects the recombination of electrons and holes in the emitting zone and results in poorer luminance efficiency. The study find that as SiO_2 thickness increased, leakage current was significantly reduced. When SiO_2 thickness increased from 0 nm to 250 nm, the luminance efficiency of device rose from 1.7 cd/A to 3.4 cd/A at 150 mA/cm^2. Lastly the study examined the effect of photoresist thickness on reducing the leaking current and finds that at 2250 nm, the luminance efficiency was enhanced further to 4.5 cd/A at 150 mA/cm^2.

1. INTRODUCTION

There remain a myriad of issues to address in the applications of organic light-emitting diode (OLED), for example, how to enhance its luminance efficiency and life-time. The cathode of OLED is typically made of Al layer collocated with CsF [2], NaCl [3], or LiF [4] to boost the injection number of electrons into the organic layer. Or a buffer layer is added to the anode [5], which also improves the luminance efficiency of device effectively. This study inserts a layer of SiO_2 or photo-resistor at the edge of patterned anode and between the cathode (ITO) and hole transport layer (NPB) to explore its effect on the leakage current of OLED and on the electron-hole recombination ratio in the emission area, thereby increasing the luminance efficiency.

2. EXPERIMENT

The pattern of anode (ITO) was first defined on a glass substrate by photo-lithography. The glass substrate was then immersed in acetone, methanol and DI water, respectively, by ultrasonic vibration cleaning. After blow drying with nitrogen gun, the glass substrate was baked in the oven for 30 minutes. Next a layer of SiO_2 was deposited by PECVD or a layer of photo-resistor (P.R.) was coated by photo-lithography on the annular area along the edge of ITO pattern as isolation layer. The device cross section is shown in Fig. 1. The emission area of both devices with and without the isolation area was fixed at 0.36 cm^2. The ITO substrate with isolation annular was then placed in O_2 plasma (90 sccm O_2 and microwave power of 150 W) for surface cleaning. Subsequently, the substrate was placed in the organic evaporator to deposit in sequence NPB (hole transport layer) and Alq_3 (emitting and electron transport layer) at the rate of 0.1–0.2 nm/sec. The substrate was then moved to metal evaporator to deposit LiF and Al at the rate of 0.2 nm/sec and 2–3 nm/sec respectively as cathode. The cross section view of the device is shown in Fig. 1. The luminance-current-voltage (L-I-V) characteristics of the device were measured after its fabrication. SpectraScan PR650 was employed to measure the luminance and spectrum, and Keithley 2400 was employed for power supply and testing of current density-voltage characteristics.

3. RESULTS AND DISCUSSION

Figure 2 shows current density (J) vs. voltage (V) of OLED using SiO_2 and photo-resistor (P.R) as isolation layer. It is found under 7 V, the total current of device without any isolation layer was significantly higher. As charges of the device without an isolation layer leak from the edge of emission area, electrons cannot effectively combine with holes in the emission area. As shown in Fig. 3 and Fig. 4 which depict respectively current density vs. luminance and current density vs. yield, under the same current density, the thicker the SiO_2 thickness, the brighter the luminance and the higher the luminance yield. This is because thicker SiO_2 provides better leakage current

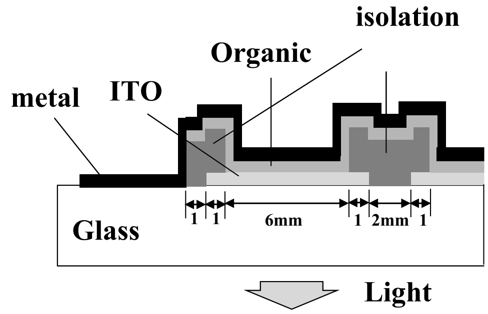


Figure 1: Cross section view of OLED with isolation layer.

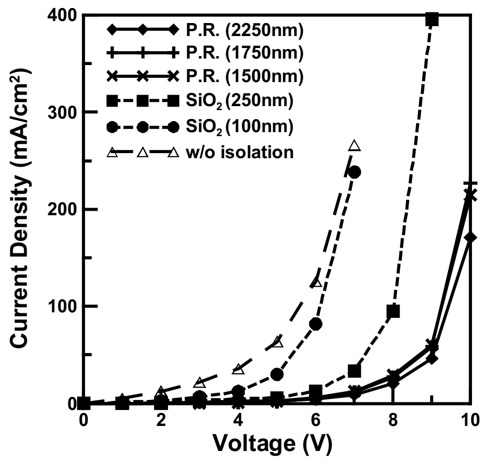


Figure 2: Current density vs. voltage of OLED with SiO_2 or P.R. as isolation layer.

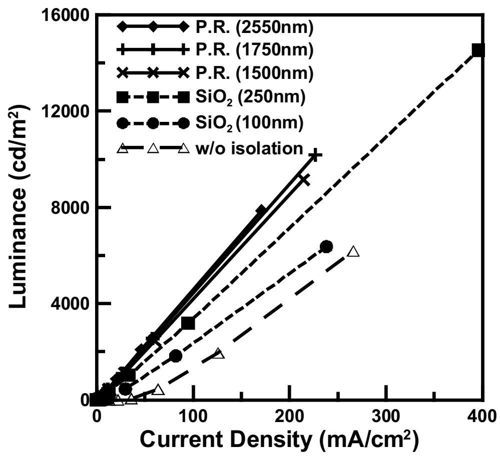


Figure 3: Luminance vs. current density of OLED with SiO_2 and P.R. as isolation layer.

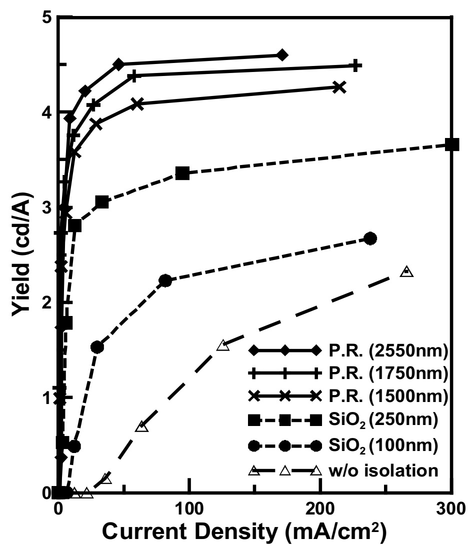


Figure 4: Luminance yield vs. current density of OLED with SiO_2 or P.R. as isolation layer.

prevention to allow effective electron-hole recombination in the emission area, thereby enhancing the luminance and yield. When the device has high leakage current, the non-radiative recombination effect would be more pronounced and the heat generated thereof might burn the device. As shown in Fig. 3 and Fig. 4, when SiO_2 from 2600 to 5230 cd/m^2, while the luminance yield increased from 1.7 to 3.4 cd/A at 150 mA/cm^2.

Aside from using SiO_2, this study also employed photo-resistor as isolation material and compared its performance with SiO_2. Fig. 2 depicts current density vs. voltage of OLED with varying photo-resistor thickness. As shown in Fig. 2, the current density was generally lower for using P.R. as isolation layer, but luminance and yield increased very well. As shown in Fig. 4, when photo-resistor thickness increased from 1500 nm to 2250 nm, the luminance yield rose from 4.1 cd/A to 4.6 cd/A at 175 mA/cm^2 current density, which was better than 3.3 cd/A with SiO_2 as isolation layer. OLED is a current-driven de-vice. Preventing leakage current to allow effective recombination of electrons and holes in the emission area is vitally important to the enhancement of luminance and luminance efficiency. As shown in Fig. 3 and Fig. 4, the thicker the photo-resistor, the better the isolation effect, and the luminance and yield of the device could be increased to 7200 cd/m^2 and 4.6 cd/A respectively under current density of 150 mA/cm^2. Furthermore, the stability of device efficiency could be maintained. In addition, the luminance efficiency of device using photo-resistor as isolation layer was better than that with SiO_2, mainly because SiO_2 was much thinner than photo-resistor.

4. CONCLUSIONS

This study inserted an isolation layer at the edge of ITO cathode and between ITO and organic layer to effectively reduce leakage current, so as to prevent the generation of excess heat, enhance luminance efficiency and maintain device stability. Regardless of the material employed for isolation layer, thicker material brings about more reduction of leakage current, thereby enhancing the luminance and yield of device. Because photo-resistor coating is much thicker than SiO_2 deposited by PECVD, it produces better leakage-current blocking effect. This study used 2250 nm photo-resistor as isolation layer against current leak and obtained maximum luminance of 8000 cd/m^2 and maximum luminance yield of 4.6 cd/A at 150 mA/cm^2, which was a significant improvement over 1.7 cd/A in the case without any isolation layer.

ACKNOWLEDGMENT

We would like to thank Prof. C. T. Lee and S. J. Chang for helpful discussions. This project was supported by the National Science Council for the funding of this project No. NSC-94-2218-E-150-003

REFERENCES

1. Tang, C. W. and S. A. VanSlyke, *Appl. Phys. Lett.*, Vol. 51, 913, 1987.
2. Piromreun, P., H. S. Oh, Y. Shen, G. G. Malliaras, J. C. Scott, and P. J. Brock, *Appl. Phys. Lett.*, Vol. 77, 2403, 2000.
3. Kang, S. J., D. S. Park, S. Y. Kim, C. N. Whang, K. Jeong, and S. Im, *Appl. Phys. Lett.*, Vol. 81, 2581, 2002.
4. Hung, L. S., C. W. Tang, and M. G. Mason, *Appl. Phys. Lett.*, Vol. 70, 152, 1996.
5. Yang, T. H., F. S. Juang, Y. S. Tsai, W. K. Kuo, and M. C. Yokoyama, *Jpn. J. Appl. Phys.*, Vol. 45, No. 4B, 2006.
6. Zhao, J. M., S. T. Zhang, X. J. Wang, Y. Q. Zhan, X. Z. Wang, G. Y. Zhong, Z. J. Wang, X. M. Ding, W. Huang, and X. Y. Hou, *Appl. Phys. Lett.*, Vol. 84, 2913, 2004.
7. Chen, S. F. and C. W. Wang, *Appl. Phys. Lett.*, Vol. 85, 765, 2004.

Beam Propagation in Laser Scattering Communication

Ying Yan and Jianguo Xin
Department of Physical Electronics, School of Information Science and Technology
Beijing Institute of Technology
Beijing 100081, China

Abstract— In this paper, a new technique of laser scattering communication is put forward; its advantages are anti-jamming ability and relatively longer corresponding distance. It is believed that the laser scattering communication technique will play an important role in the situation when the current laser communication technique is limited. The basic theory used in laser scattering communication will be discussed in this paper.

In the traditional laser free space communication, information is transferred by laser propagation in straight line. Usually, due to the requirement of unblocked laser beam propagation path, the traditional free space laser communication is limited in the distance of stadium less than 10 km. Particularly in the case of the laser beam propagation being blocked off by such as trees, mountains or buildings, the communication will not be accomplished. Therefore looking for a technique to overcome the beam propagation path blocking becomes an alternative to the laser beam free space communication. In this paper, a technique of laser scattering communication is proposed, which utilizes the laser scattering in the troposphere to transmit the laser signal from emitter to the receiver to overpass the blockings. The microwave scattering in the troposphere [1] has been well developed to be applied to the communication.

Compared to microwave scattering, because the laser wavelength is much shorter than that of the microwave, the laser scattering in troposphere will be much stronger than that of the microwave, which leads to the results of the short propagation distance and stronger scattering intensity for the laser beam. Therefore, the laser scattering communication technique may be utilized for the shorter distance communication compared with the ultra-long distance of the microwave scattering communication.

In the following analysis, the variation of air density at the different height has been taken into account. The intensity, the received laser signal in time domain and frequency domain superposition has also been discussed.

During laser transmission, the absorption and scattering are the main losses in atmosphere, which depend on air density variation, laser wavelength, transmission distance and et al. According to Beer law, the intensity of laser transmission at the point R is:

$$I(R) = I_o \cdot \exp\left[-\int_0^R \sigma(r,\lambda)dr\right] \tag{1}$$

where I_o is the initial intensity, and $\sigma(r,\lambda)$ is the losses modulus, which contain the absorption and scattering of gas molecules and aerosol. Its numerical value has been given for five kinds of wavelength in ref. [2].

Because of the existence of molecule and aerosol, laser will be scattered during transmission. According to the size of the particles, the scattering can be divided into two kinds, Rayleigh scattering and Mie scattering. In this paper, the intensity distribution of light scattered by one particle can be calculated by the existing equations [2, 4]. When laser is scattered by a number of particles, the circumstances are more complex. If the distance between particles is equal to several times of the radius of particles, it can be considered as the scattering of light through a single particle. That is to say the total intensity of the scattered light is the sum of scattered light through independent particles.

In laser scattering communication, the information is transferred though scattering in the troposphere. According to this principle, a geometrical model can be established as shown as below.

The laser beam starts at O with an angle of α and intercepted at A. From the figure it can be seen that the length of each path from O to A is different, which result in different time delay. It will strongly affect the frequency of the signal.

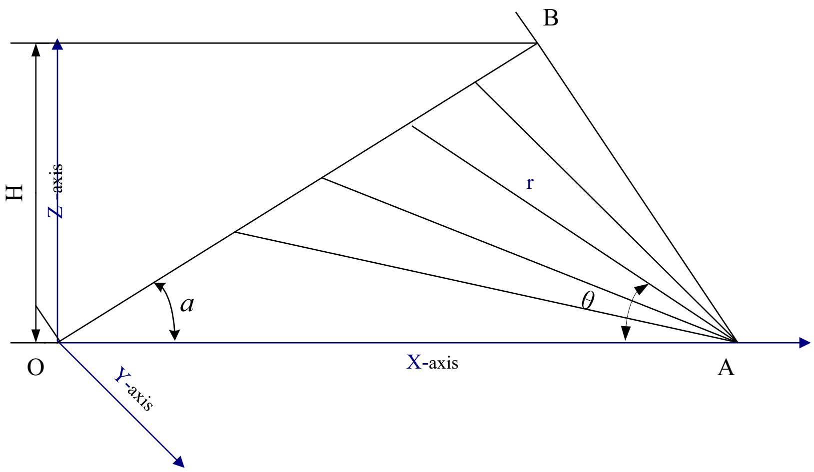


Figure 1: The geometrical model of the communication by laser scattering.

Frequency shift is another factor which should be considered. According to Doppler phenomenon, the free movement of gas molecule and aerosol can bring shift in frequency of scattered signal. In the whole process, the shift in frequency should be considered due to the scattering through the moving particles and then duo to the scattered light received at A. The equation is:

$$f_D = f_s - f_o = \frac{1}{\lambda}\left|\vec{U}\cdot(\vec{e}_s - \vec{e}_o)\right| \tag{2}$$

where f_D is shift in frequency between incident and scattered laser, f_o and f_s are the incident and scattered laser frequencies respectively. λ is wavelength of incidence laser, $\vec{U}$ is velocity vector of particles moving, $\vec{e}_s$, $\vec{e}_o$ are the unit vectors in the direction of incidence and scattering. The frequency shift may bring signal decrease in some time.

In conclusion, a new technique of the laser scattering communication has been proposed in this paper which has the advantages over the traditional laser communication. Compared with the traditional free space laser communication, it has merits in many aspects. For instance, it has longer corresponding distance and no needing for the clear transmission path as the traditional free space laser communication requires.

In the further study, an intention has been paid to the signal modulation and demodulation, laser emitter and receiver, and the influence of weather conditions to the signal modulation.

REFERENCES

1. Li, D., *Scattering Communication*, People Post and Telecommunications Publishing Company, Beijing, 1982.
2. Sun, J., *Laser Atmosphere Detection*, Science Publishing Company, Beijing, 1986.
3. Sheng, F., *Atmosphere Physics*, Beijing University Publishing Company, Beijing, 2003.
4. Song, F., *Basic of Application Atmosphere Optics*, Weather Publishing Company, Beijing, 1990.

Entanglement Theory and a q Analogue Entangled State

Huihui Zhang, A. Rauf, and Xiaoguang Zhou

Department of Engineering Optics, School of Information Science and Technology
Beijing Institute of Technology, Beijing 100081, China

Abstract— Entangled state and entanglement theory in the quantum information are presented. A q analogue entangled state is first put forward in this paper. It exhibits a much richer structure than ordinary bipartite entanglement. Thus there will be abundant information in the quantum teleportation.

1. INTRODUCTION

Quantum entanglement is not only one of the striking features in quantum mechanics but also a vital resource in some quantum information and quantum computation processes [1–4], such as teleportation, dense coding, quantum key distribution, and quantum computation. It was first noted by Einstein, Podolsky and Rosen (EPR) and Schrodinger. Entanglement describes a system composed of two or more particles, and exhibits the novel property that the result of measurement on one particle cannot be specified independently of the parameters of the measurements on other particles. Entanglement can exhibit the nature of a non-local correlation between quantum systems that have no classical interpretation. It has been noted that quantum teleportation can be viewed as an achievable experimental technique to quantitatively investigate quantum entanglement.

2. ENTANGLED STATE AND THE BELL BASIS

The entangled state can be expressed by the form:

$$|\psi\rangle_{AB} \neq |\psi\rangle_A \otimes |\psi\rangle_B. \tag{1}$$

Bell basis is the maximum entangled state in the bipartite entanglement:

$$\begin{cases} |\phi^{\pm}\rangle = \dfrac{1}{\sqrt{2}}\left(|00\rangle \pm |11\rangle\right), \\ |\psi^{\pm}\rangle = \dfrac{1}{\sqrt{2}}\left(|01\rangle \pm |10\rangle\right). \end{cases} \tag{2}$$

The Bell basis measurement can become true by the simple quantum network as shown below:

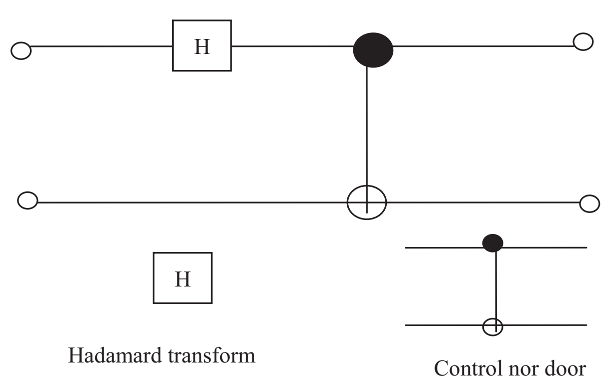


Figure 1: Quantum network of identify Bell basis.

3. ENTANGLEMENT THEORY IN QUANTUM INFORMATION

Quantum information theory studies the transmission and processing of information when information itself is carried by quantum systems and is processed according to the laws of quantum mechanics [5–9]. The recent achievements in this field include the discovery of new ways of information transmission, of secure communications and the performance of some kinds of computation faster than with classical means. A key ingredient and fundamental resource in the development of all these tasks is quantum entanglement.

Quantum teleportation, proposed by Bennett et al., is the process that transmits an unknown qubit state from a sender (Alice) to a receiver (Bob) via a quantum channel with the help of sending some classical information. In original scheme, such a quantum channel has been represented by a Bell maximally entangled state or Einstein-Podolsky-Rosen (EPR) pair. Quantum teleportation has been considered for N-dimensional, and continuous variable states. Teleportation of the polarization photon state and coherent state of light field has been demonstrated in optical experiments. Other strategy in quantum information is related to use of a single-rail logic. So, teleportation of a qubit occupying only one mode via one-photon quantum channel was studied. Quantum teleportation of N-particle entangled state via N+1-particle quantum channel was considered in Experimental. Realization of teleportation through one-photon quantum channel has reported. And problems of quantum teleportation by employing Greenberg-Horne-Zeilinger (GHZ) entanglement have been studied. More general questions of quantum teleportation of two qubits involving noisy quantum channels are involved in. The standard method for encoding qubits in optics is to use the polarization degrees of freedom of single photon. However, it was recently recognised in that spatial encoding is easier, for example, to manipulate and construct universal quantum gates. First, address the issue of qubit transfer by the two-photon mode entangled quantum channel that takes simultaneously four modes. The source of the mode entanglement may consist of two noncollinear degenerated on frequency spontaneous parametric down converters with phase matching. As pointed out by Bennett et al., in their original proposal for quantum teleportation, entanglement can be transferred through teleportation of two modes one of the particles forming the entangled state. This method, known as entanglement swapping, provides only partial teleportation of entanglement, an alternative method in which the entire mode entangled state is directly transferred from one place to another has been proposed. Use four-photon quantum channel constructed as a tensor product of two-photon mode entangled states to teleport the unknown mode entangled state. Quantum teleportation utilizing such four-photon quantum channel is not required special detectors distinguishing between one and two photon number states. The teleportation scheme of the entangled state through four-photon quantum channel may be more easily performed in practice than the teleportation schemes.

At present entanglement between two systems, i. e., bipartite entanglement, is quite well understood, but that between more than two systems, i. e., multipartite entanglement, remains still far from being satisfactorily known. In spite of that, multipartite entanglement has proven to play a superior role in recently emerging fields of quantum information processing and quantum computing since it exhibits a much richer structure than bipartite entanglement. Motivation for studying multipartite entanglement arises from many reasons some of which are listed now. First, multipartite entanglement provides a unique means to check the Einstein locality without invoking statistical arguments, contrary to the case of Bell inequalities using bipartite entanglement. Second, multipartite entanglement serves as a key ingredient for quantum computing to achieve an exponential speedup over classical computation. Third, multipartite entanglement is central to quantum error correction where it is used to encode states, to detect errors and, eventually, to allow fault-tolerant quantum computation. Fourth, multipartite entanglement helps to better characterize the critical behavior of different many-body quantum systems giving rise to a unified treatment of the quantum phase transitions. Fifth, multipartite entanglement is crucial also in condensed matter phenomena and might solve some unresolved problems Sixth, multipartite entanglement is recognized as a unreplaceable or efficient resource to perform tasks involving a large number of parties such as network teleportation, quantum cryptography, quantum secret sharing, remote entangling, quantum (tele)cloning, quantum Byzantine agreement, etc. Finally, multipartite entanglement is conjectured to yield a wealth of fascinating and unexplored physics. Current research in multipartite entanglement is progressing along two directions in parallel. One direction deals with problems such as how to classify, quantify, generate, control, distill and witness multipartite entanglement. The other direction proceeds to advance various applications exploiting the nonclassical multiway correlation inherent in multipartite entanglement. Many of the quantum information processing protocols typically involve sending.

Several quantum teleportation schemes of both discrete and continuous variables have been proposed. Continuous variable quantum teleportation of an unknown coherent states has been realized experimentally by employing a two-mode squeezed vacuum state as an entanglement resource. Recently, the teleportation schemes via the entangled coherent states have been discussed.

The entanglement of entangled coherent states in vacuum environment by employing the entanglement of formation has been studied and find that the entanglement of formation of the entangled coherent states is sensitive with the relative phase when the amplitude is very small. A scheme of probabilistic teleportation via entangled coherent states, in which the amount of classical information sent by Alice is restricted to one bit. In this probabilistic teleportation scheme, a coherent superposition state can be probabilistically perfectly teleported via a properly chosen entangled coherent state. When the interaction with the vacuum environment is addressed, the mean fidelity of the scheme is studied and the optimal amplitude of the teleported state is found. Simultaneous distance-independent correlation between different systems called entanglement is the most characteristic trait that sharply distinguishes between quantum and classical worlds.

The study of quantum teleportation protocol is not only limited to qubits and qudits (systems in d-dimensional Hilbert spaces) but also to quantum systems in infinite dimensional Hilbert spaces. In real situations sender and receiver may not have shared maximally entangled state but some form of non-maximally entangled pure state (due to some imperfection at the source). Usually if one follows the standard protocol, one will not be able to complete the teleportation process with unit fidelity and unit probability. Rather, the fidelity will depend on the parameters of the unknown state and the teleportation will not be reliable. Of course, if one has several non-maximally entangled pairs, one can first perform entanglement concentration and then recover fewer perfect maximally entangled pairs, and then use one of them to teleport an unknown state using the standard protocol. If Alice and Bob have only one pair, they can perform local filtering first, and convert a non-maximally entangled pair to maximally entangled pair with certain probability. Then they can follow standard protocol. Teleporting an unknown state using any pure entangled state but using generalized measurements has been proposed. This has been termed as conclusive teleportation. Also, there has been a qubit assisted conclusive teleportation process. It has also been mentioned that teleportation with unit fidelity but less than unit probability is possible for a qubit encoded in a coherent state [10–13].

How to characterize and to measure the entanglement is a basic problem. Although many impressive progresses has been obtained during the past decade, there is no general qualitative and quantitative theory of entanglement. Among the known criterions for characterizing entanglement, entanglement witness (EW) is an important criterion for characterizing the presence of entanglement [14]. The EWs are Hermitian but not positive operators whose expectation value is positive in every separable state. The importance of EW's stems from the fact that a given state is inseparable if and only if there exists an EW that detects it. Unlike the Peres-Horodecki criterion, which is a necessary and sufficient condition for determining entangled states on the low-dimensional quantum systems, for higher dimensions, this criterion is necessary one, the EW criterion is still sufficient and necessary one regardless of the dimension of quantum systems. In the other words, for any composite quantum systems the existence of the EW denotes the presence of entanglement. On the other hand, the EW, as operator, can be decomposed as the combination of the product projector operators, by which one can detect the entanglement experimentally. And the measurement can be implemented locally. Some authors introduced a kind of entanglement measure based on the optimal EW of states. For a given quantum state, how to construct its EW (if it has any) or how to prove it has no any EW is an interesting and hard problem. If one can find all EWs then the problem of characterizing entangled states is solved completely. Unfortunately, the determination of EWs for all states is also being computationally intractable. For some special states, however, their EWs can be constructed easily.

4. A Q ANOLOGUE ENTANGLED STATE

The q deformation of Lie algebra (the so-called quantum group) has its origin in the quantum inverse problem and in physics it is related to the solutions of integrable systems, to particular problems of statistical physics and to a conformal field theory [15, 16]. Attention has been paid to the realization of the quantum group in terms of the q analogue of the quantum harmonic oscillator (q-QHO) The quantum harmonic oscillator (QHO) is one of the most fundamental objects in quantum physics. In quantum optics the QHO describes, in particular, the single mode of the quantized cavity field.

On the other hand, up to now, there has not been an appropriate q anologue to describe quantitatively the entanglement of two and more subsystems due to the high complexity of entanglement in multi-particle system. The purpose of this paper is to propose a new theory of entanglement using q analogue of the quantum harmonic oscillator and give the basic form of entangled state.

Considering a two dimension q analogue of the quantum harmonic oscillator, x orientation is an ordinary harmonic oscillator (system A), the ground state is $|0\rangle$, the first excited state is $|1\rangle$, y orientation is a q analogue of the quantum harmonic oscillator(system B), the ground state is $|0\rangle_q$, the first excited state is $|1\rangle_q$. A form of entangled state can be proposed:

$$|\varphi\rangle = \cos\beta|0\rangle|1\rangle_q + \sin\beta|1\rangle|0\rangle_q \tag{3}$$

It is easy to prove that $|\varphi\rangle$ is entangled state by reduced density matrix and Schmidt law.

5. CONCLUSIONS

Quantum teleportation can be understood as a quantum computation and it has been suggested that quantum teleportation will play an important role as a primitive subroutine in quantum computation. Transmission of a rich state from one place to another is very important in the field of quantum information. One amazing discovery in this context is teleportation of a q analogue entangled state. In this paper entangled state and entanglement in the quantum information is presented. A q analogue entangled state is first put forward. Thus in the quantum teleportation we can have more abundant information by adjusting the value of q.

REFERENCES

1. Chiara, M., "On the role of entanglement in quantum information," *Physica A*, Vol. 338, 68–75, 2004.
2. Deutsch, D., "Information flow in entangled quantum systems," *Proceedings of the Royal Society of London, Series A (Mathematical, Physical and Engineering Sciences)*, Vol. 456, No. 1999, 1759–1774, 2000.
3. Dugic, M., "Quantum information processing: the case of vanishing interaction energy," *Physics Letters A*, Vol. 302, No. 5–6, 291–298, 2002.
4. Ham, B. S., "Dynamic quantum information processing based on quantum coherent control," *CLEO/Pacific Rim 2003. The 5th Pacific Rim Conference on Lasers and Electro-Optics (IEEE Cat. No.03TH8671)*, Vol. 1, No. 1, 198, 2003.
5. Chiu, F. L., "Exploiting randomness in quantum information processing," *Physics Letters A*, Vol. 301, No. 5–6, 343–349, 2002.
6. Poulin, D., "Classicality of quantum information processing," *Physical Review A (Atomic, Molecular, and Optical Physics)*, Vol. 65, No. 4, 042319/1–10, 2002.
7. Munro, W. J., "Efficient optical quantum information processing," *Journal of Optics B: Quantum and Semiclassical Optics*, Vol. 7, No. 7, S135–40, 2005.
8. Griffiths, R. B., "Nature and location of quantum information," *Physical Review A (Atomic, Molecular, and Optical Physics)*, Vol. 66, No. 1, 012311/1–15, 2002.
9. Simon, C., "Optimal quantum cloning via stimulated emission," *Physical Review Letters*, Vol. 84, No. 13, 2993–2996, 2000.
10. Pankaj, A. and A. K. Pati, "Probabilistic quantum teleportation," *Physics Letters A*, Vol. 305, 12–17, 2002.
11. Sergey, A. P., "Quantum teleportation of entanglement via quantum channel constructed from mode entangled states," *Optics Communications*, Vol. 249, 245–253, 2005.
12. Shangbin, L. and X. Jingbo, "Quantum probabilistic teleportation via entangled coherent states," *Physics Letters A*, Vol. 309, 321–328, 2003.
13. Xiaoguang, W. and P. Zanardi, "Quantum entanglement and Bell inequalities in Heisenberg spin chains," *Physics Letters A*, Vol. 301, 1–6, 2002.
14. Yuchun, W., H. Yongjian, and G. Guangcan, "When different entanglement witnesses can detect the same entangled states," *Physics Letters A*, Vol. 356, 402–405, 2006.
15. Isar, A. and W. Scheid, "Deformation of quantum oscillator and of its interaction with environment," *Physica A*, Vol. 335, 79–73, 2004.
16. Ballesteros, A., O. Civitarese, and M. Reboiro, "Correspondence between the q-deformed harmonic oscillator and finite range potentials," *Physical Review C*, Vol. 68, No. 4, 44307-1-5, 2003,

Influence of Electro-optic Effect on Wave-guide Efficiency of Optical Fiber with Cladding Made of Uniaxial Crystal Materials

Shanglin Hou[1,2] and Chunlian Hu[3]

[1]School of Science, Lanzhou University of Techology, Lanzhou , China
[2]Key Lab. of Optical Commu. & Lightwave Techno. of Ministry of Education
Beijing University of Posts and Telecomm., Beijing 100876, China
[3]Alloy Powder Works, Lanzhou University of Technology, Lanzhou, China,

Abstract— Influence of electric field running along the fiber on the wave-guide efficiency of the fiber with cladding made of uniaxial anisotropic and electro-optic crystal material whose optical axis is parallel to the axis of fiber was investigated. The calculated results indicate that the electric field has a strong impact on power distribution of this kind of fiber, and more optical energy was transmitted by the cladding with it increasing. These studies provide an important basis for designing new voltage sensors and studying some characteristics of electro-optic crystal.

1. INTRODUCTION

The index contribution has a great impact on the energy contribution between the core and the cladding of the fiber [1, 2]. Uniaxial anisotropic material is complex due to its indices different along its various axes. Thus, more parameters can be adjusted to design the characteristics of the fiber if its cladding is made of the Uniaxial material. Stevenson and Cozens et al. analyzed an optical fiber with a single-crystal core for the first time in 1974 [3, 4]. A doubly-clad W-type optical fiber with an inner cladding made of uniaxial crystal material has been studied in 2002 [5], and the studied results indicated that the adjustable range of zero dispersion wavelength of this kind fiber is wilder than that of common doubly-clad optical fiber. The fiber-optic polarizer with cladding made of uniaxial crystal material was also described [6]. Electro-optic effect and elasto-optic effect in a chirped fiber grating with cladding made of uniaxial crystal material are theoretically analyzed and the results indicate that the reflective spectra of the chirped grating can be changed by the electric field and the strain applied to the fiber grating cladding along z-axis [7]. The paper published in 2003 [8] predicted the characteristics of a new type of fiber Bragg grating with cladding made of uniaxial crystal material whose optical axis i. e., z-axis, is parallel to the axis of fiber Bragg grating, and the calculated results indicate that parameter K_{cl}, i. e., the ratio of the extraordinary ray refractive index to the ordinary ray index, has a strong impact on the reflectivity, Bragg wavelength, the bandwidth of the reflected wave and the dispersion.

In this work, the Influence of electric field running along the fiber on wave-guide efficiency of optical fiber with cladding made of uniaxial crystal materials whose optical axis i. e., z-axis, is parallel to the axis of fiber Bragg grating were predicted using numeric solution. The calculated results indicate that with the electric field increasing, the parameter K_{cl}, i. e., the ratio of the extraordinary to the ordinary ray refractive index, linearly decreases but its relative index difference linearly increases, and more optical energy was transmitted by the cladding. The result provides an important basis for designing new voltage sensors and studying some characteristics of electro-optic crystal.

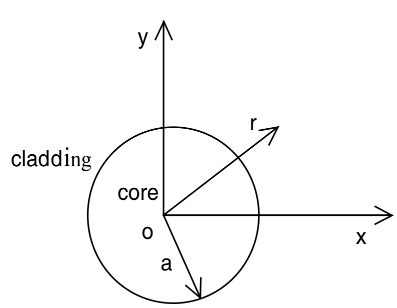


Figure 1: Cross profile of the fiber with cladding made of uniaxial electro-optic crystal material.

2. THEORETICAL ANALYSIS

2.1. The Characteristic Equation of the Vector Mode

The sketch of the fiber with cladding made of uniaxial crystal material was shown in Fig. 1. It has a core radius a and a core refractive index n_0. The cladding with an infinite diameter is made of uniaxial crystal material whose optical axis is taken to be parallel to the axis of the fiber, i. e., z-axis, and its principal axis indices are n_x, n_y and n_z respectively, which satisfy: $n_x = n_y \neq n_z$. For the optical fiber, the axial electric and magnetic field components satisfy wave equation as follows [1]:

$$\begin{cases} \left(\nabla_t^2 + k^2 n_0^2 - \beta^2\right) e_z = 0 \\ \left(\nabla_t^2 + k^2 n_0^2 - \beta^2\right) h_z = 0 \end{cases} \quad (r < a), \tag{1}$$

$$\begin{cases} \left(\nabla_t^2 + k^2 n_z^2 - \dfrac{n_z^2}{n_t^2}\beta^2\right) e_z = 0 \\ \left(\nabla_t^2 + k^2 n_t^2 - \beta^2\right) h_z = 0 \end{cases} \quad (r > a), \tag{2}$$

where $n_t = n_x = n_y$ and $n_0 > n_t$. $k = \dfrac{2\pi}{\lambda}$ is the wave number in vacuum and β is the propagation constant. If $\beta^2 < k^2 n_0^2$, the parameters are defined as

$$K_{cl} = \frac{n_z}{n_t}, \quad \Delta = \frac{n_0 - n_t}{n_0}, \quad U = a\sqrt{k^2 n_0^2 - \beta^2}, \quad W = a\sqrt{\beta^2 - k^2 n_t^2}, \quad V = ak\sqrt{n_0^2 - n_t^2},$$

where K_{cl} is the ratio of the extraordinary to the ordinary ray refractive index and V is normalized frequency. Through matching the relationship between axial and tangential field components, and using the boundary conditions of electromagnetic field, the characteristic equation of the guide mode can be obtained as follows:

$$F\left[\frac{n_0^2}{n_t^2 K_{cl}} \frac{\mathrm{J}_m'(U)}{U \mathrm{J}_m(U)} + \frac{\mathrm{K}_m'(K_{cl} W)}{W \mathrm{K}_m(K_{cl} W)}\right] = \frac{m^2 Q}{K_{cl}}\left(\frac{n_0^2}{n_t^2 U^2} + \frac{1}{W^2}\right), \tag{3}$$

where J_m, K_m are the Bessel and modified Bessel functions respectively. The parameters F and Q are defined as follows:

$$F = \frac{\mathrm{J}_m'(U)}{U \mathrm{J}_m(U)} + \frac{\mathrm{K}_m'(W)}{W \mathrm{K}_m(W)}, \quad Q = \frac{1}{U^2} + \frac{1}{W^2}.$$

2.2. Electro-optic Effect

For $3m$ point group uniaxial material of trigonal system, the relationship among the incremental matrix $[\Delta B_i]$, the electro-optic tensor $[\gamma_{ij}]$ and the adscititious electric field E_3 running along the fiber of the refractive index ellipsoid equation coefficient is expressed as [9]

$$\left(\frac{1}{n_t^2} + \gamma_{13} E_3\right)\left(x^2 + y^2\right) + \left(\frac{1}{n_z^2} + \gamma_{33} E^3\right) z^2 = 1. \tag{4}$$

The cladding is also the uniaxial material by adding the electric field E_3 running along the fiber, while its refractive index and optical axis is not changed. Thus, the new refractive indices after adding electric field the can be obtained as follows:

$$n_t' = n_t - \frac{1}{2} n_t^3 \gamma_{13} E_3, \tag{5}$$

$$n_z' = n_z - \frac{1}{2} n_z^3 \gamma_{33} E_3, \tag{6}$$

then two parameters are defined as

$$\Delta_{cl} = \frac{n_0 - n_t'}{n_0}, \tag{7}$$

$$K_{cl}' = \frac{n_z'}{n_t'}. \tag{8}$$

2.3. Wave-guide Efficiency

It is clear that a substantial amount of light power is transported by the cladding, while most of the power is confined in the core. Thus, the total power flow is the sum of power carried by both the core and by the cladding in the HE_{11} mode. The formula expressing this statement is $P_{\text{total}} = P_{\text{core}} + P_{\text{clad}}$, where P_{core} and P_{clad} is the integration of the longitudinal Poynting vector $\vec{S}_z$ with the corresponding cross-sectional area by inserting the traverse component of electric field derived in Eqs. (1) and (2). Then

$$\begin{cases} P_{\text{core}} = -\dfrac{\pi a^2\beta\omega}{U^2}\left(\varepsilon_0 n_0^2 A^2 + \mu_0 B^2\right)\left(\dfrac{1}{4}G1 + \dfrac{a^2m^2}{U^2}G2\right) \\ P_{\text{clad}} = -\dfrac{\pi a^2\beta\omega\varepsilon_0 n_t^2}{W^2}A^2\dfrac{\mathrm{J}_m^2(U)}{\mathrm{K}_m^2(k_{cl}W)}\left(\dfrac{k_{cl}^2}{4}G3 + \dfrac{a^2m^2}{W^2}G4\right) \\ \qquad -\dfrac{\pi a^2\beta\omega\mu_0}{W^2}B^2\dfrac{\mathrm{J}_m^2(U)}{\mathrm{K}_m^2(W)}\left(\dfrac{1}{2}G5 + \dfrac{am^2}{W^2}G6\right) \end{cases}, \tag{9}$$

where

$$G1 = \int_0^a \left[\mathrm{J}_{m-1}\left(\frac{U}{a}r\right) - \mathrm{J}_{m+1}\left(\frac{U}{a}r\right)\right]^2 r\mathrm{d}r, \qquad G2 = \int_0^a \frac{1}{r}\mathrm{J}_m^2\left(\frac{U}{a}r\right)\mathrm{d}r,$$

$$G3 = \int_a^\infty \left[\mathrm{K}_{m-1}\left(\frac{k_{cl}W}{a}r\right) + \mathrm{K}_{m+1}\left(\frac{k_{cl}W}{a}r\right)\right]^2 r\mathrm{d}r, \qquad G4 = \int_a^\infty \frac{1}{r}\mathrm{K}_m^2\left(\frac{k_{cl}W}{a}r\right)\mathrm{d}r,$$

$$G5 = \int_a^\infty \left[\mathrm{K}_{m-1}\left(\frac{W}{a}r\right) + \mathrm{K}_{m+1}\left(\frac{W}{a}r\right)\right]^2 r\mathrm{d}r, \qquad G6 = \int_a^\infty \frac{1}{r}\mathrm{K}_m^2\left(\frac{W}{a}r\right)\mathrm{d}r.$$

Then the wave-guide efficiency can be obtained as follows:

$$\eta = \frac{P_{\text{core}}}{P_{\text{total}}} = \frac{W^2 S_1\left(F^2 - n_0^2k^2Q^2\right)}{U^2\mathrm{J}_m^2(U)\left[\dfrac{F^2}{\mathrm{K}_m^2(W)}S_2 - \dfrac{n_t^2k^2Q^2}{\mathrm{K}_m^2(K_{cl}W)}S_3\right] + W^2S_1\left(F^2 - n_0^2k^2Q^2\right)}, \tag{10}$$

where

$$Q = \frac{\mathrm{K}_{m-1}(W)+\mathrm{K}_{m+1}(W)}{2W\mathrm{K}_m(W)} - \frac{\mathrm{J}_{m-1}(U) - \mathrm{J}_{m+1}(U)}{3U\mathrm{J}_m(U)},$$

$$F = m\beta\left(\frac{1}{U^2} + \frac{1}{W^2}\right), \qquad S_1 = \frac{1}{4}G1 + \frac{a^2m^2}{U^2}G2,$$

$$S_2 = \frac{1}{4}G5 + \frac{a^2m^2}{W^2}G6, \qquad S_3 = \frac{k_{cl}^2}{4}G3 + \frac{a^2m^2}{W^2}G4.$$

3. CALCULATED RESULTS AND ANALYSIS

$LiNbO_3$ uniaxial crystal, as a $3m$ point group conventional uniaxial material of trigonal system, was taken as the cladding material for this kind of fiber. Its principal indices n_t, n_z are 2.29 and 2.20 respectively, the electro-optical coefficient $\gamma_{13} = 8.6 \times 10^{-12}$ m/V and $\gamma_{33} = 30.8 \times 10^{-12}$ m/V at $\lambda_0 = 633$ nm, setting $n_0 = 2.33$ and $a = 500$ nm, then the calculated curve of K'_{cl} as function of the electric field E_3 and the curve of the relative index difference Δ_{cl} as function of the electric field E_3 can be obtained from Eqs. (5)–(8) and shown in Fig. 2 and Fig. 3 respectively. It can be seen that K'_{cl} linearly decreases and the relative index difference Δ_{cl} linearly decreases with it increasing at the order of magnitude of 10^7 V/m.

Thus the curve of the wave-guide efficiency η versus the electric field E_3 can be obtained from Eq. (10) due to the influence of the electric field E_3 on K'_{cl} and Δ_{cl}, as shown in Fig. 4. It can be seen that the wave-guide efficiency η approximately linearly decreases with E_3 increasing. It means that less optical energy is confined in the core with the electric field running along the fiber increasing, i. e., the more energy transmitted by the cladding, the stronger the electric field is.

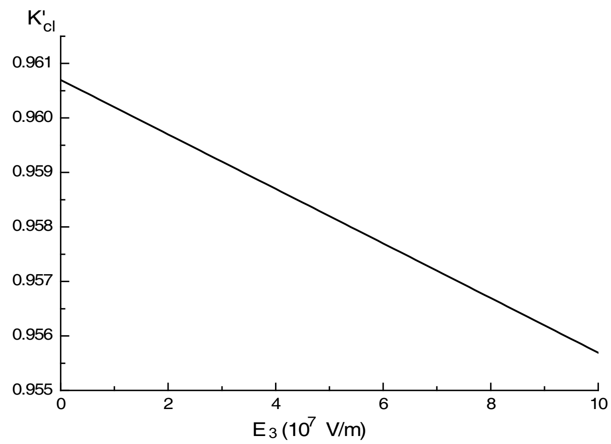


Figure 2: K_{cl} varies with E_3.

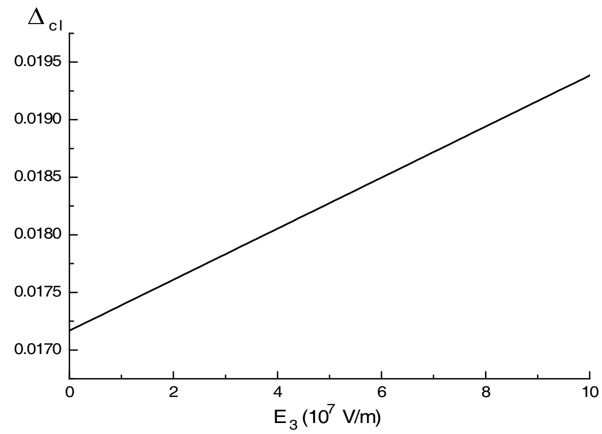


Figure 3: Δ_{cl} varies with E_3.

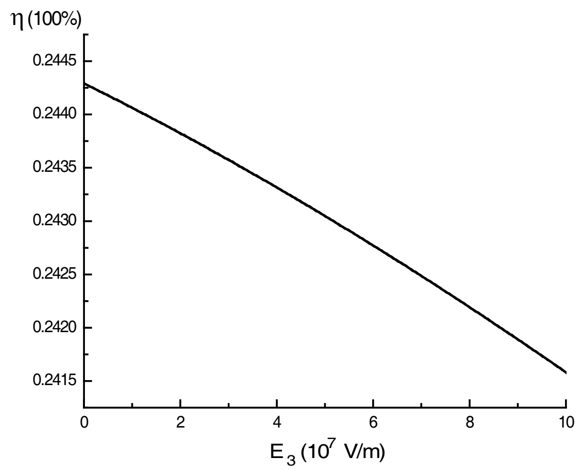


Figure 4: η varies with E_3.

It follows from the above discussion that less power is confined in the core of the fiber with cladding made of uniaxial anisotropic and electro-optic crystal material with the electric field running along the fiber increasing. The result provides an important basis for designing new voltage sensors and studying some characteristics of electro-optic crystal.

4. CONCLUSION

The influence of electric field running along the fiber on wave-guide efficiency of optical fiber with cladding made of uniaxial crystal materials whose optical axis i. e., z-axis, is parallel to the axis of fiber Bragg grating were predicted using numeric solution. The calculated results indicate that with the electric field increasing, the parameter K'_{cl} linearly decreases but its relative index difference linearly increases, and more optical energy was transmitted by the cladding. The result provides an important basis for designing new voltage sensors and studying some characteristics of electro-optic crystal.

ACKNOWLEDGMENT

This work was supported by National Basic Research Program of China (2003CB314906), National 863 High Technology Project of China (2003AA311010), the Key grant Project of Chinese Ministry of Education (NO. 104046) and the Outstanding Young Teacher Training Plan of Lanzhou University of Technology, China.

REFERENCES

1. Snydan, A. W. and J. D. Love, *Optical Waveguide Theory*, Chapman and Hall, London, 1983.
2. Hou, S. L. and X. Q. Qi, "Influence of optical parameters of doubly clad single-mode fibers on propagation characteristics," *Chinese Journal of Quantum Electronics*, Vol. 22, No. 4, 642–645, 2005.

3. Stevenson, J. L. and R. B. Dyott, "Optical fiber waveguide with a single-crystal core," *Electron. Lett.*, Vol. 10, 449–450, 1974.
4. Cozens, J., "Propagation in cylindrical fibers with anisotropic crystal cores," *Electron. Lett.*, Vol. 12, 413–415, 1974.
5. Zhang, X. P. and Z. H. Tan, "Analysis of transmission characteristics of doubly clad fibers with an inner cladding made of uniaxial crystal materials," *Optics Commu.*, Vol. 204, 127–136, 2002.
6. Wakaki, M., Y. Komachi, and H. Machida, "Fiber-optic polarizer using birefringent crystal as a cladding," *Applied Optics*, Vol. 35, No. 15, 2591–2594, 1996.
7. Zhang, X. P. and D. C. Lv, "Theoretical study of electro-optic effect and elasto-optic effect in chirped fiber grating with unixial crystal taterials cladding," *Optical and Quantum Electronics*, Vol. 36, No. 5,469–481, 2004.
8. Zhang, X. P. and S. H. Hou, "Study on characteristics of fiber Bragg grating with uniaxial crystal cladding," *Optics Commu.*, Vol. 219, No. 1–6, 193–198, 2003.
9. Yariv, A., *Introduction to Optical Electronics*, Rinehart and Winston, Holt, 1976.

Diffraction by a Kerr-type Nonlinear Dielectric Layer

Yu. V. Shestopalov[1] **and V. V. Yatsyk**[2]

[1]Karlstad University, Karlstad, Sweden
[2]IRE National Academy of Science of Ukraine, Ukraine

Abstract— The diffraction of a plane wave by a transversely inhomogeneous isotropic nonmagnetic linearly polarized dielectric layer filled with a Kerr-type nonlinear medium is considered. The diffraction problem is reduced to a cubic-nonlinear integral equation (IE) of the second kind and to a system of nonlinear operator equations of the second kind solved using iterations. Sufficient conditions of the IE unique solvability are obtained using the contraction principle.

1. STATEMENT OF THE PROBLEM OF DIFFRACTION BY A WEAKLY NONLINEAR LAYER

Denote by $\vec{E}(\vec{r}) \equiv \vec{E}(\vec{r},\omega)$ and $\vec{H}(\vec{r}) \equiv \vec{H}(\vec{r},\omega)$ the complex amplitudes of the stationary electromagnetic field; the time dependence is $\exp(-i\omega t)$. Consider the problem of diffraction of a plane stationary electromagnetic wave by a nonmagnetic, isotropic, transversely inhomogeneous, $\varepsilon^{(L)}(z) = \varepsilon^{(L)}_{xx}(z)$, and linearly polarized, $\vec{E}(\vec{r}) = (E_x(y,z),0,0)$, $\vec{H}(\vec{r}) = (0,H_y,H_z)$ (E-polarization), with vector of polarization $\vec{P}^{(NL)} = \left(P^{(NL)}_x,\ 0,\ 0\right)$, and Kerr-like weakly nonlinearity, $\left|\varepsilon^{(NL)}\right| << \left|\varepsilon^{(L)}\right|$ layered dielectric structure, see Fig. 1 and [1,2]. Here $P^{(NL)}_x = \frac{3}{4}\chi^{(3)}_{xxxx}|E_x|^2 E_x$, $\varepsilon = \varepsilon^{(L)} + \varepsilon^{(NL)}$ at $|z| \le 2\pi\delta$ is the permittivity of the nonlinear layer, $\varepsilon^{(L)} = 1 + 4\pi\chi^{(1)}_{xx}(z)$, $\varepsilon^{(NL)} = 3\pi\chi^{(3)}_{xxxx}(z)|E_x|^2 = \alpha(z)|E_x|^2$, $\alpha(z) = 3\pi\chi^{(3)}_{xxxx}(z)$, and $\chi^{(1)}_{xx}(z)$ and $\chi^{(3)}_{xxxx}(z)$ are components of the susceptibility tensor.

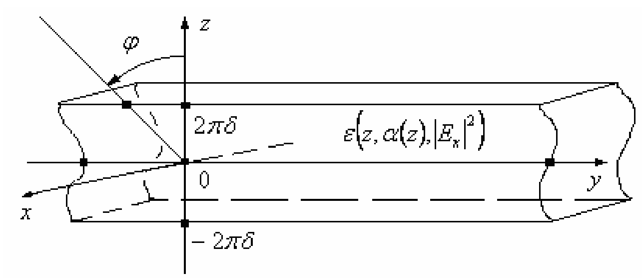


Figure 1: Weakly nonlinear dielectric layered structure.

One can show, similarly to [3,4], that the total field $E_x(y,z) = E^{inc}_x(y,z) + E^{scat}_x(y,z)$ of diffraction of the plane wave $E^{inc}_x(y,z) = a^{inc} \cdot \exp\{i\,[\phi\, y - \Gamma \cdot (z - 2\,\pi\,\delta)]\}$, $z > 2\,\pi\,\delta$, by the weakly nonlinear dielectric layer (Fig. 1) satisfies the equation

$$\nabla^2 \cdot \vec{E} + \frac{\omega^2}{c^2} \cdot \varepsilon^{(L)}(z) \cdot \vec{E} + \frac{4\pi\omega^2}{c^2} \cdot \vec{P}^{(NL)} \equiv \left[\nabla^2 + \kappa^2 \cdot \varepsilon\left(z,\ \alpha(z),\ |E_x|^2\right)\right] \cdot E_x(y,z) = 0, \tag{1}$$

and the following generalized boundary conditions: continuity of the tangential field components on the layer boundary, the spatial inhomogeneity condition

$$E_x(y,z) = U(z) \cdot \exp(i\phi\, y), \tag{2}$$

and the radiation condition for the scattered field

$$E^{scat}_x(y,z) = \left\{ \begin{array}{c} a^{scat} \\ b^{scat} \end{array} \right\} \cdot \exp\left(i \cdot (\phi\, y \pm \Gamma \cdot (z \mp 2\,\pi\,\delta))\right), \quad z \begin{array}{c} > \\ < \end{array} \pm 2\,\pi\,\delta. \tag{3}$$

Here $\omega = \kappa c$ is the dimensionless circular frequency; $\kappa = \omega/c = 2\pi/\lambda$ is the dimensionless frequency parameter such that $h/\lambda = 2\kappa\delta$, where λ is the free-space wavelength; $c = (\varepsilon_0\mu_0)^{1/2}$ is the dimensionless quantity equal to the speed of light in the medium containing the layered structure ($\mathrm{Im}\, c = 0$); ε_0 and μ_0 are the material parameters of the medium;

$$\nabla^2 = \partial^2/\partial y^2 + \partial^2/\partial z^2;$$

$$\varepsilon\left(z,\ \alpha(z), |E_x|^2\right) \equiv \varepsilon\left(z,\ \alpha(z), |U(z)|^2\right) = \begin{cases} 1, & |z| > 2\pi\delta \\ \varepsilon^{(L)}(z) + \alpha(z)\cdot|U(z)|^2, & |z| \le 2\pi\delta \end{cases},$$

where $\varepsilon^{(L)}(z)$ is piecewise continuously differentiable with respect to z; $\alpha(z) = 3\pi\chi^{(3)}_{xxxx}(z)$; $\Gamma = (\kappa^2 - \phi^2)^{1/2}$; $\phi \equiv \kappa\cdot\sin(\varphi)$; and φ is the angle of incidence of the plane wave, $|\varphi| < \pi/2$.

We look for the solution to problem (1)–(3) in the form

$$\begin{aligned} E_x(y,z) &= U(z)\cdot\exp(i\ \phi\ y) \\ &= \begin{cases} a^{inc}\cdot\exp\{i\cdot[\phi y - \Gamma\cdot(z-2\pi\delta)]\} + a^{scat}\cdot\exp\{i\cdot[\phi y + \Gamma\cdot(z-2\pi\delta)]\}, & z > 2\pi\delta, \\ U^{scat}(z)\cdot\exp(i\cdot\phi y), & |z| \le 2\pi\delta, \\ b^{scat}\cdot\exp\{i\cdot[\phi y - \Gamma\cdot(z+2\pi\delta)]\}, & z < -2\pi\delta. \end{cases} \end{aligned} \tag{4}$$

Here $U(-2\pi\delta) = b^{scat}$ and $U(2\pi\delta) = a^{inc} + a^{scat}$.

In [3] we solve problem (1)–(3) by reducing it to a one-dimensional nonlinear integral equation (IE) along the layer height $z \in [-2\pi\delta,\ 2\pi\delta]$ with respect to the scattered field component $U(z) \equiv U^{scat}(z)$

$$U(z) + \frac{i\kappa^2}{2\Gamma}\int_{-2\pi\delta}^{2\pi\delta} \exp(i\Gamma|z-z_0|)\left[1 - \left(\varepsilon^{(L)}(z_0) + \alpha(z)\,|U(z_0)|^2\right)\right]U(z_0)\,dz_0 = U^{inc}(z), \quad |z| \le 2\pi\delta, \tag{5}$$

where $U^{inc}(z) = a^{inc}\exp[-i\Gamma\cdot(z-2\pi\delta)]$.

2. SUFFICIENT CONDITION OF THE EXISTENCE OF SOLUTION TO NONLINEAR IE

Assume that the permittivity function $\varepsilon^{(L)}(z)$ (in the argument of ϵ) is positive, bounded, and continuous in the interval $[-2\pi\delta, 2\pi\delta]$ and

$$1 < \varepsilon^{(L)}(z) \le E, \quad z \in [-2\pi\delta, 2\pi\delta], \quad E > 1. \tag{6}$$

Write IE (5) in the operator form

$$U + AU - F(\alpha, U) = f, \tag{7}$$

where

$$\begin{aligned} f(z) &= U^{inc}(z) = a^{inc}\exp[-i\kappa\cos(\varphi)\cdot(z-2\pi\delta)], \\ AU &= \int_{-2\pi\delta}^{2\pi\delta} k(z-z_0)\cdot\left[1-\varepsilon^{(L)}(z_0)\right]\cdot U(z_0)dz_0 \end{aligned} \tag{8}$$

is a linear integral operator with the continuous kernel

$$k(t) = s_0\exp[2i\kappa\cos(\varphi)\cdot|t|], \quad s_0 = \frac{i\kappa}{2\cos(\varphi)}, \quad \left(-\frac{\pi}{2} < \varphi < \frac{\pi}{2}\right),$$

and

$$F(\alpha, U) = \int_{-2\pi\delta}^{2\pi\delta} k(z-z_0)\cdot\alpha(z_0)\cdot|U(z_0)|^2\,U(z_0)dz_0$$

It is easy to see that

$$\max_{z\in[-2\pi\delta,\ 2\pi\delta]}\left[\int_{-2\pi\delta}^{2\pi\delta} |k(z-z_0)|\cdot\left|1-\varepsilon^{(L)}(z_0)\right|dz_0\right] = (E-1)\,\frac{2\pi\delta\kappa}{\cos(\varphi)} = (E-1)\,q_0.$$

Assume that $|\varphi| < \pi/2$, and parameters a, κ, δ are positive and satisfy the condition

$$(E-1)q_0 < 1, \quad q_0 = \frac{2\pi\delta\kappa}{\cos(\varphi)}. \tag{9}$$

Then one can show that the operator $T(U) = -AU + F(\alpha, U) + f$ defined by (7) and (8) is a contraction with respect to the maximum-norm if the nonlinearity parameter $\tilde{\alpha} = \max\limits_{z\in[-2\pi\delta,\, 2\pi\delta]} |\alpha(z)|$ is sufficiently small, which implies that IE (5) has the unique solution $U^*(z)$ continuous in the closed interval $[-2\pi\delta,\, 2\pi\delta]$. $U^*(z)$ is a limit with respect to the maximum-norm of the function sequence $U_n(z)$ determined according to

$$U_{n+1} = T\left(U_n\right), \quad n = 0, 1, 2, \ldots, \tag{10}$$

here the initial function is chosen subject to the condition $\max\limits_{z\in[-2\pi\delta,\, 2\pi\delta]} |U_0(z)| < p, \quad p \in (p_1, p_2)$, where p_1 and p_2 are positive zeros of the polynomial $\tilde{\alpha} q_0 p^3 - [1-(E-1)q_0]p + \left|a^{inc}\right|$.

The results can be extended to the problem of diffraction by a lossy nonlinear layered structure when the permittivity $\varepsilon^{(L)}(z)$ is an arbitrary complex-valued function continuous and bounded on the line. To this end denote $\tilde{\varepsilon}(z) = \varepsilon^{(L)}(z) - 1 = \varepsilon_1(z) e^{i\varepsilon_2(z)}$, where, according to physical assumptions of the model, the real and imaginary parts of the "shifted" permittivity function $\tilde{\varepsilon}(z)$ are positive, continuous, and bounded on the line such that the modulus $\varepsilon_1(z)$ and argument $\varepsilon_2(z)$ satisfy $0 \le \varepsilon_2(z) < \pi/2$ and $\varepsilon_1(z) \ge 1$.

Assuming, similar to (6) and taking into account the conditions for the permittivity function that $0 < \varepsilon_1(z) \le E_1$ (that is, $0 < \left|\varepsilon^{(L)}(z)\right| \le E_1$) we can estimate the maximum-norm of the integral operator (8)

$$A_1 U = \int\limits_{-2\pi\delta}^{2\pi\delta} k_1\left(z, z_0\right) \varepsilon_1\left(z_0\right) U\left(z_0\right) dz_0, \quad k_1\left(z, z_0\right) = -s_0 \exp\left\{i\left[2\kappa\cos\left(\varphi\right)\left|z - z_0\right| \varepsilon_2\left(z_0\right)\right]\right\},$$

in the complex domain and show that the corresponding nonlinear operator $T_1(U) = -A_1 U + F(\alpha, U) + f$ defined using (7) and (8) is a contraction and IE (5) has the unique solution $U^*(z)$ continuous in the closed interval $[-2\pi\delta, 2\pi\delta]$. This solution is a limit with respect to the maximum-norm of the function sequence $U_n(z)$ (the fixed point of operator $T_1(U)$) determined according to (10) in which the initial function satisfies the condition $\max\limits_{z\in[-2\pi\delta, 2\pi\delta]} |U_0(z)| < q, \quad q \in (q_1, q_2)$, where q_1 and q_2 are positive zeros of the polynomial $\tilde{\alpha} q_0 p^3 - q - E_1 q_0 p + \left|a^{inc}\right|$. The rate of convergence of the fixed-point iterations can be estimated in the form $\|U_n - U^*\| < t_1^{n-1} \left\|T_1\left(U_0\right) - U^*\right\|, \quad n = 2, 3, \ldots$, using the quantity $t_1 = q_0\left(E_1 + 3\tilde{\alpha} q_1\right) < 1$.

Taking into account the equivalence of IE (5) and boundary value problem (1)–(3) we conclude that under condition (9) this problem is uniquely solvable if the nonlinearity parameter $\tilde{\alpha}$ is sufficiently small.

3. SUFFICIENT CONDITION OF THE EXISTENCE OF SOLUTION: REDUCING TO A FUNCTIONAL EQUATION SYSTEM

Here we present the proof of an alternative sufficient condition for the existence of a solution to nonlinear IE (5). The approach developed in this section enables one to create a rather efficient method of the numerical solution of the IE [4]. To this end, reduce (5) to a nonlinear functional equation system considering the system of two IEs in the domain $|z| \le 2\pi\delta$:

$$\begin{gathered} U_{n+1}(z) + \frac{i\kappa^2}{2\Gamma} \int\limits_{-2\pi\delta}^{2\pi\delta} \exp\left(i\Gamma \cdot |z - z_0|\right) \cdot \left[1 - \left(\varepsilon^{(L)}\left(z_0\right) + \alpha\left(z_0\right)\left|U_n\left(z_0\right)\right|^2\right)\right] \cdot U_n\left(z_0\right) dz_0 = U^{inc}(z), \\ \Psi_n(z) + \frac{i\kappa^2}{2\Gamma} \int\limits_{-2\pi\delta}^{2\pi\delta} \exp\left(i\Gamma \cdot |z - z_0|\right) \cdot \left[1 - \left(\varepsilon^{(L)}\left(z_0\right) + \alpha\left(z_0\right)\left|U_n\left(z_0\right)\right|^2\right)\right] \cdot \Psi_n\left(z_0\right) dz_0 = U^{inc}(z). \end{gathered} \tag{11}$$

The first equation of system (11) is an iteration scheme of solution to nonlinear equation (5). The second is a linear IE with respect to $\Psi_n(z)$ for the given $U_n(z_0)$. If $\Psi_n(z)$ is not an eigenfunction of the problem of diffraction by the layer with the permittivity $\varepsilon\left(z,\, \alpha\left(z\right), \left|U_n(z)\right|^2\right) \equiv \varepsilon^{(L)}(z) + \alpha(z)\left|U_n(z)\right|^2$, then the second equation is uniquely solvable [3, 4] and its solution can be represented in the form

$$\Psi_n(z) = \Psi\left(z,\, \alpha\left(z\right), \left|U_n(z)\right|^2\right) \cdot U^{inc}\left(z\right).$$

Here $\Psi\left(z, \alpha(z), \left|U_n(z)\right|^2\right)$ is the solution to the linear IE at $U^{inc}(z) = 1$ such that $\left|\Psi\left(z, \alpha(z), \left|U_n(z)\right|^2\right)\right| \leq 1$.

The analysis of the convergence criterion for the sequence $U_n(z)$, $\Psi_n(z)$ specified by system (11) enables one to obtain a sufficient condition for the existence of solution to nonlinear IE (5).

Kernels of IEs (11) are identical, which makes it possible to calculate and estimate the L_2-norm of the difference between $U_n(z)$ and $\Psi_n(z)$

$$\begin{aligned}
\rho\left[U_{n+1}(z),\, \Psi_n(z)\right] &= \left[\int_{-2\pi\delta}^{2\pi\delta} \left|U_{n+1}(z) - \Psi_n(z)\right|^2 dz\right]^{1/2} \\
&\leq \frac{\kappa^2}{2\Gamma} \cdot \left\{\int_{-2\pi\delta}^{2\pi\delta}\int_{-2\pi\delta}^{2\pi\delta} \left|1 - \left(\varepsilon^{(L)}(z_0) + \alpha(z_0)\left|U_n(z_0)\right|^2\right)\right|^2 dz_0 dz\right\}^{1/2} \\
&\quad \cdot \left\{\int_{-2\pi\delta}^{2\pi\delta} \left|U_n(z_0) - \Psi_n(z_0)\right|^2 dz_0\right\}^{1/2} \\
&\leq \frac{\kappa^2}{2\Gamma} \cdot 4\pi\delta \cdot \max_{|z|\leq 2\pi\delta} \left|1 - \left(\varepsilon^{(L)}(z) + \alpha(z)\left|U_n(z)\right|^2\right)\right| \cdot \rho\left[U_n\left(z\right), \Psi_n(zt)\right] \\
&\leq \frac{\kappa^2}{2\Gamma} \cdot 4\pi\delta \cdot \max_{|z|\leq 2\pi\delta} \left[\left|1 - \varepsilon^{(L)}(z)\right| + \left|\alpha(z)\right| \cdot \left|U_n(z)\right|^2\right] \cdot \rho\left[U_n(z), \Psi_n(z)\right] \\
&\leq \frac{\kappa^2}{2\Gamma} \cdot 4\pi\delta \cdot \max_{|z|\leq 2\pi\delta} \left[\left|1 - \varepsilon^{(L)}(z)\right| + \left|\alpha(z)\right| \cdot \left|U^{inc}(z)\right|^2\right] \cdot \rho\left[U_n(z), \Psi_n(z)\right].
\end{aligned} \tag{12}$$

The last inequality in (12) is obtained taking into account the condition $\max_{|z|\leq 2\pi\delta}\left|U_n\left(z\right)\right| \leq \max_{|z|\leq 2\pi\delta}\left|U^{inc}(z)\right|$ which holds for all $n = 0, 1, 2, \ldots$ and directly follows from the inequality $\max_{|z|\leq 2\pi\delta}\left|U(z)\right| \leq \max_{|z|\leq 2\pi\delta}\left|U^{inc}\left(z\right)\right|$ due to (4). We see that, according to (12), in the case under study of weakly nonlinear approximation when

$$\max_{|z|\leq 2\pi\delta}\left[\left|\alpha(z)\right| \cdot \left|U\left(z\right)\right|^2\right] \leq \max_{|z|\leq 2\pi\delta}\left[\left|\alpha(z)\right| \cdot \left|U^{inc}(z)\right|^2\right] << \max_{|z|\leq 2\pi\delta}\left|\varepsilon^{(L)}(z)\right|, \tag{13}$$

the iterations defined by the first equation of (11) converges to the unique solution determined by (11) if the term in the last inequality of (12) multiplying the norm satisfies the condition

$$\frac{\kappa^2}{2\Gamma} \cdot 4\pi\delta \cdot \max_{|z|\leq 2\pi\delta}\left[\left|1 - \varepsilon^{(L)}(z)\right| + \left|\alpha(z)\right| \cdot \left|U^{inc}(z)\right|^2\right] < 1.$$

Taking into consideration the expression for the transverse wavenumber $\Gamma = \left\{\kappa^2 - \left[\kappa\sin(\varphi)\right]^2\right\}^{1/2}$ rewrite the last inequality as

$$\kappa \cdot 2\pi\delta \cdot \max_{|z|\leq 2\pi\delta}\left[\left|1 - \varepsilon^{(L)}(z)\right| + \left|\alpha(z)\right| \cdot \left|U^{inc}(z)\right|^2\right] < \cos(\varphi). \tag{14}$$

We have proved the following statement which constitutes a sufficient condition for the existence of solution to nonlinear IE (5): Assume that the weakly nonlinear approximation (13) holds. Then nonlinear IE (5) has the unique continuous solution if condition (14) holds. This solution can be obtained using both the iterations defined by the first equation of (11) and the equivalent iteration scheme according to the second equation of (11) if to consider its solution $\Psi_n(z)$ as the $n+1$ approximation (setting $\Psi_n(z) \equiv U_{n+1}(z)$) to the sought for $U(z)$.

4. CONCLUSION

We have reduced the problem of diffraction of a plane wave by a dielectric layer filled with a Kerr-type nonlinear medium to a cubic-nonlinear IE of the second kind and to a system of nonlinear operator equations of the second kind. Sufficient conditions of the IE unique solvability have been obtained; the equations have been solved using iterations.

On the basis of these solution techniques and the IE obtained one can perform numerical investigation of the resonance effects caused by certain nonlinear properties of the object under study irradiated by an intense electromagnetic field. In particular, one can determine the critical limits of the excitation field intensity that govern applicability of the developed mathematical model. The proposed methods can be further applied to the analysis of various physical phenomena including self-influence and interaction of waves; determination of eigenfields, natural (resonance) frequencies of nonlinear objects, and dispersion amplitude-phase characteristics of the diffraction fields; description of evolution processes in the vicinities of critical points; and to the design and modeling of novel scattering, transmitting, and memory devices.

REFERENCES

1. Akhmediev, N. N. and A. Ankiewicz, *Solitons*, Moscow, Fizmatlit, 2003.
2. Smirnov, Y., H. W. Schürmann, and Y. Shestopalov, "Integral equation approach for the propagation of TE-waves in a nonlinear dielectric cylindrical waveguide," *J. Nonlin. Math. Phys.*, Vol. 11, No. 2, 256–268, 2004.
3. Shestopalov, V. P. and Y. K. Sirenko, *Dynamic Theory of Gratings*, Naukova Dumka, Kiev, 1989.
4. Yatsyk, V. V., "Diffraction by a layer and layered structure with positive and negative susceptibilities of Kerr-nonlinear media," *Uspehi Sovremennoj Radioelektroniki*, No. 8, 68-80, 2006.

1D Canonical and Perturbed Quantum Potential-well Problem: A Universal Function Approach

I. Ahmed and A. R. Baghai-Wadji

School of Electrical and Computer Engineering, RMIT University, Melbourne, Victoria 3001, Australia

Abstract— We compute the eigenfunctions of canonical and associated perturbed quantum systems and utilize them as co-ordinate functions for solving more complex problems. We start with the infinite quantum potential-well problem because of the availability of the corresponding complete set of eigensolutions in closed-form. Consecutively, we consider perturbed systems defined by various added potentials, and set up the wavefunction equation of the perturbed system as a linear combination of the original eigensolutions, introducing an infinite number of *a priori* unknown coefficients. We utilize the notion of Universal Functions previously introduced for solving complex boundary value problems. An easy-to-use software package for the calculation of eigenpairs has been created. A glimpse of the numerical results will illuminate details of our theory.

1. INTRODUCTION

This work is the 1D analog of the theory and results presented in [1], following the same logical and procedural steps. Given a complex Boundary Value Problem (BVP), which will be referred to as the original problem, our method consists of constructing an auxiliary problem, which resembles the original problem but is much simpler. The simplicity assumption allows us to construct BVPs which are numerically more tractable or have closed-form solutions. We shall illuminate details of our technique by examples.

2. ON THE CONSTRUCTION OF THE AUXILIARY PROBLEMS

Consider the following Ordinary Differential Equation (ODE) for the function $\widetilde{\psi}(x)$ (actual problem):

$$-\frac{\hbar^2}{2M}\frac{d^2}{dx^2}\widetilde{\psi}(x)+V(x)\widetilde{\psi}(x)=\widetilde{E}\widetilde{\psi}(x) \tag{1}$$

The positive real-valued function $V(x)$ is defined in the interval $[0, L_x]$, with $\hbar$ and M being given positive real constants. The fact that the operator $\widetilde{\mathfrak{I}}=-\frac{\hbar^2}{2M}\frac{d^2}{dx^2}+V(x)$ is positive definite (PD) can be seen immediately:

$$\begin{aligned}\left\langle\widetilde{\mathfrak{I}}f(x)\middle|f(x)\right\rangle &= \left\langle-\frac{\hbar^2}{2M}\frac{d^2}{dx^2}f(x)\middle|f(x)\right\rangle+\langle V(x)f(x)|f(x)\rangle\\ &= -\frac{\hbar^2}{2M}\int_0^{L_x}dx\left(\frac{d^2}{dx^2}f(x)\right)f(x)+\int_0^{L_x}dxV(x)f(x)f(x)\\ &= \frac{\hbar^2}{2M}\left(\frac{df(x)}{dx}\right)f(x)\Big|_0^{L_x}+\frac{\hbar^2}{2M}\int_0^{L_x}dx\left(\frac{df(x)}{dx}\right)\left(\frac{df(x)}{dx}\right)+\int_0^{L_x}dxV(x)f^2(x). \end{aligned} \tag{2}$$

Restricting ourselves to the class of functions which vanish at the boundary points $x = 0$ and $x = L_x$, we obtain:

$$\left\langle\widetilde{\mathfrak{I}}f(x),f(x)\right\rangle=\frac{\hbar^2}{2M}\int_0^{L_x}dx\left(\frac{df(x)}{dx}\right)^2+\int_0^{L_x}dxV(x)f^2(x)>0 \tag{3}$$

Having shown the PD property of $\widetilde{\mathfrak{I}}$ we next construct an auxiliary problem related to our original problem by simplifying the operator $\widetilde{\mathfrak{I}}$. We set, for example, $V(x)\equiv 0$ for $x\in[0,L_x]$ (auxiliary problem):

$$-\frac{\hbar^2}{2M}\frac{d^2}{dx^2}\psi(x)=E\psi(x) \tag{4}$$

Since $-\frac{\hbar^2}{2M}\frac{d^2}{dx^2}$, is positive definite, the eigenvalues and the corresponding eigenfunctions of (2) are strictly positive. By substitution it can be verified that the solutions of the type ($n \in N$)

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right) \Longleftrightarrow E_n = \frac{\hbar^2}{2M}\frac{\pi^2}{L^2}n^2 \tag{5}$$

satisfy (4). The infinite set of the solution functions $\{\psi_n(x)\}$ constitutes a complete set of coordinate functions in $L_2(R)$ which can be utilized for expanding any arbitrary function in their span. Stated more precisely, we have the orthonormality condition $\langle\psi_n|\psi_m\rangle = \delta_{nm}$ where as, the completeness property, the expression for the resolution of identity takes the form:

$$I = \sum_{n=1}^{\infty} |\psi_n\rangle\langle\psi_n| \tag{6}$$

In (6) we made use of the Dirac's "bra" and "ket" notation. To see the latter relationship we apply (5) to the function $f(x) \in \mathrm{span}\{\psi_n\}$:

$$If(x) = f(x) = \sum_{n=1}^{\infty} |\psi_n\rangle \underbrace{\langle\psi_n|f\rangle}_{f_n} \tag{7}$$

where the "generalized Fourier coefficients" f_n are defined in (7) in the obvious fashion. Thus we arrive at

$$f(x) = \sum_{n=1}^{\infty} f_n|\psi_n\rangle. \tag{8}$$

This result shows that any function in the span of $\{\psi_n\}$ can be expressed in terms of a linear combination of $\{\psi_n\}$. In particular the solution for our original problem can be expressed as:

$$\widetilde{\psi}(x) = \sum_{n=1}^{\infty} \alpha_n|\psi_n(x)\rangle. \tag{9}$$

Next we determine the "action" of $\widetilde{\psi}$ on the auxiliary system and the "action" of ψ on the original system and subtract the two "actions." Minimization of the weighted residual "action" leads to an algebraic system of equations for the determination of the expansion coefficients α_n in (9). The above mentioned computational recipe consists of the following steps:

Step I: Multiply (4) and (1) by $\widetilde{\psi}(x)$ and $\psi(x)$ respectively, and subtract the resulting two "actions" to obtain:

$$-\frac{\hbar^2}{2M}\left[\widetilde{\psi}(x)\frac{d^2}{dx^2}\psi(x) - \psi(x)\frac{d^2}{dx^2}\widetilde{\psi}(x)\right] - V(x)\psi(x)\widetilde{\psi}(x) = \left(E - \widetilde{E}\right)\psi(x)\widetilde{\psi}(x) \tag{10}$$

Step II: Let $\langle f(x)|g(x)\rangle = \int_0^{L_x} dx f(x)g(x)$ denote the inner-product of the two real-valued functions $f(x)$ and $g(x)$. Integrate (10) over the interval $[0, L_x]$. Integrating by parts shows that the first term on LHS vanishes. This is because the contributions of I and J, as stated below, cancel out and thus result in (11):

$$I = \left\langle \widetilde{\psi}(x) \middle| \frac{d}{dx}\left(\frac{d\psi(x)}{dx}\right)\right\rangle \qquad \text{and} \qquad J = \left\langle \psi(x) \middle| \frac{d}{dx}\left(\frac{d\widetilde{\psi}(x)}{dx}\right)\right\rangle$$

$$-\left\langle \psi_m(x)\middle|V(x)\middle|\widetilde{\psi}(x)\right\rangle = \left(E_m - \widetilde{E}\right)\left\langle \psi_m(x)\middle|\widetilde{\psi}(x)\right\rangle \tag{11}$$

Here we have chosen the particular solution-pair $\{\psi_m, E_m\}$ for $\{\psi, E\}$. Using the expansion formula (9) for $\widetilde{\psi}(x)$, into (11) and exchanging the order of integration and summation we obtain:

$$-\sum_{n=1}^{\infty}\alpha_n\left\langle\psi_m|V(x)|\psi_n\right\rangle = \left(E_m-\widetilde{E}\right)\sum_{n=1}^{\infty}\alpha_n\left\langle\psi_m|\psi_n\right\rangle \tag{12}$$

$$A_{mn} = \left\langle\psi_m|V(x)|\psi_n\right\rangle = \int_0^{L_x} dx\psi_m V(x)\psi_n \tag{13}$$

Using the orthonormality relationship, we obtain (14), with δ_{mn} denoting the Kronecker delta symbol:

$$\sum_{n=1}^{\infty}\left(A_{mn}+\delta_{mn}\varepsilon_m\right)\alpha_n = \widetilde{\widetilde{E}}\alpha_m \tag{14}$$

Varying m from 1 to N, a suitably-chosen upper limit, we obtain an $N\times N$ algebraic eigenvalue problem. In (14) $\widetilde{\widetilde{E}}$ stands for $\widetilde{E}$ normalized by $\frac{\hbar^2}{2M}\frac{\pi^2}{L^2}$. Furthermore, $\varepsilon_m = m^2$.

3. ON THE CONSTRUCTION OF "UNIVERSAL FUNCTIONS"

Next we construct "Universal Function" associated with the class of problems characterized by $V(x)$ in (15) leading to the "mutual" inter-action terms A_{mn}, given in (16).

$$V(x) = \begin{cases} V_0 & x\in[a,b]\subset[0,L_x] \\ 0 & \text{elsewhere,} \end{cases} \tag{15}$$

$$A_{mn} = \left\langle\psi_m|V(x)|\psi_n\right\rangle = V_0\int_a^b dx\psi_m(x)\psi_n(x) \tag{16}$$

Rewriting the integral at the right-hand side we obtain:

$$A_{mn} = V_0\int_0^b dx\psi_m(x)\psi_n(x) - V_0\int_0^a dx\psi_m(x)\psi_n(x) \tag{17}$$

Using the definition of $\psi_m(x)$ and $\psi_n(x)$ as given in (5) leads to:

$$A_{mn}=V_0\left[b\frac{\sin(m-n)\pi b/L}{(m-n)\pi b/L}-\alpha\frac{\sin(m-n)\pi\alpha/L}{(m-n)\pi\alpha/L}\right]-V_0\left[b\frac{\sin(m+n)\pi b/L}{(m+n)\pi b/L}-\alpha\frac{\sin(m+n)\pi\alpha/L}{(m+n)\pi a/L}\right] \tag{18}$$

It is instructive to introduce the notion of Universal Functions [2]:

$$U(x) = \operatorname{sinc}(x) \tag{19}$$

Therefore, for the "inter-action" terms $A_{mn}(m\neq n)$ we obtain:

$$A_{mn} = V_0\left\{bU\left[(m-n)\frac{\pi}{L}b\right]-aU\left[(m-n)\frac{\pi}{L}a\right]-bU\left[(m+n)\frac{\pi}{L}b\right]+aU\left[(m+n)\frac{\pi}{L}a\right]\right\} \tag{20}$$

Based on the fact that $U(0)=1$, for the "self-action" terms A_{mm} we obtain:

$$A_{mm} = V_0\left\{(b-a)-bU\left(2m\frac{\pi}{L}b\right)+aU\left(2m\frac{\pi}{L}a\right)\right\} \tag{21}$$

4. NUMERICAL RESULTS

We discuss three characteristic cases where we perturb our original canonical system with $V(x)$ which support extends from a to b in the interval $[0,L_x]$. Case 1: We consider $V(x)$ defined on $[a,b]=[0,1/3]$ (Figure 1). Case 2: $[a,b]=[1/3,2/3]$ (Figure 2). Furthermore, we assume that in both cases we have $V(x)=V_0=100$.

We have obtained our numerical results by assigning the above-mentioned values to a and b in (20) and (21). In addition, we have utilized (14) for calculating the perturbed energy $\widetilde{\widetilde{E}}$. Perturbed wave functions $\widetilde{\psi}(x)$ in case 1 are given in Figures 3 and 4, and in case 2 in Figures 5 and 6.

Figures 3 and 5 show that for $\widetilde{\widetilde{E}}$ values less than V_0, the wavefunction $\widetilde{\psi}(x)$ decays exponentially. Figures 4 and 6 illustrate that where $\widetilde{\widetilde{E}}$ is greater than V_0 we have oscillation.

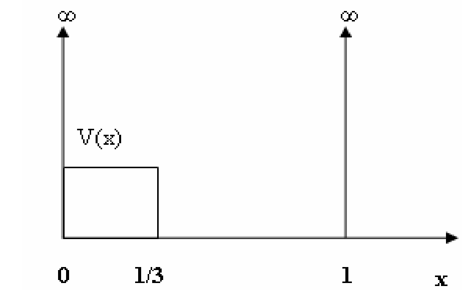


Figure 1: Perturbation function for case 1.

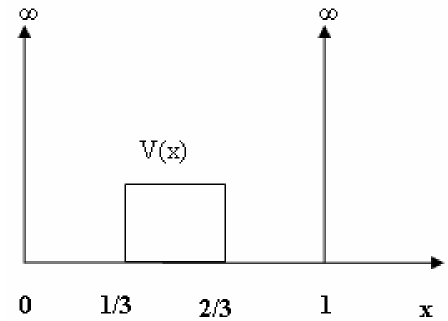


Figure 2: Perturbation function for case 2.

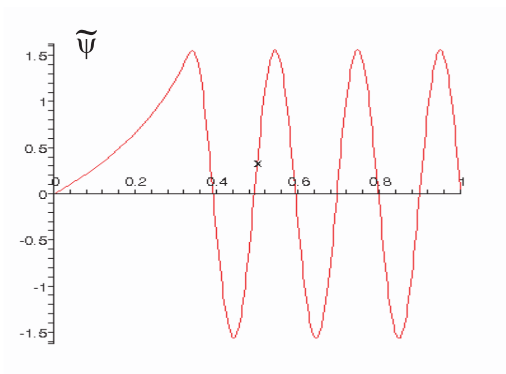


Figure 3: $\widetilde{\psi}(x)$ for $\widetilde{\widetilde{E}} = 96.826622133 < V_0$.

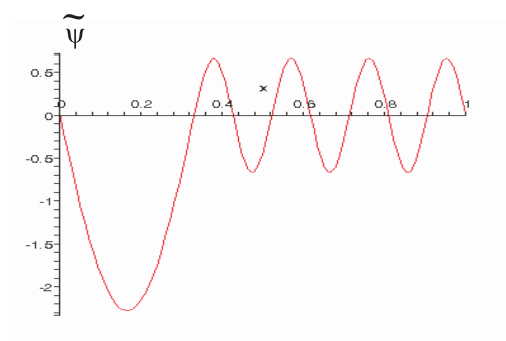


Figure 4: $\widetilde{\psi}(x)$ for $\widetilde{\widetilde{E}} = 109.1793820 > V_0$.

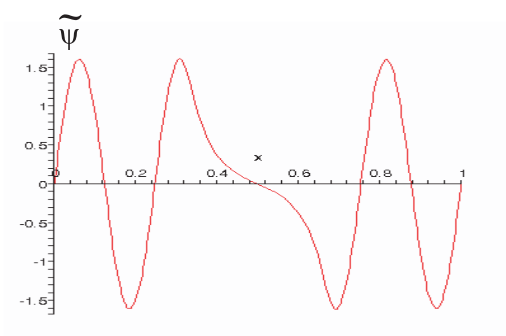


Figure 5: $\widetilde{\psi}(x)$ for $\widetilde{\widetilde{E}} = 98.95719841 < V_0$.

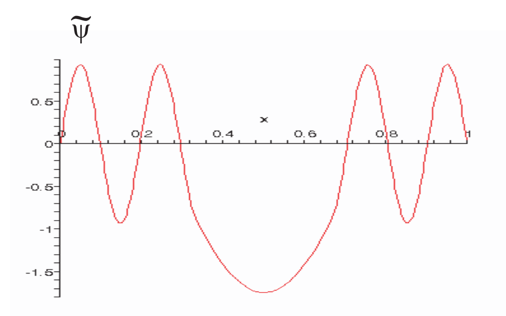


Figure 6: $\widetilde{\psi}(x)$ for $\widetilde{\widetilde{E}} = 137.2751778 > V_0$.

5. EIGENVALUE DISTRIBUTION USING ARBITRARY POTENTIAL FUNCTIONS

Case 3: Our formulation allows assuming arbitrary potential functions $V(x)$ and thus attaining control over the distribution of eigenvalues. For these cases, the upper limit b would be iL_x/N_S, where i refers to the number of added potentials, L_x stands for the length of the infinite potential well and N_S is the number of equal sectors in the interval $[0, L_x]$. The lower limit a is equivalent to $(i-1)L_x/N_S$. Therefore,

$$A_{mn} = \sum_{i=1}^{N_S} V_i \int_{(i-1)L_x/N_S}^{iL_x/N_S} dx \psi_m(x)\psi_n(x) \tag{22}$$

Figure 7 is an example where five arbitrary potentials in the form of $V1$ to $V5$ are taken in the interval $[0, L_x]$. The values used for $V1$ to $V5$ are respectively, 1000, 1500, 2750, 3500 and 1250.

A fifty by fifty matrix was incorporated and the highest number of solution i. e., the 50th solution for perturbed wave function $\widetilde{\psi}_{50}(x)$ is shown in Figure 8. The energy eigenvalue $\widetilde{\widetilde{E}}$ was measured

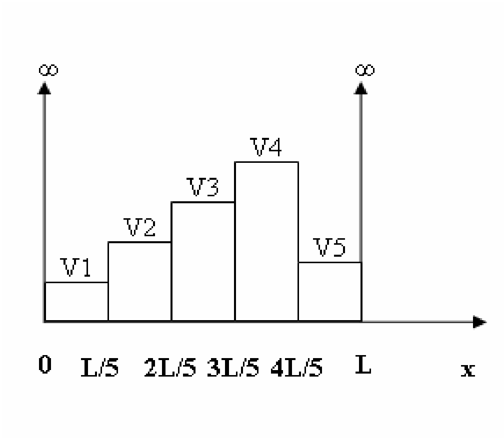


Figure 7: Arbitrary potential functions in $[0, L_x]$.

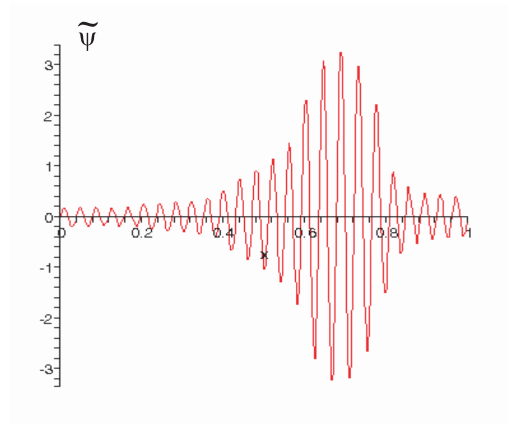


Figure 8: 50th eigensolution of $\widetilde{\psi}_{50}(x)$.

at 5484.814105. Figure 8 shows oscillation for energy values greater than the largest potential. It is also important to note that for higher-order eigensolutions, the region with the highest potential always has the maximum amplitude and contains minimum wave numbers.

6. CONCLUSION

Utilizing our methodology we succeeded in tuning the arbitrary potential function $V(x)$ for controlling the distribution of eigenvalues and thus achieving a certain degree of localization.

ACKNOWLEDGMENT

This work of IA was supported by a research grant from the School of Electrical and Computer Engineering, RMIT University. ARBW is grateful to the Victorian Partnership for Advanced Computing (VPAC) for support.

REFERENCES

1. Ahmed, I. and A. R. Baghai-Wadji, "2D canonical and perturbed quantum potential-well problem: A universal function approch," *Proceedings of ACES, Applied Computational Electromagnetics Society*, Verona, Italy, 19–23 March, 2007.
2. Baghai-Wadji, A. R., "Universal functions in computational electromagnetics," *Proceedings of ACES, Applied Computational Electromagnetics Society*, Monterey, USA, 24–28 March, 2003.

On the Calculation of Polynomially Perturbed Harmonic Oscillators

P. Peidaee and A. R. Baghai-Wadji

School of Electrical and Computer Engineering, RMIT University, Melbourne, Victoria, Australia

Abstract— we propose an easy-to-implement method for the calculation of eigenfunctions of polynomially perturbed higher-order quantum harmonic oscillators. Our method is based upon using the eigensolutions of the unperturbed harmonic oscillator as auxiliary gauging functions. The obtained eigenfunctions of perturbed harmonic oscillators can be used as a complete set of functions for analyzing related boundary value problems. Alternatively, the eigenfunctions can be utilized as the starting point for the calculation of even-order polynomial perturbations.

1. INTRODUCTION

It is known that the eigenvalues and the corresponding eigenfunctions of the quantum harmonic oscillator can be calculated in closed-form using Hermite polynomials [1]. It is also known that the determination of the eigenpairs of linearly perturbed harmonic oscillators can be reduced to the solution of the unperturbed problem [1]. In [2] we introduced a numerical procedure for the calculation of linearly perturbed harmonic oscillator. In this contribution we extend the theory presented in [1] and demonstrate the applicability of our method to more complex problems. Thereby, the key idea is the construction of an auxiliary problem the eigenpairs of which are known to us. As will be discussed shortly, we shall choose as our auxiliary problem the ideal unperturbed harmonic oscillator. As it will also be seen clearly the applicability range of the concept presented here is quite large. More challenging cases will be treated elsewhere.

2. PROPOSED METHODOLOGY

The Auxiliary Problem: The Ordinary Differential Equation ODE for the quantum harmonic oscillator is:

$$-\frac{\hbar^2}{2M}\frac{d^2\psi(x)}{dx^2}+kx^2\psi(x)=E\psi(x) \tag{1}$$

Here, $\hbar$ refers to the reduced Planck's constant, M is the electron mass, and k a parameter specifying the magnitude of the quadratic potential function. We shall consider the eigenpairs $\{\psi_m, E_m\}$ of this problem as our gauging system. Therefore, we will refer to (1) as our auxiliary problem.

The Actual Problem: Our actual (original) problem is a perturbation of the harmonic oscillator by introducing an additive polynomial potential function. Here, we consider general monomials of the form βx^r with r being an integer and β a positive constant specifying the magnitude of the perturbation. The ODE characterizing our actual problem has therefore the form:

$$-\frac{\hbar^2}{2M}\frac{d^2\widetilde{\psi}(x)}{dx^2}+kx^2\widetilde{\psi}(x)+\beta x^r\widetilde{\psi}(x)=\widetilde{E}\widetilde{\psi}(x) \tag{2}$$

Our objective is the solution of (2) and thus the determination of eigenfunctions $\widetilde{\psi}_m$ and the corresponding eigenvalues $\widetilde{E}_m$. To this end we proceed as follows:

Multiply the auxiliary system by $\widetilde{\psi}(x)$ and the actual problem by $\psi(x)$ and substract the resulting "actions" to obtain:

$$-\frac{\hbar^2}{2M}\left(\widetilde{\psi}(x)\frac{d^2}{dx^2}\psi(x)-\psi(x)\frac{d^2}{dx^2}\widetilde{\psi}(x)\right)-\psi(x)\beta x^r\widetilde{\psi}(x)=\left(E-\widetilde{E}\right)\psi(x)\widetilde{\psi}(x) \tag{3}$$

Integrate both sides from $-\infty$ to ∞:

$$-\frac{\hbar^2}{2M}\left(\int_{-\infty}^{\infty}dx\widetilde{\psi}(x)\frac{d^2}{dx^2}\psi(x)-\int_{-\infty}^{\infty}dx\psi(x)\frac{d^2}{dx^2}\widetilde{\psi}(x)\right)-\beta\int_{-\infty}^{\infty}dx\psi(x)x^r\widetilde{\psi}(x)$$

$$=\left(E-\widetilde{E}\right)\int_{-\infty}^{\infty}dx\psi(x)\widetilde{\psi}(x) \tag{4}$$

Introduce Dirac's notation for the inner product to obtain:

$$-\frac{\hbar^2}{2M}\left\{\langle\widetilde{\psi}(x)\left|\frac{d^2}{dx^2}\psi(x)\rangle-\langle\psi(x)\right|\frac{d^2}{dx^2}\widetilde{\psi}(x)\rangle\right\}-\beta\langle\psi(x)|x^r|\widetilde{\psi}(x)\rangle=\left(E-\widetilde{E}\right)\langle\psi(x)|\widetilde{\psi}(x)\rangle \qquad (5)$$

Integration by parts results in:

$$I=\langle\widetilde{\psi}(x)|\frac{d^2}{dx^2}\psi(x)\rangle=\widetilde{\psi}(x)\frac{d}{dx}\psi(x)\Big|_{-\infty}^{+\infty}-\langle\frac{d}{dx}\widetilde{\psi}(x)|\frac{d}{dx}\psi(x)\rangle \qquad (6)$$

Since $\psi(x)$ and $\widetilde{\psi}(x)\in L^2(R)$ (quadratically Lebesgue integrable) the first term at the right-hand side vanishes and we obtain

$$I=-\langle\frac{d}{dx}\widetilde{\psi}(x)|\frac{d}{dx}\psi(x)\rangle \qquad (7)$$

Similarly, we have

$$J=\langle\psi(x)|\frac{d^2}{dx^2}\widetilde{\psi}(x)\rangle=-\langle\frac{d}{dx}\widetilde{\psi}(x)|\frac{d}{dx}\psi(x)\rangle \qquad (8)$$

Because of the equality of I and J the first two terms in the curly brackets in (5) cancel out and we are left with:

$$-\beta\langle\psi(x)|x^r|\widetilde{\psi}(x)\rangle=\left(E-\widetilde{E}\right)\langle\psi(x)|\widetilde{\psi}(x)\rangle \qquad (9)$$

It is known that the "ground state" for the harmonic oscillator has the form:

$$\psi_0(x)=\frac{1}{\sqrt[4]{2\pi}}e^{-\frac{1}{2}x^2} \qquad (10)$$

By applying the annihilation operator $a=\left(\frac{d}{dx}+x\right)$ m-times onto the expression for $\psi_0(x)$ we obtain the expression for m-th "excited state" $\psi_m(x)$ along with its corresponding eigenvalue E_m:

$$\psi_m=\frac{1}{\sqrt[4]{2\pi}}H_m(x)e^{-\frac{1}{2}x^2}\Longleftrightarrow E_m=m\hbar\omega \qquad (11)$$

We substitute $\psi(x)=\psi_m(x)$ and $E=E_m$ into (9):

$$-\beta\langle\psi_m(x)|x^r|\psi(x)\rangle=\left(E_m-\widetilde{E}\right)\langle\psi_m(x)|\psi(x)\rangle \qquad (12)$$

Next we exploit the fact that the infinite set of eigenfunction of $\psi_m(x)$ is complete in the space $L_2(R)$. This property allows us to expand $\widetilde{\psi}(x)$ in terms of $\{\psi_m(x)\}$:

$$\widetilde{\psi}(x)=\sum_{n=1}^{\infty}\alpha_n\psi_n(x) \qquad (13)$$

Substituting (13) into (12), and using the linearity property of the inner-product we obtain:

$$-\beta\sum_{n=1}^{\infty}\alpha_n\langle\psi_m(x)|x^r|\psi_n(x)\rangle=\left(E_m-\widetilde{E}\right)\sum_{n=1}^{\infty}\alpha_n\langle\psi_m(x)|\psi_n(x)\rangle \qquad (14)$$

Using the orthonormality condition $\langle\psi_m(x)|\widetilde{\psi}_n(x)\rangle=\delta_{mn}2^n n!\sqrt{\pi}$ in (14), and introducing A_{mn} as given in (15-a) we obtain:

$$\sum_{n=1}^{\infty}A_{mn}\alpha_n=\frac{\left(\widetilde{E}-m\hbar\omega\right)2^m m!\sqrt{\pi}}{\beta}\alpha_m \qquad (15\text{-a})$$

$$A_{mn}=\langle\psi_m(x)|x^r|\psi_n(x)\rangle \qquad (15\text{-a})$$

In (15-a) we also have substituted the value for E_m as given in (11).

A_{mn} can be solved in closed form [3]:

$$A_{mn} = \begin{cases} 0 & (r-m-n) \text{ odd} \\ \sqrt{\frac{1}{2}m!n!2^{m+n}}\frac{r!}{2^r}\sqrt{\frac{2^{m+n}}{m!n!}}\sum_{P=\max(0,-S)}^{\min(m,n)}\binom{n}{P}\binom{m}{P}\frac{P!}{2^P(S+P)!} & \text{otherwise} \end{cases} \tag{16}$$

with $S = \dfrac{r-m-n}{2}$.

3. NUMERICAL RESULTS

To obtain numerical results we first need to truncate the sum in (15-a) by taking into account N terms only $(n = 1 \dots N)$:

$$\sum_{n=1}^{N} A_{mn}\alpha_n + \sum_{n=1}^{N}\frac{n\hbar\omega 2^n n!\sqrt{\pi}}{\beta}\delta_{mn}\alpha_n = \tilde{\tilde{E}}\sum_{n=1}^{N} 2^n n!\sqrt{\pi}\delta_{mn}\alpha_n \tag{17}$$

Here we have introduced $\tilde{\tilde{E}} = \frac{\tilde{E}}{\beta}$. Note that for every given m (17) represents one equation for the unknown expansion coefficients $\{\alpha_1, \dots, \alpha_N\}$. Letting m to vary from 1 to N we obtain an $N \times N$ matrix, the eigenvalues and the corresponding eigenvectors of which are the desired $\tilde{\tilde{E}}_k$ and the associated $\{\alpha_{lk} | l = 1, \dots, N\}$ with $k = 1, \dots, N$.

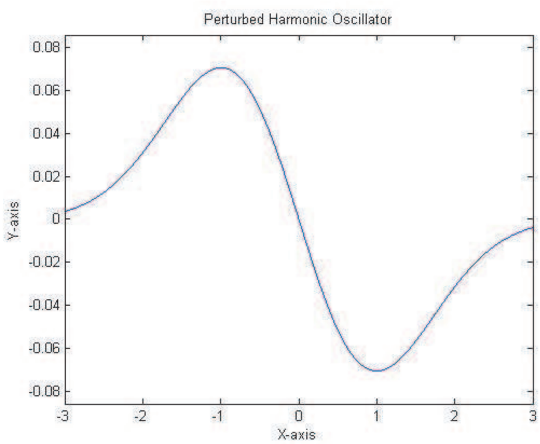


Figure 1: $N = 5$, $r = 3$, for the first eigensolution.

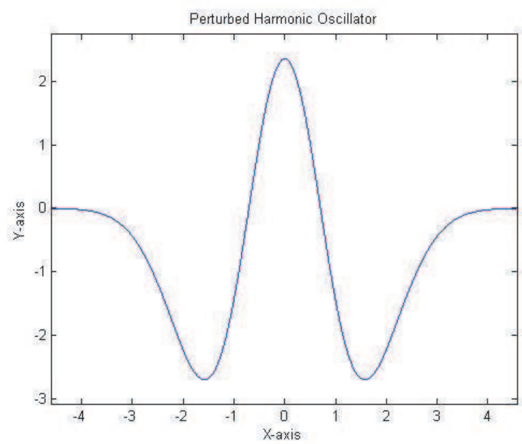


Figure 2: $N = 5$, $r = 4$, for the second eigensolution.

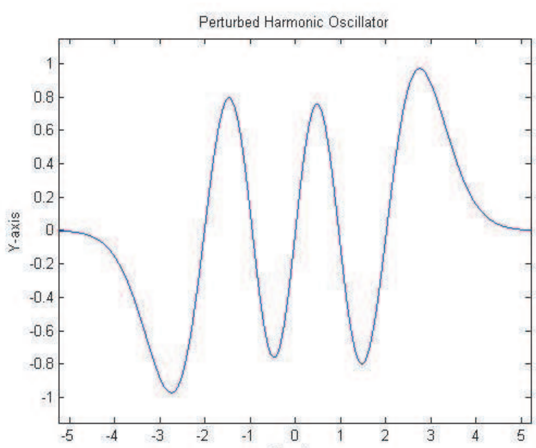


Figure 3: $N = 5$, $r = 5$, for the fifth eigensolution.

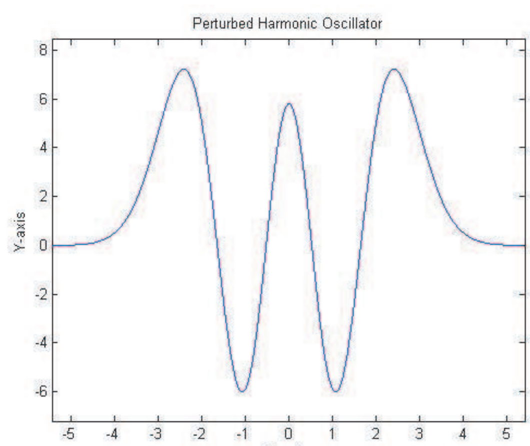


Figure 4: $N = 5$, $r = 8$, for the fourth eigensolution.

4. CONCLUSION

An easy-to-implement method for the determination of eigenpairs of polynomially perturbed harmonic oscillators has been purposed. The method is based on using the unperturbed harmonic oscillator as an auxiliary problem. We have shown that our solution concept leads to an algebraic system for the determination of unknowns of the problem with closed-form formula for the resulting matrix elements.

ACKNOWLEDGMENT

Partial support provided by the Victorian Partnership for Advanced Computing (VPAC), Australia, is gratefully appreciated.

REFERENCES

1. Louisell, W. H., *Quantum Statistical Properties of Radiation*, John Wiley & Sons, 1973.
2. Peidaee, P. and A. R. Baghai-Wadji, "On the calculation of linearly-perturbed harmonic oscillator," *Proceedings of ACES, Applied Computational Electromagnetic Society Conference*, Verona, Italy, 2007.
3. http://mathworld.wolfram.com

On the Determination of Eigenpairs of 2D Positive Differential Operators with Periodic Boundary Conditions

A. Rezaee and A. R. Baghai-Wadji

School of Electrical and Computer Engineering
RMIT University, Victoria 3001, Australia

Abstract— An easy-to-implement and versatile method for computing the behaviour of an electron propagating in 2D periodic quantum structures has been introduced. Numerical results for various parameters of the problem has been presented. The resulting eigenvalues directly lead to the Brillouine dispersion diagrams of the periodic structures studied.

1. INTRODUCTION

We consider the properties of an electron propagating in 2D periodic structure. Even though this problem is important for its own sake, our ultimate goal, however, is the solution of the so-called defect problem. Defects are characterized by localized violation of the otherwise periodic structures. The result will be presented elsewhere. In this contribution, which is an extension of [1] we present a simply-by-inspection and easy-to-implement technique for ideal 2D periodic structures.

2. STATEMENT OF THE PROBLEM

Consider the 2D Schrödinger equation for an electron propagating in a doubly-periodic structure characterized by a doubly-periodic potential function $V_P(x, y)$, [2]:

$$-\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\psi(x,y)+V_P(x,y)\psi(x,y)=E\psi(x,y) \tag{1}$$

Here $\hbar$ is the reduced Planck constant, M the mass of the electron, and E the total energy of the electron.

Our aim is the solution of (1) under the periodicity condition

$$V_P\left(x+P_x, y+P_y\right)=V_P(x,y). \tag{2}$$

2.1. Solution Technique

Floquent theorem suggests the solution ansatz:

$$\psi(x,y)=\psi_P(x,y)e^{ik_1x}e^{ik_2y} \tag{3}$$

Here, $e^{ik_1x}e^{ik_2y}=e^{i\boldsymbol{kr}}$ is a propagation phasing factor characterized by the wave numbers k_1 and k_2 in the x and y directions, respectively. A consequence of (3) is that:

$$\begin{aligned}\psi\left(x+P_x, y+P_y\right) &= \psi_P\left(x+P_x, y+P_y\right)e^{ik_1(x+P_x)}e^{ik_2(y+P_y)}\\ &= \psi_P(x,y)e^{ik_1x}e^{ik_1P_x}e^{ik_2y}e^{ik_2P_y}\\ &= \psi(x,y)e^{ik_1P_x}e^{ik_2P_y}\end{aligned} \tag{4}$$

By substituting (3) into (1) we obtain:

$$\begin{aligned}&-\frac{\hbar^2}{2M}\left\{\left[\frac{d^2\psi_P(x,y)}{dx^2}\right]e^{ik_1x}e^{ik_2y}+2ik_1\left[\frac{d\psi_P(x,y)}{dx}\right]e^{ik_1x}e^{ik_2y}+(ik_1)^2\psi_P(x,y)e^{ik_1x}e^{ik_2y}\right.\\ &\left.+\left[\frac{d^2\psi_P(x,y)}{dy^2}\right]e^{ik_1x}e^{ik_2y}+2ik_2\left[\frac{d\psi_P(x,y)}{dy}\right]e^{ik_1x}e^{ik_2y}+(ik_2)^2\psi_P(x,y)e^{ik_1x}e^{ik_2y}\right\}\\ &+V_P(x,y)\psi_P(x,y)e^{ik_1x}e^{ik_2y}=E\psi_P(x,y)e^{ik_1x}e^{ik_2y}\end{aligned} \tag{5}$$

Since $e^{ik_1x}e^{ik_2y}\neq 0$, we obtain the following Partial Differential Equation (PDE) for $\psi_P(x,y)$:

$$\begin{aligned}&-\frac{\hbar^2}{2M}\left\{\frac{\partial^2\psi_P(x,y)}{\partial x^2}+\frac{\partial^2\psi_P(x,y)}{\partial y^2}+2ik_1\frac{\partial\psi_P(x,y)}{\partial x}+2ik_2\frac{\partial\psi_P(x,y)}{\partial y}-\left(k_1^2+k_2^2\right)\psi_P(x,y)\right\}\\ &+V_P(x,y)\psi_P(x,y)=E\psi_P(x,y)\end{aligned} \tag{6}$$

Since $V_P(x,y)$ is a doubly-periodic function, it can be represented by a 2D Fourier series expansion:

$$V_P(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} V_{m,n} e^{im\frac{2\pi}{P_x}x} e^{in\frac{2\pi}{P_y}y} \tag{7}$$

The Fourier coefficients are:

$$V_{m,n} = \frac{1}{P_x P_y} \int_{-\frac{P_x}{2}}^{\frac{P_x}{2}} dx \int_{-\frac{P_y}{2}}^{\frac{P_y}{2}} dy e^{-im\frac{2\pi}{P_x}x} e^{-in\frac{2\pi}{P_y}y} V_P(x,y) = \langle e^{im\frac{2\pi}{P_x}x} e^{in\frac{2\pi}{P_y}y} | V_P(x,y) \rangle \tag{8}$$

where the inner product $\langle . | . \rangle$ of the functions $f(x,y)$ and $g(x,y)$ is defined by

$$\langle f(x,y) | g(x,y) \rangle = \frac{1}{P_x p_y} \int_{-\frac{P_x}{2}}^{\frac{P_x}{2}} dx \int_{-\frac{P_y}{2}}^{\frac{P_y}{2}} dy f^*(x,y) g(x,y). \tag{9}$$

Here, the star indicates complex conjugation. Next, we multiply (6) by $\exp\left(-im\frac{2\pi}{P_x}x\right) \exp\left(-in\frac{2\pi}{P_y}y\right)$ and integrate $\frac{1}{P_x P_y} \int_{-\frac{P_x}{2}}^{\frac{P_x}{2}} dx \int_{-\frac{P_y}{2}}^{\frac{P_y}{2}} dy \cdots$:

$$\begin{aligned}
&\frac{1}{P_x P_y} \int_{-\frac{P_x}{2}}^{\frac{P_x}{2}} dx \int_{-\frac{P_y}{2}}^{\frac{P_y}{2}} dy \frac{d^2\psi_P(x,y)}{dx^2} e^{-im\frac{2\pi}{P_x}x} e^{-in\frac{2\pi}{P_y}y} \\
&+\frac{1}{P_x P_y} \int_{-\frac{P_x}{2}}^{\frac{P_x}{2}} dx \int_{-\frac{P_y}{2}}^{\frac{P_y}{2}} dy \frac{d^2\psi_P(x,y)}{dy^2} e^{-im\frac{2\pi}{P_x}x} e^{-in\frac{2\pi}{P_y}y} \\
&+2ik_1 \frac{1}{P_x P_y} \int_{-\frac{P_x}{2}}^{\frac{P_x}{2}} dx \int_{-\frac{P_y}{2}}^{\frac{P_y}{2}} dy \frac{d\psi_P(x,y)}{dx} e^{-im\frac{2\pi}{P_x}x} e^{-in\frac{2\pi}{P_y}y} \\
&+2ik_2 \frac{1}{P_x P_y} \int_{-\frac{P_x}{2}}^{\frac{P_x}{2}} dx \int_{-\frac{P_y}{2}}^{\frac{P_y}{2}} dy \frac{d\psi_P(x,y)}{dy} e^{-im\frac{2\pi}{P_x}x} e^{-in\frac{2\pi}{P_y}y} \\
&-\left(k_1^2 + k_2^2\right) \frac{1}{P_x P_y} \int_{-\frac{P_x}{2}}^{\frac{P_x}{2}} dx \int_{-\frac{P_y}{2}}^{\frac{P_y}{2}} dy \psi_P(x,y) e^{-im\frac{2\pi}{P_x}x} e^{-in\frac{2\pi}{P_y}y} \Bigg\} \\
&+\frac{1}{P_x P_y} \int_{-\frac{P_x}{2}}^{\frac{P_x}{2}} dx \int_{-\frac{P_y}{2}}^{\frac{P_y}{2}} dy V(x,y) \psi_P(x,y) e^{-im\frac{2\pi}{P_x}x} e^{-in\frac{2\pi}{P_y}y} \\
&= E \frac{1}{P_x P_y} \int_{-\frac{P_x}{2}}^{\frac{P_x}{2}} dx \int_{-\frac{P_y}{2}}^{\frac{P_y}{2}} dy \psi_P(x,y) e^{-im\frac{2\pi}{P_x}x} e^{-in\frac{2\pi}{P_y}y}
\end{aligned} \tag{10}$$

Since $\psi_P(x,y)$ is also a doubly-periodic function we can write

$$\psi_P(x,y) = \sum_{\hat{m}=-\infty}^{\infty}\sum_{\hat{n}=-\infty}^{\infty}\alpha_{\hat{m}\hat{n}}e^{i\hat{m}\frac{2\pi}{P_x}x}e^{i\hat{n}\frac{2\pi}{P_y}y} \tag{11}$$

$$\alpha_{\hat{m}\hat{n}} = \frac{1}{P_xP_y}\int_{-\frac{P_x}{2}}^{\frac{P_x}{2}}dx\int_{-\frac{P_y}{2}}^{\frac{P_y}{2}}dy\psi_P(x,y)e^{\left(-i\hat{m}\frac{2\pi}{P_x}x\right)}e^{\left(-i\hat{n}\frac{2\pi}{P_y}y\right)} \tag{12}$$

Or, more compactly,

$$\alpha_{\hat{m}\hat{n}} = \langle e^{i\hat{m}\frac{2\pi}{P_x}x}e^{i\hat{n}\frac{2\pi}{P_y}y}\Big|\psi_P(x,y)\rangle. \tag{13}$$

Furthermore, we obtain

$$\langle e^{im\frac{2\pi}{p_x}x}e^{in\frac{2\pi}{p_y}y}\Big|\frac{d\psi_P(x,y)}{dx}\rangle = im\frac{2\pi}{P_x}\alpha_{mn} \tag{14}$$

$$\langle e^{im\frac{2\pi}{p_x}x}e^{in\frac{2\pi}{p_y}y}\Big|\frac{d\psi_P(x,y)}{dy}\rangle = in\frac{2\pi}{P_y}\alpha_{mn} \tag{15}$$

$$\langle e^{im\frac{2\pi}{p_x}x}e^{in\frac{2\pi}{p_y}y}\Big|\frac{d^2\psi_P(x,y)}{dx^2}\rangle = \left(im\frac{2\pi}{P_x}\right)^2\alpha_{mn} \tag{16}$$

$$\langle e^{im\frac{2\pi}{p_x}x}e^{in\frac{2\pi}{p_y}y}\Big|\frac{d^2\psi_P(x,y)}{dy^2}\rangle = \left(in\frac{2\pi}{P_y}\right)^2\alpha_{mn} \tag{17}$$

$$\langle e^{im\frac{2\pi}{p_x}x}e^{in\frac{2\pi}{p_y}y}\Big|V(x,y)\psi_P(x,y)\rangle = \sum_{\hat{m}=-\infty}^{\infty}\sum_{\hat{n}=-\infty}^{\infty}\alpha_{\hat{m}\hat{n}}V_{m-\hat{m},n-\hat{n}} \tag{18}$$

Substituting these results into (10) we obtain:

$$-\frac{\hbar^2}{2M}\sum_{\hat{m}=-\infty}^{\infty}\sum_{\hat{n}=-\infty}^{\infty}\left[\left(k_1+\hat{m}\frac{2\pi}{P_x}\right)^2+\left(k_2+\hat{n}\frac{2\pi}{P_y}\right)^2\right]\delta_{\hat{m}-m}\delta_{\hat{n}-n}\alpha_{\hat{m}\hat{n}}+\sum_{\hat{m}=-\infty}^{\infty}\sum_{\hat{n}=-\infty}^{\infty}V_{m-\hat{m},n-\hat{n}}\alpha_{\hat{m}\hat{n}}$$
$$= E\sum_{\hat{m}=-\infty}^{\infty}\sum_{\hat{n}=-\infty}^{\infty}\delta_{\hat{m}-m}\delta_{\hat{n}-n}\alpha_{\hat{m}\hat{n}} \tag{19}$$

It is instructive to introduce the non-dimensional variables $\hat{k}_1 = \frac{k_1}{\frac{2\pi}{P_x}}$ and $\hat{k}_2 = \frac{k_2}{\frac{2\pi}{P_y}}$.

For numerical calculations the series expansions in (19) need to be truncated by letting $\hat{m}$ and $\hat{n}$ vary from, say, $-N$ to N. Then for every given pair (m, n), Eq. (19) establishes a relationship between $(2N+1)\times(2N+1)$ unknown coefficients $\alpha_{\hat{m}\hat{n}}$. Varying m, n from $-N$ to N we generate a $(2N+1)\times(2N+1)$ matrix eigenvalue problem for $\alpha_{\hat{m}\hat{n}}$ and the corresponding eigenvalues.

3. NUMERICAL RESULTS

In the following we assume the simplest possible function for $V_P(x,y)$, namely a box function:

$$V_P(x,y) = \begin{cases} V_0 & (x,y)\in\left[-\frac{L_x}{2},\frac{L_x}{2}\right]\times\left[-\frac{L_y}{2},\frac{L_y}{2}\right]\subset\left[-\frac{P_x}{2},\frac{P_x}{2}\right]\times\left[-\frac{P_y}{2},\frac{P_y}{2}\right] \\ 0 & \text{elsewhere}\end{cases} \tag{20}$$

with this assumption we have

$$A_{mn} = \left[\left(\hat{k}_1+\hat{m}\right)^2+\gamma\left(\hat{k}_2+\hat{n}\right)^2\right]\delta_{\hat{m}-m}\delta_{\hat{n}-n}+\beta\alpha_1\alpha_2\sin c\left[(m-\hat{m})\,\pi\alpha_1\right]\sin c\left[(n-\hat{n})\,\pi\alpha_2\right] \tag{21}$$

with

$$0<\alpha_1=\frac{L_x}{P_x}<1, 0<\alpha_2=\frac{L_y}{P_y}<1, \quad \beta=\frac{V_0}{\frac{\hbar^2}{2M}\left(\frac{2\pi}{P_x}\right)^2}, \quad \gamma=\left(\frac{P_x}{P_y}\right)^2.$$

In the following figures the first few branches have been shown for $\alpha_1 = 0.4$, $\alpha_2 = 0.6$, $\beta = 1$, $\gamma = 1$, and $k_2 = 0.2$.

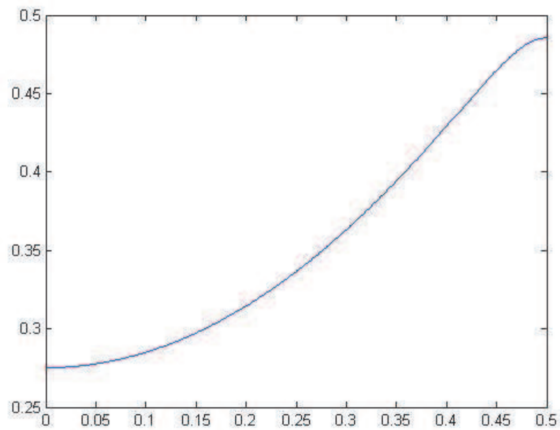


Figure 1: The first branch in the Brillouin diagram for $\alpha_1 = 0.4$, $\alpha_2 = 0.6$, $\beta = 1$, $\gamma = 1$, $k_2 = 0.2$.

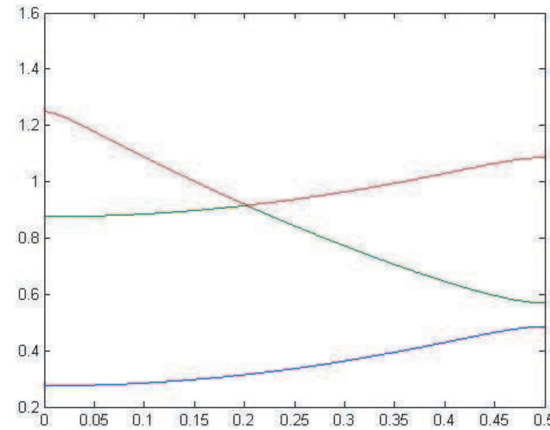


Figure 2: The first three branches in the Brillouin diagram for $\alpha_1 = 0.4$, $\alpha_2 = 0.6$, $\beta = 1$, $\gamma = 1$, $k_2 = 0.2$.

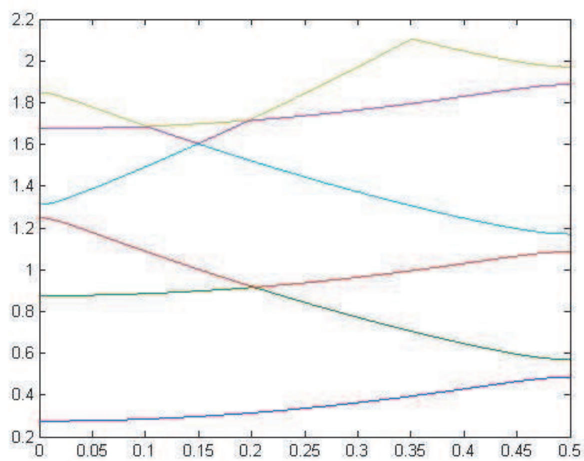


Figure 3: The first six branches in the Brillouin diagram for $\alpha_1 = 0.4$, $\alpha_2 = 0.6$, $\beta = 1$, $\gamma = 1$, $k_2 = 0.2$.

4. CONCLUSION

In the above calculations we assumed a box function $B(x, y)$ for the potential function $V_P(x, y) = V_0 B(x, y)$. We can show that more complicated functions for $V_P(x, y)$ such as $B(x, y) \otimes B(x, y) = B^2(x, y)$ (convolution of $B(x, y)$ function with itself), or, $B^2(x, y) \otimes B(x, y) = B^3(x, y)$, or, convolutions to any order (B-spline functions) lead to closed form expressions for the so-called Universal Functions from which the matrix elements A_{mn} can be calculated. Furthermore, we can construct Universal Functions associated with B-spline wavelets.

ACKNOWLEDGMENT

Partial support provided by the Victorian Partnership for Advanced Computing (VPAC), Australia, is gratefully appreciated.

REFERENCES

1. Rezaee, A. and A. R. Baghai-Wadji, "On the determination of the eigenpairs of 1d positive differential operators with periodic boundary conditions," *Proceedings ACES*, Verona, Italy, 2007.
2. Louisell, W. H., *Quantum Statistical Properties of Radiation*, John Wiley & Sons, 1973.

A Theoretical Study of Particle Plasmons for Single and Multiple Metallic Nanospheres

Ruey-Lin Chern[1], **Chih-Yu Kuo**[2], **and Chien-Cheng Chang**[2,1]

[1]Institute of Applied Mechanics & Center of Nanostorage Research, Taiwan University
Taipei 106, Taiwan, China
[2]Division of Mechanics, Research Center for Applied Sciences, Academia Sinica
Taipei 115, Taiwan, China

Abstract— In this paper, particle plasmons for metallic nanopspheres are investigated using the method of multiple scattering. The extinction, absorption, and scattering efficiencies for single and multiple nanopspheres are calculated for silver in the visible and ultraviolet regimes, including where the electronic interband transition occurs. Both the real metal effect and the Drude model are considered. In particular, multiple peaks, redshift, and broadening of the extinction spectrum for single and multiple nanopspheres are identified and explained. Moreover, we investigate the field patterns and the associated enhancements at the extinction peaks due to particle plasmon resonance, and the extinction dip due to interband transition. Physical significance is discussed regarding the effects of some important parameters such as particle size, number of particles, interparticle spacing, and particle orientation.

The study of optical properties for metal spheres dates back to 1970s [1–3]. One of the most important features is appearance of particle plasmons (PP), which are resonant modes of the metal sphere coupled with the incident light. PP differ from surface plasmon polaritons (SPP) on planar surfaces [4] in two aspects. First, they are a special type of localized surface plasmons (LSP) associated with metal particles, yielding significant field enhancement near or on the particle surfaces. Their resonant behaviors are strongly affected by geometry of the particles rather than intrinsic properties of the metal such as the bulk plasma frequency ω_p or surface plasma frequency ω_{sp} for planar interface. Second, there is no need of coupling techniques [5] such as prism, grating, aperture, or lattice coupling for providing extra momenta to excite surface plasmons on particles. In fact, the field patterns for PP resonance can be decomposed as multipolar oscillations and eddy currents [6], both of which are fundamental modes attributed to spherical particles. For metal particles of size from micrometers down to tens or hundreds nanometers, PP resonance would occur in the infrared, visible, or even ultraviolet regime.

In the past, PP of single metal nanospheres has been studied intensively using the Mie theory. The points of interest include an infinite number of resonant modes [7], plasmon resonance broadening [8], and influence of the size and shape [9]. Recently, there is an increasing number of works devoted to the study of interaction between the incident field and many nanospheres. Additional new physics has been addressed regarding interaction between particles and effect of the polarization of incident field [10, 11], electromagnetic energy transport and optical pulse propagation [12, 13], and surface plasmon dispersion relations [14]. However, the nature of resonant modes and their dependence on various geometric factors were apparently not identified and classified. On the other hand, the real metal effect that takes account of the size correction received relatively little attention. The present study is aimed to investigate PP for single and multiple metallic nanospheres with particular emphasis on the characteristics of resonant modes and the associated field enhancements. The dielectric function of the metal nanospheres is taken from experiments [15], and correction of the surface effects on the nanoscales [16] is also considered.

The multiple scattering method built upon the Mie theory and the addition theorem for spherical harmonics is developed to compute the field distributions. The method is semi-analytical and accurate, and is much more efficient in computing the fields for many particles than the commonly used finite-difference time-domain method for the particular geometry of spheres. These special advantages enable us to examine closely the field distributions of many as well as single metal spheres, which help uncover some features not observed before. Basic features of PP are identified through the absorption, scattering, and extinction efficiencies, for the enhanced field at resonance may radiate a substantial amount of energy (due to electric charge oscillation) to the far field, and produces peaks in the spectra of these efficiencies [3]. In general, these peaks are located at much lower frequencies than ω_p of the metal or ω_{sp} for planar interface. In particular, we shall identify PP as electric dipole, quadrupole, octopole, and so forth, and show how the characteristics of resonant PP modes, such as redshift and broadening of the resonant peaks and the associated field enhancements, alter due to coupling of neighboring resonance, with increasing the number of particles and the interparticle spacing as well as the orientation of the particle chain. A comparison of the Drude model with real dielectric function is also made to illustrate the discrepancies of results between the ideal (Drude model) and real metals (experimental constants with or without size correction).

It will be shown that the number of extinction peaks and their redshift and broadening are dependent upon the size, and in particular strongly upon the number and the interparticle spacing of the nanoparticles. For a chain of nanoparticles, we distinguish three cases: normal incidence (with the chain oriented in the polarization direction of the incident wave); parallel incidence (with the chain oriented normally to the polarization); and oblique incidence. A larger nanoparticle can sustain more resonant modes because it allows more degrees of freedom to accommodate the field pattern. At normal incidence, by decreasing the spacing between two nanospheres the number of extinction peaks increases from one to two or even three. Moreover, the number of resonant peaks increases also with increasing the number of nanospheres. In contrast, at parallel incidence there are always only two peaks, irrespective of the number of particles in the chain. At normal incidence, the field enhancement associated with the first peak is greatest in the midgap of a chain of nanospheres. The field enhancement associated with other peaks may occur in other gaps of the particle chain. As the number of spheres is increased, the field enhancement becomes less significant. In the case of oblique incidences, the greatest enhancement does not necessarily occur in the midgap of the chain, but depends on the orientation angle θ. It was found for a chain of two nanospheres that the field enhancement in the midgap decreases *log-linearly* in the power 1.38 with increasing the interparticle spacing. Moreover, in a range of θ near normal incidence, the field enhancement of the first peak in the midgap of the chain was found to follow the *sine formula*: $a \sin\theta + b$ with a fixed a and varying b which decreases with increasing the number of particles.

REFERENCES

1. Ruppin, R., "Optical properties of small metal spheres," *Phys. Rev. B*, Vol. 11, No. 8, 2871–2876, 1975.
2. Bohren, C. and D. Huffman, *Absorption and Scattering by Small Particles*, Wiley, New York, 1983.
3. Kreibig, U. and M. Vollmer, *Optical Properties of Metal Clusters*, Springer, 1995.
4. Raether, H., *Surface Plasmons on Smooth and Rough Surfaces and on Grating*, Springer-Verlag, 1988.
5. Peyghambarian, N., S. Koch, and A. Mysyrowicz, *Introduction to Semiconductor Optics*, Prentice-Hall, 1994.
6. Kreibig, U., B. Schmitz, and H. D. Breuer, "Separation of plasmon-polariton modes of small metal particles," *Physical Review B*, Vol. 36, No. 9, 5027–5030, 1987.
7. Fujimoto, F. and K. Komaki, "Plasma oscillations excited by a fast electron in a metallic particle," *J. Phys. Soc. Japan*, Vol. 25, 1679, 1968.
8. Kraus, W. and G. Schatz, "Plasmon resonance broadening in small metal particles," *J. Chem. Phys.*, Vol. 79, 6130, 1983.
9. Kelly, K. L., E. Coronado, L. L. Zhao, and G. C. Schatz, "The optical properties of metal nanoparticles: the influence of size, shape, and dielectric environment," *J. Phys. Chem. B*, Vol. 107, 668, 2003.
10. Su, K. H., Q. H. Wei, X. Zhang, J. J. Mock, D. R. Smith, and S. Schultz, "Interparticle coupling effects on plasmon resonances of nanogold particles," *Nano Lett.*, Vol. 3, 1087, 2003.
11. Rechberger, W., A. Hohenau, A. Leitner, J. R. Krenn, B. Lamprecht, and F. Aussenegg, "Optical properties of two interacting gold nanoparticles," *Opt. Commun.*, Vol. 220, 137, 2003.
12. Quinten, M., A. Leitner, J. R. Krenn, and F. R. Aussenegg, "Electromagnetic energy transport via linear chains of silver nanoparticles," *Opt. Lett.*, Vol. 23, 1331, 1998.
13. Maier, S. A., P. G. Kik, and H. A. Atwater, "Optical pulse propagation in metal nanoparticle chain waveguides," *Phys. Rev. B*, Vol. 67, 205, 402, 2003.
14. Park, S. Y. and D. Stroud, "Surface-plasmon dispersion relations in chains of metallic nanoparticles: An exact quasistatic calculation," *Phys. Rev. B*, Vol. 69, No. 12, 125, 418, 2004.
15. Johnson, P. B. and R. W. Christy, "Optical constants of the noble metals," *Physical Review B*, Vol. 6, No. 12, 4370, 1972.

16. Kawata, S., *Near-Field Optics and Surface Plasmon Polaritons*, Springer, 2001.

Spectrum Compression of a Short Pulse from a Central Obstructed Circular Aperture in the Far-field

Pin Han
Institute of Precision Engineering, Chung Hsing University
250, Kuo Kuang Road, Taichung 402, Taiwan, China

Abstract— The spectrum changes of a Gaussian pulse in the far-field is studied with Fresnel diffraction integral when it is incident on a circular aperture with a central obstruction. It can be shown that the central obstruction will cause diffracted spectrum compressed, compared with that of the normal circular aperture without a central obstruction. This novel effect could be useful for optical engineering or optical communications.

1. INTRODUCTION

Recently there is increasing interests in so called "aperture dispersion", that is, the spectral changes of a short pulse resulting from the aperture diffraction. It also includes the red or blue shift of the spectrum maximum or the distortion of the incident pulse's spectrum [1–5]. It is well known that an aperture with central obstruction can reduce the diffraction pattern width for constant intensity light wave; hence it is used to enhance the resolving power of image-forming systems [6, 7]. In this work, we study the spectral characteristics of a time-dependent Gaussian pulse when it is incident on a circular aperture width a central obstruction, which is not investigated before. It can be shown that the central obstruction of the aperture will have compression effect for the diffracted pulse spectrum.

2. THEORY

Consider an incoming wave with a spectral scalar field $U'(p',\omega)$ incident on an annular aperture, which has outer radius a and inner radius εa, $(0 \le \varepsilon \le 1)$, as show in Figs. 1.(a)(b). The diffraction field $U(p,\omega)$ on the observation plane can be obtained by using the Fresnel diffraction integral [8]:

$$U(p,\omega) = \frac{1}{j\lambda}\iint\limits_{\Sigma'} U'(p',\omega)\frac{\exp(j\omega r/c)}{r}\chi(\theta)\mathrm{d}\sigma', \tag{1}$$

where $\chi(\theta)$ is the obliquity factor, λ is the wavelength, ω is the angular frequency, c is the velocity of the light wave, and r is the distance from point $p'(x',y',0)$ on the aperture plane to point $p(x,y,z)$ on the observation (image) plane. The Cartesian (cylindrical) coordinate systems used for the incident (aperture) plane and the observation (image) plane are $x'o'y'(\rho',\phi')$ and $xoy(\rho,\phi)$ coordinate planes, respectively. Σ' is the aperture function and $\mathrm{d}\sigma'$ is the integration to it, as shown in Fig. 1(a). The factor $1/j\lambda$ in front of the integral of Eq. (1) is important because $1/\lambda(=\omega/2\pi c)$ includes the ω dependence and the diffraction spectrum analysis will not be correct if it is omitted [9, 10]. It is also noted that Eq. (1) is usually used for a monochromatic incident field, $U'(p',t) = U'(p',\omega)e^{-j\omega t}$, with a single frequency ω and the constant complex amplitude $U'(p',\omega)$, but it is also applicable for a broad-band incident pulse [1–5], which can be superposed by monochromatic fields via the Fourier integral [11].

The aperture function for the annular aperture in Fig. 1(b) can be written as:

$$g(\rho') = circ\left(\frac{\rho'}{a}\right) - circ\left(\frac{\rho'}{\varepsilon a}\right), \tag{2}$$

where $circ(\rho')$ is the circle function and is defined as: $circ(\rho'/2a) = 1$ for $0 \le \rho' \le a$ and $circ(\rho'/2a) = 0$ for $\rho' > a$. Let $u(p',t) = \exp\left[-1/2(t/\tau)^2 + j\omega_0 t\right]$ is the incident time-dependent Gaussian pulse where ω_0 is the pulse central frequency and τ is its duration time. Its spectral amplitude $U'(p',\omega) = \tau/\sqrt{2\pi}\cdot\exp\{-1/2[(\omega-\omega_0)\tau]^2\}$. Substituting $U'(p',\omega)$ into (1) and using the far-field approximation, the diffraction field $U(p,\omega)$ is

$$U(p,\omega) = \frac{1}{j\lambda R}\exp[jk\cdot R]\cdot U'(p',\omega)\cdot F(g(\rho')), \tag{3}$$

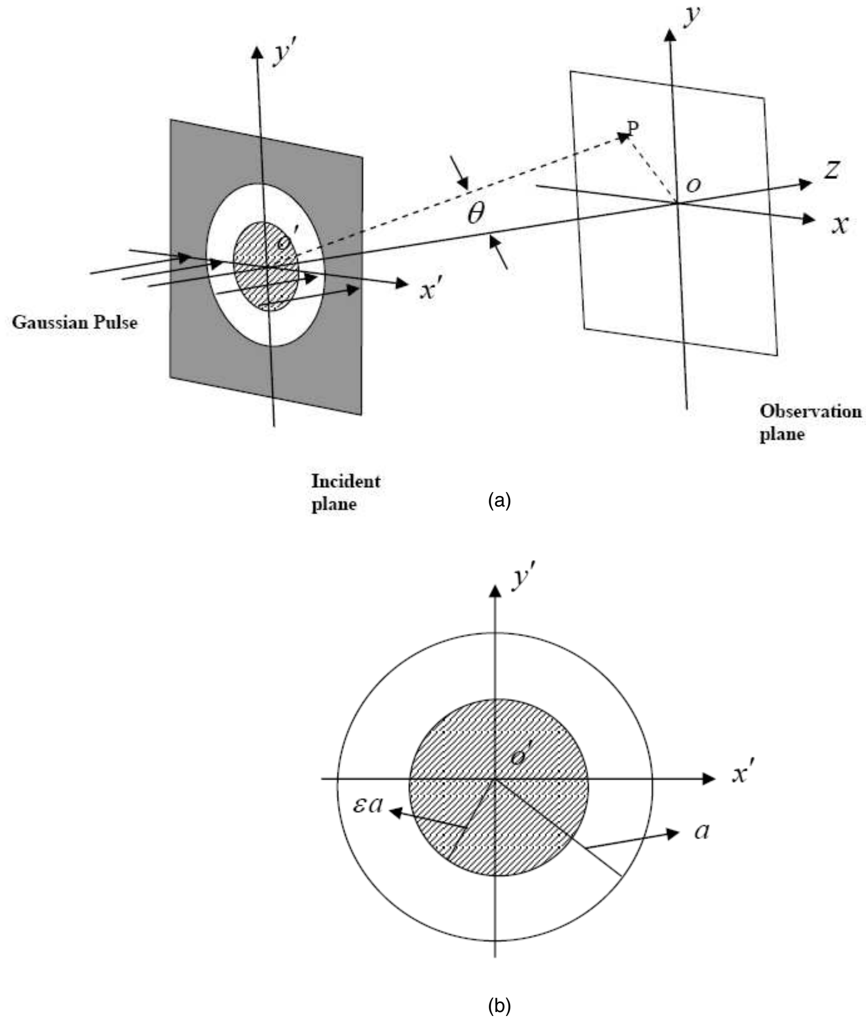


Figure 1: (a) The geometry of the set up. (b) The structure of the circular aperture with a central obstruction.

where R is the distance for $\overline{o'p}$, $k = \omega/c = 2\pi/\lambda$ and the last term $F(g(\rho'))$ is the Fourier transform of the aperture function $g(\rho')$ with the spatial frequency $f_\rho = \rho/\lambda R$ as The Fourier Bessel transform of this aperture function $F(g(\rho'))$ is [6]:

$$F(g(\rho')) = 2\pi a^2 \left[\frac{J_1(2\pi a f_\rho)}{2\pi a f_\rho} - \varepsilon^2 \left(\frac{J_1(\varepsilon 2\pi a f_\rho)}{\varepsilon 2\pi a f_\rho} \right) \right], \tag{4}$$

where the $J_1(x)$ is the Bessel function of the first kind and of the first order and f_ρ is the spatial frequency.

With substitution Eq. (4) into Eq. (3) and using the relations $1/\lambda = \omega/2\pi c$, $f_\rho = \omega\rho/2\pi c R = \omega \sin\theta/2\pi c$, the diffraction field $U(p, \omega)$, can be rewritten as

$$\begin{aligned} U(\theta,\omega) \;=\; & \frac{1}{jR}\left(\frac{\omega}{2\pi c}\right)\exp[j(k\cdot R)]\left(\frac{\tau}{\sqrt{2\pi}}\right)\exp\left\{-\frac{1}{2}\left[(\omega-\omega_0)\,\tau\right]^2\right\} \\ & \cdot\left\{(2\pi a^2)\left[\frac{J_1(a\omega\sin\theta/c)}{a\omega\sin\theta/c} - \varepsilon^2\left(\frac{J_1(\varepsilon a\omega\sin\theta/c)}{\varepsilon a\omega\sin\theta/c}\right)\right]\right\}, \end{aligned} \tag{5}$$

where $\sin\theta = \rho/R$ and θ is the angle between $\overline{o'o}$ and $\overline{o'p}$ [see Fig. 1(a)].

With the formula $I(\theta,\omega) = |U(\theta,\omega)|^2 = U(\theta,\omega)\cdot U(\theta,\omega)^*$, the spectral intensity $I(\theta,\omega)$ of the diffraction field, can be obtained as:

$$I(\theta,\omega) = A\cdot\omega^2\exp[-(\omega-\omega_0)\tau]^2\cdot\left\{\frac{J_1(a\omega\sin\theta/c)}{a\omega\sin\theta/c} - \varepsilon^2\left(\frac{J_1(\varepsilon a\omega\sin\theta/c)}{\varepsilon a\omega\sin\theta/c}\right)\right\}^2, \tag{6}$$

where $A = \tau^2 a^4 / 2\pi c^2 R^2$.

3. NUMERICAL RESULTS

Figs. 2(a) and (b) show the numerical results of $I(\theta, \omega)$ according to Eq. (6) with two different values of ε. It is noted that, if we define the first two zeros on the right and left as the width of the $I(\theta, \omega)$, the spectrum width is compressed to a narrower extent for an annular aperture ($\varepsilon = 0.5$ in (a) and $\varepsilon = 0.8$ in (b)) than that of a normal circular aperture ($\varepsilon = 0$) and that the wider of the central obstruction, the narrower of the diffracted spectrum, as illustrated in the figure. The compression effect can also be found by comparing FWHM (Full Width at Half Maximum) of $I(\theta, \omega)$ for the central obstructed and the normal circular aperture. The numerical calculations give that FWHM for a normal circular aperture ($\varepsilon = 0$) is $0.73\omega_0$ and it is $0.67\omega_0$ for ($\varepsilon = 0.5$) and it is $0.60\omega_0$ for ($\varepsilon = 0.8$), respectively.

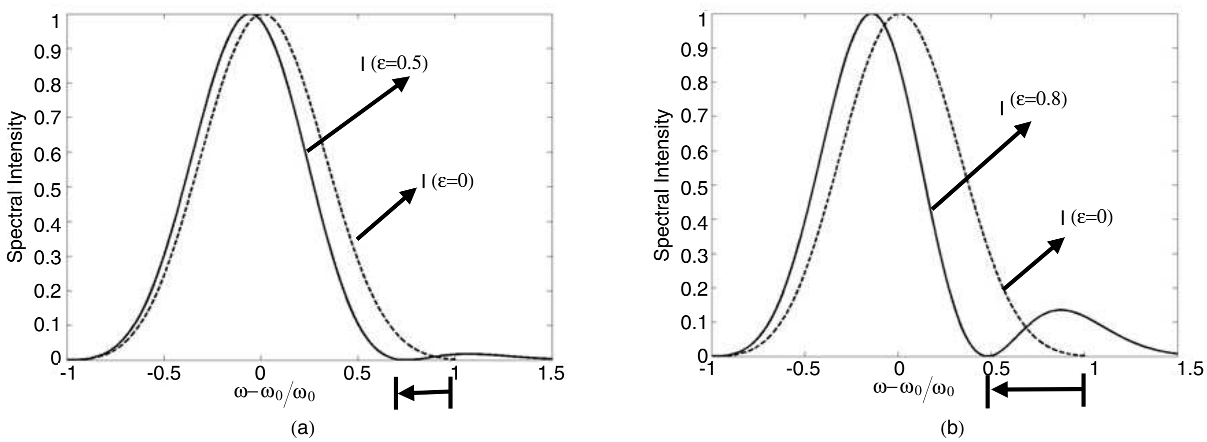


Figure 2: (a) The spectral intensity for $\varepsilon = 0, 0.5$. (b) The spectral intensity for $\varepsilon = 0, 0.8$.

This novel effect can be useful when the diffracted spectral intensity compression is needed. However, it is at the cost of less energy received on the observation plane, because for a point p on the observation plane with angle θ. the total energy received is $E(p) = \int_{-\infty}^{\infty} I(\theta, \omega) d\omega$, based on parseval's theorem [12].

4. CONCLUSIONS

In this work, the aperture dispersion of a short Gaussian pulse is studied when it is diffracted by a circular aperture with a central obstruction. The novel effect called "spectrum compression" can be found under such a situation, which is considered to be useful when the diffracted spectral contribution needs to have a narrower bandwidth. Also this effect occurred in spectral domain for a short pulse can be compared with its counterpart for a constant intensity light wave when a narrower diffraction pattern is wanted in spatial domain [6, 7].

Besides its research interests, this novel effect may have important applications on the optical engineering fields like short pulse lasers or optical communications where the bandwidth of the pulse needs to be compressed or controlled.

ACKNOWLEDGMENT

This study was supported by the National Chung Hsing University Taiwan, R.O.C. The authors also like to thank the colleagues for their thoughtful comments and useful suggestions. This work was also supported by the NSC of R.O.C. under the contract nos. NSC 95-2221-E-005-116.

REFERENCES

1. Han, P., "Far-field spectral intensity characteristics of a time-dependent gaussian pulse from two types of apodized slits," *Jpn. J. Appl. Phys.*, Vol. 43, 3386–3393, 2004.
2. Ponomarenko, S. A. and E. Wolf, "Spectral anomalies in a fraunhofer diffraction pattern," *Opt. Let.*, Vol. 27, 1211–1213, 2002.

3. Han, P. and H. E. Hwang, "Far-field spectral intensity characteristics of a time-dependent gaussian pulse from two types of apodized slits," *Optik*, Vol. 115, 499–506, 2004.
4. Han, P., G. H. Yang, and H. E. Hwang, "Far-field diffraction characteristics of a gaussian pulse incident on a sinusoidal amplitude grating," *J. Opt. A: Pure Appl. Opt.*, Vol. 5, 442, 2003.
5. Hwang, H. E., G. H. Yang, and P. Han, "Near-field diffraction characteristics of a time dependent gaussian-shaped pulsed beam from a circular aperture," *Opt. Eng.*, Vol. 42, 686–695, 2003.
6. Born, M. and E. Wolf, *Principles of Optics*, Cambridge University Press, New York, 7th ed., 461, 1999.
7. Iizuka, K., *Engineering Optics*, Kyoritsu, Tokyo, 2nd ed., Chap. 3, 78, 1983.
8. Hecht, E., *Optics*, Addison Wesley, New York, 4th ed., 511, 2002.
9. Sereda, L., A. Ferrari, and M. Bertolotti, "Diffraction of a time gaussian-shaped pulsed plane wave from a slit," *Pure Appl. Opt.*, Vol. 5, 349–353, 1996.
10. Liu, Z. Y. and D. Y. Fan, "Diffraction characteristics of an ultra-short pulsed beam in the far field," *Pure Appl. Opt.*, Vol. 6, L43–47, 1997.
11. Goodman, J. W., *Introduction to Fourier Optics*, Mcgraw-Hill, New York, 2nd ed., 53, 1996.
12. Hecht, E., *Optics*, Addison Wesley, New York, 4th ed., 544, 2002.

Optimization of the Coupling of the Surface Plasmon Re-emitted from Silver Nanorods

Sheng Chung Chen[1], **Zheng Yu Lin**[2], **Dong Lin Li**[2], **Huai Yi Xie**[2], **Jun You Iv**[2]
Kuo Pin Chiu[2], **and Din Ping Tsai**[2]

[1]Department of Electrical Engineering, Far East University, Tainan, Taiwan, China
[2]Department of Physics, Center of Nanostorage Research
Taiwan University, Taipei, Taiwan, China

Abstract— If a polarized optical wave is incident from the top of a silver nanorod array, the surface plasmons will be generated and surround the nanorods at the bottom of the array. As they shift a little bit from the bottom, the localized effect of the silver nanorod will decrease, and the surface plasmons can couple together if the spacing between the plasmons is small enough. The coupled surface plasmons look like being re-emmitted by the silver nanorod array. The point of this research is to study how to optimize the re-emitted surface plasmon, and investigate the availability of the applications.

The interaction of the silver nanoparticles with optical waves is one of attractive issues in surface plasmons. Under the illumination of optical light, the free electrons inside the silver nanoparticle will be driven by the electromagnetic filed of the light. The collective motion of the free electrons leads to generate the surface plasmon on the nanoparticle. On the other point of view, the silver nanoparticle can localize the electromagnetic fields on its surface. Especially, the localized electromagnetic field has an intensity which is larger than the incident wave. However, the enhanced

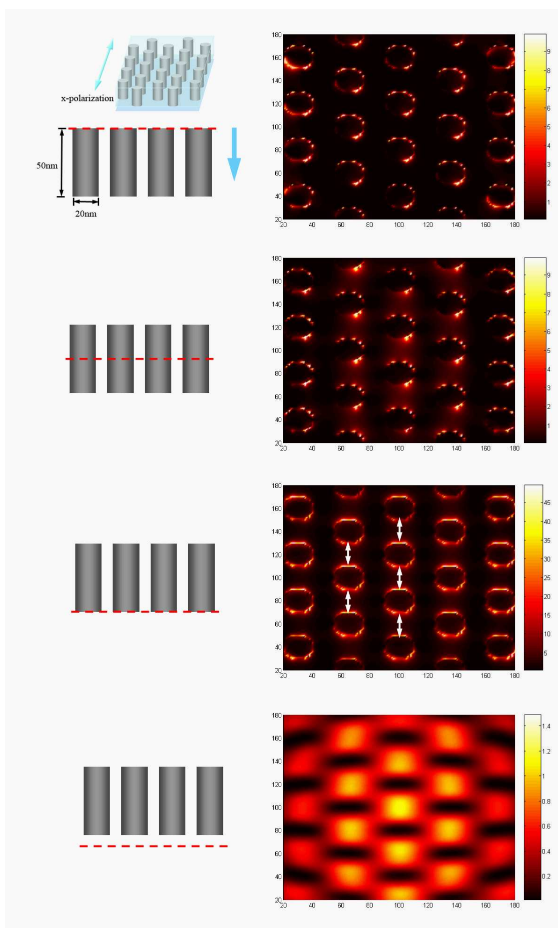


Figure 1: The re-emitted surface plasmons from a hexagolly arranged silver nanorod array.

surface plasmon is an evanescent wave, that is, it can not radiate from the nanopaticle, and its intensity decay significantly as it leaves away from the surface of the particle. Under such situation, one can not use the enhanced surface plasmon without approaching it.

If the nanoparticles are replaced by a compact nanorod array, it seems that the distribution of the surface plasmon will be more regular and controllable. Consider an array of silver nanorods which are arranged hexagonally, if a polarized optical wave is incident from the top of the array, the surface plasmons will be generated and surround the nanorods from top to bottom of the nanorods. As they shift a little bit from the bottom, the localized effect of the silver nanorod will decrease, and the surface plasmons can couple together if the spacing between the plasmons is small enough, as shown in Figure 1. The coupled surface plasmons look like being re-emmitted by the silver nanorod array.

The coupled surface plasmons are very useful in nanoimaging, nanodetector, signal processor and so forth. This coupling is influenced by many factors: wavelength, polarization, size of nanorod, arrangement of nanorod array and so on. The point of this research is to numerically study how to optimize the re-emitted surface plasmon, and investigate the availability of the applications.

This research support by the Ministry of Economic Affairs of Taiwan, R. O. C. through grant number 95-EC-17-A-08-S1-0006.

REFERENCES

1. Raether, H., *Sufrace Plasmons on Smooth and Rough Surfaces and on Gratings*, Springer-Verlag, Berling, 1988.
2. Sullivan, D. M., "Electromagnetics simulation using the FDTD method," *IEEE Press*, New York, 2000.

Improvement of PAE in Doherty Amplifier Using Dual Bias Control and PBG Structure

Hyoungjun Kim and Chulhun Seo
University of Soongsil, Korea

Abstract— This research improved the PAE(Power Added Efficiency) of a Doherty amplifier using dual bias control and PBG structure. A PBG(Photonic BandGap) structure was used to implement on output matching circuit and dual bias control was applied to improve the efficiency of a Doherty amplifier at a low input level by applying it to a carrier amplifier. A Doherty amplifier using the proposed structure improved PAE by 9% and 5 dBc of IMD_3 (3rd Inter-Modulation Distortion) compared to that of a conventional power amplifier.

1. INTRODUCTION

The significance of high output amplifiers in wireless communication has been emphasized due to the rapid distribution of mobile communication systems. The current high power amplifiers (HPA) require high power added efficiency and linearity [1]. In a solution for the linearity of conventional power amplifiers, various methods, such as back-off, feedback, predistortion, feedforward, and PBG, have been used. Among these methods, feedforward and predistortion methods have exhibited the disadvantage of requiring additional elements. Thus, a PBG based method has been used to implement linearization [2, 3]. Doherty amplifier exhibits problems in linearity although it improves PAE [4, 5].

This paper implemented a DC voltage control for drain and gate and PBG structure for the output matching circuit of a Doherty amplifier. The circuit has been based on a dual bias control under the conditions that a class-A or class-AB carrier amplifier had certain degradations in linearity due to saturation when a class-B or class-C peaking amplifier was operated. It is possible to improve the problem of PAE and IMD_3 in a power amplifier and characteristics occurring in non-linearity, simultaneously. In addition, there is a possibility that it may improve performance by more than 30% on all bandwidth for input signals.

2. THEORY AND DESIGN OF DUAL BIAS CIRCUITS

PAE can be expressed as Equation (1)

$$PAE = \frac{(RF_{\text{output\_power}}) - (RF_{\text{input\_power}})}{(V_{gs} \times I_{gs}) + (V_{ds} \times I_{ds})} \tag{1}$$

A method that controls bias voltages relative to the scale of input RF signals in order to improve PAE. There are many kinds of improving linearity of HPA as follow : The voltage of drain or gate can be controlled and both voltages can be controlled, simultaneously [6, 7]. This paper has been used a dual bias control circuit that controlled both drain and gate voltages in which an envelope detector using an AD8313 Analog Device with an excellent linearity and temperature insensitivity.

As shown in Figure 1(a), (b) although the output voltage was produced according to the scale of input RF signals through an envelope detector, the voltage was adjusted through OP-Amps as illustrated in Figure 1(c). As shown in Figure 2, according to the increase in input power, a bias control circuit was fabricated in this research using a directional coupler due to the decrease in the final gate and drain DC value from a class-AB to class-C voltage and class-B to class-C.

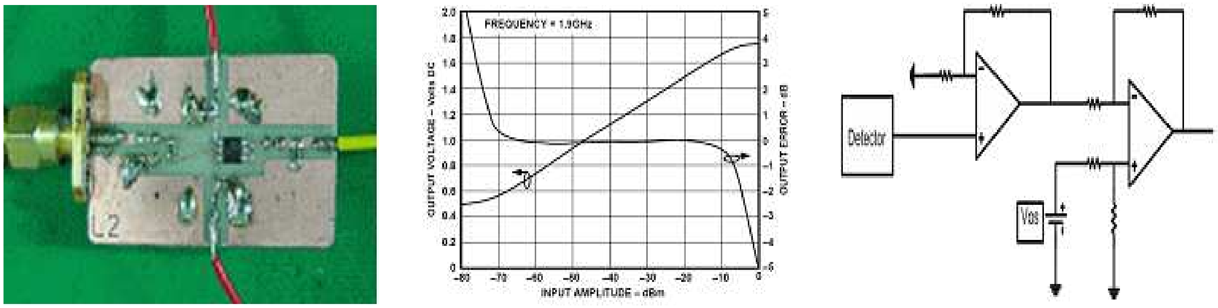
3. THEORY AND DESIGN OF PBG

Using a PBG structure, the forming of a stop band could be estimated based on the frequency that corresponded to 2Λ in the Bragg grating principle. It was possible to form a stop band at a desired point based on this PBG structure. The grating phase Λ can be noted as Equation (2)

$$\Lambda = \lambda_g/2 \tag{2}$$

where λ_g is the wavelength of the wave induced from a microstrip line structure and calculated using the effective permittivity and center frequency of a desired stop band as expressed in Equation (3).

$$\lambda_g(f) = \frac{\upsilon_p(f)}{f} = \frac{c}{f\sqrt{\mu_r \varepsilon_{r,eff}(f)}} \tag{3}$$



(a) (b) (c)

Figure 1: Envelope detector with AD8313. (a) Configuration, (b) Response and (c) Circuit diagram for dual bias circuit.

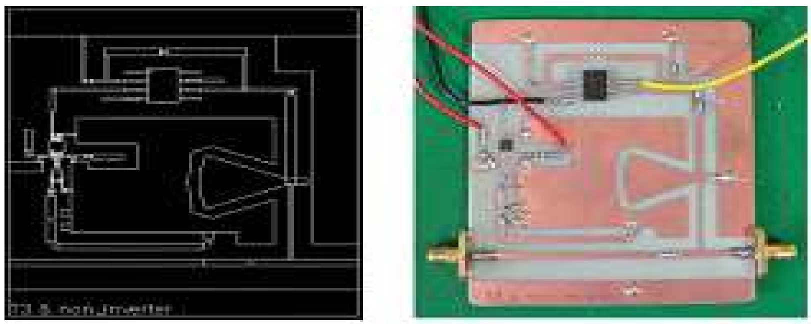


Figure 2: Dual bias circuit with directional coupler.

where $\varepsilon_{r,\,eff}(f)$ is the effective refraction index of the center frequency of a stop band in a microstrip structure. HFSS 8.0 by Ansoft was used as a simulation tool in order to design the PBG.

As illustrated in Figure 3, the PBG was designed to obtain a minimum signal decreasing at a center frequency of 2.14 GHz that was exhibited as $S_{21} = -1.2$ dB. In addition, it was designed to obtain $S_{21} = -39.4$ dB levels at a secondary harmonic frequency of 4.28 GHz in order to decrease a harmonic frequency that significantly affects nonlinearity.

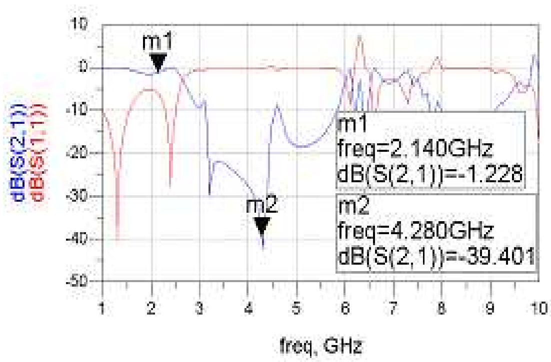
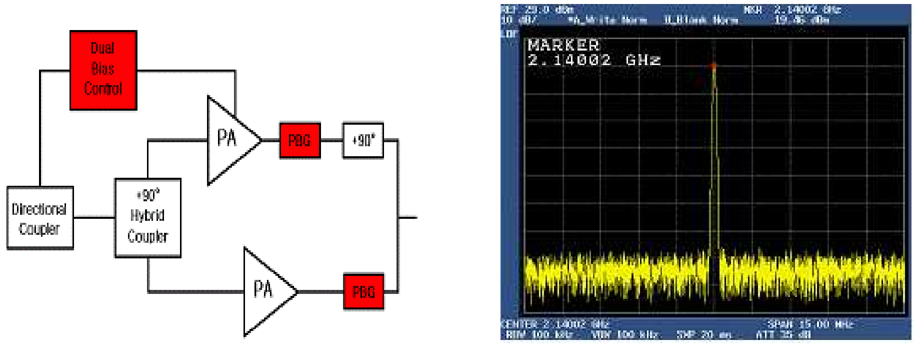


Figure 3: Result of the PBG simulation.

4. DESIGN OF A DOHERTY AMPLIFIER USING A DUAL BIAS CIRCUIT AND PBG STRUCTURE

A Teflon board with a permittivity of 3.2 was used in this research. In addition, ATF34143 of Agilent has been used in HPA. A loadpull simulation using ADS2005 was applied to determine an output matching point. Then, a Doherty amplifier has been designed using output matching according to this output matching point [8]. Figure 4(a) illustrates the implementation of an output matching circuit including the PBG in an offset-line when the output matching circuit was produced by determining the loadpull matching point of the power amplifier. Based on these processes, a Doherty amplifier was designed to using the PBG in order to improve linearity. At first, a class-AB power amplifier was designed as a reference amplifier in order to compare it with this research.

As shown in Figure 4(b), an output of 19.45 dBm was obtained from the 1-tone measurement in this reference amplifier in which the power efficiency and IMD_3 characteristics were 27.12% and −27.45 dBc, respectively.



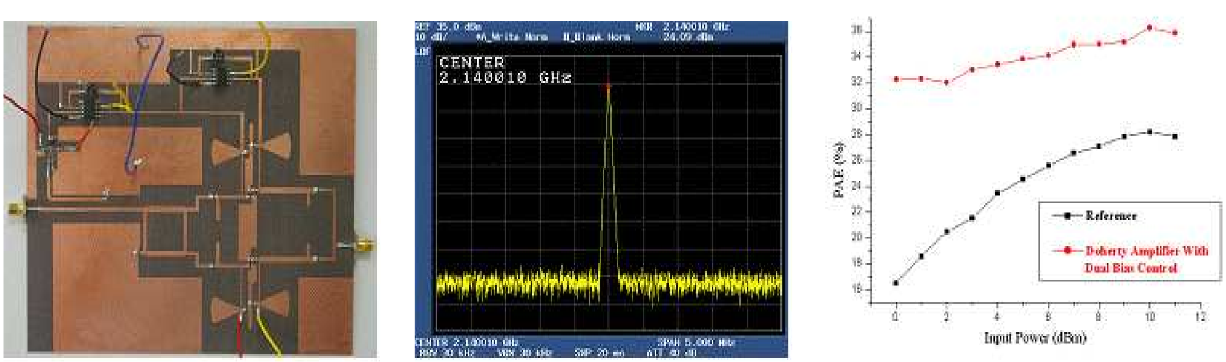
(a) (b)

Figure 4: (a) Structure of a Doherty amplifier using the proposed dual bias Control and PBG structure and (b) Output power of the reference power amplifier (class-AB).

This paper proposes a structure that improves the PAE of a Doherty amplifier by applying dual bias control designed using OP-Amps to control the voltage of an envelope detector based on a Doherty amplifier that uses the previously proposed PBG. Figure 5(a) illustrates the fabricated Doherty amplifier using a dual bias circuit and PBG structure.

DC voltages were generated through an envelope detector according to each input level and were applied to a power amplifier with different amplification rates through OP-Amps in order to configure the bias of gate and drain. In addition, it improved linearity by removing the 2nd harmonic using an additional PBG structure as an offset-line. As shown in Figure 5(b), the Doherty amplifier using the proposed dual bias control and PBG structure exhibited an output of 24.09 dBm, efficiency of 36.28%, and IMD_3 characteristic of −32.47 dBc. In addition, it was evident that if an class-AB was used as a reference amplifier at an input of 10dBm as illustrated in Figure 5(c), the Doherty amplifier using the proposed dual bias control and PBG structure showed a roughly 9% increase in output compared to that of the reference power amplifier.

In addition, this research achieved the flatness in PAE for all input power level. Because gate and drain voltage could be changed by controlling dual bias at the bias point of a carrier amplifier according to the input power, it was possible to obtain more than 30% efficiency at a low input power level compared to three other power amplifiers from the operation point of an actual amplifier. Table 1 show the voltage variation in each gate and drain bias according to the input power level with a value from 0 to 10 dBm. Table 2 notes the comparison between output power, PAE and IMD_3 for each power amplifier.



(a) (b) (c)

Figure 5: (a) Configuration of a Doherty amplifier using a dual bias circuit and PBG structure, (b) Output power of the power amplifier using the proposed configuration and (c) Input power vs PAE.

Table 1: Bias voltage variation according to the input power level.

	Bias Voltage	Input power level
Gate	$0\,\mathrm{V} \sim 0.5\,\mathrm{V}$	$0 \sim 10\,\mathrm{dBm}$
Drain	$2\,\mathrm{V} \sim 4\,\mathrm{V}$	

Table 2: Comparison of measured data for each power amplifier.

	Reference (Class-AB)	Doherty (Classical)	Doherty (With PBG)	Doherty (Dual Bias and PBG)
Output power	19.45	21.14	22.24	24.09
PAE (%)	27.12	30.18	32.85	36.28
IMD_3 (dBc)	−27.45	−28.24	−31.84	−32.47

5. CONCLUSION

In order to increase the power efficiency of a power amplifier, this paper proposed an operational power control method for an amplifier to control gate and drain voltage according to the scale of input signals. This was achieved by employing a proposed dual bias control and structure that improved linearity by applying a PBG structure played the role of a wideband stop band filter at the output port of a power amplifier. Output of the amplifier designed by the proposed method exhibited more improved power efficiency than that of the class-AB, classical Doherty amplifier, and Doherty amplifier using a PBG structure. In addition, it was more than 30% efficiency at a low input power. Furthermore, power efficiency increased approximately 9% according to an input power of 10 dBm compared to the proposed class-AB based structure. The IMD_3 characteristic improved approximately 5 dBc. Moreover, it was evident that the designed structure showed more than a 30% increase in efficiency for flatness and high characteristics for all input power level. In the future, this method will be applied to a power amplifier with high output power.

REFERENCES

1. Cripps, S. C., *RF Power Amplifiers for wireless Communications*, Artech House, 1999.
2. Kim, T. S. and C. H. Seo, "A novel photonic bandgap structure for lowpass filter of wide stopband," *IEEE Microwave Guided Wave Lett.*, January 2000.
3. Rumsey, I., P. M. Melinda, and P. K. Kelly, "Photonic bandgap structures used as filter in microstrip circuits," *IEEE Microwave Guided Wave Lett.*, Vol. 8, 336–338, 1998.
4. Doherty, W. H., "A new high efficiency power amplifier for modulated waves," *Proceeding of the IRE*, Vol. 24, No. 9, 1163–1182, 1936.
5. Upton, D. M., P. R. Maloney, et al., "A microwave doherty amplifier," *Applied Microwaves and RF*, Winter 1994.
6. Hanington, G., "High-efficiency power amplifier using dynamic power-supply voltage for CDMA applications," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, 1471–1476, Aug. 1999.
7. Kenington, P. B., *High-linearity RF Amplifier Design*, Artech House, 2000.
8. Iwamoto, M., "An extended Doherty amplifier with high efficiency over a wide power range," *IEEE Trans. Microwave Theory Tech.*, Vol. 49, 2472–2479, Dec. 2001.

Very Low Power Single-ended Cross-coupled Oscillator in CMOS Technology

Jiho Ryu and Sangwook Nam

School of Electrical Engineering and INMC, Seoul National University, Korea

Abstract— A low power single-ended cross-coupled oscillator (SCO) topology is introduced and the minimum oscillation current is derived from a one port model. The proposed SCO has been fabricated using 0.18 μm CMOS technology. Measurement shows that the phase noise is −103 dBc/Hz at 100 kHz offset from 440 MHz oscillation while dissipating 154 μA of current from a 1.5 V supply.

1. INTRODUCTION

Having been aggressively scaled down, CMOS technology becomes an attractive choice to implement low-cost integrated wireless transceiver systems. And a low power design is also an important issue for RF transceiver designers considering long operating time.

Most of the previously reported papers about a voltage-controlled oscillator (VCO) using CMOS process have described the differential $-G_m$ oscillator topology [1]. However, in case of the single-ended transceiver with a monopole antenna, a differential topology is not recommended because of its unnecessary complexity and an additional balun. This balun is usually implemented with passive elements that increase chip size and power loss.

In this letter, a single-ended cross-coupled oscillator (SCO) topology is introduced that is composed of a complementary NMOS and PMOS transistor pair. The PMOS of the SCO operates as an active buffer which constructs a positive feedback path and also increases transconductance. Moreover, since the inductor of the LC resonator provides DC bias path, the SCO does not require an additional circuit for bias. Therefore, the proposed SCO, which is simple and has a complementary structure, is suitable for low noise and low current oscillator design.

2. PROPOSED OSCILLATOR

Figure 1(a) shows the proposed single-ended cross-coupled oscillator (SCO) structure. Since the oscillation frequency and the resonance frequency of the LC resonator are almost same, the small-signal equivalent circuit to calculate the negative input impedance of the SCO at the oscillation frequency can be represented without the LC resonator as shown in Fig. 1(b). Here g_{mn} and g_{mp} represent the transconductance of the MOS transistor M_1 and M_2, respectively. Note that the

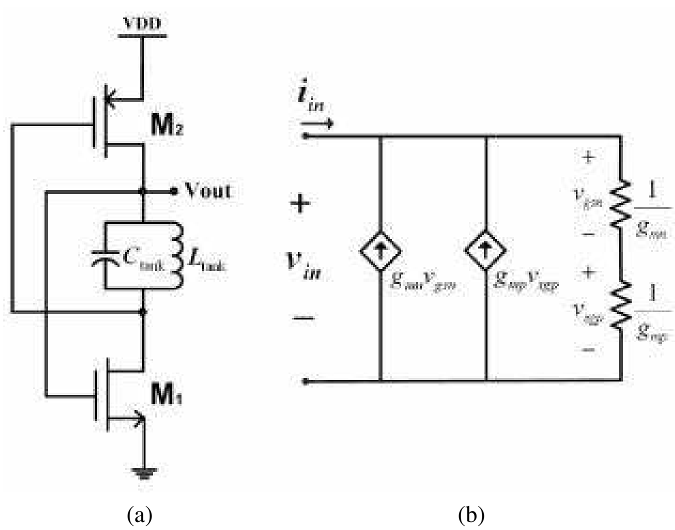


Figure 1: (a) Proposed SCO, (b) Equivalent circuit of the SCO without LC resonator.

parasitic capacitors of M_1 and M_2 are ignored for simplicity. An expression for the current that flows into the circuit can be written as

$$i_{in} = \frac{v_{in}}{\frac{1}{g_{mn}} + \frac{1}{g_{mp}}} - g_{mn}v_{gsn} - g_{mp}v_{sgp}. \tag{1}$$

If it is assumed that the transconductances of the both transistors (M_1 and M_2) are made identical, then $g_{mn} = g_{mp} = g_m$, and the input impedance is derived from (1) as

$$Z_{in} = \frac{v_{in}}{i_{in}} = -\frac{2}{g_m}. \tag{2}$$

From (2), the small-signal equivalent model for Fig. 1(a) is obtained as shown in Fig. 2. In this circuit, the necessary condition for oscillation is

$$g_m > \frac{2}{R_p} \tag{3}$$

where R_p is the equivalent parallel resistance of the LC resonator. This condition is exactly the same as the conventional differential $-G_m$ oscillator [2]. Thus, by using the proposed SCO, the minimum oscillation current can be cut in half, because two amplifiers (PMOS & NMOS) in a positive feedback path share the same DC current.

Colpitts oscillator is one of the most widely used oscillators of single-ended type. This has three passive feedback elements (two capacitors and one inductor), and the oscillation frequency is determined by series combination of two capacitors and one inductor. On the other hand, the oscillation frequency of the SCO determined by one capacitor, so that the capacitor size of the SCO can be smaller than a quarter total capacitor size of Colpitts oscillator when the same inductor is used. Aforementioned characteristics can be said not only to the SCO, but also to the differential $-G_m$ oscillator [2].

Furthermore, the voltage swing of the LC resonator in SCO can be larger than supply voltage, since the output voltage can be lower than ground level by the LC resonator inserted. This improves the phase noise performance of the oscillator [3].

Therefore, the proposed single-ended cross-coupled oscillator can oscillate at ultra low current bias level and have good phase noise performance when a high-Q inductor is used [4]. These are caused by the fact that the complementary structure provides high transconductance ($g_{m\_total} = g_{mn} + g_{mp}$).

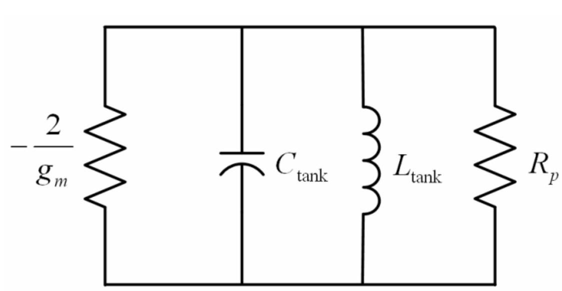


Figure 2: One port equivalent model of the SCO.

3. MEASUREMENT RESULT

The proposed low current SCO was fabricated in UMC 0.18 µm process targeting the RF transceiver of the wireless capsule endoscope application. The carrier frequency (440 MHz) of the oscillator is determined as the minimum loss frequency point considering the antenna loss and the loss property of human body, etc. At 440 MHz, a relatively low frequency, the LC resonator of the SCO needs a large inductor (40 nH), together with high-Q factor. Thus, the inductor of the LC resonator is implemented off-chip using Murata LQW15A series. In order to measure the SCO biased with a current source, an output buffer has been adopted. As shown in Fig. 3, the measured phase noise of the SCO with current of 156 mA at 1.5 V supply is −103 dBc/Hz at 100 kHz offset from the 440 MHz carrier. The calculated figure of merits (FOM) [5] of the oscillator is −182.2 dBc/Hz.

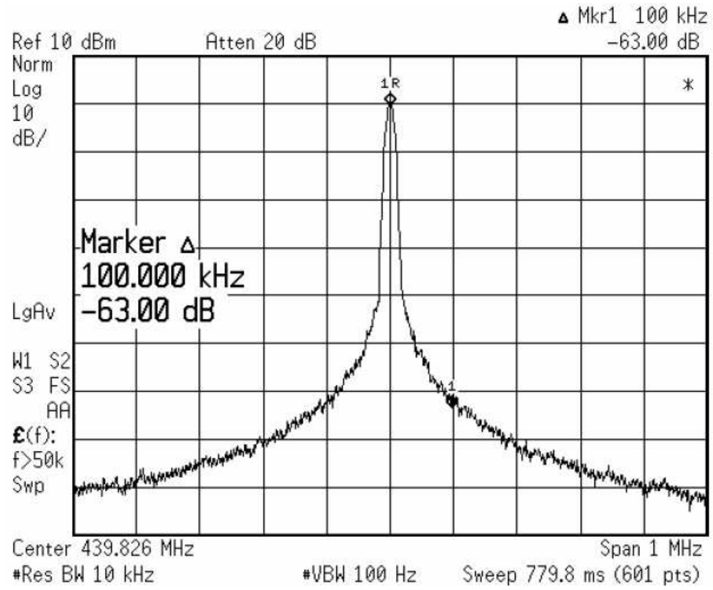


Figure 3: Output spectrum of the SCO(Center 439.826 MHz; RBW 10 kHz; VBW 100 Hz; SPAN 1 MHz; Attenuation 10 dB; RL 0 dBm; Mkr −0.22 dBm, 439.826 MHz).

4. CONCLUSIONS

A novel single-ended cross-coupled oscillator topology is proposed and the minimum oscillation current condition is analyzed based on the one-port small-signal equivalent model. Considering the simple and complementary structure, the proposed SCO is adequate for designing an ultra low power and low noise oscillator, since it can generate a relatively high negative conductance at low current. The measured phase noise of 440 MHz at 100 kHz offset is −103 dBc/Hz while dissipating only 154 μA of current from a 1.5 V supply.

ACKNOWLEDGMENT

This research has been supported by the Intelligent Microsystem Center (IMC; http://www.microsystem.re.kr), which carriers out one of the 21st century's Frontier R&D Projects sponsored by the Korea Ministry Of Commerce, Industry and Energy.

REFERENCES

1. Razavi, B., "RF microelectronics," *Upper Saddle River*, 227–231, Prentice-Hall, NJ, 1998.
2. John, R. and P. Calvin, *Radio Frequency Integrated Circuit Design*, 262–265, Artech house, Inc., 2003.
3. Ham, D. and A. Hajimiri, "Concepts and methods in optimizatiion of integrated LC VCOs," *IEEE Journal of Solid-State Circuits*, Vol. 36, No. 6, 896–909, June 2001.
4. Vendelin, G. D., *Microwave Circuit Design Using Linear and Nonlinear Techniques*, 434–442, Wiley-Interscience Publication, 1990.
5. Bhattacharjee, J., "A 5.8 GHz fully integrated low power low phase noise CMOS LC VCO for WLAN application," *IEEE MTT-S*, Vol. 1, 585–588, June 2002.

Cascode Feedback Amplifier Combined with Resonant Matching for UWB System

Pei-Zong Rao, **Yuh-Chuan Cheng**, **Ching-Piau Liang**, **and Shyh-Jong Chung**

Department of Communication Engineering, Chiao Tung University
Hsinchu, Taiwan, China

Abstract— This paper shows the design of a broadband single-end low noise amplifier for UWB system with TSMC 0.18 μm CMOS processes. The proposed amplifier achieves broadband input matching by cascode feedback topology with parallel resonate cavity and the bandwidth is from 3.1 GHz to 10.1 GHz. The simulation results are as follows: 12 ± 1 dB of gain, the input/output return loss are greater than 10 dB, minimum noise figure is about 3.3 dB, and 1 dB gain compression point is about −15 dBm. The total current consumption included bias network is 9.1 mA under a 1.8 V single supply voltage. Chip area is 0.8 mm × 0.8 mm.

1. INTRODUCTION

The Federal Communications Commission (FCC) authorizes the unlicensed use of Ultra-Wideband (UWB) technology, this technology could be applied in various kinds of commodities. Under the FCC UWB rules, Wireless UWB has a wide bandwidth with an actual transmitting range of approximately 10 meters, and rate of transmitting data between 110 Mbps and 480 Mbps. The power dissipation needed by UWB is very low, having its frequency range at 3.1–10.6 GHz [1], and limited power emission at −41.3 dBm/MHz.

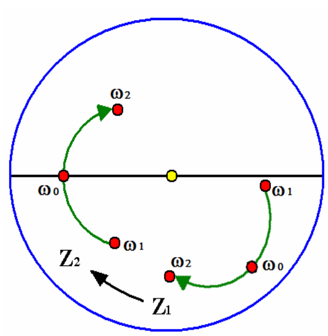
As we can see apparently, UWB technology announces the coming of the high-speed wireless technology new age. This high-speed technology can be applied in wireless network, home network connection, short distance radar, etc. Specific devices like multimedia DVDs, high quality television images, satellites, and televisions are some examples of its application. Moreover, in the coming future, digital cameras, scanners, printers and MP3 players can be connected to computers wirelessly, rather than the wired connection used presently. Through the brief description of UWB technology above, we shall foresee the future prevalence of high-speed wireless consuming multimedia facilities, not to mention the immense commercial profit it will produce.

The feedback architecture is widely used in wideband amplifiers, and performs well in wideband matching and flat gain. Based on the Feedback topology, this paper shows with the combination of parallel resonate matching [2–4], using of inductance of the input matching network can be reduced and leading to smaller chip area.

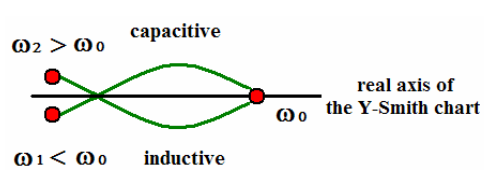
The paper is organized as the following. After the introduction, comes the Section 2 describing how we design wideband LNA; the Section 3 explains the simulating result, and the last chapter is the conclusion.

2. DESIGN OF AMPLIFIER

A load is placed at Z_2 which can be plotted on the Smith chart and it's trace as shown in Fig. 1(a) ($\omega_2 > \omega_1$). Once we transfer Z_2 into admittance Y_2($Y_2 = G_1 - jB_1$ at $\omega > \omega_0$ or $Y_2 = G_1 + jB_2$ at



(a)



(b)

Figure 1: The matching process. (a) after adding a component, Z_1 move to Z_2 and (b) Y_2 connect with parallel resonant cavity.

$\omega < \omega_0$), connect it with parallel resonate cavity, and design an appropriate resonate cavity L and C value, making resonate frequency fall on ω_0, and utilizes $-jB$ offered at low frequency resonate cavity and $+jB$ offered at high frequency, $-jB_1$ and $+jB_2$ at Y_2 could be eliminated, as showed in Fig. 1(b), thus we achieve wideband matching. Mostly the bandwidth of the wideband amplifier does not always go through the real axis precisely, as Z_1 in Fig. 1(a) or make the central point of the bandwidth falls accurately on the real axis. We can move the frequency load to the real axis by adding a "component" to it.

$$C_M = C_2 (1 - A_v) \tag{1}$$

$$R_M = \frac{R_1}{1 - A_v} \tag{2}$$

Figure 2 is the LNA topology we designed. The first step is to choose the size of transistor M3 in order to achieve minimal noise figure. In next step, we decide the device value of the input matching network which is composed of L1, C_M, R_M and L3.

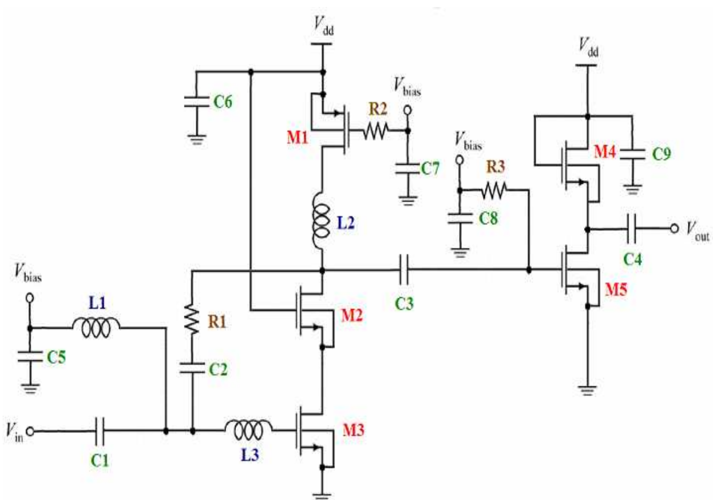


Figure 2: Proposed wideband amplifier.

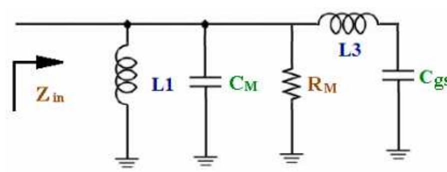
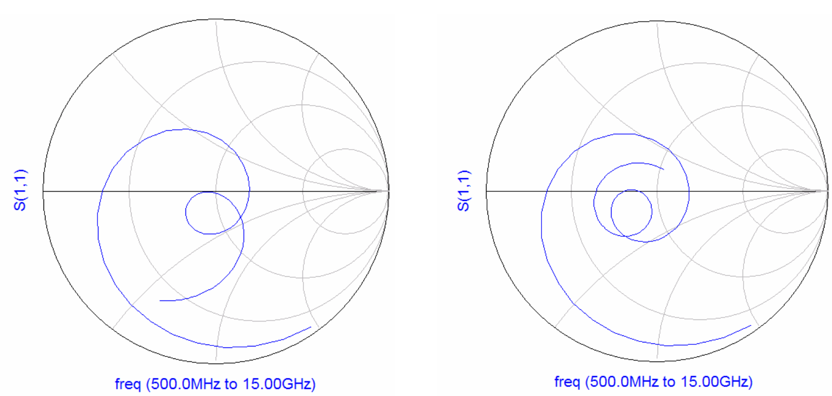


Figure 3: Small-Signal equivalent circuit at the input.

A_v is the open loop voltage gain of the LNA [5], while C2 can resist the output stage current from Cascode [6] and achieve optimization of transistor (M3) bias and increase the gm value by separating the first stage output DC level from the first stage input DC level. By doing so, power dissipation is decreased and the gain of the amplifier is increased. Also, appropriate regulation of R1 can enable the input resistance to approach the matching to $50\,\Omega$.

L1 and C_M will cause parallel resonate by using the inductive reactance and capacitive reactance appeared before and after the resonating, we can eliminate the reactance of the circuit we want to match. Then, the impedance in certain frequency range will approach $50\,\Omega$; thus the broadband matching achieved. Inductor L3 can move the matching bandwidth to the real axis. As showed in



(a) (b)

Figure 4: Input matching network (a) without L3 and (b) with L3.

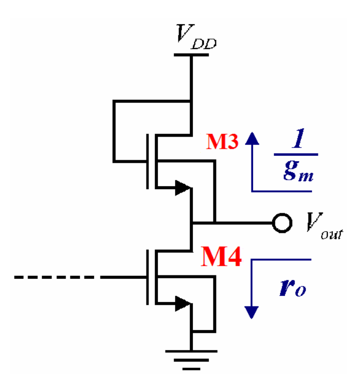


Figure 5: Output stage circuit.

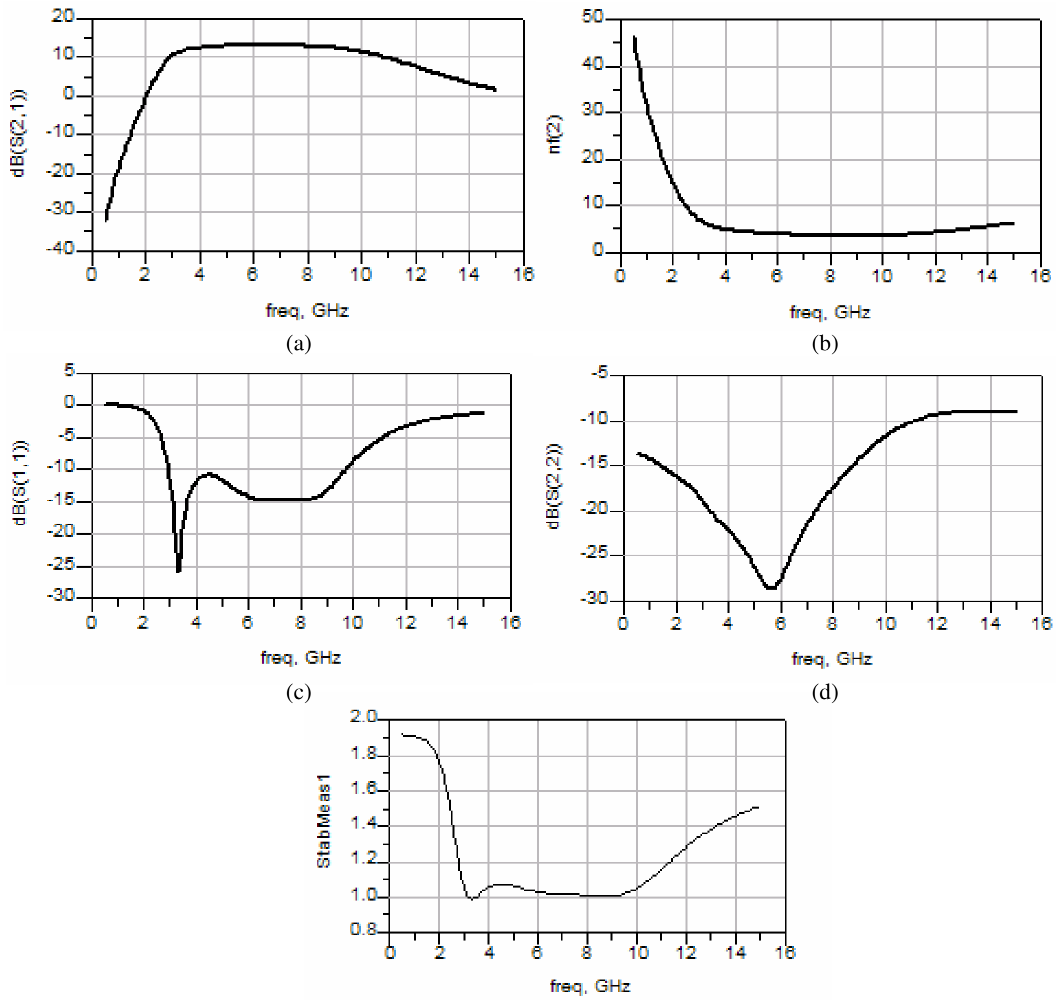


Figure 6: Simulation results of the proposed LNA. (a) gain, (b) noise figure, (c) input matching (S11), (d) output matching (S22), (e) stability factor.

Figs. 4(a) and (b), we find that the closer the high frequency approaches the central of the circle, the better matching effect achieves. With better matching effects, the gain will improve without extra power dissipation.

There is an amplifying effect at the output stage also, which can increase the total gain. As we can see in Fig. 5, by using active load, the output impedance equals $R_{\text{out}} = 1/g_{m3} \,//\, r_{o4} \cong 1/g_{m3}$. Then, regulate $1/g_{m3}$ as approximately $50\,\Omega$, and we achieve matching at the output.

In order to extend the Bandwidth, the output load of cascode amplifier will be done with inductive-peaking technique as showed in Fig. 2. Inductor L2 will give the cascode amplifier output an extra zero to eliminate the pole, and extend the bandwidth. Because the Cascode stage is the main amplifier, the current being nearly 6 mA, we will parallel 6–7 resistors to prevent the wire be broken in actual application. Knowing that process variations usually occur in semiconductor process, this circuit adopts PMOS to complete the resister device of inductive-peaking. When operating PMOS at the deep triode region [7], the MOS device performs like a resistor (see M1 in Fig. 2), keeping the gate voltage of M1 low, so that all loads of the voltage output amplitudes remain in the deep triode region. Resistance RON is showed in Equation (3).

$$R_{\text{on}} = \frac{1}{\mu_p C_{ox} \left(\frac{W}{L}\right) \left(V_{DD} - V_b - |V_{THP}|\right)} \tag{3}$$

3. SIMULATION RESULTS

The proposed architecture shown in Fig. 2 is applied to a 3.1∼10.1 GHz broadband amplifier based on 0.18 µm CMOS technology. The simulated maximum gain is 13.1 dB at 6.6 GHz and 1 dB of gain flatness as shown in Fig. 6(a).The simulated NF shows a minimum value of 3.3 dB at 8.9 GHz as shown in Fig. 6(b). The input/output matching simulation results are shown in Figs. 6(c) and (d) respectively. S11 and S22 of LNA are less than −10 dB in the frequency ranges of interest. The stability of LNA was simulated as shown in Fig. 6(e), and it is satisfied in the operation frequency range. The input third intercept point is performed at 5.5 GHz and 10.5 GHz. The performance simulation results for proposed LNA are summarized in Table 1.

Table 1: Simulation results of wideband LNA.

Process	TSMC 0.18 μm CMOS
Bandwidth (GHz)	3.1∼10.1
S11 (dB)	> 10 dB
S21 (dB)	12 ± 1
S12 (dB)	> 30 dB
S22 (dB)	> 10 dB
NF (dB)	$3.3 \sim 5$
P1dB (dBm)	−15.7@ 5.5 GHz −9 @ 10.5 GHz
IIP3 (dBm)	−6.6@ 5.5 GHz −1.7@ 10.5 GHz
Power Consumption (mW)	16.4
Chip Size (mm×mm)	0.8×0.8

4. CONCLUSIONS

The novel broadband amplifier is achieved by combining the cascode feedback topology and parallel resonant cavity. The designed LNA is applied to a $3.1 \sim 10.1$ GHz UWB amplifier implementation based on 0.18 μm CMOS technology. The simulation results show 12 dB gain, minimum noise figure 3.3 dB, and less than −10 dB input/output matching within bandwidth with 16.4 mW power consumption. This proposed broadband LNA could be used for UWB system.

REFERENCES

1. Purton, P., *Ultra Wideband: Opportunities and Threats for the Mobile Communications Sector*, Baskerville, London, 2004.
2. Gonzalez, G., *Microwave Transistor Amplifiers: Analysis and Design*, 2nd ed., Prentice-Hall, 1997.
3. Rizzi, P. A., *Microwave Engineering: Passive Circuit*, Prentice-Hall, 1988.
4. Cripps, S. C., *Microwave Power Amplifiers for Wireless Communications*, Artech.
5. Kim, C. W., M. S. Kang, P. T. Anh, H. T. Kimand, and S. G. Lee, "An ultra-wideband CMOS low noise amplifier for 3–5-GHz UWB system," *IEEE J. Solid-State Circuits*, Vol. 40, 544–547, 2 Feb., 2005.
6. Vishwakarma, S., S. Jung, and Y. Joo, "Ultra wideband CMOS low noise amplifier with active input matching," *2004 International Workshop*, 415–419, 18–21 May, 2004.
7. Razavi, B., *Design of Analog CMOS Integrated Circuit*, McGraw Hill, 1996.

Compact Surface-mount Wideband and Multi-band Internal Chip Antenna for Mobile Handset

H. C. Choi, D. S. Shin, S. U. Park, and H. D. Kim

Hanyang University, Korea

Abstract— A novel and broadband spiral chip antenna is presented in this paper. The size of the proposed antenna is $20 \times 3.2 \times 8\,\text{mm}^3$ and is suitable to be embedded in a mobile phone as an internal antenna. Also, this antenna is designed to be installed at the newest mobile handset whose size is $40 \times 93\,\text{mm}^3$. The minimization of this antenna was realized by using spiral line structure on FR-4 of dielectric permittivity $\varepsilon_r = 4.4$. The proposed antenna is SMD type to be easily installed in the practical mobile handset. We use PCB technology which can allow this antenna to be produced massively with low cost. We can get the wideband operation in the upper band by overlapping high order resonances of branch structure. The measured bandwidth (VSWR< 3) of the designed antenna is 160 MHz (980–1140 MHz) in the lower band and 1230 MHz (1910–3140 MHz) in the higher band. The antenna has been designed by a commercial software HFSS.

1. INTRODUCTION

Mobile phone companies make their efforts to get better designs by installing the antenna inside mobile handset instead of outside of it. Accordingly, miniaturization research of antenna has been pursued for many years. As a result of it, many internal type antennas have been studied [1]. Because the mobile phone business is in active in export as well as domestic demand, interest in the development of the global roaming phone which can be sold in several countries by one development is rapidly growing. For the global roaming phone development the company needs the technology of multi-band antenna that can accommodate the bands more than quad-band. Also, the multi-band antenna is the essential condition of mobile phone as consumers prefer the multi-media functions.

Most antennas, studied now, are PIFA type that can accommodate more than quad-band. But these antennas can work properly only under the limited condition that they should be installed high from the ground of mobile phone [2, 3], which makes them unfit for the cell phone of thin slices shape. It is hard for mobile handset with PIFA type antenna to lead the fashion. There were attempts to manufacture small antenna which take advantage of chip type antenna technology to solve this problem. Until now small dual-band antennas that use structure such as meander line or spiral line on dielectric substance have been studied much [4, 5]. This paper proposes the small dual wide-band antenna to be built in global roaming phone using dual-band chip antenna technology.

Printing Circuit Board (PCB) technology, which is used in the proposed antenna, can reduce production cost and manufacturing time. Because this antenna is manufactured as SMD type, it can be easily installed in mobile handset. The antenna with spiral lines is constructed on FR-4 dielectric substrate ($\varepsilon_r = 4.4$). The size of antenna is $20 \times 3.2 \times 8\,\text{mm}^3$ and test board is $40 \times 93\,\text{mm}^2$. Although the larger ground plane helps to get the wider bandwidth, in this paper the size of test board is limited to that of common mobile handset. The measured bandwidth (VSWR 3:1) of the proposed antenna is 160 MHz (980–1140 MHz) at low frequency range and 1230 MHz (1910–3140 MHz) at high frequency range.

2. DESIGN OF THE CHIP ANTENNA

This proposed penta-band chip antenna is designed to act in 880–960 MHz (GSM900) and 1850–2480 MHz (USPCS/WCDMA/WiBro/Bluetooth). The resonance frequency given in test jig generally falls by permittivity of phone case when the antenna is installed in a real mobile phone [6]. So we designed resonance frequency of this antenna a little higher. Also, considering the variation of resonant frequency by the users hand or head [6], the proposed antenna is designed to have the wider bandwidth than the target bandwidth.

The proposed antenna is manufactured by taking advantage of PCB technology. In making a small size antenna by using this PCB technology, meander line structure is more effective [4, 7]. But the bandwidth will be decreased if one uses too many meander sections in limited height. Also we cannot expect linear falling of resonance frequency [7, 8]. The branch structure that consists of

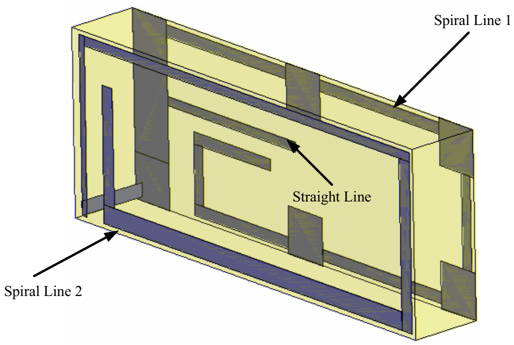


Figure 1: Proposed antenna.

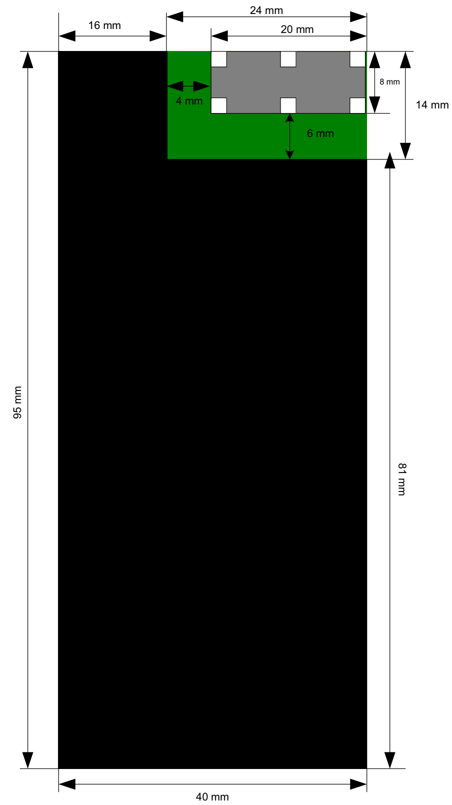


Figure 2: Geometry of test board.

two meander line to get multi-band is affected by distance between two structures [4]. Therefore it is difficult to design the antenna under the condition that mobile phone's thickness is fixed. In addition to the capacitance between branch structure makes bandwidth at low frequency narrow [9]. So, in this paper we manufacture the branch structure antenna using two spiral lines and one straight line. Fig. 1 shows the entire antenna structure whose two spiral patterns are connected by via component. The proposed antenna is simulated on test board seen in Fig. 2. The test board is designed to be similar to the board of the general mobile handset.

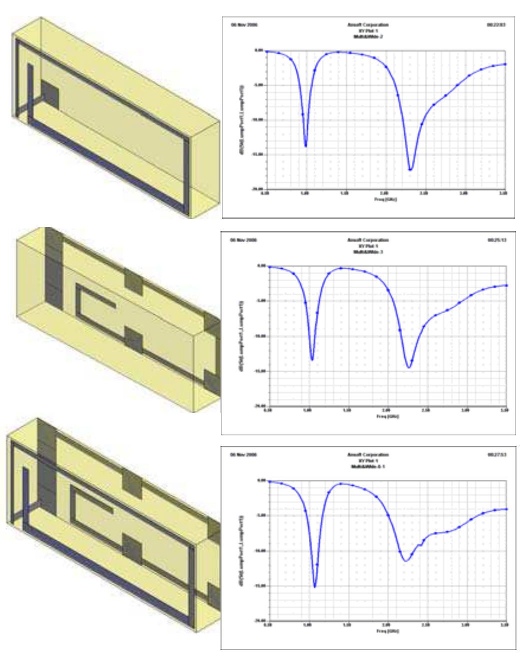


Figure 3: Effect of branch (two spiral lines).

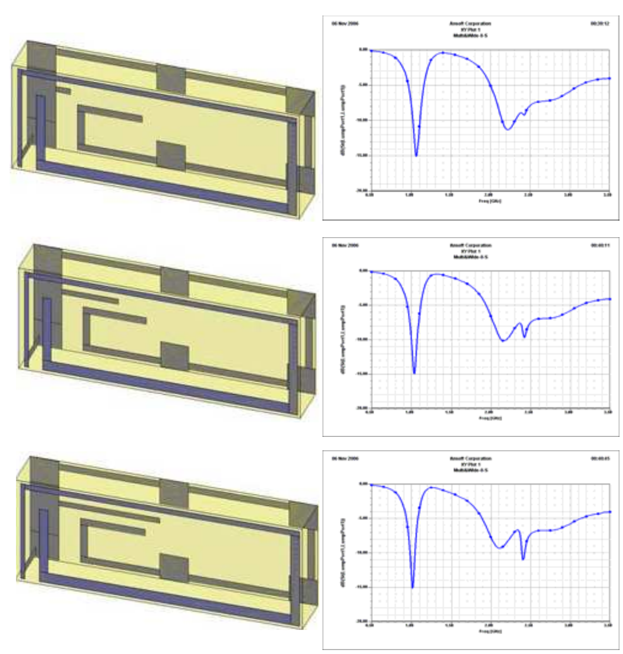


Figure 4: Effect of branch (one straight line).

The proposed antenna is designed by adjusting the length and width of two spiral lines to overlap the harmonic resonance components made by each structure. The characteristic of overlapping resonance components is observed at Fig. 3. Also, Fig. 4 shows the effect of the straight line. By adjusting three branch lines the designed antenna has wide-band operation.

Figure 5: Manufactured antenna.

Figure 6: Antennas manufactured on one PCB board.

3. ANTENNA MANUFACTURED AND RESULTS MEASURED

The manufactured antenna is shown in Fig. 5. The type of the proposed antenna has an advantage that one can reduce product development time because instantly many models can be produced which has various resonance frequencies by using PCB technology shown in Fig. 6. Fig. 7 shows that the proposed antenna is attached by SMD Type to test jig to measure with network analyzer. In Fig. 8, the measured and simulated VSWR are compared which are agreeable in tendency. The measured bandwidth (VSWR< 3) of the proposed antenna is 160 MHz (980–1140 MHz) in the lower band and 1230 MHz (1910–3140 MHz) in the higher band. The measurement of VSWR data was achieved by using Agilent E8361 network analyzer.

Figure 7: Installation to test board.

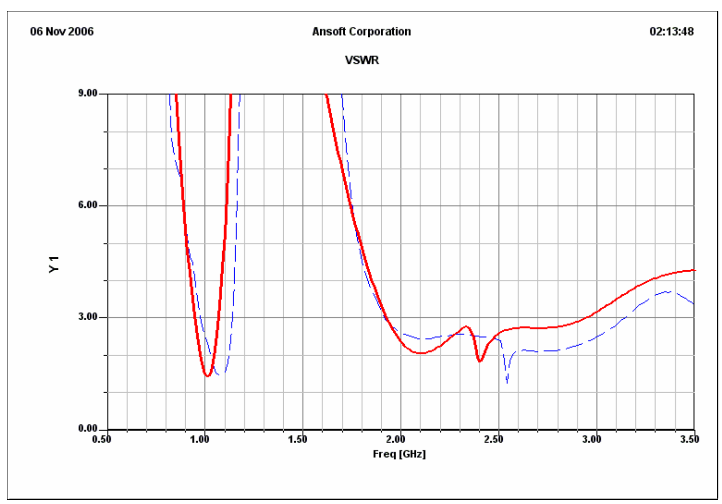


Figure 8: Simulation VSWR data (solid line) and Measured VSWR data (dashed line).

Table 1: Radiation efficiency of the proposed antenna.

(a) in low frequency range.

	Start	Center	Stop
Frequency (MHz)	980	1020	1140
Efficiency (dB)	−1.83	−1.53	−1.62
Efficiency (%)	65.58	70.33	68.80

(b) in high frequency range.

	Start	Center	Stop
Frequency (MHz)	1910	2525	3140
Efficiency (dB)	−1.92	−1.33	−1.91
Efficiency (%)	64.33	88.26	64.42

The measured radiation efficiency in VSWR 3:1 is summarized in Table 1. These values were measured in A PLUS tech CTIA OTA test chamber. Table 1 illustrates that the radiation efficiency within (VSWR 3:1) bandwidth is over 60%.

4. CONCLUSIONS

A dual wide-band monopole antenna using branch structure with two rectangular spiral lines and one straight line has been proposed, fabricated and tested. We improved productivity taking advantage of PCB method. And we could achieve broadband characteristic by overlapping dual resonances in high frequency band. The bandwidth of this proposed antenna is wider than the necessary band of GSM900/USPCS/WCDMA/WiBro/Bluetooth. Therefore this antenna is robust for the variation of resonant frequency by hand or head. In conclusion, this proposed antenna is expected to be efficiently used for the global roaming phone in many corporations.

ACKNOWLEDGMENT

This work was supported by HY-SDR Research Center at Hanyang University, Seoul, Korea, the ITRC program of MIC, Korea.

REFERENCES

1. Wong, K. L., *Planar Antenna for Wireless Communication*, John Wiley and Sons, New York, 2003.
2. Guo, Y. X., M. Y. W. Chia, and Z. N. Chen, "Miniature built-in multiband antennas for mobile handsets," *IEEE Transactions on Antennas and Propagation*, Vol. 52, Issue 8, 1936–1944, Aug. 2004.
3. Nashaat, D. M., H. A. Elsadek, and H. Ghali, "Single feed compact quad-band PIFA antenna for wireless communication applications," *IEEE Transactions on Antennas and Propagation*, Vol. 53, Issue 8, Part 2, 2631–2635, Aug. 2005.
4. Kim, Y. D., M. S. Kim, H. Y. Kim, and H. M. Lee, "Dual-band chip antenna using LTCC multilayer technology for mobile communication applications," *Antennas and Propagation Society International Symposium, IEEE*, Vol. 3, 20–25, 3115-3118, June 2004.
5. Tung, H. C., W. S. Chen, and K. L. Wong, "Integrated rectangular spiral monopole antenna for 2.4/5.2 GHz dual-band operation," *Antennas and Propagation Society International Symposium, IEEE*, Vol. 3, 16–21, 496–499, June 2002.
6. Sim, D. U. and S.-O. Park, "The effects of the handset case, battery, and human head on the performance of a triple-band internal antenna," *Antennas and Propagation Society International Symposium, IEEE*, Vol. 2, 20-25, 1951–1954, June 2004.
7. Best, S. R. and J. D. Morrow, "Limitations of inductive circuit model representations of meander line antennas," *Antennas and Propagation Society International Symposium, IEEE*, Vol. 1, 22–27, 852–855, June 2003.
8. Best, S. R., "On the performance properties of the Koch fractal and other bent wire monopoles," *IEEE Transactions on Antennas and Propagation*, Vol. 51, Issue 6, 1292–1300, June 2003.
9. Makimoto, M. and S. Yamashita, *Microwave Resonators and Filters for Wireless Communication*, Chap. 4, Springer Verlag, New York, 2000.

Broadband Internal Antenna for Mobile DTV Handsets

Seung-Gil Jeon, Dong-Hyun Seo, Yeon-Sik Yu, and Jae-Hoon Choi

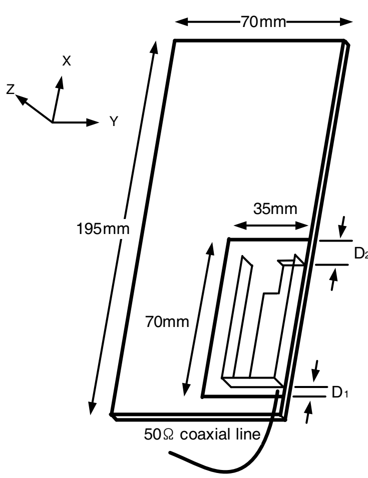
Department of Electrical and Computer Engineering, Hanyang University
17 Haengdang-Dong, Seongdong-Gu, Seoul 133-791, Republic of Korea

Abstract— A novel low profile metal plate monopole antenna with L-shape silts for mobile handsets operating DTV (470–740 MHz) band is proposed. The antenna can be easily constructed by folding the planar patch with the silt and occupies a small volume of $56 \times 23 \times 10$ mm. The measured return losses for operating frequencies over the DTV band are better than -10 dB. The proposed antenna shows omni-directional or monopole-like radiation patterns and has good antenna gain over the frequencies of interest. The antenna can be applied to DTV application.

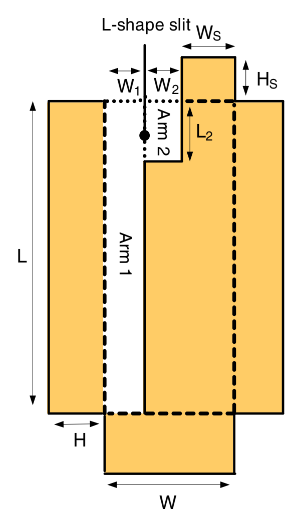
1. INTRODUCTION

Digital television (DTV) broadcasting is expected to provide data service and replace the analog TV broadcasting. It is very attractive for wireless communication device based on digital communication technology such as mobile handset. Many wireless communication devices have internal antenna for several advantages over the conventional external antennas [1]. It makes the mobile handset look aesthetic and compact. For this reason an internal antenna for DTV application should be low profile and have wideband (470∼740 MHz) characteristic [2]. Especially the key issue in the design of an internal antenna is the size reduction while maintaining the antenna performance the same as that of an external antenna. A low profile planar metal plate monopole antenna for DTV signal reception in the UHF band for laptops has been reported recently [3].

In this paper, we present an embedded DTV antenna for mobile handset application. The proposed antenna has a narrow width of 23 mm and is very promising to be embedded within the narrow space between the display and the casing of the mobile handset to perform as a concealed or embedded antenna. In addition, the antenna can provide a wide bandwidth (about 50% centered at 625 MHz) for operating in UHF band.



(a)

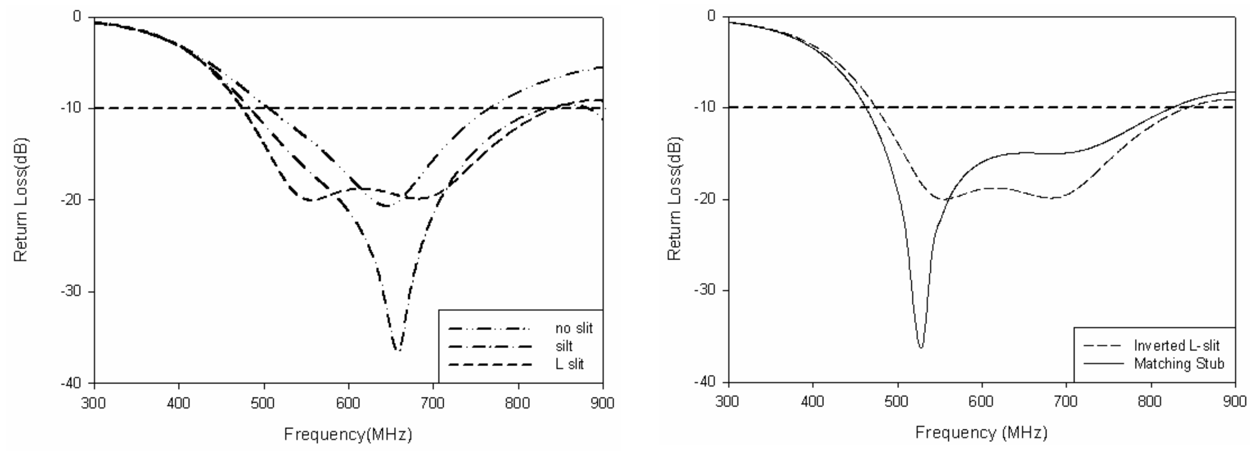


(b)

Figure 1: The structure and geometry of the proposed antenna, (a) overall view, (b) detailed geometry of the radiation element.

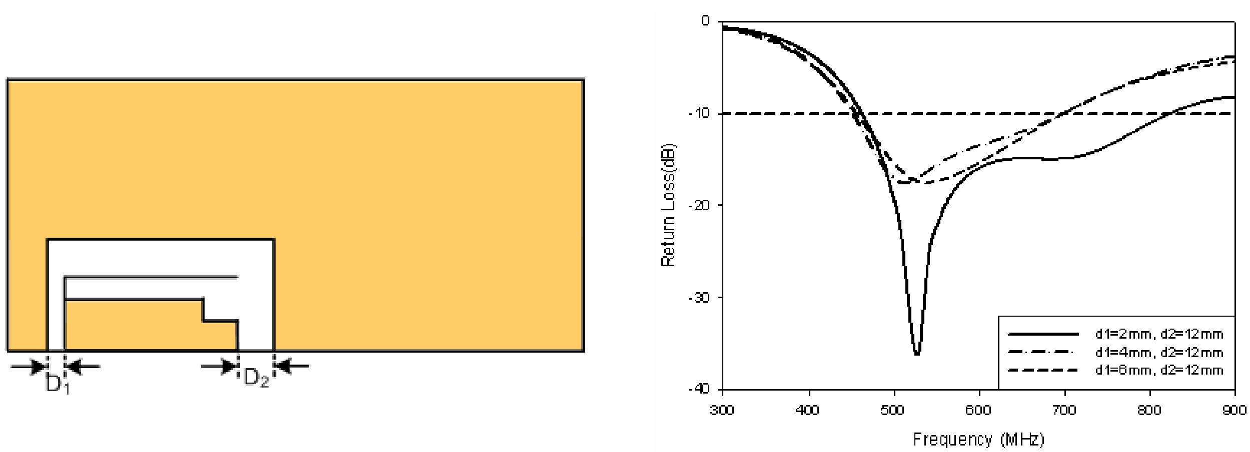
2. ANTENNA DESIGN

Figure 1 shows the structure and geometry of the proposed antenna. The antenna is placed on the FR4 substrate with height (H) of 0.8 mm and relative permittivity of 4.4 and has a dimension of 195 mm×70 mm. The proposed antenna is a folded metal plate monopole with L-shape slit. The folded metal plate monopole, the antenna ground plane and the system ground plane were all fabricated using a 0.2 mm thick brass sheet. There is a small gap of 2 mm between the folded structure and the ground plane and the feeding point is located at the edge of the monopole structure. 50 Ω coaxial line is connected to feed point to excite the monopole structure and its grounding sheath connected to the antenna ground plane. The length of folded monopole (L) is 56 mm. The L-shaped slit consists of two arms that control the impedance bandwidth. Each length of the two arms is a prime element to decide the impedance bandwidth of a suggested antenna.



(a) (b)

Figure 2: Simulated return loss for different structures, (a) for various slit structure, (b) for various matching structure.

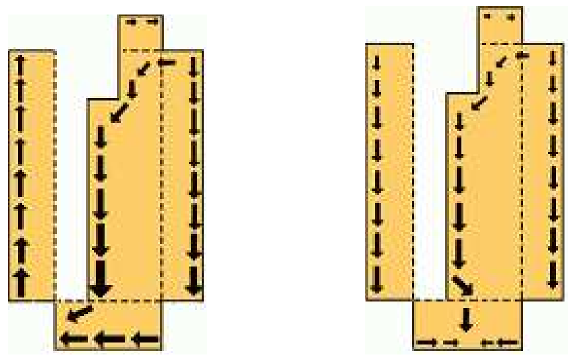


(a) (b)

Figure 3: Return loss characteristics for various values of D_1 and D_2, (a) definition of geometric parameter D_1 and D_2 (top view), (b) simulated return loss for various values.

Table 1: The optimum design parameters.

Parameter	Value (mm)	Parameter	Value (mm)
L	56	D_2	12
L_2	11	W	23
H	10	W_S	8
H_S	8	W_1	10
D_1	2	W_2	7



(a) (b)

Figure 4: Current paths on the folded antenna for different frequencies, (a) 530 MHz, (b) 740 MHz.

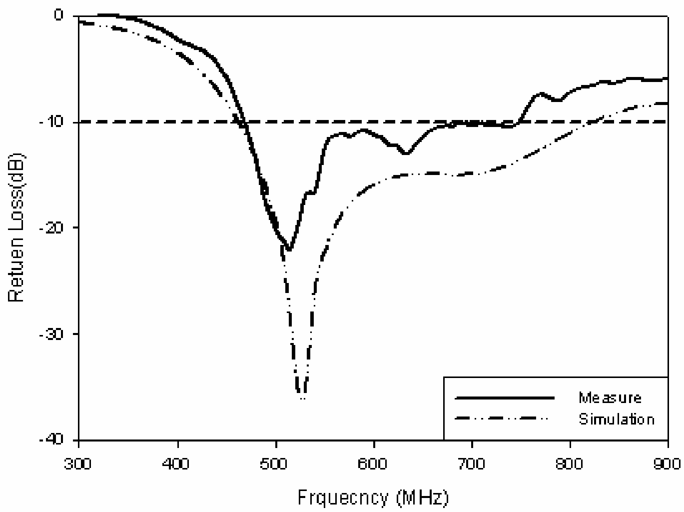
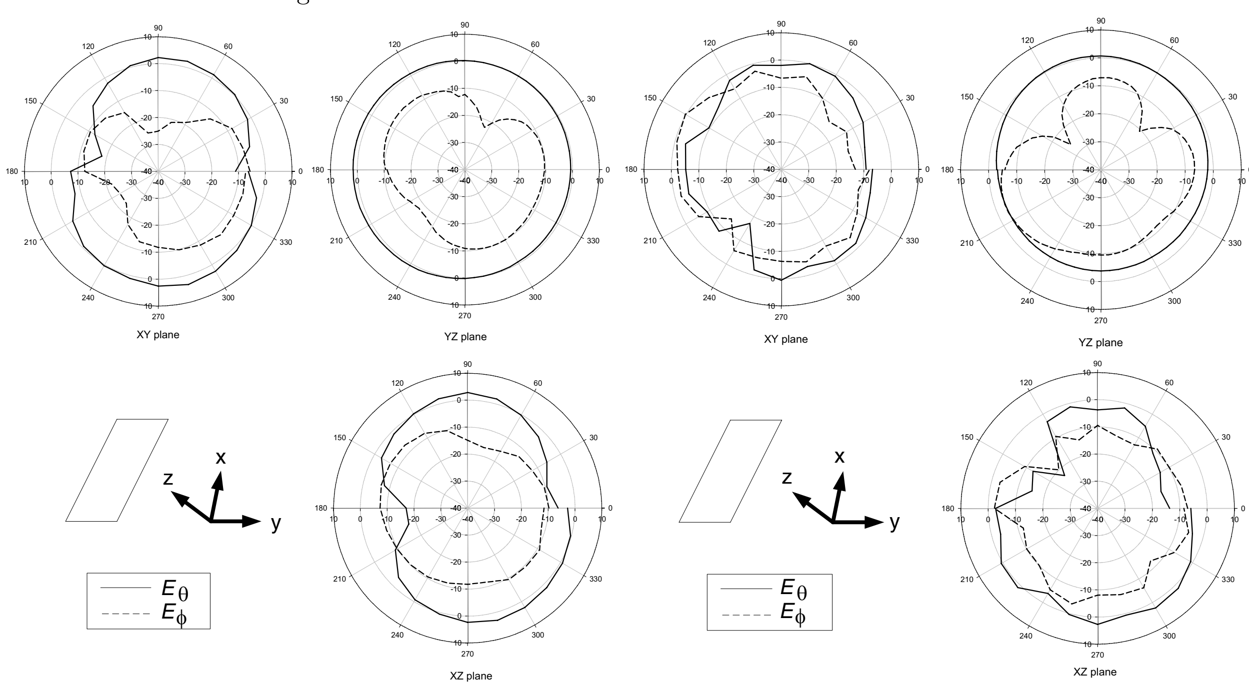


Figure 5: Simulated and measured return loss characteristics.



(a) (b)

Figure 6: Measured radiation patterns, (a) 520 MHz, (b) 740 MHz.

3. SIMULATION AND MEASUREMENT

Figure 2 shows the simulated return loss characteristic of a proposed antenna with L-shape slit and matching stub (Hs). The arms and matching stub enhance the wideband impedance performance by providing an additional resonance. Figure 3 shows the input return loss characteristic for various values of D1 and D2. The two values of gaps between the folded monopole structure and ground plane make differences in matching performance. The current paths of folded antenna are shown in Figure 4. The length of the longest current path is about $\lambda/4$ at 560 MHz and 740 MHz. To achieve low profile structure on the finite ground, the monopole antenna is folded. Figure 5 shows the simulated and measured return loss of the proposed antenna. Frequency tuning can be achieved by adjusting the lengths of the two arms. The optimum design parameters are listed in Table 1. Figure 6 presents the measured radiation patterns in the x-y plane, y-z plane and z-x plane at 520 MHz and 740 MHz are illustrated. Nearly omni-directional radiation patterns are observed at each frequency. Measured antenna gain is varying from -1 dBi to 4 dBi as shown in Figure 7. The antenna is modeled numerically by using Ansoft's High-Frequency structure simulator (HFSS) [4].

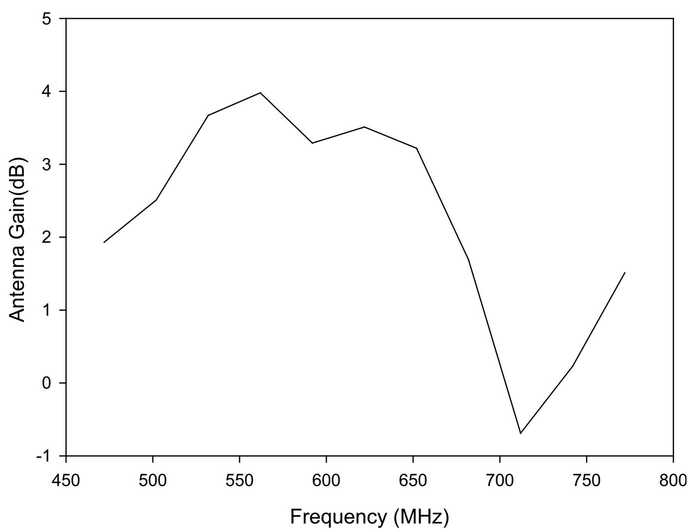


Figure 7: Measured antenna gain.

4. CONCLUSION

A novel low profile metal-plate monopole antenna with wideband characteristic has been proposed and implemented for DTV applications. The proposed antenna has a folded structure and L-shape slit. To obtain the wide bandwidth characteristic, the gap and L-shape slit with two arms are used. The dimensions of an L-shape slit and the value of the gap are optimized by the parametric analysis and the wide bandwidth performance from 465 MHz to 745 MHz are obtained. Design parameters of the antenna are experimentally optimized and the measured radiation and return loss characteristics of the antenna are good enough for DTV application. This antenna can be a good candidate for mobile DTV handset due to its low profile, small size and wideband characteristics.

ACKNOWLEDGMENT

This work was supported in port by MIC & IITA of Korea through IT Leading R&D Support Project.

REFERENCES

1. Ang, I., Y. X. Guo, and Y. W. Chia, "Compact internal quad-band antenna for mobile phones," *Micro. Opt. Technol. Lett.*, Vol. 38, No. 3, 217–223, 2003.
2. Su, C. M., L. C. Chou, C. I. Lin, and K. I. Wong, "Embaded DTV antenna for laptop application," *2005 IEEE Antennas and Propagation Society International Symposium*, Vol. 4B, 68–71, 3–8 Jul., 2005.
3. Su, C. M., L. C. Chou, C. I. Lin, and K. I. Wong, "Internal DTV receiving antenna for laptop application," *Micro. Opt. Technol. Lett.*, Vol. 44, No. 1, 4–6 Jan., 2005.
4. Ansoft High-Frequency Structure Simulator (HFSS), Ver. 10, Ansoft Corporation.

A Compact Band-selective Filter and Antenna for UWB Application

Yohan Jang, Hoon Park, Sangwook Jung, and Jaehoon Choi
Department of Electrical and Computer Engineering, Hanyang University
17 Haengdang-Dong, Seongdong-Gu, Seoul 133-791, Republic of Korea

Abstract— A novel compact CPW ultra-wideband (UWB) antenna combined with a compact band-selective filter is proposed. The proposed antenna without filter satisfies the return loss requirement of less than -10 dB over the frequency range of 3.0 GHz to 12.1 GHz. The optimized antenna has dimensions of $10 \times 16\,\text{mm}^2$ on a Teflon substrate ($\varepsilon_r = 3.48$) with good radiation characteristics and is easy to fabricate. The proposed antenna with filter has the frequency band of 3 GHz to 11 GHz for VSWR less than 2.0 with a rejection band around 5.0 to 5.9 GHz.

1. INTRODUCTION

Since the recent approval of UWB radio system, many researchers have extensively investigated on the modern indoor wireless communication system with high data transmission rate such as sensors, radar, and tracking applications. For the reliable use of these communication services, the design of UWB devices such as antennas, filters, and LNAs is required. Various types of UWB antennas operating from 3.1 GHz to 10.6 GHz have been studied [1–3]. However, due to the coexistence of the UWB frequency band with Wireless LAN and Hiper LAN service band from 5.15 GHz to 5.825 GHz, UWB radio signal can be interfered with those services. Also, the effect of unwanted noise increases Noise Figure (NF) of the entire receiving system and can cause performance degradation.

In this letter, novel compact CPW ultra-wideband (UWB) antenna using notches and stubs combined with a compact band-selective filter is proposed. To enhance impedance bandwidth, notches and stubs at the rectangular radiation patch were used. To realize pseudo-highpass filter transmission response over UWB frequency band, the distributed highpass filter scheme is used. The narrow bandstop function is achieved using coupled resonators. To miniaturize the total filter size, the shape of resonators was modified to have small dimension. Three attenuation poles for the filter can be tuned by controlling the length of resonators to have narrow rejection band around Wireless LAN and Hiper LAN service band.

2. COPLANAR WAVEGUIDE-FED UWB ANTENNA DESIGN AND ANALYSIS

In UWB application, the antenna operates from 3.1 GHz to 10.6 GHz. Figure 1 shows the proposed antenna for UWB application. The proposed antenna structure consists of a main rectangular radiation patch with notches and stubs at side corners of the patch. They are printed on an RO4350b substrate with thickness of 0.762 mm and relative permittivity of 3.48. The thickness of copper coating on the top side of the substrate is approximately 0.0175 mm. A 50 Ω CPW feed line, having a metal strip width $W_f = 4$ mm and a gap distance $G = 0.28$ mm, is used to excite the proposed antenna.

The dimensions of rectangular patch are optimized to have resonant frequency at 3.7 GHz. By optimizing parameters of $L2$, $W2$, and $T1$ affecting the length and width of stub and notch attached

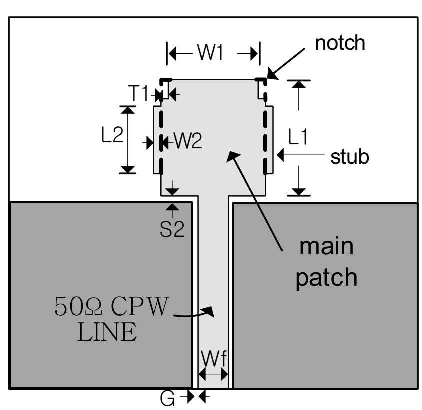
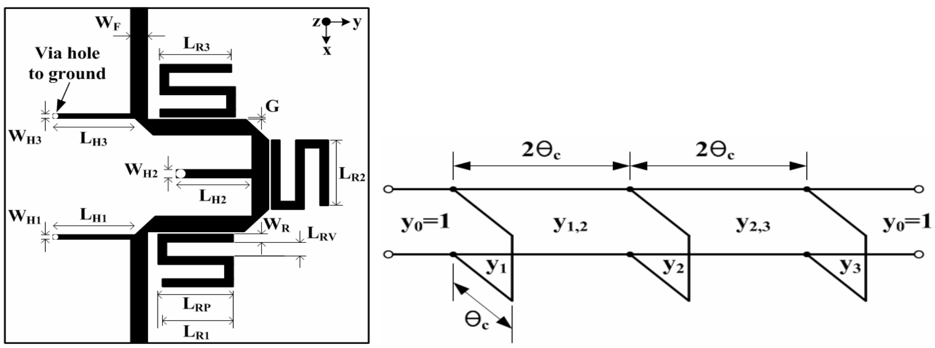


Figure 1: The proposed antenna.

to the rectangular radiation patch, improved impedance bandwidth performance can be achieved for the proposed antenna. That is largely due to the fact that the two stubs and two notches affect the reactive coupling between the rectangular patch and the ground plane and input impedance of main patch. The impedance bandwidth of the antenna covers from 3.0 GHz to 12.1 GHz for the return loss of less than −10 dB. The final design parameters are as follows: $W_1 = 10$ mm, $W_2 = 1$ mm, $L_1 = 16$ mm, $L_2 = 8.5$ mm, $T_1 = 2$ mm, $G = 0.32$ mm, $W_f = 4$ mm , and $S_2 = 1$ mm.

3. FILTER DESIGN AND ANALYSIS

Figure 2 shows the geometry and equivalent models of the proposed filter. This filter can be divided into a conventional highpass and a bandstop filter. To realize band-selective characteristic within UWB frequency band, two filters are integrated on both sides of the 50-ohm microstrip line and design parameters are adjusted to obtain superior frequency response throughout the operating frequency band. The filter is printed on 30-mil Rogers RO4350 substrate with dielectric constant (ε_r) of 3.48.



(a) (b)

Figure 2: (a) Geometry of the proposed filter, (b) Equivalent transmission line model of the highpass filter.

All simulations are carried out by using Ansoft High Frequency Structure Simulator (HFSS) [4]. To design highpass filter for wideband application up to 10.6 GHz, the mixed lumped/distributed (L/D) highpass filter scheme is chosen [5]. The filter consists of a cascade of three shunt short-circuited stubs of electrical length of θ_c at the cutoff frequency f_c, separated by connecting lines of electrical length of $2\theta_c$. The transfer function and filtering function can be described by the following equations [6].

$$|S_{21}(\theta)|^2 = \frac{1}{1+\varepsilon^2 F_N^2(\theta)}, \tag{1}$$

$$F_N(\theta) = \frac{(1+\sqrt{1-x_c^2})T_{2n-1}\left(\frac{x}{x_c}\right) - (1-\sqrt{1-x_c^2})T_{2n-3}\left(\frac{x}{x_c}\right)}{2\cos\left(\frac{\pi}{2}-\theta\right)}, \tag{2}$$

where ε is the passband ripple constant, θ is the electrical length, and $T_n(x)$ is the Chebyshev function of the first kind of degree n. The associated characteristic line impedance for the given terminating impedance Z_0 are determined by

$$Z_i = Z_0/y_i. \tag{3}$$

$$Z_{i,i+1} = Z_0/y_{i,i+1}. \tag{4}$$

The calculated characteristic impedances for line elements are $Z_1 = Z_3 = 82.7$ ohms, $Z_2 = 58.5$ ohms, and $Z_{1,2} = 49.1$ ohms. To realize the extremely wide passband, the line width is modified using full-wave EM simulation and determined to have much better performance throughout UWB frequency band. The initially calculated electrical length (θ_c) of each filter is 40.68°.

A three-pole Chebyshev lowpass prototype with a passband ripple of 0.1 dB is selected to design narrowband bandstop filter as introduced in [5]. For simplicity, the line width of the meandered resonators is fixed at 0.5 mm because the impedance of the resonators does not significantly affect the stop-band characteristic. The separation between the resonators is a quarter-guide wavelength ($\lambda_g/4$) at the center frequency of 5.49 GHz and resonator length is approximately a half-guide wavelength ($\lambda_g/2$). The designed filter has the maximum attenuation of 49.2 dB at 5.65 GHz with three resonators and the 3-dB rejection bandwidth from 5.29 GHz to 5.85 GHz. Design parameters of the final filter are summarized in Table 1.

Table 1: Parameters of the designed band-selective filter [Unit: mm].

Parameters	Values	Parameters	Values
W_{H1}	0.2	L_{R1}	5.26
W_{H2}	0.4	L_{R2}	5.46
W_{H2}	0.2	L_{R3}	5.36
L_{H1}	7.5	L_{RV}	1
L_{H2}	7	L_{RP}	5.76
L_{H3}	7.5	G	0.1
W_R	0.5	W_F	1.7

4. RESULTS

Figure 3 shows the proposed antenna combined with a compact band-selective filter for UWB application.

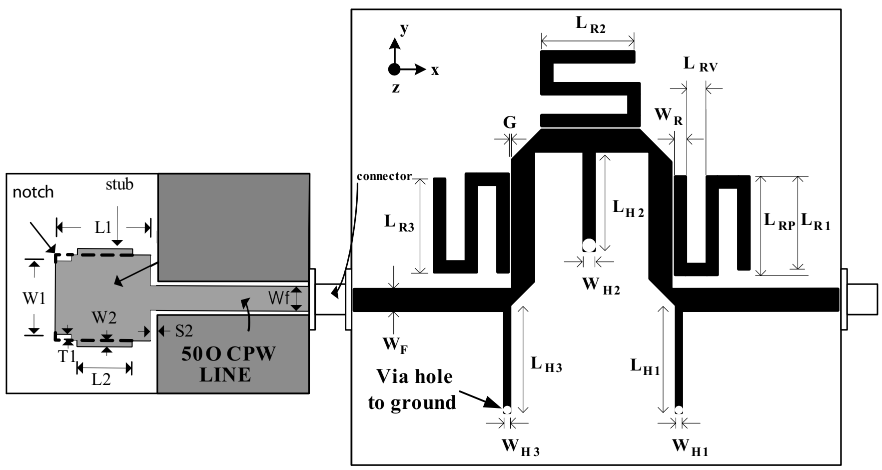


Figure 3: The proposed antenna combined with a compact band-selective filter.

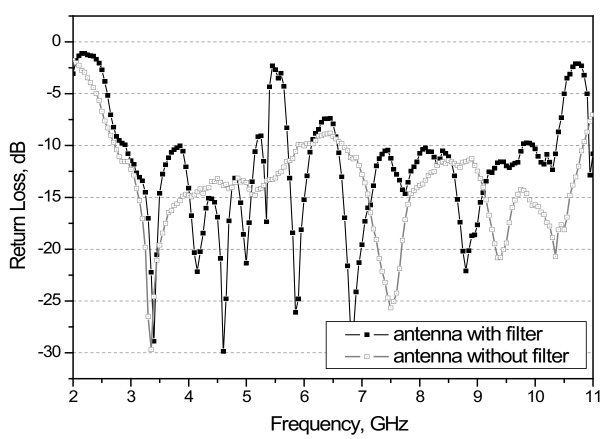
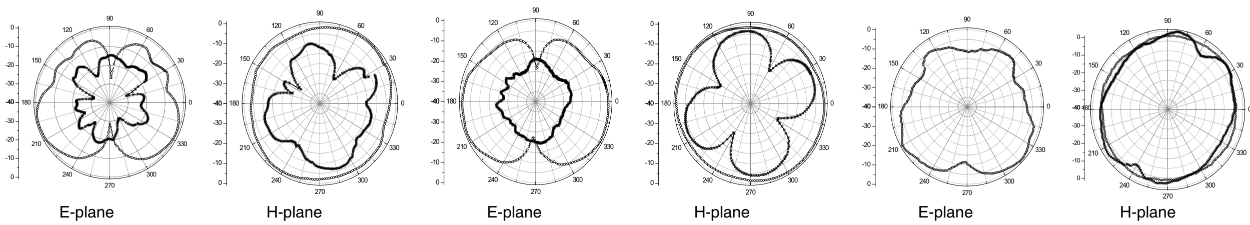


Figure 4: Measured return loss of the antenna with and without filter.

The measured return loss characteristic of proposed antenna is shown in Figure 4. The impedance bandwidth of the proposed antenna reaches 9.1 GHz (3.0∼12.1 GHz) for the return loss of less than −10 dB, which is enough to cover the entire UWB system. By using the filter, the frequency stop band is created while maintaining the wide bandwidth performance.

Figures 5(a) to 5(c) show measured radiation patterns, including E-and H-planes and co-and cross–polariza tions at 3.5, 5, and 9 GHz. Good omni-directional performance at lower frequency band with a cross-polarization level of less than −15 dB are assumed.



(a) (b) (c)

Figure 5: Measured radiation patterns of E- and H-planes and co-and cross-polarizations at (a) 3.5 GHz, (b) 5 GHz, and (c) 9 GHz. (-- ○ --: co-pol, — * —: x-pol)

Figure 6 illustrates the measured antenna gains of the designed antenna. From the results illustrated above, one can conclude that the suggested antenna system can be readily utilized for UWB application.

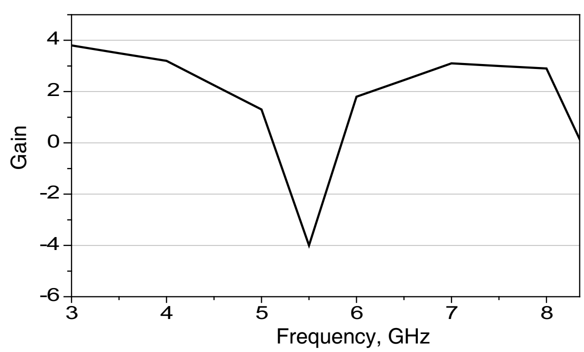


Figure 6: Measured antenna gain.

5. CONCLUSION

A novel and compact CPW-fed UWB antenna with a microstrip band-selective filter is proposed. To obtain the wideband characteristic for UWB frequency band, notches and stubs are utilized. Combining an UWB antenna with a filter having notch-function, the antenna with ultra-wideband performance over UWB band and band-rejection characteristic over WLAN band are obtained.

ACKNOWLEDGMENT

This work was supported by HY-SDR Research Center at Han-yang University, Seoul, Korea, under the ITRC Program of IITA, Korea.

REFERENCES

1. Ammann, M. J. and Z. N. Chen, "A wide-band shorted planar monopole with bevel," *IEEE Trans. on Antennas and Propagation*, Vol. 51, No. 4, 901–903, April 2003.
2. Chung, K., T. Yun, and J. Choi, "Wideband CPW-fed monopole antenna with parasitic elements and slots," *Electronic Letters*, Vol. 40, No. 17, August 2004.
3. Suh, S. Y., W. L. Stutzman, and W. A. Davis, "A new ultrawideband printed monopole antenna: the Planar Inverted Cone Antenna (PICA)," *IEEE Trans. on Antennas and Propagation*, Vol. 52, No. 5, May 2004.
4. Ansoft High Frequency Structure Simulator (HFSS), Ver. 9.2, Ansoft Corporation.
5. Levy, R., "A new class of distributed prototype filters with applications to mixed lumped/distributed component design," *IEEE Trans.*, Vol. MTT-18, 1064–1071, December 1970.
6. Hong, J. S. and M. J. Lancaster, *Microstrip Filters for RF/Microwave Applications*, Wiley, New York, 2001.

Optical Generation of mm-Wave Signal with Wide Linewidth Lasers for Broadband Communications

B. N. Biswas
Academy of Technology, Hooghly 712 121, India

Abstract— In future generation space based phased array antennas, pico-cellular broadband communication systems and antenna remoting, optical techniques using optical fiber, laser diodes and photo-detectors for transport electrical signals are widely used. Interest in the deployment of optical fibers arises from the low transmission loss compared to electrical media. Aside this, optical fibers are small, flexible, lightweight, and they have good phase accuracy and very large bandwidth. In view of the use of laser diodes and photo-detectors in the link, it is expected that the system costs can be considerably reduced if these optical devices are used for the generation of mm-wave signals so that no mm-wave oscillator and modulator are required. The problem is to use large linewidth lasers to generate mm-wave signal with desired spectral purity. This paper suggests a solution.

1. INTRODUCTION

The simplest way to transport electrical signal at the remote site is to intensity modulated the laser beam with the electrical signal, send it through an optical fiber and recover the electrical signal by the method of photo-detection. The signal to noise output at the receiving end is given by

$$\left(\frac{S}{N}\right)_{\text{out}} = \frac{aP_s}{b + cP_s + dP_s^2}$$

where "a" depends on the photo-detector responsivity, depth of intensity modulation and transmission loss; "b" depends on thermal noise characteristic of the link and "d" depends on the noise power already in the transmitter. The variation of the signal to noise power output is shown in Fig. 1. From the figure it is seen that the maximum SNR output is limited by the noise in the laser source, such as RIN, laser phase noise, etc. By this method it is possible to realize fiber links with gain and signal to noise ratio of few dB without the use of electrical amplifiers. In-line optical amplifiers can be used to compensate losses due to fiber attenuation and other optical components. Although the in-line optical amplifiers have many desirable advantages, such as, high gain, wide bandwidth, etc., cascaded optical amplifiers inevitably degrades the performances of it, because the accumulated amplified spontaneous emission (ASE) noise increases post-detector noises to the receivers.

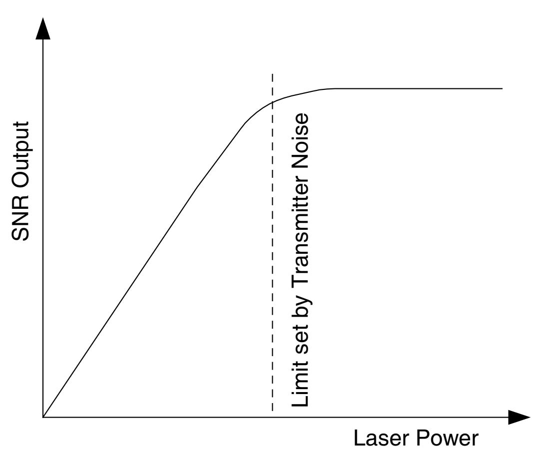


Figure 1: Limitation of direct laser diode modulation.

2. TOWARDS THE GOAL

At frequencies, above which direct modulation or external modulation has the limitation, one method [1, 2] of accomplishing optical generation of electrical signal is obtained by mixing the outputs from two very narrow line-width lasers. But it is not commercially viable because of prohibitive costs. Use of commercially available DFB lasers, although eliminates the cost, ends up with the generation of electrical signal with large line-width due to inherent laser phase noise of the each source. The spectral purity of the electrical signal can be improved if the noise terms of the optical waves are correlated. The easiest way to achieve this is to utilize the technique of sideband injection locking of two commercial laser diodes with linewidths in range of few megahertz. This is illustrated in Fig. 2. The master laser is frequency modulated to create sidebands in the laser optical spectrum. The slave lasers are synchronized to sideband components of the FM signal. The important point is that two slave lasers, used as active high-Q circuits [7], are thus phase coherent with the master laser and each other, while their frequency separation equals the integer multiples of master laser frequency modulation. The main practical limitations on optical sideband injection locking are: (1) the locking range is small (typically few MHz) so that the laser temperature must be controlled with milli-Kelvin precision, as the temperature sensitivity is typically $1.0\,\mathrm{GHz}/^{0}\mathrm{K}$, (2) increase of locking range beyond the limit set by the locking ratio of $-50\,\mathrm{dB}$ [6] after which the system becomes unstable and (3) the disturbing effect of pulling and pushing effects due to the neighbouring sidebands, that are away from the main locking sideband component by the modulating frequency only. Moreover, the increase of the strength of the injection signal decreases the effective Q-value of the injection locked oscillator [7].

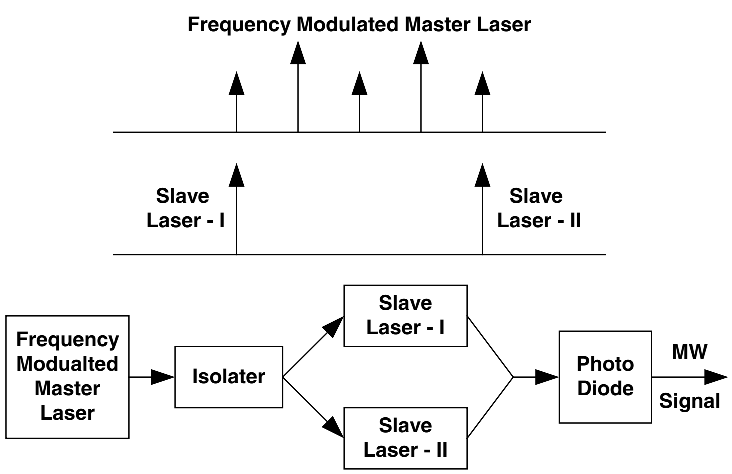


Figure 2: Coherent mixing of two optical sources.

In order to overcome this instability problem, the traditional alternative of optical injection synchronization [9], i.e., an optical phase locked loop (OPLL) is recommended. A conceptual arrangement is depicted in Fig. 3. Had there been no loop propagation delay in an optical phase locked loop, the required phase error variance can be easily realized by increasing the locking range of the OPLL. Unfortunately the presence of the finite loop propagation delay does not only increases the phase error variance but also makes the system unstable. In order to overcome the instability problem a modified OPLL has been suggested [5]. At this point it is worthwhile to note that when a semiconductor laser is used as LO laser, the frequency is obtained by changing its bias current. Therefore, RMS phase error will cause a bias current fluctuation. This bias current fluctuation in turn induces a photo-detector current fluctuation. Moreover, the photocurrent fluctuates inevitably due to shot noise effect. In most cases this residual AM noise due to bias current predominated over the shot noise. Unless this is suitably controlled through balanced-mixer-type detection, this could be hazardous in system performance.

Thus we note that though there is limitation on the locking range of the injection synchronization yet it does not add loop delay. Whereas the OPLL apparently does not have limitation on the locking range till the AM noise due to bias current predominates over the shot noise. But it is prone to instability due to the unavoidable loop delay, which in turn increases the phase error variance. Thus the solution is to combine the principle of injection synchronization and phase locking technique, which the author [3] advocated in 1971 and in 1996 [4] and later Berdonali et al, launched a detailed investigation in 1999 [6]. Thus it is expected that the combination will have an improvement in the performance of the system. The combination results in a system with low values of phase error variance (though not up to the desired value) over a much wider stable locking range. But detailed analysis shows that the effect of the loop delay, although minimized, still remains. It is known that wider the locking range, longer the mean time between cycle-slips in a tracking system. As a result, the restrictions imposed on the linewidth of the lasers are relaxed and the use of commercially available DFB lasers in place of very expensive and bulky narrow linewidth lasers is seen commercially viable. However, the need of a very low value of phase error variance for the (semi) coherent reception of digital signals with the error probability of 10^{-9} is not fulfilled with this arrangement. It is easily appreciated from the above discussion that the addition of the injection synchronization path effectively increases the loop bandwidth. As a result the tracking system becomes more faithful in following phase variation of the input signal and because of the increase of the loop bandwidth noise rejection capability becomes weak.

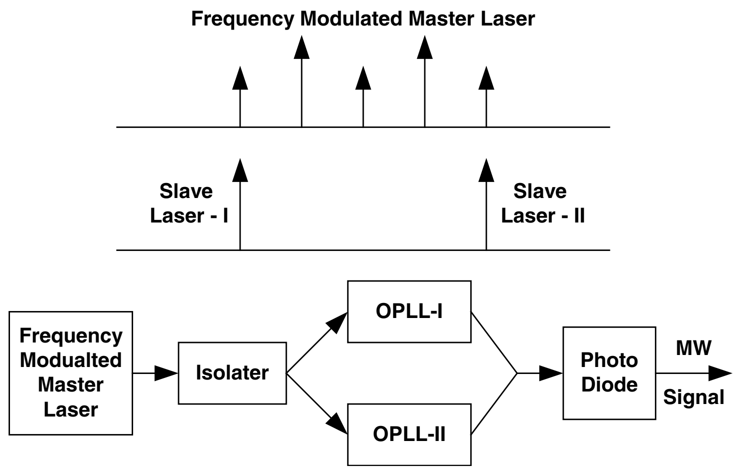


Figure 3: Coherent mixing of two optical sources using OPLL.

3. ULTIMATE NEW SYSTEM

Thus we observe two opposing tendencies, namely,

(i) Requirement of the larger locking range for faithful tracking of the phase variation of the input signal and

(ii) The need of smaller locking bandwidth for achieving better noise squelching capability of the tracking system,

an additional phase modulator is incorporated in the loop as shown in Fig. 5. The addition of the phase modulator at the output of the VCO laser reduces the effective phase error and consequently the improvements are:

1. The effective loop noise bandwidth becomes narrower resulting better noise rejection capability.
2. Due to the reduction of the effective phase error, the possibility of exceeding a phase error of 90° becomes less and so the frequency of cycle-slips is considerably reduced.
3. The instability of the system due to the loop delay is considerably reduced.

The second modification, to be observed here, is the replacement of the FM laser source by an IM laser source. The purpose is to get rid of the disturbing effect of pulling and pushing effects due to the neighbouring sideband, that are away from the main locking sideband component by the modulating frequency only. Cascaded MZ modulators are judiciously modulated so that the optical field takes character of DSB modulation and modulation is chosen in such a way that the disturbing pulling and pushing effects be considerably reduced increasing the locking capability of the signal.

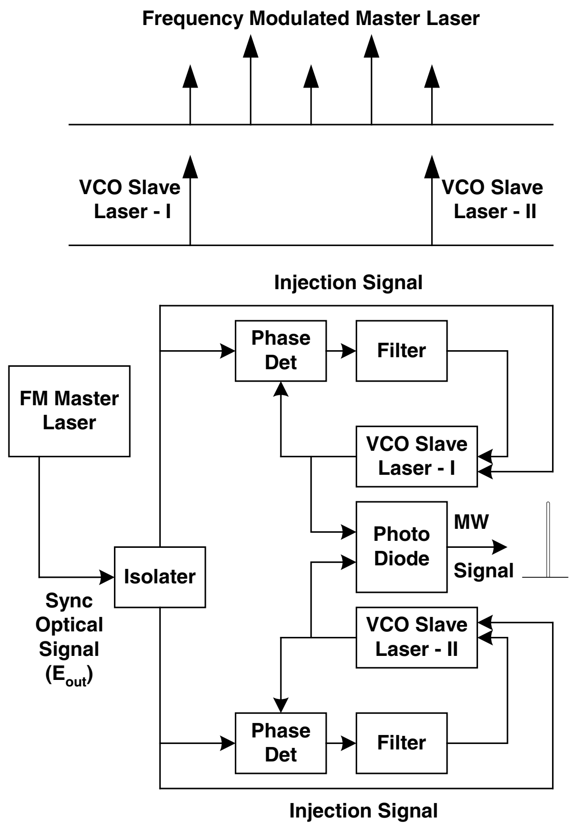


Figure 4: VCO slave lasers with injection inputs.

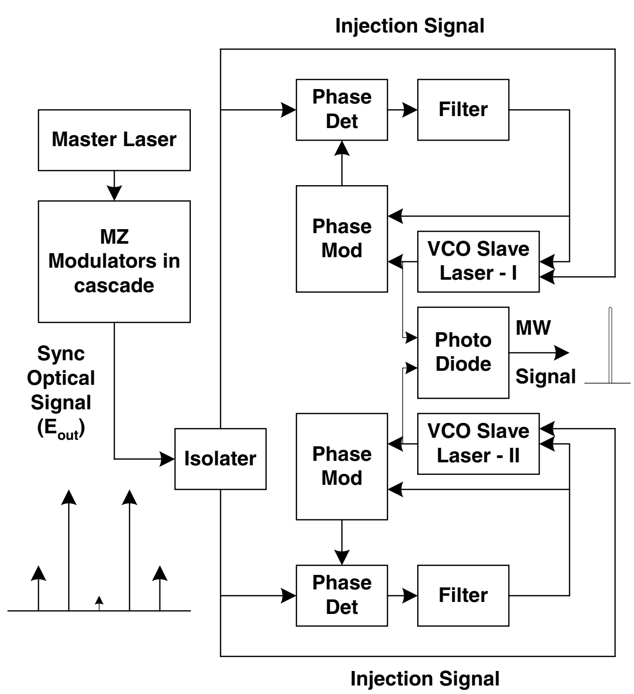


Figure 5: Injection PM optical phase locked loop.

4. NUMERICAL RESULTS

We illustrate the feasibility of a PSK scheme using DFB lasers with wide linewidth. Referring to Kikuchi's work [8], it is seen that the required rms phase error variance is 0.2 rad in order to suppress the receiver sensitivity degradation at BER of 10^{-9} below 3 dB. We presume that the summed linewidth of the lasers is 6.0 MHz. Controlling the level of injection signal to the maximum value that keeps the injection locked laser unconditionally stable [6], the variation (*as shown by dash-dot*) of the phase error variance with the delay in the optical locking system is shown in Fig. 6. For obvious reasons the variation is seen to independent of the optical delay. But it is much above the required value of the variance of the phase error (*as shown by the dash line*).

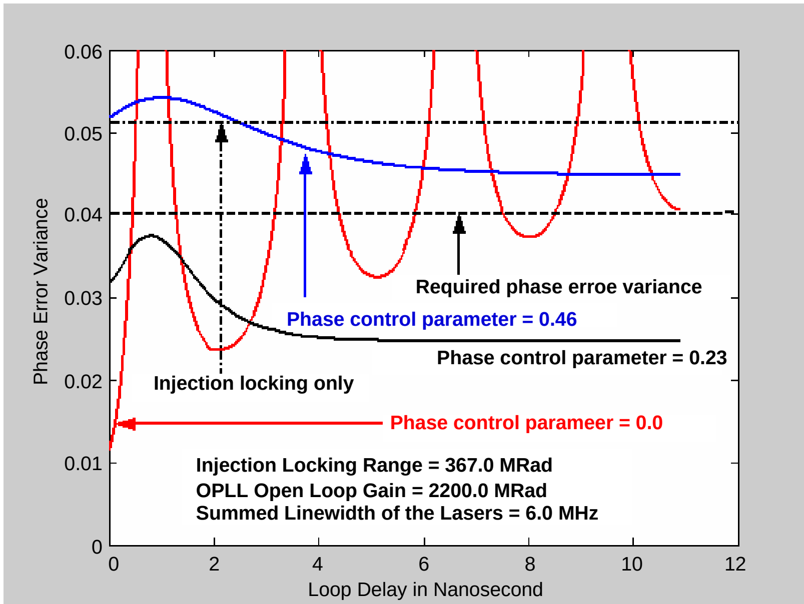


Figure 6: Dependence of phase error variance on loop delay.

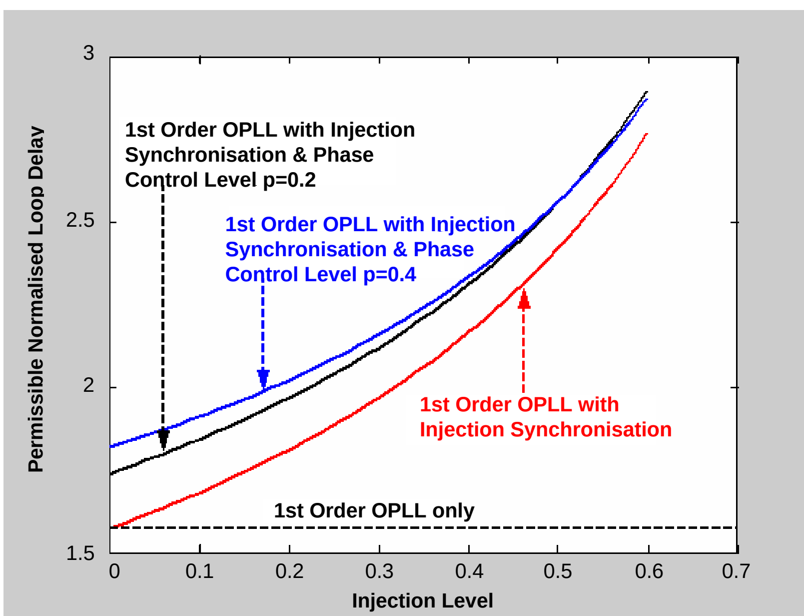


Figure 7: Permissible loop delay versus injection level.

The next we consider the optical phase locking system in place of the optical injection locking mechanism. Here we consider a first order optical phase locked loop. Since a first order loop is unconditionally stable for low value of loop delay, we choose an open loop gain of 1.468 Grad and with no injection signal. The *red line plot* shows the variation of the phase error variation with the delay of the loop. It is seen that this arrangement is applicable only up to about of 0.6 nano-second loop-delay. The figure also shows the effect of the phase control, which is seen to have very little effect. *The blue and the black lines* show the variation of the phase error variance with the phase control parameters 0.46 and 0.23 respectively. This result clearly reflects that by choosing appropriate value of the phase control parameter the phase error variance can be kept much below the required level. Moreover, the proposed system can tolerate very large value of the delay in the optical phase locked loop. Fig. 7 indicates the improvement of the loop stability when injection synchronization and additional phase control are incorporated. It also indicates the choice of the injection locking range in relation to the OPLL loop gain.

ACKNOWLEDGMENT

The author is grateful to Professor J. Banerjee, Chairman of the Academy of Technology for his encouragement and financial assistance during the course of the work.

REFERENCES

1. Braun, R. P, G. Grossskopf, D. Rhode, and F. Schmidt, "Fiber optic millimeter-wave generation and bandwidth efficient data transmission for broadband mobile 18–20 and 60 GHz band communication," *IEEE Intl. Microwave Photon Topical Meeting*, 235–238, Essen, UK, 1997.
2. O'Reilly, J. J, P. M. Lane, and M. H. Capstick, "Optical generation and delivery of modulated mm-wave for mobile communication," *Analog Optical Fiber Communications*, IEE, 229–256, 1995.
3. Biswas, B. N, "Combination injection locking with indirect synchronization technique," *IEEE Trans. Commn.*, Vol. COM-19, 574–576, August 1971.
4. Biswas, B. N, P. Lahiri, D. Mondal, and S. Biswas, "Injection synchronized optical phase locked loop," *Proceedings, XXV General Assembly URSI*, 643, Lille, France, August 25–September 05, 1996.
5. Biswas, B. N, A. P. Sinha, P. Lahiri, and D. Mondal, "A novel optical phase locked loop," *Proceedings, XXV General Assembly URSI*, 164, Kyoto, Japan, August 25–September 02, 1993.
6. Berdonalli, A. S, C. Walton, and A. J. Seeds, "High performance phase locking of wide linewidth semiconductor lasers by combined use of optical injection locking and optical phase locked loop," *J Lightwave Technology*, Vol. 17, No. 2, 328–342, February 1999.
7. Biswas, B. N., *Phase Lock Theories and Applications*, Oxford and IBH, New Delhi, 1988.
8. Kakuchi, K., T. Okoshi, M. Nagamatsu, and M. Henmi, "Degradation of bit error rate in coherent communication due to spectral spread of the transmitter and local oscillator," *J Lightwave Technology*, Vol. LT-2, No. 6, 1024–1033, December 1984.
9. Biswas, B. N., "On the injection synchronization," *Proc. IEEE (Letters)*, Vol. 40, 880, June 1966.

Characterization of a Self-complementary Sierpinski Gasket Microstrip Antenna

B. N. Biswas[1], **Rowdra Ghatak**[2], **Rabindra K. Mishra**[3], **and Dipak R. Poddar**[4]

[1]Academy of Technology, P.O Aedconagar, Adisaptagram, Hoogly, West Bengal-712121, India
[2]Physics Department, Burdwan University, Burdwan-713104, West Bengal, India
[3]ALCEAST, Sambalpur University, Burla, Sambalpur-768018, Orissa, India
[4]Electronics and Telecommunication Engineering Department, Jadavpur University, Kolkata-32, India

Abstract— Fractal antennae are recent topic of research interest. This work reports a self-complementary Sierpinski Gasket fractal microstrip patch antenna. The first iteration is named CSMASLI1, which stands for Complementary Sierpinski Microstrip Antenna Same Layer Iteration 1 and similarly the second iteration is named as CSMASLI2. The antennas are analyzed for two iterations using commercially available electromagnetic simulation software IE3D$^{\text{TM}}$ and CST Microwave Studio$^{\text{TM}}$. The antenna radiates prominently at all resonant frequencies. The impedance bandwidth increases at higher harmonics. The radiation pattern is consistent for the different resonant frequencies and for different iterations. However some undulations are observed at higher resonance. Thus this new design can have usage in vehicle-mounted antenna for multiband wireless systems. A prototype was experimented which matches simulated results.

1. INTRODUCTION

Fractal antenna provides a paradigm shift to the conventional planner antenna technology. While the later is in commercial practice, the former is in the process of descending to this practice from the novelty of academics. In general, they exhibit multi resonant frequencies like a log periodic antenna. They can be thought of as a special case of log periodic antenna which folds inward. Sierpinski fractal monopole antenna has been reported in [1]. The self-complementary Sierpinski gasket monopole was reported in [2]. Empirical formulae of resonant frequency and input impedances for some geometry are available in literature [1, 3]. Some other fractal structures reported in [4] include self-complementary versions of the Koch tie dipole and the Gosper Island dipole. The dual layer complementary Sierpinski microstrip antenna has been reported by the present authors in [5].

Here we have tried to implement the benefits of the complementary structure on the same plane and came up with the Sierpinski quasi-complementary microstrip pre-fractal patch antenna. The feed point was judiciously chosen to provide good match at the resonant frequencies. However the behavior replicates with higher iteration, as they are basically self-similar structures.

2. DESIGN

The antenna structure is generated by an initial right-angled triangle with base s and height h as shown in Figure 1. The right half plane is the complementary of the left half plane. The fundamental triangle is tessellated to form the entire geometry.

The parameters h and s becomes half with every iteration. Thus the number of replications increases to form the geometry. For example the first iteration is made of 8 such right-angled triangles. This gives a surface area for first iteration as $a_1 = 4hs$. For the second iteration the new height and base is half of h and s respectively. But now we require 30 such triangles to generate the structure. So area of the second iteration geometry is $a_2 = \frac{1}{8}hs \times 30 = 3.75hs$ and for third iteration it is $a_3 = 3.5hs$. Thus in general the area for various iteration is given as

$$a \cong \left.\frac{4}{1.076^{I-1}}hs\right\} I = 1, 2, 3 \dots \tag{1}$$

In the Equation 1, I denotes the iteration. For increasing iteration the surface area decreases but the number of corners and edges increases and the current is allowed to meander longer distances. This however increases the electrical length. The geometry is constructed on a CuCLAD substrate of $\varepsilon_r = 2.54$ and $\tan\delta = 0.0009$. The thickness of the substrate is 1.59 mm. The starting value of s and h is 7.5 mm and 12.9904 mm respectively. The feed is via a SMA coaxial probe arrangement as shown in Figure 1 with a black dot.

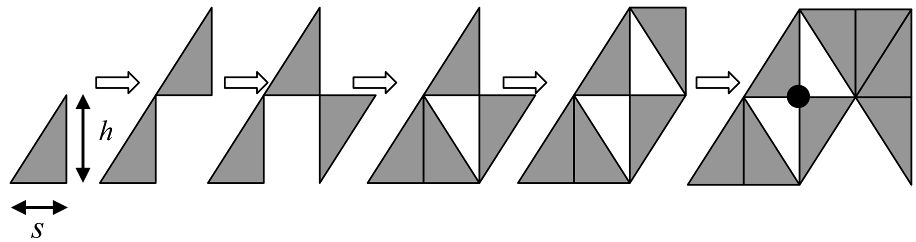


Figure 1: The starting right-angled triangle of base s and height h. The steps followed in tessellating the fundamental triangle to form the entire geometry of first iteration are also shown.

3. RETURN LOSS CHARACTERISTICS

The antenna was simulated in IE3D and CST Microwave Studio till second iteration. The same were experimented to obtain the S11 characteristics using a HP8722C vector network analyzer as shown in the Figure 2. In CST Microwave Studio we have modeled a coaxial feed arrangement to resemble

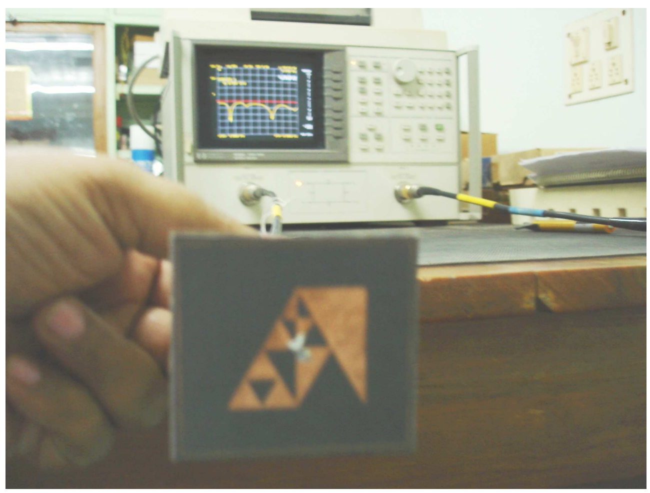


Figure 2: Setup of HP 8722C VNA for measuring return loss measurement of CSMASLI2.

the actual fabricated structure. The simulated and experimental return loss is compared for second iteration in Figure 3. This structure considerably enhances the resonance peaks at the lower frequencies for second iteration, in the band of interest, compared to a simple Sierpinski antenna. From experimental results we see that for the second iteration impedance bandwidth increases from 2% at the first resonance to 4% at the third resonance. The simulated and experimental results are in good agreement.

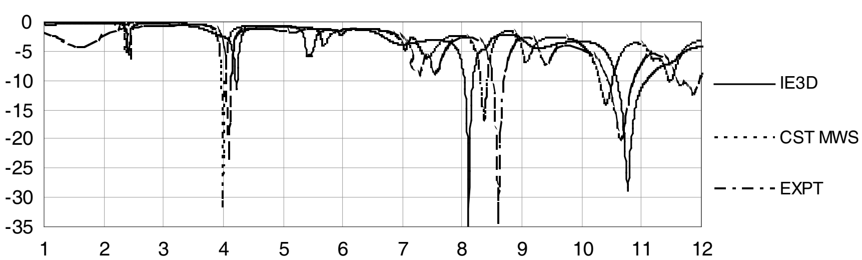
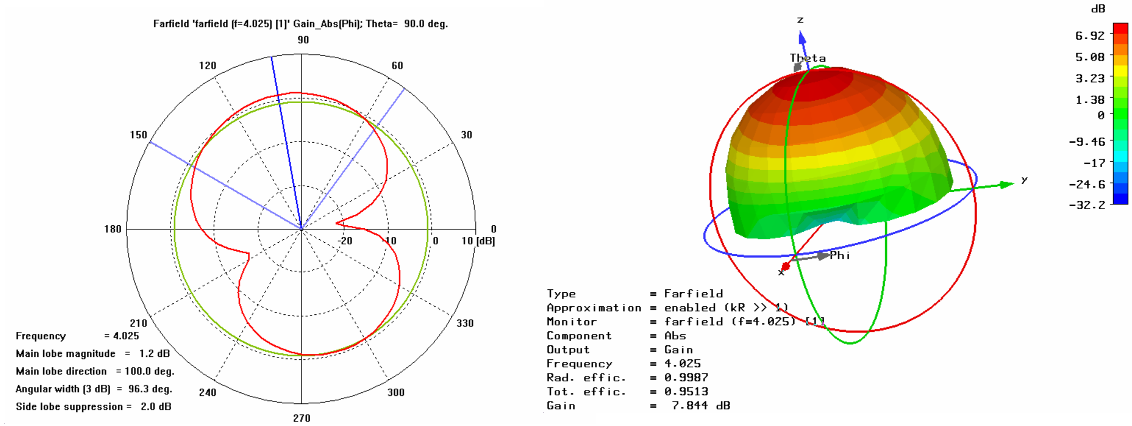


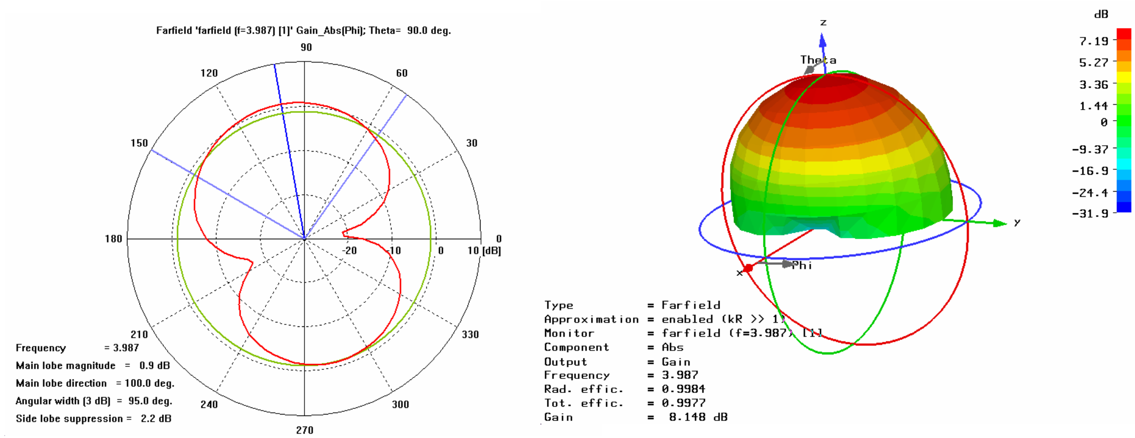
Figure 3: Return loss of self-complementary Sierpinski microstrip antenna of second iteration.

4. RADIATION CHARACTERISTICS

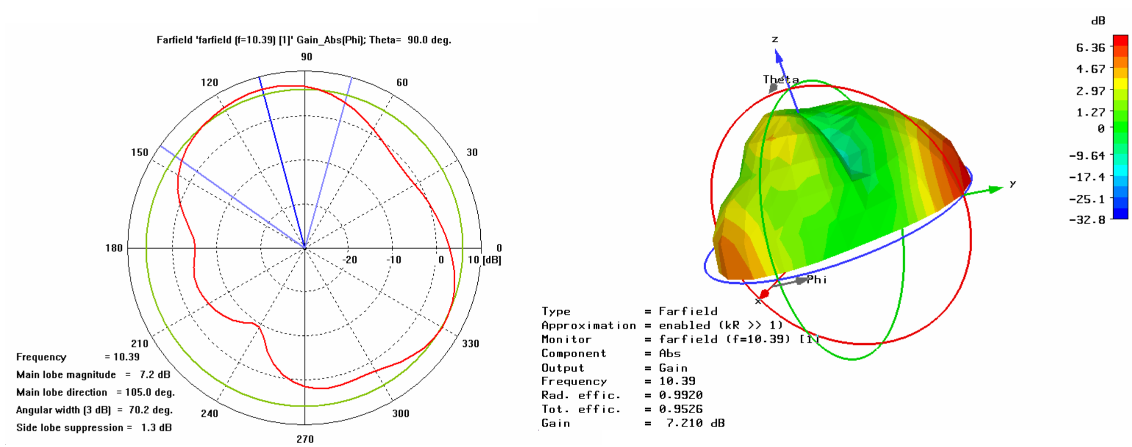
The radiation patterns of CSMASLI1 at 4.025 GHz and CSMASLI2 at 3.987 GHz are presented in Figure 4(a) and (b). The last resonance at 10.39 GHz of CSMASLI2 is shown in Figure 4(c). The radiation pattern is consistent for the different resonant frequencies and for different iterations as seen in Figure 4. An interesting phenomenon is observed that with increasing resonant frequencies the patterns show some undulations. This can be physically explained by taking in to account the diffraction effect at higher frequencies.



(a)



(b)



(c)

Figure 4: **E**-Phi and **E**-Total radiation pattern for (a) first frequency $f = 4.025$ Ghz of first iteration (b) first resonance $f = 3.987$ GHz of second resonance (c) third resonance of second iteration $f = 10.39$ GHz.

5. CONCLUSION

Experimental and simulated results of CSMASLI1 and 2 have been presented. A general approach to the design of the antenna has been reported in this paper. The resonant frequencies are prominent at all resonance. This property is however not seen in the Sierpinski geometry microstrip patch antenna as well as monopoles where the lower resonances are less emphatic. The present structure, which is a variation from conventional fractal geometry, gets focused as possible candidate for true multiband operation. The physical area of the antenna decreases with increasing iteration but the number of resonance increases. The radiation characteristics show uniformity for both the iterations and among different resonance in a given iteration. The overall characteristics support its inclusion in the family of fractal antennas and its variations.

REFERENCES

1. Baliarda, C. P., J. Romeu, R. Pous, and A. Cardama, "On the behavior of the Sierpinski multiband fractal antenna," *IEEE Transactions on Antennas and Propagation*, Vol. 46, No. 4, 517–524, April 1998.
2. Siah, E. S., B. L. Ooi, P. S. Kooi, and X. D Xhou, "Experimental investigation of several novel fractal antennas-variants of the sierpinski gasket and introducing fractal fssscreens," *IEEE CNF*, 170–173, 1999.
3. Musliake, Y., "Self-complementary antennas," *IEEE Transactions on Antennas and Propagation*, Vol. 34, No. 6, 23–29, Dec. 1992.
4. Gonzalez-Arbesu, J. M. Rius, and J. M. Romeu, "Some pre-fractal self-complementary antennas," *IEEE APS Int. Symposium*, Vol. 4, 3449–3452, 20–25 June, 2004.
5. Ghatak, R., R. K. Mishra, D. R. Poddar, and A. Patnaik, "Multilayered complementary quasi-fractal sierpinski patch antenna for wireless terminals," URSI XXVIII GA Poster Presentations Programme, 95, 2005.

Characteristics Study of Four Coplanar Waveguide Feeding Devices

Wenwen Chai[1,2], Xiaojuan Zhang[1], and Jibang Liu[1,2]

[1]Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China
[2]Graduate School of Chinese Academy of Sciences, Beijing 100039, China

Abstract— Four types of coupling slots on the coplanar waveguide (CPW) are studied and their effects on resonance frequency and bandwidth of microstrip antennas are simulated by the finite-difference time domain (FDTD) method. The simulation results are explained by the near field distributions. Feeding characteristics of four CPW coupling devices obtained by analysis and comparison greatly improve the flexibility of antennas design and will be very useful for wireless communication.

1. INTRODUCTION

The microstrip patch antennas have a lot of merits such as low-cost, low-profile, conformability, and ease of manufacture, which make them very attractive. The primary barrier to implementing these antennas in many applications, however, is their limited bandwidth — only on the order of a few percent for a typical patch radiator [1, 2]. Because of this fact, much work has been devoted to increasing the bandwidth of microstrip patches. A straightforward method is to use a thick substrate with a low dielectric constant for the antenna. Another technique which can be implemented in the aperture coupled configuration is to use a near-resonance aperture in combination with the thick antennas substrate [3, 4] or multilayer substrates to achieve wide frequency band. But the common aperture coupled structure where the slot in the ground plane is fed via a microstrip line on an additional substrate layer which complicates the antenna design. Coplanar waveguides have been suggested as an alternate to microstrip-line for feeding the microstrip antenna [5], and they have been used increasingly in the design of millimeter-wave microstrip antennas.

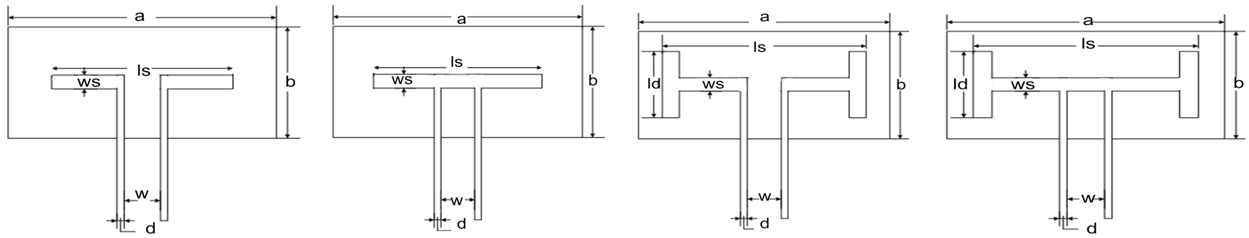
However, there haven't been yet any special papers which systematically analyze feed characteristics of the coplanar waveguide with coupling slots. Although the coplanar waveguide with a rectangular slot has been studied in some papers [6], the effects of feeding structure dimensions on the antenna performance aren't presented concretely. In this paper, the feeding performance of coplanar waveguides with rectangular, dual-T shape, and H-shape slots are studied respectively, and the relations between antennas structure and behavior are shown and explained by the near field distribution.

2. MODEL ANALYSIS

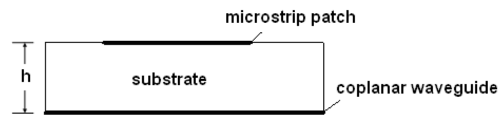
In a coplanar waveguide-fed microstrip antenna, the antenna and coplanar line are placed on the opposite sides of the same dielectric substrate and the coupling from the coplanar line to microstrip antenna is accomplished via a slot in the ground plane connected directly to the end of the coplanar line. In general, slot coupling may involve an electric polarisability, a magnetic polarisability, or both. In slot coupling to a microstrip antenna, the magnetic polarisability is the dominant mechanism for a slot near the center of the patch [7]. Because the polarisabilities strongly depend on the shape of the slot as well as the size, it is desirable to improve the antenna performance by optimizing the shape and size of coupling slot for given antenna dimensions.

Microstrip patch antennas model fed by the CPW are shown in Fig. 1. Antennas A, B, C, D are fed respectively through inductive rectangular, capacitive rectangular, dual-T shape, and H-shape slots. The width ws of parallel and vertical slots which equals to 1 mm affects the coupling level very weakly, so antennas performances are analyzed only by changing the slot length ls and ld ($ld \geq ws$). The tendency for the variation of parameters is derived by FDTD method. In the FDTD simulation, the selected space step lengths are $\Delta x = 0.1\,\text{mm}$, $\Delta y = 0.2\,\text{mm}$, and $\Delta z = 0.2\,\text{mm}$, respectively, and the selected time step length is to $\Delta t = 0.27\,\text{ps}$, satisfy the courant stability condition. A modulated Gaussian pulse is used as the excitation source. The perfectly matched layer (PML), introduced by Gedney, is used to truncate the FDTD lattices.

Figure 2 shows the evolution of resonant frequency and impedance bandwidth ($VSWR < 2$) of *antenna A* as a function of the slot length (ls). The resonant frequency decreased nearly linearly with the increase of ls. The bandwidth was greater than 0 only when ls was limited to a span ($ls_{\min}, ls_{\max}$), and the maximum bandwidth was achieved as ls was around $(ls_{\min} + ls_{\max})/2$.

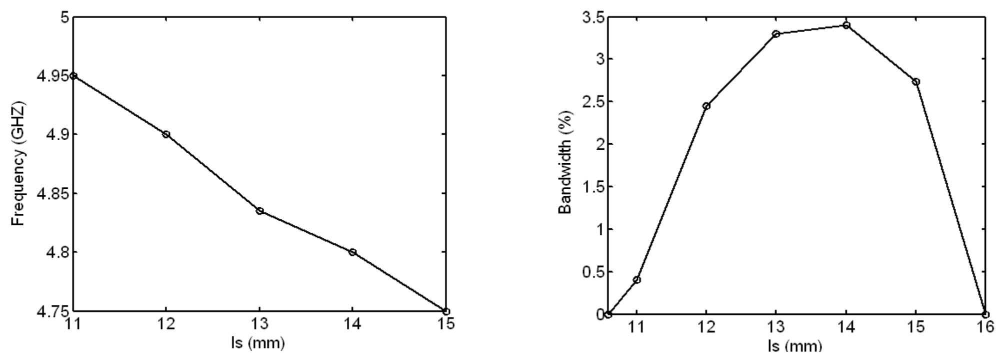


(a) Antenna A (b) Antenna B (c) Antenna C (d) Antenna D



(e) Side view

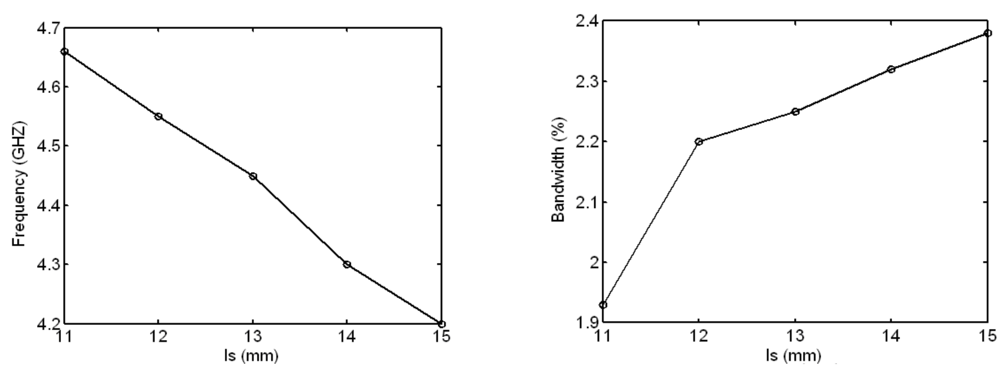
Figure 1: Microstrip patch antenna fed respectively by CPW with different shape slots.
Fixed dimensions: $a = 23$, $b = 17.6$, $\varepsilon = 2.2$, $h = 1.6$, $ws = 1$, $w = 3$ unit: mm.



(a) Resonant frequency (b) Impedance bandwidth

Figure 2: Behavior evolution of *antenna A* as a function of the slot length (ls).

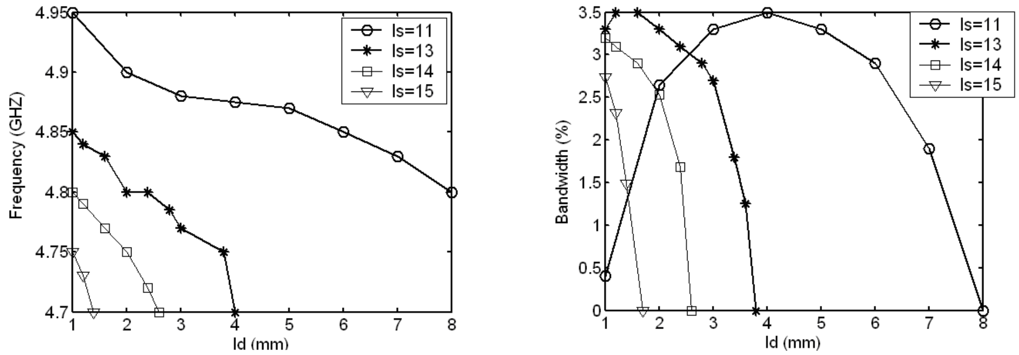
Figure 3 shows the influence of the length of the slot (ls) on resonant frequency and impedance bandwidth of *antenna B*. As ls increased, the resonant frequency decreased linearly which was similar to *antenna A*. But bandwidth curves of two antennas were different obviously. From Fig. 3(b) it was seen that the bandwidth varied relatively weakly, and ls could guarantee good antenna bandwidth in a large span.



(a) Resonant frequency (b) Impedance bandwidth

Figure 3: Behavior evolution of *antenna B* as a function of the slot length (ls).

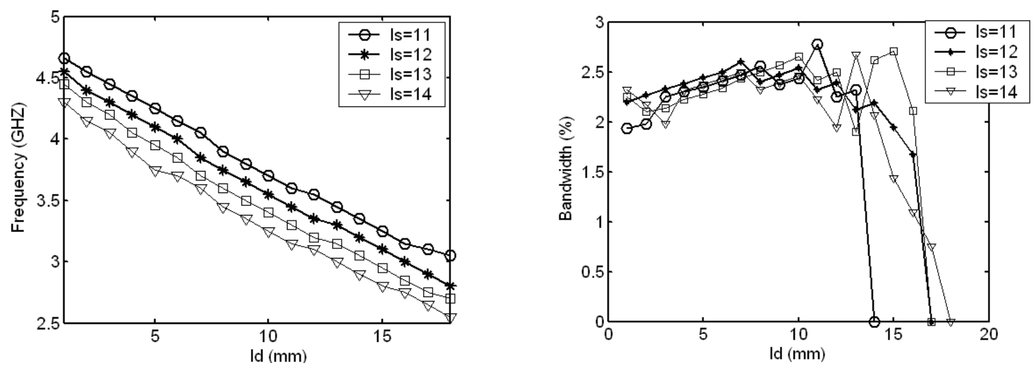
The influence of the length of the vertical feeding slot (ld) of *antenna C* at several ls value points is represented in Fig. 4. Resonant frequency was reduced as ls or ld increased, and it was seen that the resonant frequency changed with ld more strongly if ls became larger. When ls was fixed, ld had a max value $ld_{\max}$ that could assure bandwidth greater than 0, and ld_{max} became smaller with the increase of ls. The maximum bandwidth value achieved about at $ld_{\max}/2$ varied very little when $ls < 0.22\lambda_0$ ($\lambda_0 \approx 62.5\,\mathrm{mm}$), but when $ls \geq 0.22\lambda_0$ the maximum bandwidth which would be achieved at $ld = ws$ was obviously reduced with the increase of ls.



(a) Resonant frequency (b) Impedance bandwidth

Figure 4: Behavior evolution of *antenna C* as a function of the vertical slot length (ld).

Figure 5 presents the resonance and bandwidth characteristics of *antenna D* as a function of the length of vertical feeding slot (ld) at several ls value points. As ls or ld was increased, the resonant frequency decreased linearly. The bandwidth curve differed sharply with that of *antenna C* because ls and ld could assure good antenna bandwidth in a large span as shown in Fig. 5(b).



(a) Resonant frequency (b) Impedance bandwidth

Figure 5: Behavior evolution of *antenna D* as a function of the vertical slot length (ld).

3. RESULT COMPARISON AND DISCUSSION

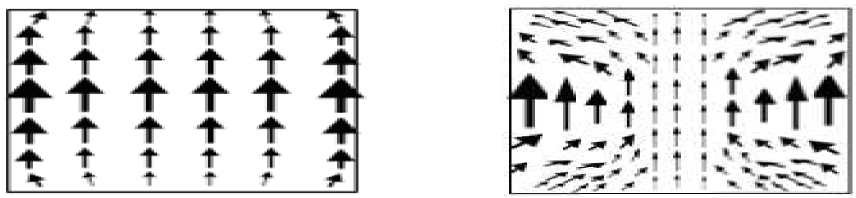
Dual-T shape slot and H shape slot are both formed by loading the rectangular slot, while the difference lies in their connection way with coplanar waveguide: the former is inductive, while the latter capacitive. Based on above result curves, it can be also concluded:

(1) with the same feed slot, the bandwidth and gain of inductively fed antennas are better than that of capacitively fed antennas, but the resonant frequency of the former is greater.

(2) when the length of feed slot (ls) is fixed, dual-T shape slot and H shape slot on coplanar waveguide can attain far greater coupling level than rectangular slot by appropriately tuning the length ld of vertical feed slot, therefore resulting in a wider impedance band.

(3) For a given antenna, different feeding types can lead to different resonant frequencies. For example, when $ls = 13\,\mathrm{mm}$ and $ld = 1\,\mathrm{mm}$, 400 MHZ frequency shift can be achieved from *antennas A* to *antennas B*, while 500 MHZ from *antennas C* to *antennas D*. Therefore, it is reasonable to tune or switch the frequency of operating using a varactor diode that can modify the coupling level in order for the antenna optimization.

It is visualized to analyze the effect of coupling slot on resonant frequency and bandwidth of antennas by the near field distribution. The current distributions on the patch without slot and with slot on the coplanar waveguide are presented respectively in Fig. 6(a) and (b). It is obvious that the existence of feed slot makes patch current mainly concentrate between the slot and patch edge, and the equivalent current path deviate more seriously as ls becomes larger. The deviation could lengthen the equivalent path and hence reduce the resonant frequency. On the other hand, for inductively fed antennas, oversize and undersize slots both lead to impedance mismatch, decreasing coupling degree and increasing energy loss, so antennas impedance bandwidth gradually reduced to 0 with the variation of slot size. Moreover, oversize slot also increases back radiation, consequently reducing antennas efficiency.



(a) Without slot (b) With slot on CPW

Figure 6: Current distribution on the patch.

4. CONCLUSIONS

Resonant frequency and impedance bandwidth of antennas fed by four types of coupling slots on coplanar waveguide are studied and compared. Feed rules of these coupling devices are attained and explained by near field distribution. The antennas can be designed flexibly based on the experimental conclusion. The freedom in modifying the coupling simply by tuning the varactor diode helps to realize the optimum impedance characteristics in practice.

REFERENCES

1. Carver, K. and J. Mink, "Microstrip antenna technology," *IEEE Trans. on AP*, Vol. 29, No. 1, 2–24, 1981.
2. Lo, Y., D. Solomon, and W. Richards, "Theory and experiment on microstrip antennas," *IEEE Trans. on AP*, Vol. 7, No. 2, 137–145, 1979.
3. Croq, F., A. Papiernik, and P. Brachat, "Wideband aperture coupled microstrip subarray," *IEEE Antennas and Propagation Symp. Dig.*, Vol. 3, 1128–1131, 1990.
4. Sullivan, P. and D. Schaubert, "Analysis of an aperture coupled microstrip antenna," *IEEE Trans. on AP*, Vol. 34, No. 8, 977–984, 1986.
5. Aksun, M. I., S. L. Chuang, and Y. T. Lo, "On slot coupled microstrip antennas and their applications to cp operation-theory and experiment," *IEEE Trans. Antennas Propagat.*, Vol. 38, No. 8, 1224–1230, 1990.
6. Deng, S.-M. and M.-D. Wu, "Impedance characteristics of microstrip antennas excited by coplanar waveguides with inductive or capacitive coupling slots," *IEEE Microwave and Guide Wave Letters*, Vol. 5, No. 11, 391–393, 1995.
7. Pozar, D. M., "A microstrip antenna aperture coupled to a microstrip line," *Electron. Lett.*, Vol. 21, 49–50, 1985.

Broadband Microstrip Patch Antenna Fed by a Novel Coupling Device

Wenwen Chai[1,2], **Xiaojuan Zhang**[1], **and Jibang Liu**[1,2]

[1]Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China
[2]Graduate School of Chinese Academy of Sciences, Beijing 100039, China

Abstract— Two broadband microstrip patch antennas are designed here. Both are fed by the CPW with an H-shape coupling slot which is modified by the conversion idea. The former has a bandwidth of 26.5% and a gain of more than 8 dBi in the operating frequency band, while the latter attain 40% impedance bandwidth and 20% gain bandwidth above 6 dBi. All result data prove the validity of designed models.

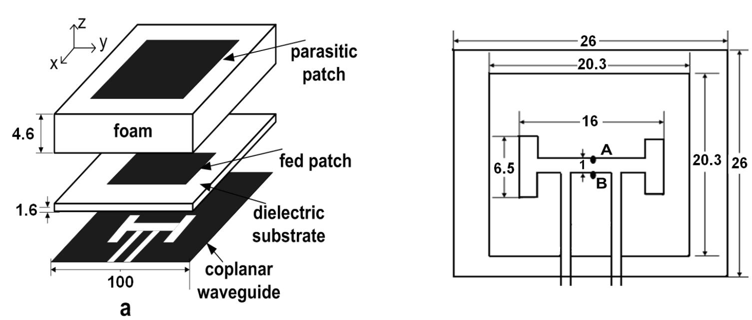
1. INTRODUCTION

Microstrip antennas have been extensively used in the past because of their several advantages (low weight, low cost, conformability, and low profile) over conventional antennas. However, their small impedance bandwidth has always been a major constraint as it limits the frequency range over which the antenna can perform satisfactorily. For broadening the bandwidth, in this paper we propose two broadband antennas where a coplanar waveguide with H-shape coupling slot and a varactor diode is used as a novel feeding system. This type of antennas not only possesses the same technological advantages as conventional coplanar waveguide-fed microstrip antennas [1–3] such as low radiation loss, less dispersion and unipolar configuration, but has two extremely attractive advantages: First, various experiments are carried out to show that H-shaped aperture has a higher resonant impedance implying larger coupling than rectangular aperture and "bowtie"-shaped aperture of the same length. Therefore, it is possible to obtain maximum coupling for the smallest slot area to reduce back radiation; Secondly, the freedom in modifying the coupling simply by tuning the varactor diode helps to realize the optimum impedance characteristics in practice. And because this tuning process doesn't involve any change of antenna structure, good antennas performance can be obtained very easily.

2. STACKED *ANTENNAS A* DESIGN AND RESULT

Dual-patch [4] *antennas A* fed by the conversion coupling device is designed here. The model is shown in Fig. 1. The fed patch is printed on one side of the substrate material with relative permittivity of 2.2 and thickness of 1.6 mm, while the coplanar waveguide with a H-shape slot is arranged on the opposite side of the substrate. Two pins of a varactor diode are connected to points A and B on the slot, respectively. The parasitic patch is separated by 4.6 mm from the fed patch. The middle dielectric layer is made of foam with relative permittivity approximately equal to 1 and thickness equal to the separation between the fed patch and parasitic patch.

The return-loss characteristics of the antenna are show in Fig. 2. Two minima are apparent, where the first resonance is almost completely due to interactions between the lower patch and the aperture, while the second one is due to a coupled resonance between the two patches. This



(a) Three dimensional view (b) Top view

Figure 1: Geometry of proposed broadband *antenna A*.

postulation is supported by the fact that the second resonance is least affected by tuning the varactor diode, while the first resonance develops with the change in the capacitance value. Therefore, the frequency of operation can be switched over a wide range by modifying the coupling between the lower patch and the aperture using the varactor diode. As a result, a bandwidth ($VSWR \leq 2$) of 26.5% in the frequency range of 3.6–4.7 GHZ is achieved. It is shown that this feeding approach results in a large increasing in bandwidth over the typical stacked patch.

The gain characteristics of *antennas A* are shown in Fig. 3 as a function of frequency. The gain is more than 8 dB and varies less than 1 dB over the broadband frequency region of 25% as shown in the figure. Surface wave efficiency is about 80% over the entire band.

The radiation patterns on E- and H-plane of *antenna A* are shown in Fig. 4 at edge frequency points. The back-radiation levers are all suppressed to less than -20 dB.

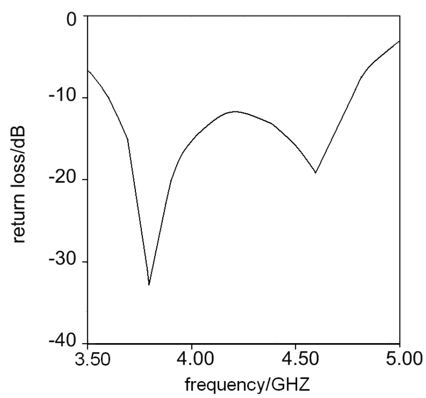


Figure 2: Return loss characteristics.

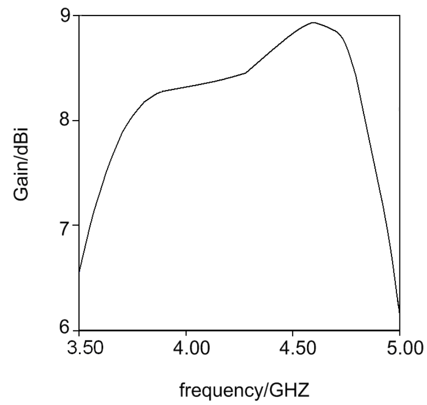
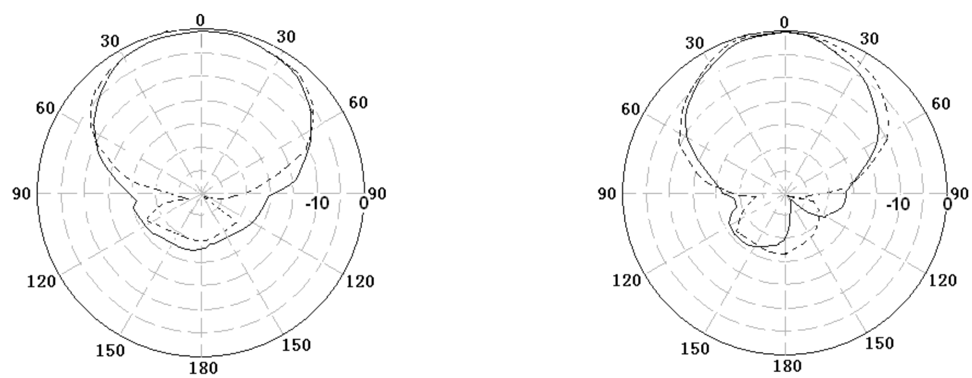


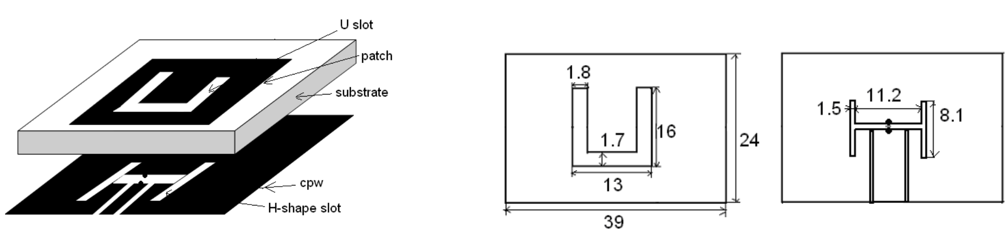
Figure 3: Gain characteristics.



(a) 3.7 GHz

(b) 4.7 GHz

Figure 4: Radiation patterns —E-plane —H-plane.



(a) Three dimensional view

(b) Top view

Figure 5: Geometry of proposed broadband *antenna B*.

3. U-SLOT PATCH *ANTENNAS B* DESIGN AND RESULT

U-shaped patch [5] *antennas B* fed by the conversion coupling device is shown in Fig. 5(a). The patch is printed on one side of the substrate material with relative permittivity of 2.2 and thickness of 6.6 mm. The detailed dimensions of the patch and feed device are presented in Fig. 5(b). The size of the patch and feed slot and capacity value are tuned appropriately to optimize antennas performance.

The *VSWR* curve of fabricated wideband *antenna B* is shown in Fig. 6. The bandwidth ($VSWR \leq 2$) is about 40% in the frequency range of 1.85–2.8 GHZ. The measured gain versus frequency curve is shown in Fig. 7 with maximum gain of 6.8 dBi, and the bandwidth of gain above 6 dBi is about 20%. Fig. 8 shows the radiation patterns of *antennas B* in E-plane and H-plane. All these data verify the success of developing a wideband U-slot patch antenna.

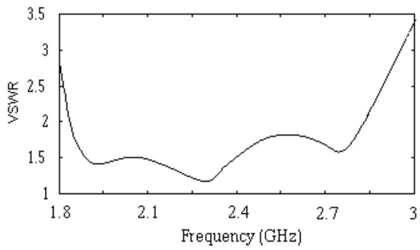


Figure 6: Return loss characteristics.

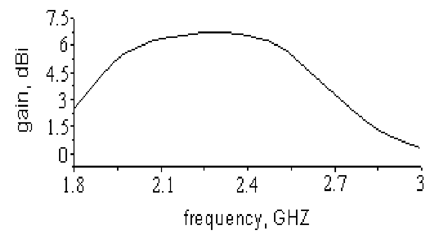
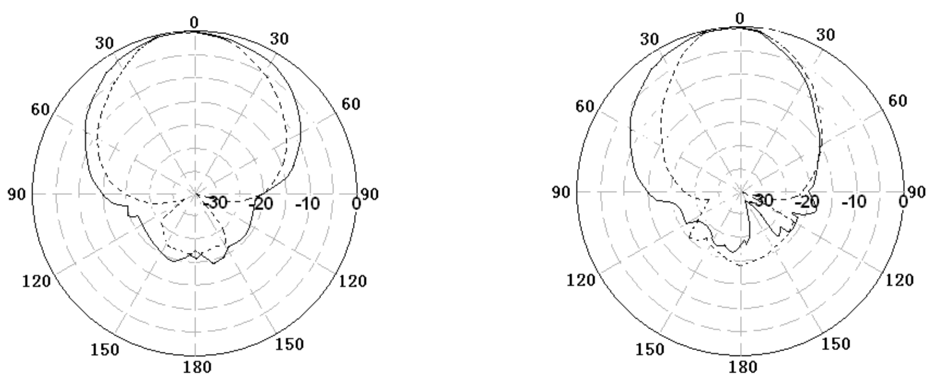


Figure 7: Gain characteristics.



(a) $f = 2\,\mathrm{GHz}$ (b) $f = 2.6\,\mathrm{GHZ}$

Figure 8: Radiation patterns on E- and H-plane of antenna.

4. CONCLUSIONS

Two broadband microstrip antenna fed by coplanar waveguide with a H-shape slot and a varactor diode is designed. The experimental results indicate that the two microstrip antenna have extremely good broadband performance in return loss, gain characteristics and radiation patterns. Therefore, this novel type of broadband antenna which can be optimized conveniently, is considered to be applicable as a candidate in mobile communication systems.

REFERENCES

1. Liu, H. C., T. S. Horng, and N. G. Alexopoulos, "Radiation of printed antennas with a coplanar waveguide feed," *IEEE Trans. Antennas Propagat.*, Vol. 43, 1143–1148, 1995.
2. Sierra-Garcia, S. and J. J. Laurin, "Study of a CPW inductively coupled slot antenna," *IEEE Trans. Antennas Propagat.*, Vol. 47, 58–64, 1999.
3. Tong, K. F., T. Matsui, and M. Izutsu, "Wideband coplanar waveguide fed coplanar patch antenna," *IEEE MTT-S Int. Microwave Symp. Dig.*, 406–409, 2001.
4. Nishiyama, E., M. Aikawa, and S. Egashira, "Stacked microstrip antenna with wide bandwidth and high gain," *IEEE Trans Antennas Propagat.*, Vol. 151, No. 2, 143–148, 2004.
5. Huynh, T. and K. F. Lee, "Single-layer single-patch wide band microstrip antenna," *Electronic. Lett.*, Vol. 31, No. 16, 1310–1312, 1995.

A Novel Wideband Antenna Design Using U-slot

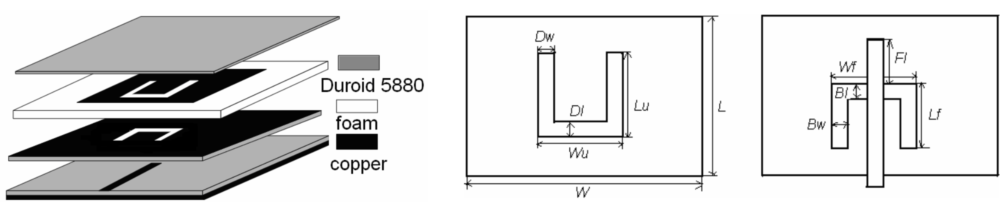
Wenwen Chai[1,2], **Xiaojuan Zhang**[1], **and Jibang Liu**[1,2]

[1] Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China
[2] Graduate School of Chinese Academy of Sciences, Beijing 100039, China

Abstract— U-slot patch antennas with Π-shaped feed slot are studied and numerical results based on the FDTD method are presented. The effects of varying physical parameters of the structure are investigated with a goal of understanding the coupling among different resonators. It is found that the U-slot patch antenna can be designed to attain 50% impedance as well as 30–40% gain bandwidths. By altering the size of U-slot and feed slot, the wideband characteristic can be changed into a dual-frequency characteristic.

1. INTRODUCTION

It is well known that microstrip antennas have very narrow impedance bandwidth, typically a few percent. One of the methods of widening the bandwidth is to cut a U-shaped slot on the patch of the coaxially fed rectangular patch antenna [1]. The other method is to use a U-slot patch proximity coupled by microstrip feed line terminated with a novel Π-shaped stub [2]. The design of microstrip antennas at microwave frequencies is closely related to the feeding technique. There are several problems associated with classical feeding techniques, such as coaxial probe or proximity feeds: 1) performance can be severely degraded by the size of the feed, which can be comparable to the size of the patch itself, and 2) soldering of probe-feeds is prone to repeatability problems. On the other hand, the aperture coupled feeding technique [3] has intrinsic properties which make it an attractive feature for microwave applications [4]. In this paper, U-slot aperture-coupled microstrip antenna which makes three resonances is presented in Fig. 1. The relationships between antenna structure and behavior are analyzed using its resonant frequencies. Based on this study, the feasibility of wideband U-slot patch antenna is demonstrated at Ku-band both theoretically and experimentally.



(a) Three-dimensional view (b) Top view

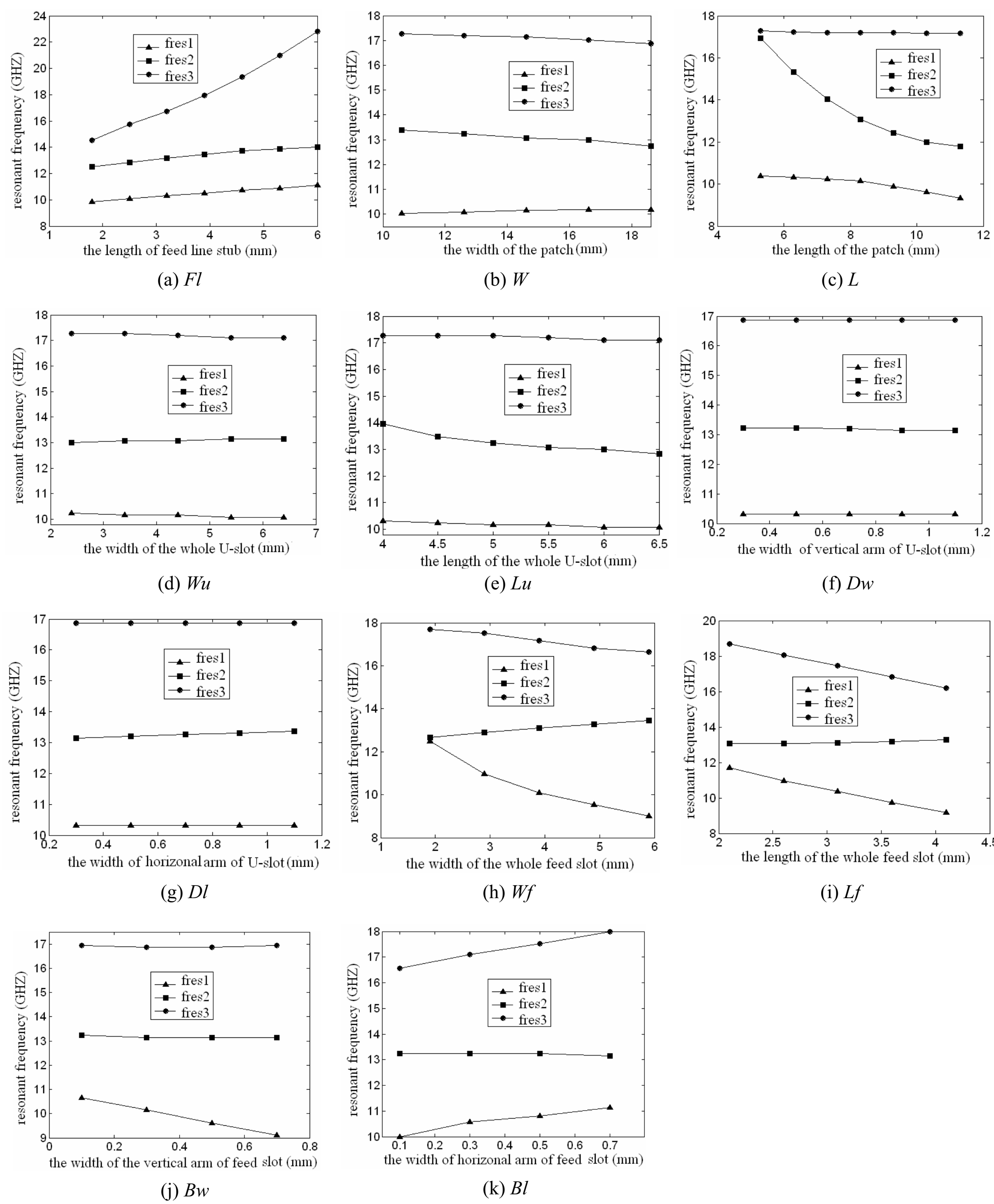
Figure 1: Geometry of U-slot patch antenna.

2. ANTENNA MODEL

The geometry and detailed dimensions of proposed antenna are shown in Fig. 1 which presents that this antenna element is a sandwich of four dielectric layers. The top dielectric layer and two feed dielectric layers at the bottom are made of the same material with relative permittivity of 2.2 and thickness of 0.8 mm, which are separated by the foam of 1.6 mm. The foam has the relative permittivity approximately equal to 1 and is used as the supporting material to reduce the excitation of surface waves among the other layers. The U-slot rectangular patch is located at the lower side of the top dielectric layer, which acts as a cover. The strip line is printed between the two feed dielectric layers. The energy is electromagnetically coupled to the patch by the feed slot on the upper side of the first feed dielectric layer.

3. PARAMETER STUDY

A FDTD code is developed, validated and applied to simulate the proposed wideband antenna structure. The surface of microstrip conductor and ground plane is applied to perfect conductor and all of the boundaries are enclosed by 8 PML layers. Total matrix size is $45 \times 30 \times 20$ and cell size is $0.1 \times 0.1 \times 0.2\,\mathrm{mm}^3$. The time step is $\Delta t = 0.22\,\mathrm{ps}$ so that the Courant stability condition is satisfied. A modulated Gaussian pulse is used as the excitation source. The center frequency of the pulse is 15 GHZ and the bandwidth is 10 GHZ. So the main frequency component of the pulse is in the frequency band of 10 ∼ 20 GHZ. U-slot aperture coupled microstrip antenna with wide-band characteristics has three resonances which are regulated by patch size and slot aspects. The sensitivity of the geometric parameters such as W, L, Wu, Lu, Dw, Dl, Wf, Lf, Bw, Bl are studied in Fig. 2.



(a) *Fl* (b) *W* (c) *L*

(d) *Wu* (e) *Lu* (f) *Dw*

(g) *Dl* (h) *Wf* (i) *Lf*

(j) *Bw* (k) *Bl*

Figure 2: The transition of design parameters.

4. WIDEBAND U-SLOT PATCH ANTENNA A

From above results, it is seen that alterations of parameters can dramatically change the behavior of antennas. Through repeated experiments, all parameters are acquired like Table 1.

Table 1: Dimensions of antenna A in millimeter.

W	L	Wu	Lu	Dw	Dl	Wf	Lf	Bw	Bl
14.5	8.2	2.4	5.4	0.5	0.5	4	3.2	0.2	0.2

The VSWR curve of wideband antenna A is shown in Fig. 3(a). The bandwidth (VSWR$\leq$2) is about 50% in the frequency range of 10.16–16.85 GHZ. The gain versus frequency curve is shown in Fig. 3(b) with maximum gain of 9 dBi, and the bandwidth of gain above 7 dBi is about 33%. Fig. 4 shows the radiation patterns of co-polar and cross-polar in E-plane and H-plane at the center frequency. All these data verify the success of developing an ultra wideband U-slot patch antenna.

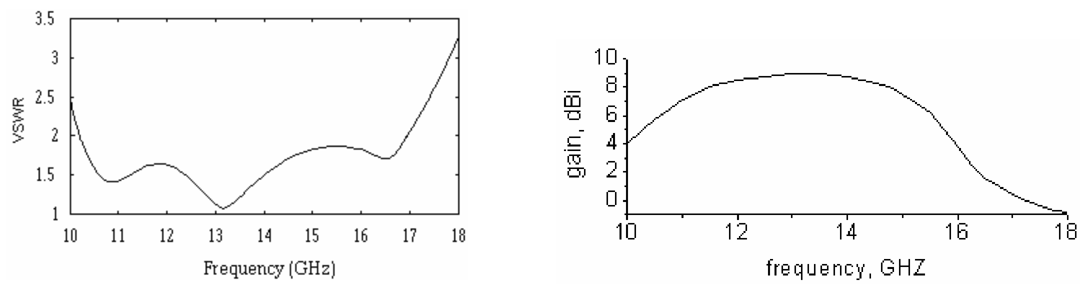


Figure 3: (a) Bandwidth characteristics, (b) Gain characteristics.

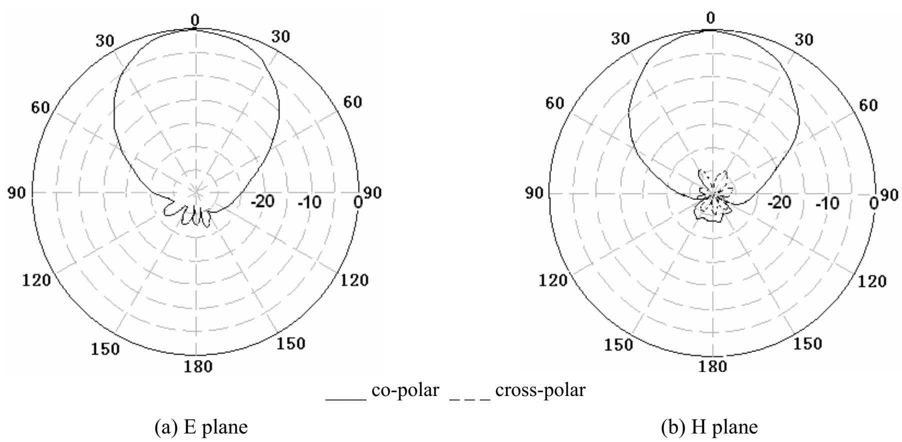


Figure 4: Radiation patterns at 13.5 GHZ.

5. DUAL-FREQUENCY U-SLOT PATCH ANTENNA B

In addition to wideband operation, the U-slot patch can also function as dual-frequency antenna. It appears that, starting with a wideband design, if one of the parameters such as the size of U-slot or the size of feeding slot is varied, the broadband characteristic is changed into a dual-frequency characteristic. An example is antenna B shown in Table 2.

Table 2: Dimensions of antenna B in millimeter.

W	L	Wu	Lu	Dw	Dl	Wf	Lf	Bw	Bl
14.5	8.2	4.2	5.4	0.5	0.5	4	2.5	0.2	0.2

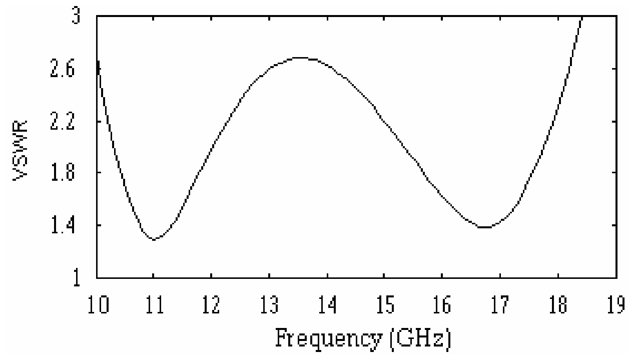


Figure 5: VSWR/frequency curve for antenna B.

The VSWR/frequency curve for antenna B is shown in Fig. 5, which exhibits a dual-frequency characteristic. The separation of the center frequencies of the bands is 5.9 GHZ, which is about 51% when measured with respect to the lower centre frequency. The upper and lower pass-bands are mainly, but not entirely, determined by the dimensions of U-slot and feeding slot.

6. CONCLUSIONS

Wide-band operation of aperture coupled U-slot patch was studied. A study of the principal parameters of the structure has given a more complete understanding of the coupling mechanism between the different resonators. Based on this study, two different antennas were designed and tested. Good results confirm the high frequency capabilities of this aperture feeding technique, and the validity of the model.

REFERENCES

1. Huynh, T. and K. F. Lee, "Single-layer single-patch wide band microstrip antenna," *Electronic. Lett.*, Vol. 31, No. 16, 1310–1312, 1995.
2. Mak, C. L., K. M. Kuk, and K. F. Lee, "Proximity-coupled U-slot patch antenna," *Electronic. Lett.*, Vol. 34, No. 8, 715–716, 1998.
3. Pozar, D. M., "A microstrip antenna aperture coupled to a microstripline," *Electronic. Lett.*, Vol. 21, No. 2, 49–50, 1985.
4. Pozar, D. M. and D. H. Schaubert, "Comparison of architectures for monolithic phased array antennas," *Microwave J.*, 93–103, 1986.

Couple-fed Circular Polarization Bow Tie Microstrip Antenna

Huan-Cheng Lien, Yung-Cheng Lee, and Huei-Chiou Tsai

Wu Feng Institute of Technology
117 Chian-Ku Rd., Sec.2, Ming-Hsiung (621), Chiayi, Taiwan, China

Abstract— This paper presents a novel signal couple fed Circular Polarization (CP) Bow Tie shape Microstrip Antenna (Bow-Ti MA) based on the concept of electromagnetic coupled. The feeding technique and radiation performance of the Bow-Ti MA has been analyzed. The objective of the present design is not limited to both the improvement on the impedance and Axial Ratio (AR) bandwidths of the traditional CP antenna but also the radiation characteristics such as CP gain. The gain bandwidth and 3dB beam widths are also examined. Simulations and measurements show a good matching and a well-behaved CP radiation pattern can be achieved.

1. INTRODUCTION

Microstrip patch antennas are used in a variety of applications due to their many salient features [1]. Reducing the number of antennas has been strongly required because of the physical limits for installation space. In many areas of wireless communications, a considerable interest in reduced-size antennas. One of the methods to miniaturize the antenna systems is the miniaturization of the antenna, to achieve the miniaturization of antenna, the planer antenna has some characteristics such as simple, small, light, low profile, and so on, have been proposed and investigated.

Microstrip patch antennas can be designed to radiate CP with a single- fed are described in [2, 3], however, the major drawback of these kinds of antennas inherently has limited impedance bandwidth (VSWR$\leq$ 2) and narrow 3dB AR bandwidth, the traditional CP microstrip antenna has bandwidth of only a few percent, which restricts its wide applications. To overcome its inherent limitation of narrow impedance and AR bandwidth, to achieve the high purity CP over an increasing operational bandwidth, both radiation properties (AR) as well as the antenna matching need to be optimized simultaneously. Therefore, develop broadband techniques to enhance the bandwidths of the microstrip antennas is very important.

Research and development in the area of microstrip antenna has been devoted to various techniques for the enhancement of microstrip antenna bandwidth [4–9]. Increased bandwidth can be achieved with [10, 11] or [12, 13]. The AR bandwidth can be enhanced by [14, 15] concept. For achieving bandwidth improvement, in this paper, a new structure of the CP Bow-Ti MA has been studied. In the proposed antenna we are using: (1) a Bow-Tie radiation patch generate a CP; (2) a parasitically couple-patch over Bow-Tie radiation patch, which is based on concept of the two-layers couple by making use of layer separation (air gap) tuning [16, 17] to improve the AR bandwidth; (3) a less than quarter-wave length of special impedance transformer, and (4) a circularly signal couple patch. The third and fourth above the mentions are using to improve the impedance bandwidth.

The objective of the proposed design is to generate a CP and to improve the bandwidth of the traditional CP microstrip antenna, at the same time radiation and other characteristics of the proposed antenna have also been investigated.

2. ANTENNA DESIGN AND ANTENNA STRUCTURE

The geometry of the present antenna is consisted of five patches, which stacking on the different three layers of the dielectric substrate each other, is illustrated in Figure 1. This configuration were included mainly a microstrip patch radiating element, a parasitically couple-patch, an impedance matching transformer, a signal couple element and a ground plane. The configuration side view showing in Figure 1(b), first, we employ low permittivity dielectric substrates situated both the third layer for the Bow-Tie shape radiation element and the rectangular parasitically couple-patch. And second, using a high permittivity dielectric substrate situated the second layer for a rectangular microstrip feed line. Finally, using a high permittivity dielectric substrate situated the first layer for a circularly signal couple element and the ground plane. All of we mention antecedently, the high permittivity is FR4 substrate and the low permittivity is RO substrate.

A rectangular parasitically couple-patch over Bow-Tie radiation element, and they are concentrically placed through the dielectric substrate with a thickness 0.508 mm, for mutual coupling between each other at the resonant frequency of both elements, as shown in Figure 1.

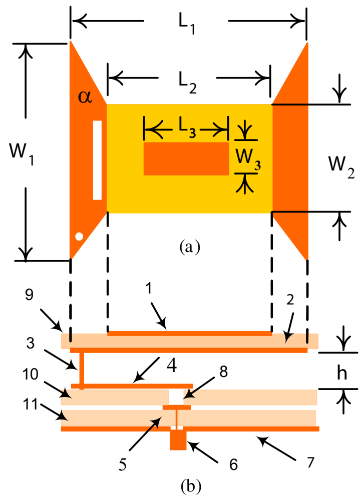


Figure 1: Geometry of the proposed antenna (a) Top view, (b) Side view.

Feeding mechanism plays an important role in the design of microstrip patch antennas. To overcome its inherent limitation of narrow impedance bandwidth due to feed networks with quarter-wave transformer or hybrid circuit, in the present paper, the impedance transformer is consists of two structures of the rectangular microstrip feed line and a cylindrical conductor. A cylindrical conductor connection the Bow-Tie radiating element of the third layer with the second layer rectangular microstrip feed line form the impedance transformer. The signal through a 50 ohms SMA connector to feed at the circularly signal couple element (radius r_i) center located on the z-axis, by electromagnetic couple to upside rectangular microstrip feed line. The probe position, the radius of the cylindrical connector and the length of the rectangular feed line should be carefully chosen to match the input impedance of the antenna.

By properly adjusting above the mentions sizes (including rectangular parasitically couple-patch, Bow-Tie radiate element, rectangular microstrip feed line, cylindrical conductor and circularly signal couple element) and the probe feed position, then we can obtain a better and wider both the impedance and AR bandwidth of the CP antenna. Figure 1 depicts the detailed antenna structure of the propose antenna, while the simulative optimized parameters are list in Table 1, respectively.

3. SIMULATION AND MEASUREMENT RESULTS

The design was analyzed with *Zeland Software*'s IE3D simulation package [18] by using an infinite ground plane and fabricated. The Return Loss (RL) of the antenna was measured by using an

Table 1: The relative parameters of the Bow-Ti MA (unit: mm).

1	Rectangular parasitically element	$L_2 = 35$, $W_2 = 24$, $L_3 = 10.5$, $W_3 = 3.5$
2	Bow-Tie shape radiating element	$L_1 = 63$, $W_1 = 57$, $\alpha = 35°$
3	Cylindrical conductor	Radius= 1.6
4	Rectangular microstrip feed line	Length= 25, Width= 3
5	Circularly couple patch	Radius= 1.6
6	SMA connector	
7	Ground plane	65×60
8	Slot	
9	RO substrate	$\varepsilon_r = 3.38$, $\tan\delta = 0.0025$, thickness= 0.508
10	FR4 substrate	$\varepsilon_r = 4.4$, $\tan\delta = 0.022$, thickness= 1.6
11	FR4 substrate	$\varepsilon_r = 4.4$, $\tan\delta = 0.022$, thickness= 1.6
h	Polymer	$\varepsilon_r = \varepsilon_o = 1$, $h = 4$

HP8720D vector network analyzer on an antenna by using a ground plane of 65 mm×60 mm. An operating frequency of 2.1 GHz and left hand CP where chosen. Figure 2 show the variation of simulated and measured RL with frequency, the simulated and measured 10 dB RL bandwidths were 35% and 31%, respectively. The comparison between the measurements and software predictions are very close. The difference between the two graphs is due to the fact that the proposed antenna was built on a very small ground plane (65 mm×60 mm), while the computations assume an infinite ground plane.

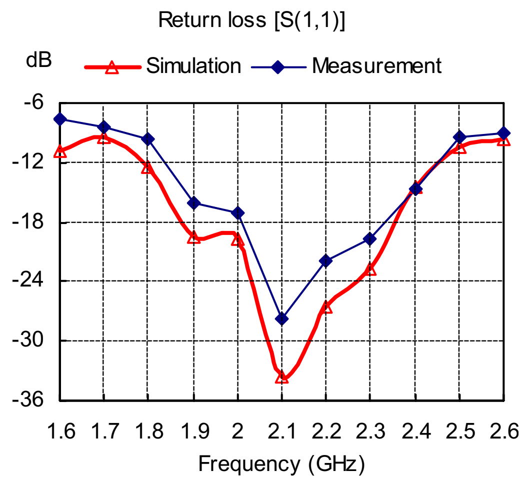


Figure 2: The variation of simulated and measured RL with frequency.

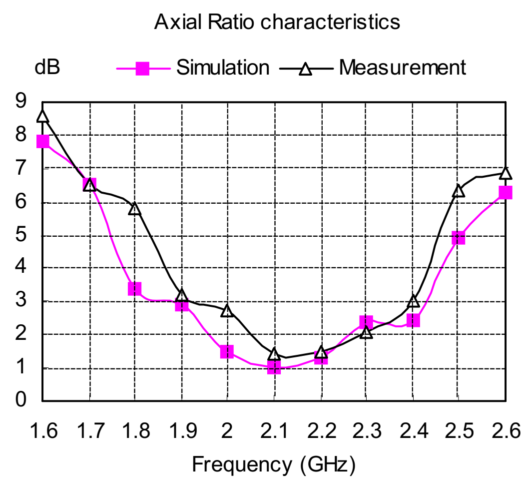


Figure 3: The variation of simulated and measured AR with frequency.

A good agreement between the simulated and measured results has been observed. The new matching technology and by electromagnetic couple method feed the signal has been confirm that an optimum matching impedance bandwidth for in excess of 30% (VSWR$\leq$ 2).

The antenna gain and the radiation patterns of the Bow-Ti MA have been measured in impulse time domain antenna measurement system. The AR results obtained from simulation and measurement are shown together in Figure 3 as a comparison. The measurement results confirm that the AR of less than 3 dB is in excess of 22% bandwidth in the range of 1.9 to 2.4 GHz was approximate simulation purpose. The measurement results show a frequency shift and a degrading in AR. This means that due to the Bow-Ti MA radiating element of the third layer and the second layer rectangular microstrip feed line are incomplete parallel. The measured AR pattern for CP in both principal $E(x-z)$- and $H(y-z)$-planes are shown in Figure 4, the 3-dB AR bandwidth are about 32 and 42 degree at the operating frequency 2.1 GHz, respectively. In the present investigation,

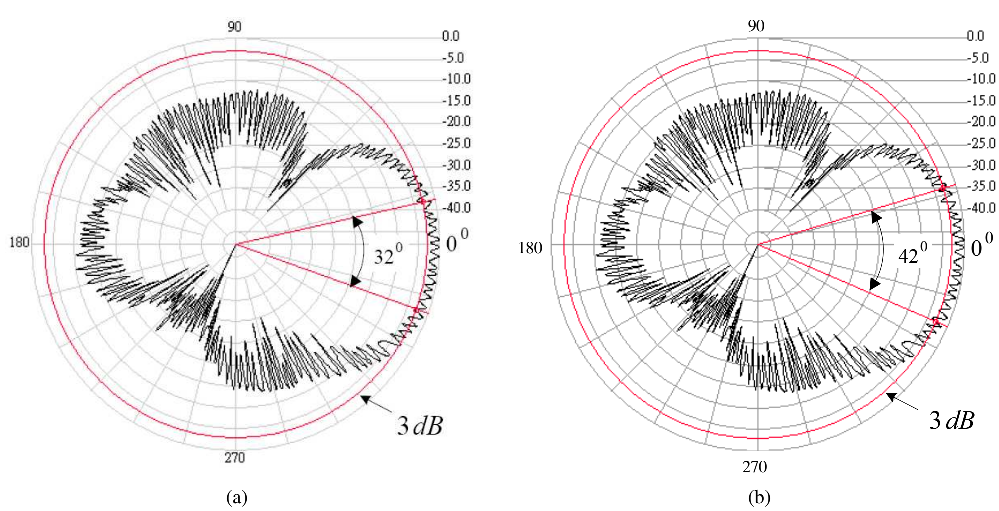


Figure 4: The measured AR pattern for CP (a) E-Cut, (b) H-Cut.

the CP gain is over 6 dBi across a frequency band between 1.9 to 2.4 GHz from simulation and measurement as shown together in Figure 5.

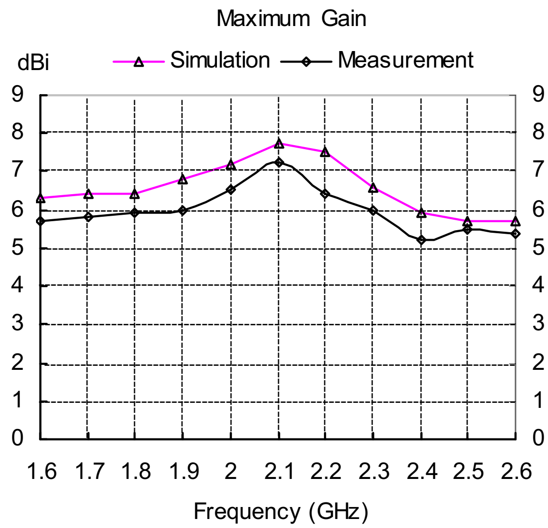


Figure 5: Antenna gain obtained from simulation and experiment.

4. CONCLUSIONS

By a parasitical couple-patch stacking over a Bow-Tie radiating element, employing the proposed matched technology and by electromagnetic couple method feed the signal has been confirm that Bow-Ti MA can be radiated good CP waves, and to achieve the impedance bandwidth and an AR bandwidth broadening.

The present design is not limited to the improvement on the impedance and AR bandwidths of the conventional but also the radiation characteristics such as CP gain. Furthermore, due to its compactness and broad bandwidth more applications can be anticipated.

REFERENCES

1. Pozar, D. M., "Microstrip antennas," *Proc. IEEE*, Vol. 80, 79–91, Jan. 1992.
2. Edwards, T. C., *Foundations for Microstrip Circuit Design*, John Wiley & Sons, USA, 1981.
3. Kumar, G. and K. C. Gupta, "Directly coupled multiple resonator wideband microstrip antennas," *IEEE Trans. Antennas Propagat.*, Vol. 33, 588–593, June 1985.
4. Griffin, J. M. and J. R. Forrest, "Broadband circular disc microstrip antenna," *Electronics Letter*, Vol. 18, 266–269, 1982.
5. Herscovici, N., "A wide band single layer patch antenna," *IEEE Trans. on Antenna and Propagations*, Vol. 46, No. 4, 471–474, Copyright SBMO ISSN 1516–7399, April 1998.
6. Chen, Z. N. and Y. W. M. Chia, "Broadband probe fed L-shaped plate antenna," *Microwave and Optical Technology Letters*, Vol. 26, No. 3, 204-206, Aug. 2002.
7. Sze, J. Y. and K. L. Wong, "Broadband rectangular microstrip antenna with of toothbrush shaped slots," *Electronics Lett.*, Vol. 34, 9230, 2186–2187, 1998.
8. Guo, Y. X., K. M. Luk, K. F. Lee, and Y. L. Chow, "Double U-slot rectangular patch antenna," *Electronics Lett.*, Vol. 34, No. 19, 1805–1806, 1998.
9. Chen, Z. N., "Broadband probe-fed L-shaped plate antenna," *Microwave and Optical Letters*, Aug. 2000.
10. Croq, F. and A. Papiernik, "Wideband aperture coupled microstrip subarray," *Electron. Lett.*, Vol. 26, 1293–1294, Aug. 1990.
11. Targonski, S. D. and D. M. Pozar, "Design of wideband circularly polarized aperture coupled microstrip antennas," *IEEE Trans. Antennas Propagat.*, Vol. 41, 214–220, Feb. 1993.
12. Wang, J., R. Fralich, C. Wu, and J. Litva, "Multifunctional aperture coupled stack antenna," *Electron. Lett.*, Vol. 26, 2067–2068, Dec. 1990.
13. Croq, F. and D. M. Pozar, "Millimeter wave design of wide-band aperture coupled stacked microstrip antennas," *IEEE Trans. Antennas Propagat.*, Vol. 39, 1770–1776, Dec. 1991.

14. Waterhouse, R. B., "Stacked patches using high and low dielectric constant material combinations," *IEEE Trans. Antennas and Propagat.*, Vol. 47, 1767–1771, Dec. 1999.
15. Chen, C. H., A. Tulintseff, and R. M. Sorbello, "Broadband two-layer microstrip antenna," *IEEE AP-S International Symposium Digest*, 251–254, 1984.
16. Lee, R. Q. and K. F. Lee, "Experimental study of the two-layer electromagnetically coupled rectangular patch antenna," *IEEE Trans. Antennas and Propagat.*, Vol. 38, 1298–1302, 1990.
17. Lee, R. Q., T. Talty, and K. F. Lee, "Circular Polarisation characteristics of stacked microstrip antennas," *Electronics Letters*, Vol. 26, 2109–2110, December 1990.
18. IE3D Software Release 8 developed by M/S Zeland Software, Inc.

Lightwave Technique of mm-Wave Generation for Broadband Mobile Communication

B. N. Biswas[1], **A. Banerjee**[1], **A. Mukherjee**[1], **and S. Kar**[2]

[1]Academy of Technology, Hooghly 712121, India
[2]Institute of Radio Physics and Electronics, Kolkata 700009, India

Abstract— In future generation pico-cellular mobile communications FM-sideband injection of two laser sources are used to generate mm-wave signals. This paper considers some overlooked issues with particular emphasis on the loss of coherence between the two lasers due to cycle slipping phenomenon and suggests a new tracking system, that considerably improves the performance.

1. INTRODUCTION

Increasing demand for broadband mobile communication and limited atmospheric propagation at mm-waves has resulted in the need for high density of pico-cells. And as such future cellular broadband mobile communication systems will comprise mm-wave components for radio link between the mobile station (MS) and the numerous base stations (BS), which are remotely controlled by the central station (CS). Moreover, the base stations are widely separated from the central station, optical transport of the mm-wave signal is the choice owing to the inherently low transmission loss coefficient of the optical fibers. This, in turn, requires the use of lasers and photo-detectors. The cost of numerous BS's should be kept as low as possible. Therefore, generation and control of mm-wave signals should be optically carried out at the control station, making use of the proposed optical devices needed for the purpose of transport of the mm-wave signals. This avoids the need for mm-wave oscillators and modulators in the numerous base stations. In achieving this purpose two approaches are adopted, viz., (1) single optical source technique and (2) multiple optical source technique. In this we present a method based on multiple optical source technique using sideband separation principle. This is depicted in Fig. 1.

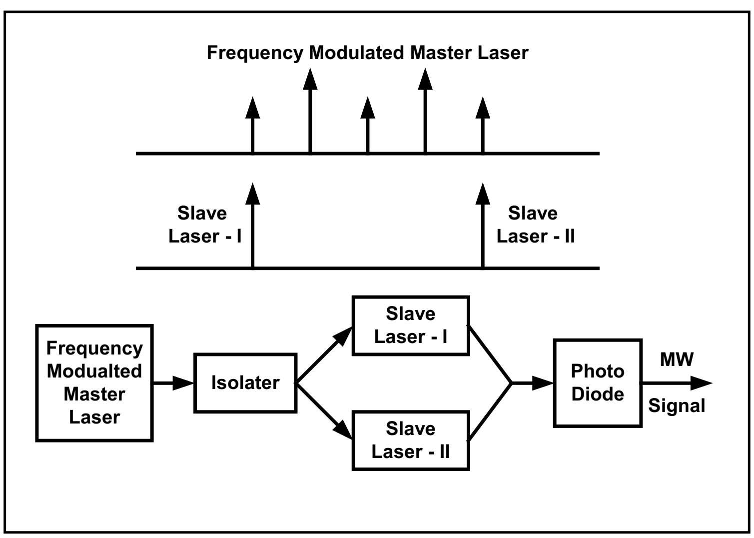


Figure 1: Coherent mixina of two optical sources.

The single optical source technique is the simplest approach for impressing microwave signal on an optical carrier. It can be realized either by direct current modulation of semiconductor laser or with an electro-optic modulator (Mach Zehnder Modulator MZM). Direct current modulation is limited to frequency range below 15.0 GHz and accompanied by a large microwave noise floor due laser intensity noise RIN and large harmonic content due to laser diode non-linearity. Another disadvantage is non-flat frequency response. Although indirect intensity modulation method MZM enjoys the advantage of larger bandwidth over the direct modulation scheme, it suffers from non-linear response, limited modulation depth, optical insertion loss, cost and complexity. Moreover, the link loss is dependent on the half-wave voltage V_π [typically $17 + 20\log(V_\pi)$ dB]. Incidentally for a polymer MZM, built at the University of Southern California, V_π has a value of 1.2 V at 1310 nm whereas it is 1.8 V at 1550 nm. This indicates that the minimum link loss is on the order of 23.0 dB (the laboratory stage).

2. TOWARDS A SOLUTION

In order to overcome the limitations of direct and external intensity modulation methods, heterodyning of two frequency-offset-low-intensity-noise lasers in a fast photo-detector has been proposed [1, 2]. A typical version of this concept is depicted in Fig. 2. A DFB laser is modulated and its output consisting of spectral lines separated by the modulating frequency is fed to two optical filters to isolate two sideband components (Alcatel within the EU FRANS Project). After this the two outputs from the two optical filters are heterodyned to generate the required microwave or millimeter wave signal (say, 60.0 GHz). This is possible only when a double sideband signal whose sideband components are separated by at least 50.0 GHz as the presently available optical filters based on Bragg Grating technology has a bandwidth of about 0.2 nm at 1550 nm, i.e., 25.0 GHz. Even then mm-wave signal generated at the output the photo-detector will a large spectral width, as the outputs from the two optical filters are phase incoherent. Moreover, double modulation at 25.0 GHz is difficult to generate.

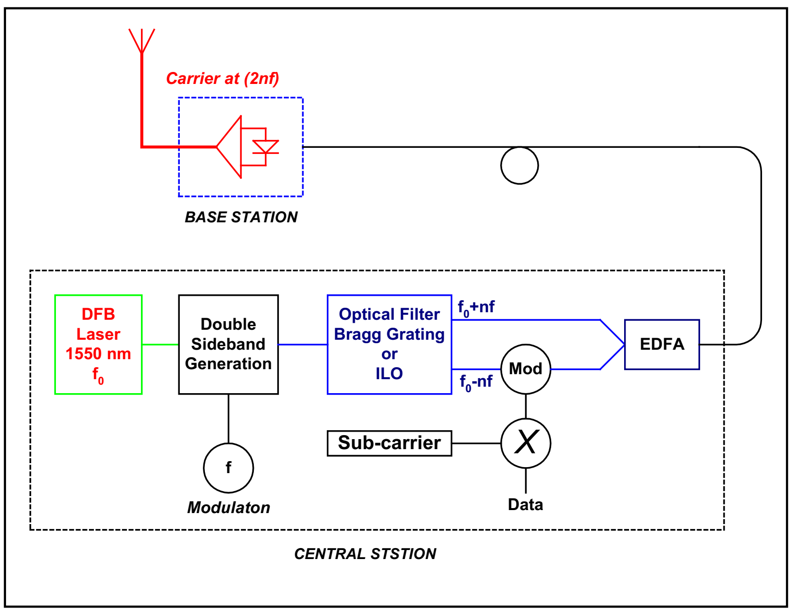


Figure 2: Optical generation of mm-wave signal.

This problems and limitations can be overcome by replacing the FG optical filters by injection locked lasers working as narrow band tunable filters. The injection locked optical filters (ILOF) have the following advantages over the Bragg Grating optical filters, being much larger and more complex than ILOF.

1. An injection locked optical filter (ILOF) can have a bandwidth on the order of 300.0 MHz only as compared to 10.0 to 25.0 GHz for a BG optical filter.
2. It is difficult to maintain phase coherence between the signals at the outputs of the two separate BG optical filters. Whereas the outputs of the two injection locked tunable filters (slave lasers SL) are are phase coherent.
3. Since an ILOF is narrow band, the master laser (ML) can be modulated at a much lower microwave frequency making it easier to modulate as well as to isolate the required higher sideband components.

Notwithstanding these possibilities, there still remains an important question that needs to be critically examined, and this is: Is it possible to realize the required phase coherence between the two outputs of the slave lasers with help commercial DFB lasers with linewidth on the order of few MHz (in order to make the system economically viable)?

3. A NEW APPROACH

This paper presents a modified optical injection phase frequency locked oscillator (OIPFLO) shown in Fig. 3. It consists of the arrangements by means of which instantaneous frequency of the slave laser is controlled by two mechanisms, namely, through injection locking and optical phase locked principle using an additional arrangement for controlling the output phase of the VCO laser in correspondence to a measure of the instantaneous phase error. The outputs from such two systems are heterodyned to generate the required mm-wave signal (cf. details on such novel system has been given in a companion paper by B N Biswas entitled "*Optical generation of mm-wave signal with wide linewidth lasers for broadband communications*").

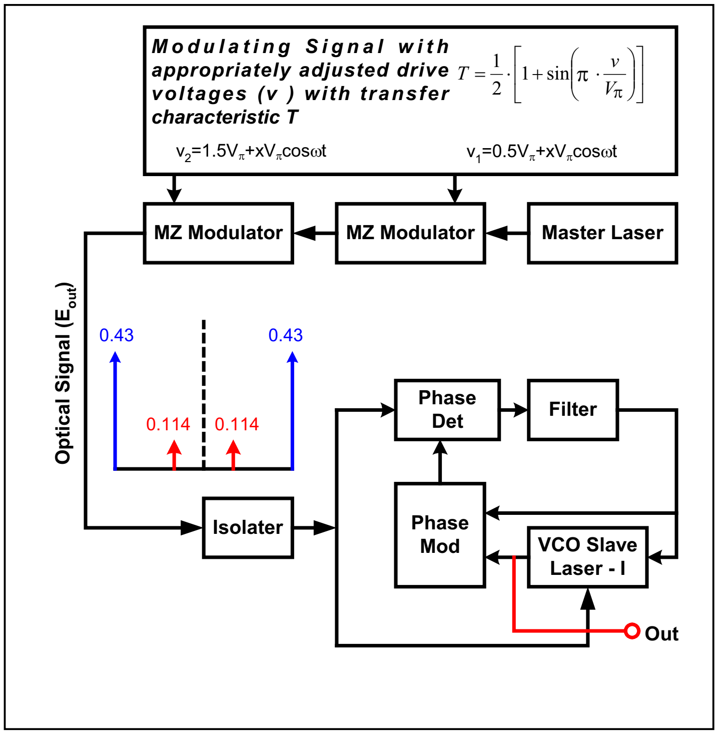


Figure 3: Injection phase frequency locked oscillator.

Phase coherence between the two outputs from the slave lasers means phase coherence between the master laser and the slave laser also, and cycle slipping is an annoying problem. Improving the phase coherence means two things simultaneously, namely, increasing the tracking bandwidth of the locked oscillator and improving the noise squelching property of the system. Unfortunately an attempt to increase one deteriorates the other [3–5]. However, the tracking system as shown in Fig. 3, achieves this property (i.e., to increase the locking range and to decrease the noise bandwidth at the same time) to a great extent. Before we come to the question of phase noises of the master laser and the slave laser, that disturb phase tracking, it is also necessary that there is no unwanted modulation component close to the center frequency of the slave laser. Otherwise, this exerts pulling and pushing force on the slave laser. Thus this will cause serious disturbances in the synchronization process.

4. REDUCTION OF UNWANTED PULLING AND PUSHING EFFECT

To solve this problem let us take the following example. Let the modulation bandwidth of the master laser is close to 10.6 Hz, and it is desired to generate 60.0 GHz mm-wave signal.

Referring to the new modulation scheme depicted in Fig. 3 and using the drive voltages with $x\pi = 3.843$ it is observed that the modulated laser field is [$\alpha(t)$ being laser phase noise].

$$\begin{aligned} E(t) = &-0.43\cos\left[(\omega_o+3\omega)\,t+\alpha\right]-0.43\cos\left[(\omega_o-3\omega)\,t+\alpha\right] \\ &+0.114\cos\left[(\omega_o+5\omega)\,t+\alpha\right]+0.114\cos\left[(\omega_o-5\omega)\,t+\alpha\right] \end{aligned} \tag{1}$$

The third harmonic components have been picked up, because the output mm-wave signal is required to be of 60.0 GHz with modulating frequency of 10.0 GHz. Disturbing components are away by 20.0 GHz and 6.0 dB less in amplitude causing almost no deleterious effect.

5. MINIMIZING FREQUENCY OF SLIPPING CYCLES

In the following we neglect the shot noise contribution to the phase error, as it is small over the range of the laser linewidth (say, 10.0 MHz) with the laser power on the order of 100 microwatt and the detector sensitivity of about 0.5 [6]. In view of this noise bandwidth gives an estimate of the phase noise contribution at laser output. In this connection it is important to note that noise output is considered at the output of the laser VCO from where the reference signal is taken and not at the output of the phase modulator after the laser VCO (Fig. 3). Here K_i is the injection locking range and K is the open loop gain of the OPLL. T_P is the time constant of the phase control loop and T is the loop delay parameter.

To demonstrate the validity of the concept in the simplest form, we consider only first order system. In this connection it is important to note that and injection locked oscillator is a first order phase locked system and its locking range K_i is limited to about 360.0 Mrad by the stability condition [5]. The loop phase error variance is inversely related to K_i. Because of the phase noise, the slave lasers will slip cycles even when the phase error variance is small, and in doing so the phase coherence is lost during the period of slipping cycles. The average time between cycle slips is approximately given by [4, 6].

$$T_{av} = \frac{\pi}{4B_n} \exp\left[\frac{2}{\sigma_\varphi^2}\right], \text{ where } \sigma_\varphi^2 = \int_{-\infty}^{\infty} |[1 - H(j2\pi f)]|^2 \left(\frac{\Delta f}{\pi f^2}\right) df \text{ and } B_n = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H_1(j2\pi f)|^2 \, d\omega \quad (2)$$

The relation clearly demonstrates that in order to increase T_{av}(a strong function of σ_φ^2) it is necessary that B_n as well as σ_φ^2 should be reduced. Incidentally, $H(s)$ denotes the closed loop transfer function and $H_1(s)$ denotes the transfer function when the output is observed at the VCO output. Let us see how does this concept apply to the PLL system. If one ignores the loop delay then these parameters are given by

$$\sigma_\varphi^2 = \frac{\pi(\Delta\alpha + \Delta\beta)}{K_i + K} \cdot \frac{1}{1 + KT_p} \text{ and } B_n = \frac{K_i + K}{4} \cdot \frac{1}{1 + KT_p} \quad (3)$$

The beauty of inclusion of the phase modulator (i.e., T_p) is that B_n as well as σ_φ^2 decreases with the increase of T_p. This is in contrast to the operation conventional OPLL.

Where $\Delta\alpha$ is FWHM (full width at half maximum) linewidth of the master laser and $\Delta\beta$ is that of the slave laser. Had there been no limitation on the value of K, the summed linewidth can be increased to a large value. But due to presence of the loop delay the maximum permissible value of K is limited.

6. MAXIMUM PERMISSIBLE VALUE OF 'K'

For the sake of simplicity we consider the first order system with delay. We write the characteristic equation of the system as

$$j\Omega + b + (1 + j\Omega p) \exp(-j\Omega d) = 0 \quad (4)$$

where Ω is the frequency normalized with respect to K, 'b' is the injection locking range normalized with respect to K, 'T_p' is the time constant of the phase modulator and d (KT) is the normalized delay time of the loop. The plot of critical value of KT against 'b' is shown in Fig. 4. An appropriate choice of the phase modulation index '$p = KT_p$' can improve the value of KT to a large extent. That is, either higher value of 'K' can be used or larger value of loop delay can be accommodated. It is seen that the introduction of the injection signal for direct synchronization and the additional phase control loop, aside directly decreasing the phase error variance and noise bandwidth, it also increases the stability of the tracking system.

To illustrate let us consider an OPLL with a loop delay of 1600 ps. Let us choose a value of 2.2 for KT with p = 0.35. Therefore, the permitted value of the K (open loop gain) is 1375.0 Mrad. Thus, b = 0.2618. Hence to realize the same phase error variance for the two cases, namely with p = 0.0 and p = 0.35. The ratio of the summed linewidths of the lasers with and without phase

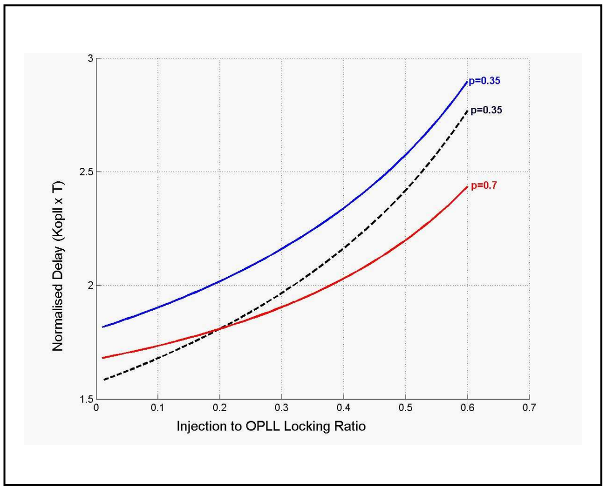


Figure 4: Variation of normalised delay with Kinj/Kopll.

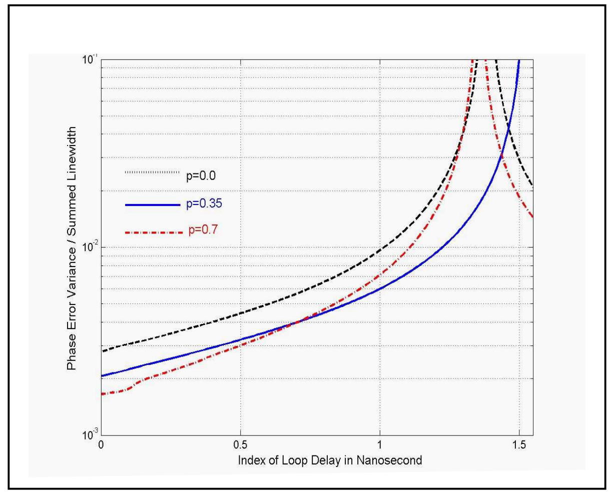


Figure 5: Variation of phase error variation with loop delay.

control can be found to 1.545. That is, with the OIPFPLO, 50% more linewidth of the laser can be allowed.

When the loop propagation delay is taken into account the noise bandwdith and the phase error variance increase rapidly after a certain values of the loop delay. Simple relations for the noise bandwidth and phase error variance do not hold good. However, when the delay is small, say up to about 700.0 pico seconds, the variation nearly obeys the rule as given in (3). But it is important to observe that by properly adjusting the value of the phase modulation index, both the noise bandwidth and phase error variance can be kept small over a larger value of the loop propagation delay (Fig. 5). It is also interesting to note that the performance of the loop will become worse than that of the simple OPLL if a large value of the phase modulation (p) is taken.

ACKNOWLEDGMENT

The author is grateful to Professor J. Banerjee, Chairman of the Academy of Technology for his encouragement and financial assistance during the course of the work.

REFERENCES

1. Braun, R. P., G. Grossskopf, D. Rhode, and F. Schmidt, "Fiber optic millimeter-wave generation and bandwidth efficient data transmission for broadband mobile 18–20 and 60 GHz band communication," *IEEE Intl Microwave Photon Topical Meeting*, 235–238, Essen, UK, 1997.
2. O'Reilly, J. J., P. M. Lane, and M. H. Capstick, "Optical generation and delivery of modulated mm-wave for mobile communication," *Analog Optical Fiber Communications*, 229–256, IEE, 1995.
3. Kroupa, V. F., *Phase Lock Loops and Frequency Synthesis*, John Wiley, England, 2003.
4. Biswas, B. N., *Phase Lock Theories and Applications*, Oxford and IBH, New Delhi, 1988.
5. Berdonalli, A. S, C. Walton, and A. J. Seeds, "High performance phase locking of wide linewidth semiconductor lasers by combined use of optical injection locking and optical phase locked loop," *J. Lightwave Technology*, Vol. 17, No. 2, 328–342, February 1999.
6. Ramo, R. T. and A. J. Seeds, "Delay, linewidth and bandwidth limitations in optical phase locked loop design," Vol. 26, No. 6, 389–390, 15 March, 1990.

Investigate Rectangular Slot Antenna with L-shaped Strip

G. Khunead[1,2], **J. Nakasuwan**[1], **N. Songthanapitak**[1], **and N. Anantrasirichai**[3]
[1]Department of Electronic and Telecommunication Engineering, Faculty of Engineering
Rajamangala University of Technology Thanyaburi (RMUTT)
Klong 6, Thanyaburi, Pathumthanee, Thailand
[2]Department of Electronic and Telecommunication Engineering, Faculty of Engineering and Architecture
Rajamangala University of Technology Suvarnahumi (RMUTSB), Thailand
[3]ReCCIT, Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang
Bangkok 10520, Thailand

Abstract— Rectangular Slot antenna fed by microstrip line is designed for standard of IEEE 802.11b/g (2.4–2.4835 GHz), IEEE 802.11a (5.15–5.35 GHz), and IEEE 802.16d (5.7–5.9 GHz). A rectangular slot antenna which comprises of two conductor strips located in a slot cut in the ground plane was investigated for dual frequency. This antenna was analyzed by using IE3D commercial software from Zeland [1]. In this case, the frequency range at return loss -10 dB is 2.26–2.61 GHz for low frequency and is 5.05–8.18 GHz for high frequency.

1. INTRODUCTION

Microstrip antenna is a type of antennas which can be used for transmitting and receiving signals. Microstrip or printed antennas are low profile, small size, light weight and widely used in wireless and mobile communications, as well as radar applications. Microstrip antennas can be divided into two basic types by their structure, namely microstrip patch antenna and microstrip slot antenna. For later case, the microstrip slot antennas can be fed by microstrip line, slot line and CPW. In this paper, we propose the slot antenna fed by microstrip line at a design frequency of 2.4 GHz which cover a frequency band of 2.4–2.4835 GHz and 5.15–5.9 GHz, respectively. Therefore, the proposed slot antenna can achieves the requirement of the IEEE standard. This antenna is designed on RT/Duroid 5880 substrate with thickness of 1.575 mm and dielectric constant of 2.2.

To describe the performance of the slot antenna, only the interrelated parameters are specified to complete description. The parameters in the characteristic of slot antenna for this analytical algorithm are input impedance, return loss (S-parameter), VSWR and radiation patterns. The main purpose of the antenna designing is to decrease the amplitude signal which returns from load (antenna). To minimize the unwanted effects of the reflected signal as much as possible, the process of matching impedance has to be concerned.

2. ANTENNA DESIGN AND SIMULATION RESULTS

In the designing of a microstrip slot antenna, it was necessary to specify the fundamental frequency that would be used in both receiving and transmitting signals. This slot antenna is designed on RT/duroid 5880 substrate with dielectric constant (ε_r) 2.2 and 1.575 mm of thickness of the dielectric substrate (h). The first designing is rectangular slot antenna cut on the ground plane and other plane is microstrip line fed slot antenna as shown in Figure 1(a). The dimension of this slot antenna is referred to guide wavelength (λ_g), which can calculated from:

$$\lambda_g = \frac{c/f}{\sqrt{\varepsilon_{eff}}} \tag{1}$$

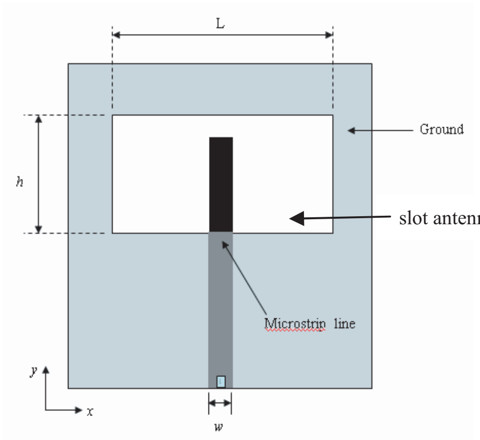
where ε_{eff} is the effective dielectric constant:

$$\varepsilon_{eff} \approx \frac{\varepsilon_r + 1}{2} \tag{2}$$

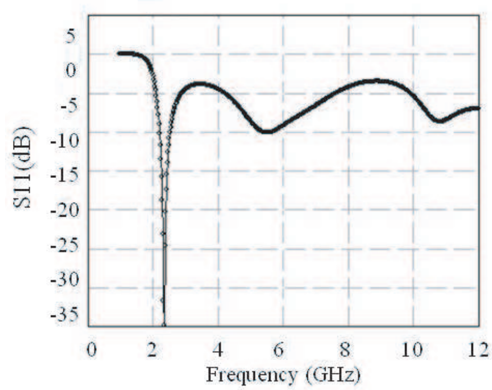
In this case, ε_{eff} is 1.6 and λ_g at frequency 2.4 GHz is 98.8 mm.

The width of the microstrip line (w) approximate 4.7 mm which calculated from:

$$\frac{w}{h} = \frac{2}{\pi}\left\{B - 1 - \ln(2B-1) + \frac{\varepsilon_r - 1}{2\varepsilon_r}[\ln(B-1)] + 0.39 - \frac{0.61}{\varepsilon_r}\right\} \tag{3}$$



(a)



(b)

Figure 1: The structure of slot antenna and simulation result of S_{11}. (a) Structure of rectangular slot antenna. (b) Characteristic of return loss.

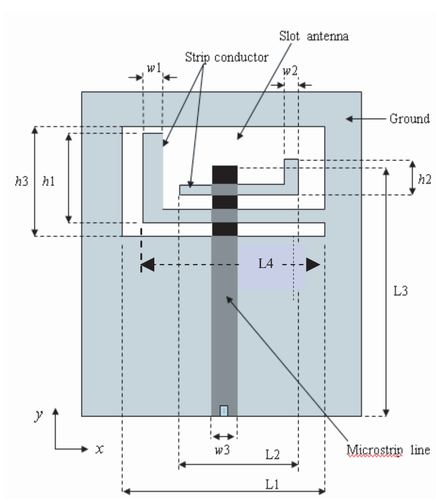
where $B = \dfrac{60\pi^2}{Z_0\sqrt{\varepsilon_r}}$.

The parameters in the structure of rectangular slot antenna are L and h, which $L = 37.5$ mm ($0.38\lambda_g$) and $h = 18$ mm ($0.18\lambda_g$). The microstrip line is adjusted until achieving the match impedance with 50 ohms transmission line. The simulation result of return loss (S_{11}) is shown in Figure 1(b). The S_{11} is -34.85 dB at resonant frequency 2.36 GHz, and bandwidth is 13.56% (2.22–2.54 GHz) at -10 dB of S_{11}. Obviously, this characteristic of the return loss covers the IEEE 802.11b/g (2.4–2.4835 GHz).

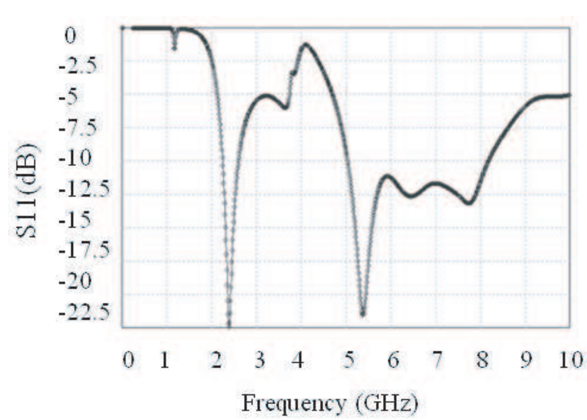
To achieve the dual frequency, the L-shaped conductor strip is inserted into rectangular slot antenna, as shown in Figure 2(a). The parameters of the rectangular slot antenna with L-shaped conductor strip are follows:

$$\begin{aligned} L_1 &= 37.5\,\text{mm},\ h_3 = 18\,\text{mm},\ L_2 = 23.55\,\text{mm},\ h_2 = 6.4\,\text{mm},\ L_4 = 33.5\,\text{mm},\ h_1 = 14.5\,mm, \\ w_1 &= 2.7\,\text{mm},\ \ w_2 = 2.5\,\text{mm},\ w_3 = 4.7\,\text{mm},\quad L_3 = 33.7\,\text{mm} \end{aligned}$$

The length of microstrip line and the dimension parameters are shown above, while the characteristic of the return loss is illustrated in Figure 2(b). It shows that the designed antenna can achieve the dual frequency and, moreover, high resonance frequency is wideband. The dual band appears at frequency band 2.26–2.61 GHz and 5.05–8.18 GHz. The S_{11} at lower resonant frequency 2.40 GHz is -22.44 dB and this bandwidth coverage IEEE 802.11b/g (2.4–2.4835 GHz). At high



(a)



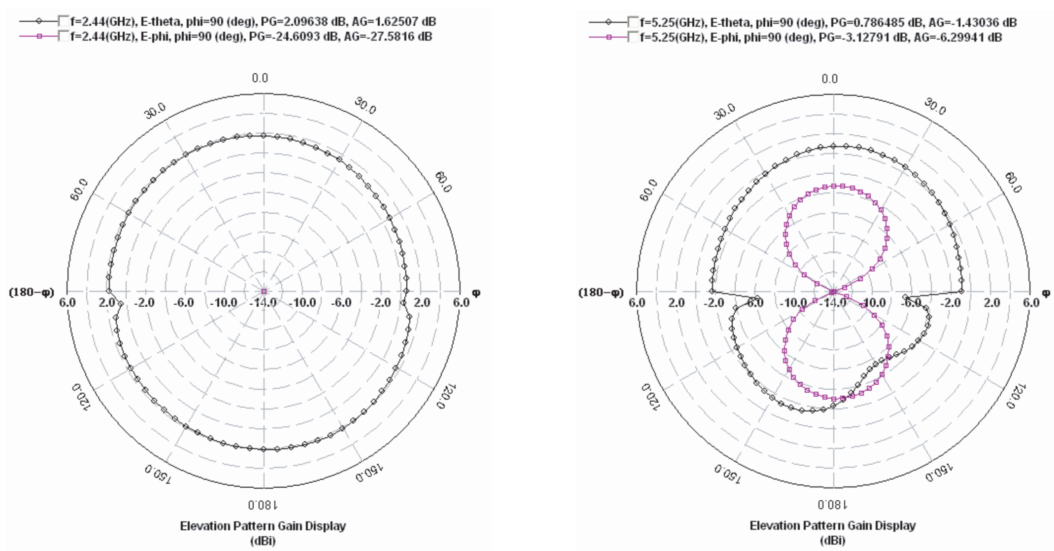
(b)

Figure 2: Configuration of slot antenna for dual frequency and characteristic of return loss. Configuration of slot antenna for dual frequency and characteristic of return loss. (b) Characteristic of return loss.

frequency, it shows that the wideband from 5.05 GHz to 8.18 GHz coverage IEEE 802.11a (5.15–5.35 GHz) and IEEE 802.16d (5.7–5.9 GHz). The simulation results are shown in Table 1.

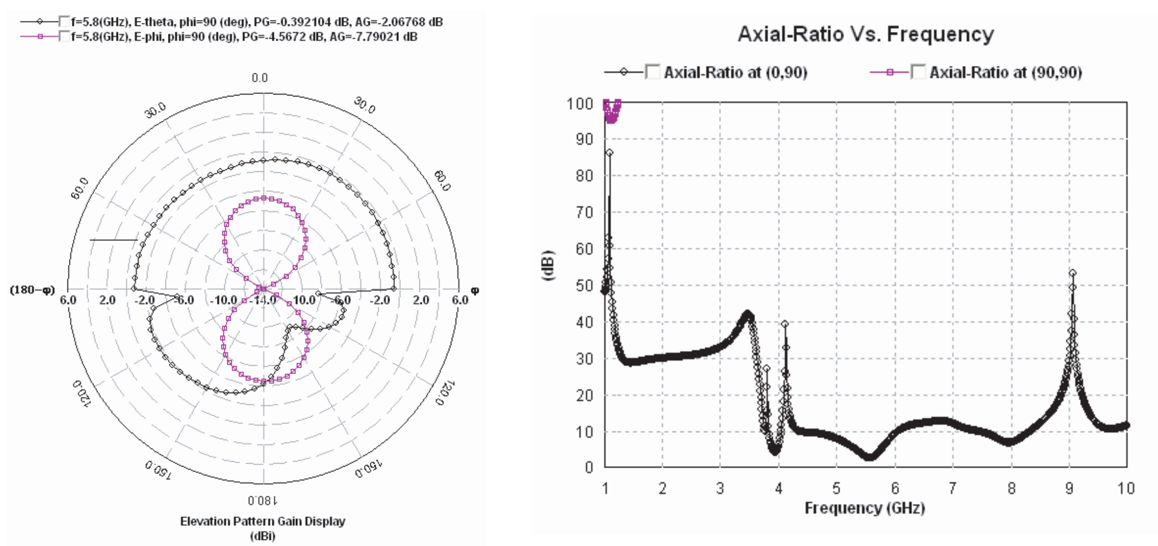
Table 1: Simulation results of rectangular slot antenna with L-shape conductor strip.

Resonance frequency (GHz)	S_{11} (dB)	Input Impedance		VSWR	Bandwidth (GHz)	Bandwidth (%)
		Real	Imaginary			
2.40	-22.44	52.89	-7.235	1.163	0.35(2.26-2.61)	14.16
5.38	-21.58	46.36	-7.187	1.182	3.13(5.05-8.18)	47.32



(a) (b)

Figure 3: The simulation results of radiation pattern on y-z plane ($\phi = 90°$). (a) At frequency 2.44 GHz. (b) At frequency 5.25 GHz.



(a) (b)

Figure 4: The simulation results of radiation pattern on y-z plane and axial ratio represent. (a) Radiation pattern on y-z plane ($\phi = 90°$) at 5.8 GHz. (b) Axial-Ratio Vs. frequency.

3. RADIATION PATTERNS

The radiation patterns at center band of three different frequency bands: 2.44 GHz, 5.25 GHz and 5.8 GHz, of rectangular slot antenna with L-shape conductor strip are presented in Figures 3(a), (b) and Figure 4(a), respectively.

From the radiation patterns, low frequency band 2.26–2.61 GHz is co-polarization and high frequency band 5.05–8.18 GHz is co-polarization and cross-polarization.

4. CONCLUSION

The designing slot antenna for WLAN applications is considered. The rectangular slot antenna with and without L-shaped conductor strip in the same size are proposed in this paper. The rectangular slot antenna without L-shaped conductor strip can achieved bandwidth of 13.56% cover standard of IEEE 802.11b/g (2.4–2.4835 GHz). But the rectangular slot antenna with L-shaped conductor strip can achieve dual band, low frequency bandwidth of 14.16% cover standard of IEEE 802.11b/g and high frequency band of 47.32% cover standard IEEE 802.11a (5.15–5.35 GHz) and IEEE 802.16d (5.7–5.9 GHz).

REFERENCES

1. Balanis, C. A., *Antenna Theory Analysis and Design*, John Wiley & Sons, Inc., 1997.
2. IE3D User's Manual Release 9, Zeland software, Inc., U.S.A., 2002.
3. Benson, F. A. and T. M. Benson, *Fields Waves and Transmission Line*, Chapman & Hall, 1991.
4. Anantrasirichai, N., P. Rakluea, and T. Wakabayachi, "Slot antenna coupled by microstrip line for dual frequency," *ISITA/NOLTA 2002*, October 7–11, 2002.

Design of high-Q Cavities in Photosensitive Material-based Photonic Crystal Slab Heterostructures

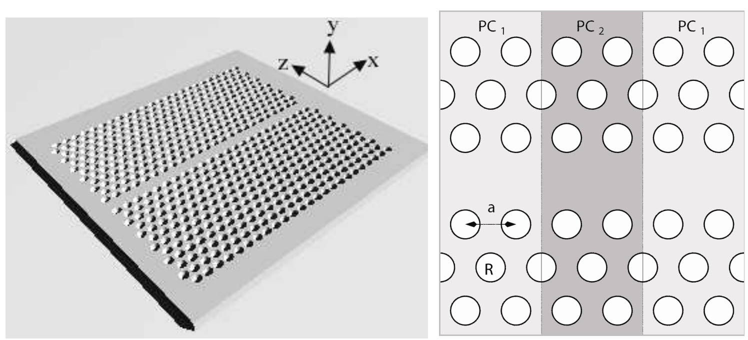
S. Tomljenovic-Hanic[1], **C. M. de Sterke**[1], **M. J. Steel**[1,2], **and D. J. Moss**[1]

[1]ARC Center of Excellence for Ultrahigh-bandwidth Devices for Optical Systems (CUDOS)
School of Physics, University of Sydney, Camperdown, NSW 2006, Australia
[2]RSoft Design Group, Inc., 65 O'Connor St, Chippendale, NSW 2008, Australia

Abstract— We propose a novel concept for creating high-Q cavities in photonic crystal slabs (PCS). We show that photonic crystal slab-based double heterostructure cavities, formed by variations in the refractive index, can have large a Q-factor (up to $Q = 1 \times 10^6$), and that such cavities can be implemented in chalcogenide glasses using their photosensitive properties.

In the last few years the study of optical microcavities based on photonic crystal slabs has attracted much attention [1–10]. Almost all of these studies consider a PCS composed of a hexagonal array of cylindrical air holes in a high-index semiconductor slab. There are many possible device applications of compact and efficient PCS nanocavities, such as channel drop filters [1], low-treshold laser [5], and cavity QED experiments [6, 7]. The principal design aim for all these applications is to obtain a high quality factor within a small modal volume.

A cavity is usually formed in either of two ways: forming a point cavity or forming a "hetero-structure". Microcavities with the highest Q values achieved to date, have been realised through the use of photonic crystal double-heterostructures [9, 10], where regions of slightly different lattice constant are combined in a single slab to create a cavity. Song et al., constructed double hetero-structure PCS, in which a short length of crystal (PC_2) with a lattice constant stretched in one direction, interrupts the main crystal (PC_1) [9, 10] (see Fig. 1).

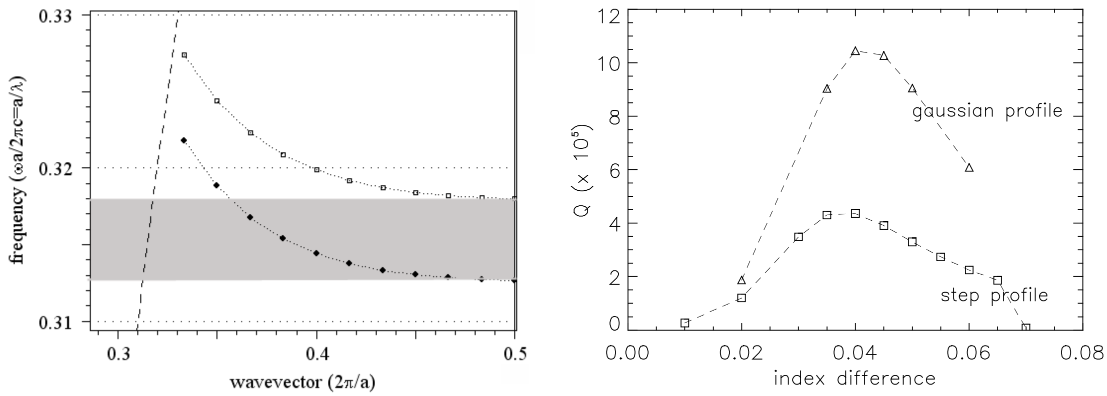


(a) (b)

Figure 1: (a) Schematic of PCS with a W1 waveguide in the Γ-K direction and (b) refractive index distribution in the plane of the structure considered. The central darker region indicates the increased index.

It is possible to form heterostructures exploiting material properties, rather than the geometry of the structure. In this paper we consider a chalcogenide glass-based PCS. It has been already shown that high quality PCS can be fabricated in this material [11]. The key property here is that chalcogenide glasses are photosensitive. This means that the refractive index of the material can change by 1 to 8%, depending on the type of glass, when it is illuminated by light, typically in the visible part of the spectrum [12].

The concept of the cavity design in hetero-structures relies on the mode-gap effect, a narrow frequency range for which PC_2 supports a mode, but not PC_1. Therefore, first we determine if there is a sufficient mode-gap to support a localized state between the structures having different refractive indices. We introduce a W1 waveguide in these structures: $W1_1$ for PC_1 with $n = 2.7$, and $W1_2$ for PC_2 with $n = 2.75$. Using the Plane Wave Expansion method, we obtain the dispersion curves. The results are shown in Fig. 2(a). The size of the mode-gap is comparable to those of the hetero-structures formed by geometric variation [10]. This suggests that hetero-structures formed by photosensitive index enhancement should also be capable of supporting localized states.



(a) (b)

Figure 2: (a) Dispersion curves for W1 of PC_1 (empty square) and W1 of PC_2,(full circle) within the region of the lowest gap; the dashed line represents the light line and (b) quality factor Q as a function of the refractive index difference between PC_2 and PC_1 for a step profile(square) and for a Gaussian profile (triangles).

Now we calculate quality factors for the structure illustrated in Fig. 1, for different cases where the refractive index change of the central part varies from $\Delta n = 0.01$ to $\Delta n = 0.07$. The results are shown in Fig. 2(b), where we see that a refractive index change as small as $\Delta n = 0.01$ is enough to obtain a quality factor $Q = 3 \times 10^4$, whilst increasing the refractive index difference to $\Delta n = 0.02$ increases the quality by a factor four. The maximum quality factor that we calculated— $Q = 4.3 \times 10^5$—appears at $\Delta n = 0.04$. The associated resonant frequency $\tilde{\o} = 0.3157$ is located in the middle of the mode gap in Fig. 2(a). A further increase in the refractive index decreases the quality factor, but the Q factors are still of the order of 105. It is interesting to note that quality factors of $Q \sim 10^5$ are achieved with large refractive index changes of $0.02 < \Delta n < 0.065$. As expected, the quality factor in chalcogenide glass-based PCS is smaller than for the silicon-based PCS, due to the smaller refractive index [7].

We also modelled a cavity with a Gaussian index variation (full-width at half maximum $2a$) along the waveguide direction, with identical maximum refractive index differences to those for the step profile. The results are plotted in Fig. 3. The maximum quality factor, $Q = 1 \times 10^6$, appears when the maximum refractive index change $\Delta n = 0.04$, just as for the step profile. Therefore changing the refractive index profile from a step to a Gaussian profile more than doubles the quality factor. This is consistent with findings of Song et al., who reported that a gradual structural change produces higher quality factors [9].

In conclusion we suggest a novel way of obtaining hetero-structures that can be applied to the PCSs made from photosensitive materials, such as chalcogenide glass and polymer. The hetero-structure is composed of elements that slightly differ in the refractive index; in practice this index difference can be easily photoinduced. We demonstrate that the chalcogenide-based PCS hetero-structure, designed in this way, can exhibit extremely high cavity mode localization due to the mode-gap effect. We also demonstrate that the nanocavities in these structures can reach ultrahigh-Q quality factors.

ACKNOWLEDGMENT

This work was produced with the assistance of the Australian Research Council (ARC) under the ARC Centers of Excellence Program. CUDOS (the Centre for Ultrahigh-bandwidth Devices for Optical Systems) is an ARC Centre of Excellence.

REFERENCES

1. Song, B. S., S. Noda, and T. Asano, "Photonic devices based on in-plane hetero photonic crystals," *Science*, Vol. 300, 1537, 2003.
2. Akahane, Y., T. Asano, B. S. Song, and S. Noda, "High-Q photonic nanocavity in a two-dimensional photonic crystal," *Nature*, Vol. 425, 944–947, 2003.
3. Tanabe, T., M. Notomi, S. Mitsugi, A. Shinya, and E. Kuramochi, "Fast bistable all-optical switch and memory on a silicon photonic crystal on-chip," *Opt. Lett.*, Vol. 30, 19, 2005.

4. Akahane, Y., T. Asano, B. S. Song, and S. Noda, "Fine-tuned high-Q photonic-crystal nanocavity," *Opt. Express*, Vol. 13, 1202–1214, 2005.
5. Park, H.-G., J.-K. Hwang, J. Huh, H.-Y. Ryu, Y.-H. Lee, and J.-S. Kim, "Nondegenerate monopole-mode two-dimensional photonic band gap laser," *Appl. Phys. Lett.*, Vol. 79, 3032–3034, 2001.
6. Hughes, S., "Quantum emission dynamics from a single quantum dot in a planar photonic crystal nanocavity," *Opt. Lett.*, Vol. 30, 1393–1395, 2005.
7. Tomljenovic-Hanic, S., M. J. Steel, C. M. de Sterke, and J. Salzman, "Diamond based photonic crystal microcavities," *Opt. Express*, Vol. 18, 3556–3562, 2006.
8. Kuramochi, E., M. Notomi, S. Mitsugi, A. Shinya, T. Tanabe, and T. Watanabe, "Ultrahigh-Q photonic crystal nanocavities realised by the local width modulation of a line defect," *Appl. Phys. Lett.*, Vol. 88, 041112, 2006.
9. Song, B. S., S. Noda, T. Asano, and Y. Akahane, "Ultra-high-Q photonic double-heterostructure nanocavity," *Nature Mater.*, Vol. 4, 207–210, 2005.
10. Song, B. S., T. Asano, Y. Akahane, Y. Tanaka, and S. Noda, "Transmission and reflection characteristics of in-plane hetero-photonic crystals," *Appl. Phys. Lett.*, Vol. 85, 4591–4593, 2004.
11. Freeman, D., S. Madden, and B. Luther-Davies, "Fabrication of planar photonic crystals in a chalcogenide glass using a focused ion beam," *Opt. Express*, Vol. 13, 3079–3086, 2005.
12. Zakery, A. and S. R. Elliot, "Optical properties and applications of chalcogenide glasses: a review," *J. Non-Cryst. Sol.*, Vol. 330, 1–12, 2003.

Tunable Photonic Crystal Based on SOI

Caihua Chen, **Binglin Miao**, **and Dennis W. Prather**
Electrical and Computer Engineering, University of Delaware, Newark, Delaware 19716, USA

Abstract— Self-collimation in photonic crystals (PhCs) has been demonstrated providing a very promising light-guiding mechanism. The fact that self-collimation allows light-guiding without any physical boundary is beneficial in high-density photonic integrated circuits (PICs) in terms of efficient coupling and arbitrary beam routing with no crosstalk. In this paper, we demonstrate a tunable photonic crystal device by combining the self-collimation lattice and band-gap lattice.

It has been well known that photonic crystals (PhCs) may exhibit very different dispersion properties with that of either host material or guest material [1–3]. For example, while the equal-frequency contours (EFCs) of an unpatterned Si slab are circular, EFCs of a Si slab with different periodic patterns may exhibit different shapes [4]. A EFC is a cross section of dispersion surfaces and a dispersion surface is a surface which characterizes the relationship between all allowed wave vectors in the structure and their corresponding frequencies. Recently, there has been a growing interest in engineering the dispersion property of photonic crystals for possible applications. One of the very interesting phenomenons found in planar photonic crystals (PhCs) during these explorations is the self-collimation phenomenon [4, 5]. This behavior in which incident waves with a certain angular range are naturally collimated along only certain directions has been utilized to efficiently guide electromagnetic waves within a planar PhC without the use of channel defects or structural waveguides. Compared to its alternatives, namely dielectric waveguides and PhC line defect waveguides, this type of waveguides offers many advantages since it does not require physical boundaries to confine light. For instance, it releases the strict alignment requirements imposed by the coupling efficiency in the case of dielectric or PhC line defect waveguides. As such, it enables high efficient in-plane coupling. On the other hand, due to lack of structural interaction, light path can arbitrarily cross each other without any crosstalk, which is very important for high density PICs to achieve arbitrary routing. To allow more functionalities integrated into the self-collimation lattice and thus a self-collimation based PICs, in this paper we demonstrated tunability of self-collimation photonic crystals by free carrier injection [6, 7].

The studied device is illustrated in Fig. 1. This device consists of the self-guiding region and the tunable region. The self-guiding region is the square lattice of air holes patterned on the silicon slab. The radius of air holes is $0.3a$, where a is the lattice constant. The dispersion surface and the dispersion contour of this lattice are obtained with plane wave method (PWM). An effective index of 3 is used for Silicon to take the finite thickness of the Silicon slab into consideration. The self-guiding lattice has an approximately square EFC at the frequency of $a/\lambda = 0.3$, where a is the lattice constant and λ is the guiding wavelength. Since $a = 0.45\,\mu\text{m}$, electromagnetic wave with wavelength of $1.5\,\mu\text{m}$ can thus be self-guided in the lattice. The tunable region is a square lattice of air holes with the same lattice constant in order to align with the self-guiding lattice to minimize the leakage along the boundary between the two lattices. The tunable lattice is designed such that it has a band gap; the self-guiding frequency is located at its band edge of the band gap without injected free carriers. Under this condition, the self-guiding beam completely passes through this region. On the other hand, when injecting free carriers into this region to slightly reduce the effective index of Silicon slab in this region, the band is pulled up and the self-guiding frequency is thus shifted into the band gap. Depending on how deep the self-guiding frequency is pulled into the band gap, a various percentage of the self-guiding beam passes through this region to port A, the rest of it is reflected to port B. Fig. 2 shows the band diagrams of the tunable lattice with different effective indices.

To validate the overall device design, we simulate the device shown in Fig. 1 using the finite-difference time-domain (FDTD) method. Fig. 4 shows the simulation results.

To demonstrate the performance of the device experimentally, it was fabricated on a Silicon-on-insulator (SOI) wafer. The Silicon device layer has a thickness of 260 nm. The thickness of the SiO_2buffer layer is $1\,\mu\text{m}$. The photonic crystal lattice was written on 200-nm-thick PMMA resist by electron-beam lithography. The patterned structures were transferred into the Silicon layer subsequently using inductive coupled plasma (ICP) etching. The reaction gases were SF_6 and

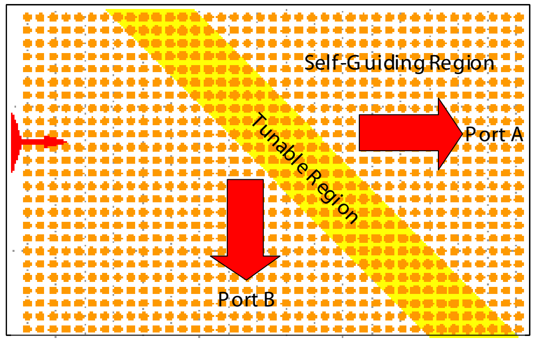


Figure 1: The Tunable photonic crystal device.

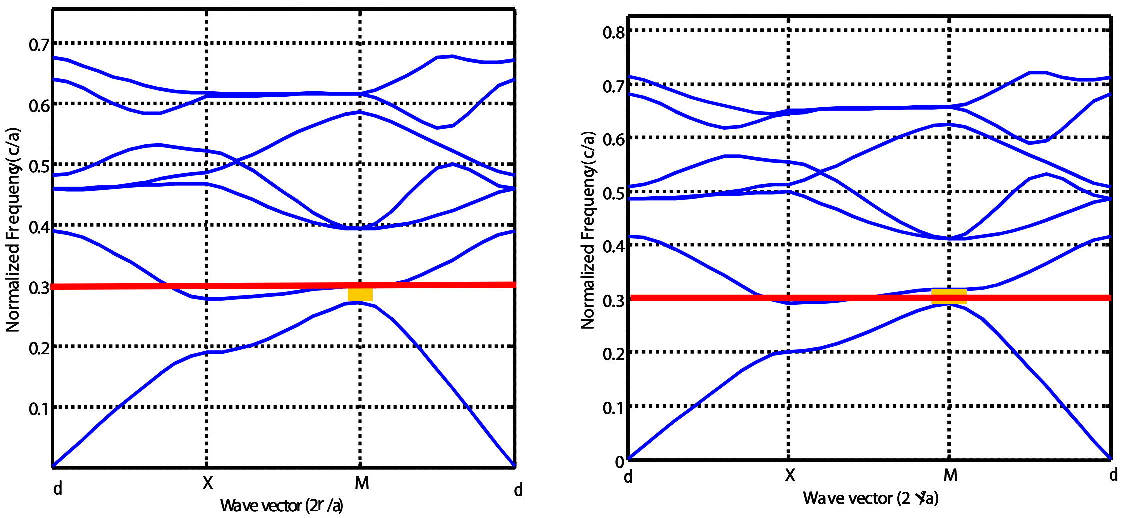


Figure 2: Dispersion diagrams of the tunable lattice: (a) with the effective index of Silicon as 3. In this case, the self-guiding frequency of $a/\lambda = 0.3$ is at the edge of the stop band around the Γ-M direction shown as the small shadow area. Therefore, the self-guiding beam passes through the tunable region to Port A as shown in Fig. 1, (b) with the effective index of Silicon as 2.8. In this case, the self-guiding frequency of $a/\lambda = 0.3$ is inside the stop band around the Γ-M direction shown as the small shadow area. Therefore, the self-guiding beam is mostly reflected to Port B as shown in Fig. 1.

C_4F_8. Fig. 4 shows the scanning electron microscope images of the etched photonic crystal. The self-collimation photonic crystal consists of a square array of air holes with hole radii of 125 nm and a lattice constant of 450 nm.

In order to inject free carriers, a p-i-p junction was formed along the diagonal direction of the self-guiding photonic crystal. At first, a 500-nm-thick SiO_2 layer was deposited onto the etched sample using plasma enhanced chemical vapour deposition (PECVD). Next, conventional UV lithography and buffered oxide etch (BOE) were employed to open doping windows via the SiO_2 layer. After spinning a boron film on the sample, it was slid in a RTA furnace of 1050°C for diffusion. After the diffusion process, the residue of the dopant was removed by soaking the sample in the BOE solution for a couple of minutes. The sheet resistance of doped regions was measured by a four-point probe. From the sheet resistance, the doping concentration was calculated as $3 \times 10^{19}\,\text{cm}^{-3}$.

To deposit the electrodes, the regions for metal electrodes were patterned using conventional photolithography. The Ge/Au electrodes were deposited using an electron-beam evaporator followed by the lift-off process. The electrodes were connected with external connectors using gold wires by wire bonding. Fig. 5 shows the microscope image of the fabricated device. The p-i-p diode injects holes only [8], and the measured I-V curve is approximately linear at -40 to 40 V ranges. The measured current is $I = 3.1$ mA at $V = 40$ V. The devices were tested on an optical measurement bench. A tapered Doping regions Electrode Electrode Photonic crystals and polarization maintained fiber launched the input light from a tunable laser to the facet of the input waveguide. The 8 μm wide input waveguide was tapered to 200 nm, narrowing the input wave to a single mode

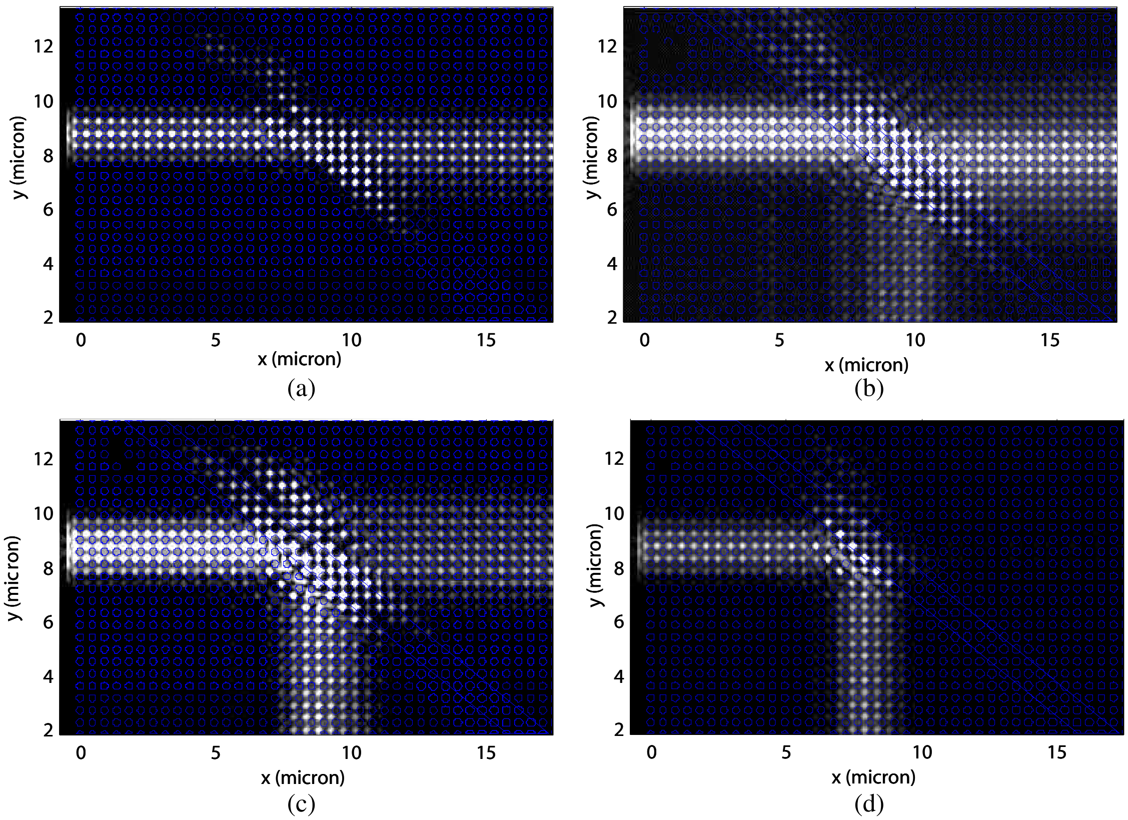


Figure 3: The FDTD simulation results of the device shown in Fig. 1. Silicon in the tunable region has an effective index of (a) 3, (b) 2.95, (c) 2.9, and (d) 2.8.

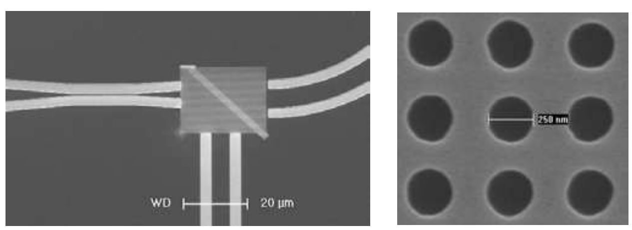


Figure 4: Left: SEM images of the etched self-collimation photonic crystal. The waveguide interfacing with the self-collimation lattice along the Γ-X1 direction is taped down first to narrow the wide input light to a single mode and is then inversely tapered for better coupling. Right: Zoom-in view of the photonic crystal lattice.

wave, and launching it to the self-guiding photonic crystal. Another tapered fiber collected light at facets of output waveguides, and sent it to a photo detector. An external voltage source was connected to the electrodes.

To test the carrier injection effect, we launched light with a wavelength of $\lambda = 1428\,\text{nm}$ from the tunable laser. The output light was collected and measured. The switching effect was observed when the external voltage of 40 V was switch on and off. The graph of the measured output light versus the applied external voltage is shown in Fig. 6. An IR CCD camera was mounted over the testing sample to observe the far field light radiated from the sample surface. Fig. 7(a) shows the captured top view of the device with no input light and no externally applied voltage. The right small images are the images captured at the output facet. Figure 7(b)–(d) show images when we launched the input light with a wavelength of $\lambda = 1428\,\text{nm}$ and the applied external voltages of $V = 0\,\text{V}$, 20 V, and 40 V, respectively.

As we can see from Fig. 7, when the applied voltage increases, the output light decreases. The radiation light from the sample surface decreases too.

The effective refractive index shift due to the free carrier injection and the free carrier absorption contribute to this phenomenon. From the captured images, the strong absorption can be observed. The output spectra with varying applied voltages are plotted in Fig. 8. However, no output light

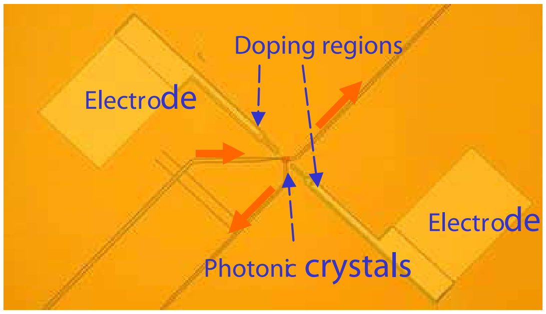


Figure 5: Microscope image of the fabricated device. The electrodes are aligned along the diagonal direction of the self-guiding photonic crystal.

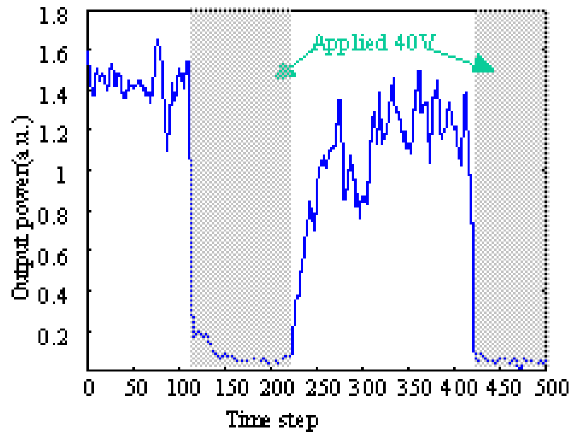


Figure 6: The measured output light versus the applied external voltage. The shaded areas indicate the time steps when the external voltage of 40 V is applied.

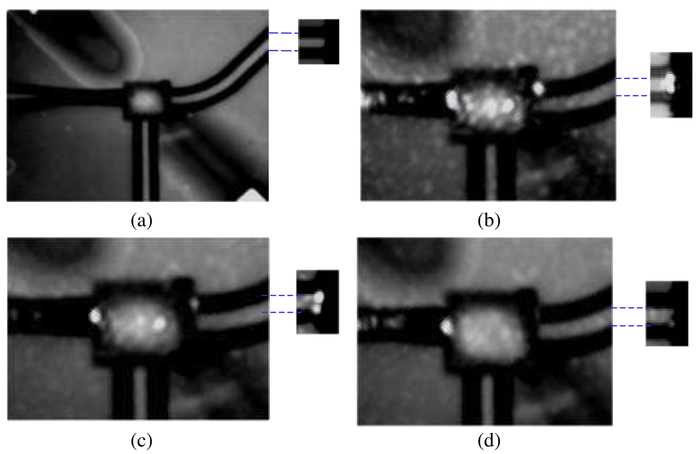


Figure 7: Captured top views (a) when no input light in launched, (b) when input light is launched but no external voltage is applied, (c) when input light is launched and the external voltage of 20 V is applied, and (d) when input light is launched and the external voltage of 40 V is applied.

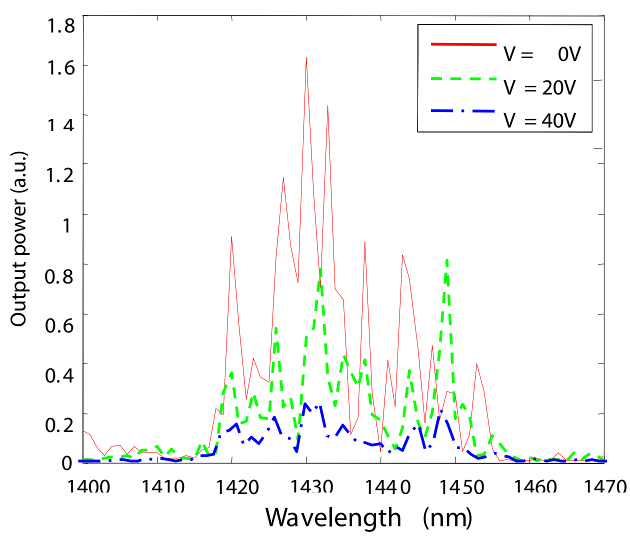


Figure 8: Measured spectra of output light at applied voltages of 0 V, 20 V, and 40 V, respectively.

could be observed at the reflection port. This might be caused by the unconfined free carriers. As we can see from Fig. 7, the free carriers pass not only the tunable region, but also other open regions. Therefore, the effective refractive index of these regions is changed too, which changes the dispersion property of the self-guiding photonic crystal. Moreover, free carriers outside the tunable region increase the overall absorption loss. We are trying to improve the free carrier confinement, and expect to observe the reflected light and faster switching speed.

In summary, we demonstrate a tunable photonic crystal device. The proposed device consists of the self-guiding region and the tunable region. The tunable region has the same lattice constant as the self-guiding lattice to minimize the leakage along the boundary between the two lattices. The tunable lattice is designed such that it has a band gap and the self-guiding frequency is located at its bottom band edge of the band gap. Therefore, in principle, when injecting free carriers into the tunable region to slightly reduce the effective index of this region, the band is pulled up and the self-guiding frequency is thus shifted into the band gap and the self-guiding beam is partly reflected to the vertical direction. Depending on how deep the self-guiding frequency is pulled into the band gap, a various percentage of the self-guiding beam is reflected. Experimentally, the free carriers are injected into the tunable region by highly p-doping two separated regions on both sides of and along the 45 degree direction to the original self-guiding beam path. It is observed that when a voltage of 40 V is applied on these two highly p-doped regions, the pass-though port has reduced power output comparing to the value when no external voltage is applied.

REFERENCES

1. Ochiai, T. and J. Sánchez-Dehesa, "Superprism effect in opal-based photonic crystals," *Phys. Rev. B*, Vol. 64, No. 24, 5113–5119, 2001.
2. Prasad, T., V. Colvin, and D. Mittleman, "Superprism phenomenon in three-dimensional macroporous polymer photonic crystals," *Phys. Rev. B*, Vol. 67, No. 16, 165103–165109, 2003.
3. Chen, C., et al., "Engineering dispersion properties of photonic crystals for spatial beam routing and non-channel waveguiding," *Integrated Photonics Research*, *OSA Technical Digest*, Washington, DC, 2003.
4. Witzens, J., M. Loncar, and A. Scherer, "Self-collimation in planar photonic crystals," *IEEE J. of Sel. Topics in Quan. Electro.*, Vol. 8, No. 6, 1246–1257, 2002.
5. Baba, T. and T. Matsumoto, "Resolution of photonic crystal superprism," *Appl. Phys. Lett.*, Vol. 81, No. 13, 2325–2327, 2002.
6. Liu, A., et al., "A high-speed silicon optical modulator based on a metal-oxide-semiconductor capacitor," *Nature*, Vol. 427, 615–618, 2004.
7. Xu, Q., et al., "Micrometer-scale silicon electro-optic modulator," *Nature*, Vol. 435, 325–327, 2005.
8. Jiang, Y., et al., "80 micron interaction length silicon photonic crystal waveguide modulator," *Appl. Phys. Lett.*, Vol. 87, 221105–221107, 2005.

Optical Properties of Mesoscopic Systems of Coupled Microspheres

V. N. Astratov[1], **S. P. Ashili**[1], **and A. M. Kapitonov**[2]

[1]Department of Physics and Optical Science, Center for Optoelectronics and Optical Communications
University of North Carolina at Charlotte
9201 University City Blvd., Charlotte, NC 28223, USA
[2]Center for Optoelectronics and Optical Communications, University of North Carolina at Charlotte
9201 University City Blvd., Charlotte, NC 28223, USA

Abstract— Two mechanisms of optical coupling between spherical cavities, tight-binding between their whispering gallery modes and focusing produced by periodically coupled microlenses, are directly observed using spatially resolved scattering spectroscopy and imaging. The results can be used for developing device concepts of lasers, optical filters, microspectrometers and sensors based on mesoscopic systems of coupled microspheres.

In this paper we consider structures formed by coupled spherical cavities which can be regarded as mesoscopic systems due to the fact that the size of their building blocks (spheres) is comparable to the characteristic wavelength. In contrast to metamaterials or photonic crystals conceptualized through the process of homogenization, the optical phenomena in such mesoscopic systems are essentially based on the properties of the constituting cavities. These include their ultra high quality whispering gallery mode (WGM) resonances [1] and their ability to focus plane waves into "nanoscale photonic jets" [2] at the shadow-side. These properties lead to two different mechanisms of optical transport between the cavities: (i) tight-binding between WGMs, (ii) propagation in a series of periodically coupled microlenses. These mechanisms are directly studied in the present work.

1. TIGHT-BINDING BETWEEN WGMS

Previously we observed WGM-related propagation effects [3] in one-dimensional (1D) chains of slightly disordered spheres using evanescently coupled dye-doped spherical cavities pumped above the lasing threshold for WGMs. In the present work we integrated spherical cavities into 3D closed packed structures with the thickness varying from one monolayer up to ~ 50 monolayers. The samples were obtained by self-assembly of spheres (with sizes in 2-10 micron range with standard 3% size dispersion) directed by hydro-dynamic flow in a specially designed cuvette placed in ultrasonic bath. The spheres were dye-doped and locally excited to create a built-in source of whispering gallery modes (WGMs). In the scattering spectra of such samples we observed fringes due to light propagation via coupled WGMs, as illustrated in Fig. 1. The study of pump dependence of scattering spectra indicated that above WGM lasing threshold the emission is provided in localized modes formed by multiple spheres. This is confirmed by observation of double peak structures, the spectral signature of strong coupling regime between multiple cavities. Although size disorder plays a negative role in the efficiency of such transport, we show that due to existence of multiple paths for photons the optical transport can be very efficient. The study of thickness dependence of scattering spectra indicates that attenuation length of light in such 3D samples exceeds 50 μm. The results show that the optical transport in such systems is provided along particular configurations of cavities according to a principle of minimization of the total WGM detuning.

2. PROPAGATION IN A SERIES OF PERIODICALLY COUPLED MICROLENSES

We report on the light transport phenomena in linear chains composed of several tens of touching spherical microcavities. A new optical mode type, namely nanojet-induced modes (NIMs) is directly observed [4]. Theoretically, formation of periodic nanojets has been predicted [5] as a result of the optical coupling of microspheres acting as a series of micro-lenses, which periodically focus propagating wave into photonic nanojets.

The chains were formed by means of the self-assembly directed by micro-flows of water suspension of polystyrene microspheres. The standard size dispersion of spheres in each particular chain was below 3%, while mean size of spheres was varied in the 2–10 micron range. To couple light to NIMs we used built-in emission sources formed by several locally excited dye-doped microcavities from the same 1-D chain.

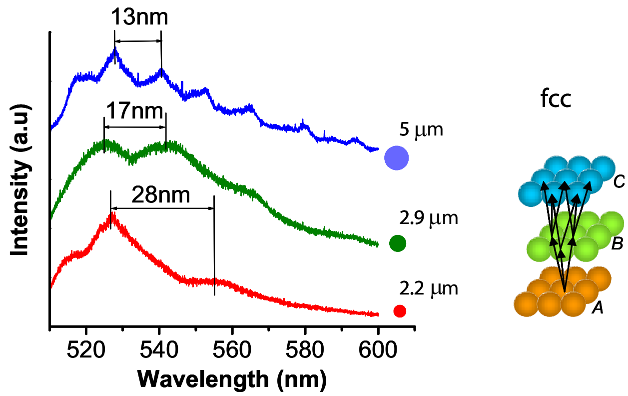


Figure 1: Scattering spectra of 3D systems of microspheres obtained using evanescently coupled sources of light (dye-doped spheres) with WGM peaks. The spectra exhibit fringes due to light propagation via coupled WGMs. Different sizes of microspheres result in different free spectral ranges. The sketch in the inset illustrates propagation between *ABC* layers along (111) direction of fcc lattice.

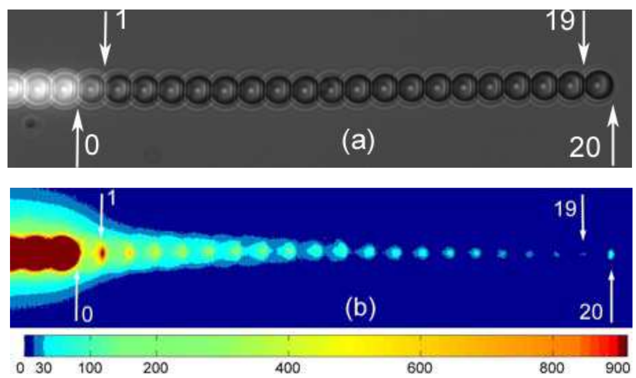


Figure 2: Visualization of NIMs in a locally excited chain of 2.9 μm spheres. (a) Image obtained with the background illumination, (b) same chain imaged due to propagation and scattering of light originating from the local fluorescence source.

The formation of NIMs as illustrated in Fig. 2. Conversion of modes emitted by the light source into the NIMs results in losses of several dB per sphere in the vicinity (first few tens of spheres) of such sources. At longer distances we found an attenuation rate as small as 0.5 dB per sphere, that reveals low intrinsic propagation loss for NIMs. The NIMs have potential applications for coupling of light in and out of spherical cavities characterized by extremely high quality (Q) whispering gallery modes and for guiding of emission in compact arrays of such microcavities.

ACKNOWLEDGMENT

This work was supported by ARO under Grant No. W911NF-05-1-0529 and by NSF under Grant No. CCF-0513179 as well as, in part by funds provided by The University of North Carolina at Charlotte. A.M.K. was partly supported by DARPA Grant No. W911NF-05-2-0053. The authors are thankful to Duke Scientific Corp. for donating microspheres for the research presented in this work.

REFERENCES

1. Ilchenko, V. S. and A. B. Matsko, "Optical resonators with whispering-gallery mode—Part II: Applications," *IEEE J. Sel. Topics Quantum Electron*, Vol. 12, No. 1, 15–32, January/February 2006.
2. Chen, Z., A. Taflove, and V. Backman, "Photonic nanojet enhancement of backscattering of light by nanoparticles: A potential novel visible-light ultramicroscopy technique," *Opt. Express*, Vol. 12, No. 7, 1214–1220, April 2004.
3. Astratov, V. N., J. P. Franchak, and S. P. Ashili, "Optical coupling and transport phenomena in chains of spherical dielectric microresonators with size disorder," *Appl. Phys. Lett.*, Vol. 85, No. 23, 5508–5510, December 2004.

4. Kapitonov, A. M. and V. N. Astratov, "Observation of nanojet-induced modes with small propagation losses in chains of coupled spherical cavities," *Opt. Lett.*, Vol. 32, No. 4, 409–411, February 2007.
5. Chen, Z., A. Taflove, and V. Backman, "Highly efficient optical coupling and transport phenomena in chains of dielectric microspheres," *Opt. Lett.*, Vol. 31, No. 3, 389–391, February 2006.

2D Photonic Crystals in Chalcogenide Glass

Cameron Smith[2], Christian Grillet[2], Eric C. Mägi[2], Yinlan Ruan[3], Darren Freeman[1]
Barry Luther-Davies[1], Steve Madden[1], Andrei Rode[1], Snjezana Tomljenovic-Hanic[2]
David Moss[2], and Benjamin J. Eggleton[2]

[1]Centre for Ultrahigh bandwidth Devices for Optical Systems, Laser Physics Centre
Australian National University, Canberra, Australia
[2]Centre for Ultrahigh bandwidth Devices for Optical Systems, School of Physics
University of Sydney, Sydney, Australia
[3]Center of Expertise in Photonics, School of Chemistry and Physics
University of Adelaide, Adelaide, Australia

Abstract— 2D photonic crystals containing waveguides and nanocavities have been fabricated in thin films of highly nonlinear chalcogenide glass. By probing the structures using evanescent coupling from a tapered optical fibre we have observed moderate resonator Q factors up to 10,000 which is adequate for all-optical switching.

Photonic crystals (PhCs) are a class of optical structure where the propagation of light is controlled by using a strong periodic modulation of the refractive index [1, 2]. Large index contrast in a PhC gives rise to a photonic band-gap within which light cannot propagate providing a possible mechanism to trap and manipulate light for photonic circuitry. Using the PhC mechanism, different types of optical circuits have been realized, including waveguides [3, 4]: a defect consisting of a row of deliberately omitted holes that act as a conduit for light (see Fig. 1(b)); and nanocavities: compact (wavelength-scale) optical structures with high Q-factors and small mode volumes [5, 6] (see Fig. 1(a)). PhC waveguides (PhCWGs) and PhC nanocavities (PhCNCs) have captivated considerable interest for all-optical signal processing [7, 8].

Chalcogenide glasses are attractive materials for all-optical signal processing. These glasses are composed of heavy elements including the chalcogens: S, Se and Te and posses a refractive index sufficiently high for PhC devices (typically 2.4–2.7); a high third-order nonlinearity (100–1000x that of silica); low two-photon absorption; and can be processed using conventional lithographic techniques. These properties suggest that devices made from chalcogenide glasses should achieve low-power ultra-fast optical switching and optical logic in photonic crystal resonators with their switching speed limited only by the resonator quality factor, Q.

Here we report on characteristics of 2-D PhC waveguides and resonators fabricated in free standing chalcogenide glass membranes by focused ion beam (FIB) milling [9, 10]. Some samples were additionally produced using conventional e-beam lithography and chemically assisted ion beam etching (CAIBE) [11]. Optical characterization was performed by resonant coupling between the modes of the photonic crystal waveguides and resonators and a tapered single mode optical fibre positioned close to the structures [12, 13].

To produce 2-D photonic crystals via ion beam milling we first coated free-standing ≈ 50 nm thick Si3N4 membranes produced by anisotropic chemical etching of nitride-coated silicon wafers with a 300 nm thick layer of $Ge_{33}As_{12}Se_{55}$ (AMTIR-1) glass. By FIB milling the membranes through the Si_3N_4 high quality 2-D photonic crystals with lattice period ≈ 600 nm and hole radius ≈ 360 nm were fabricated in a single-step process with near vertical and smooth side-walls. Since the milling time for the structures, which extended over and area $100\,\mu m \times 100\,\mu m$, was > 1 hour, custom software was used to monitor and correct any drift in the FIB system during milling to yield lattices whose structure was accurate to within our measurement limit of about 5 nm. Some samples were fabricated using a more conventional approach employing e-beam lithography to write the PhC pattern into a PMMA e-beam resist and this was then transferred into an underlying AMTIR-1 layer deposited onto an oxidized Si wafer using CAIBE. Following CAIBE the SiO_2 layer beneath the AMTIR-1 PhC was removed using buffered oxide etch (BOE).

Various structures were prepared for optical testing. These included simple waveguides formed by leaving out a row of holes from the lattice; long truncated waveguides where the structures were terminated before the edge of the lattice to create end mirrors; and nano-cavities formed by removing three adjacent holes from the lattice and shifting the positions (and in some cases the size) of the remaining holes at each end of the nanocavity to increase the resonator Q.

Testing was performed by tapering a section of conventional single mode optical fibre to a diameter of $\approx 1\,\mu$m and then forming the taper into a curve or loop 400–500 μm in diameter to restrict coupling to the region of the waveguide or nano-cavity. In conditions where the propagation constant of the mode in the taper and PhC are matched; power can couple from the fibre exciting modes of the photonic crystal and this is detectable as a dip in the transmission spectrum through the fibre taper.

When fibre coupling to a defect mode waveguide created by leaving out a single row of holes, the dispersion diagram demonstrates that coupling should occur between the fiber mode and a backward traveling wave in the photonic crystal (see Fig. 1).

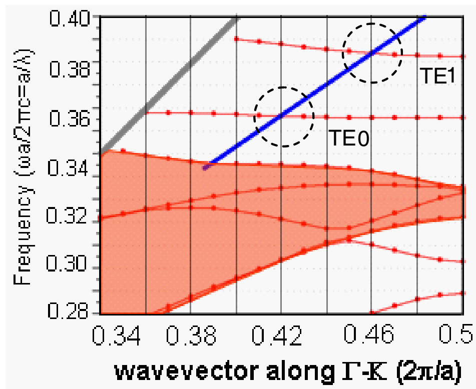


Figure 1: Dispersion diagram for the PhC structure (red line) and the nanowire (blue line). Dashed circle indicates modes for which phase-matched coupling occurs.

In the experiments the polarization of the field in the fibre was adjusted to TE relative to the surface of the lattice to allow coupling to the TE_1 and TE_0 modes of the PhC waveguide. Strong coupling could be obtained when the fibre was in contact with the PhC leading to a maximum depth in the transmission dip associated with the coupling resonance of -18 dB which corresponds to $\approx 98\%$ coupling efficiency. The variation in coupling strength with fibre to PhC separation is shown in Fig. 2. This experiment demonstrates that the chalcogenide glass platform is a promising one to realize low power all-optical switching.

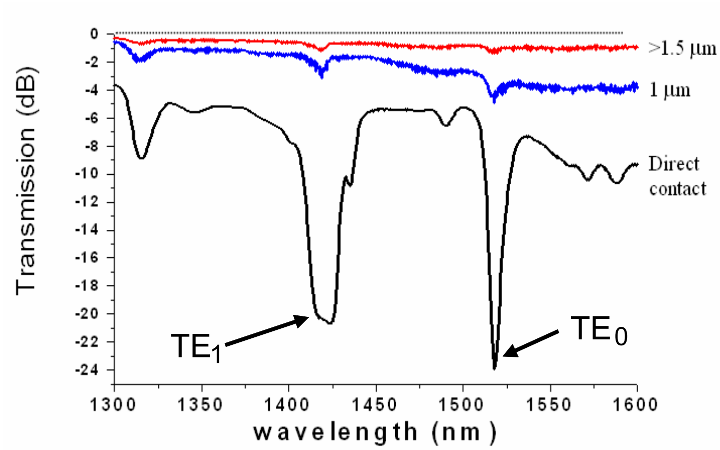


Figure 2: Transmission spectra through the tapered fibre for coupling to a PhC waveguide as a function of fibre to PhC separation on the input side of the PhC waveguide.

To this end, coupling to resonators with different end-hole shift and diameters has also been demonstrated using fibre coupling although the depth of the transmission was in this case restricted to only a few dB due to the limited intrinsic Q of the current resonant cavities. Predicted Q values obtained from 3D FDTD simulations predict intrinsic Q values of a few thousand should exist for cavities without a shift in the end-hole position increasing to > 10000 when an optimal shift of about $0.19a$ was used where a is the lattice constant. Reducing the side hole radius from

$0.33a$ to $\approx 0.15a$ approximately doubles the predicted Q. These general trends were confirmed by experimental measurements where maximum Q values as high as 10,000 were measured for samples with both a side-hole shift and diameter reduction whilst values of around 4000 were obtained for side-hole shift alone. To obtain these values the loading effect of the tapered fibre had to be minimized by separating the fibre from the PhC surface. For a fibre in contact the measured Q's were a few thousand as shown in Fig. 3. As expected the resonant frequency becomes sensitive to the shift of the end holes which results in an effective change in resonator length as is also apparent from the figure.

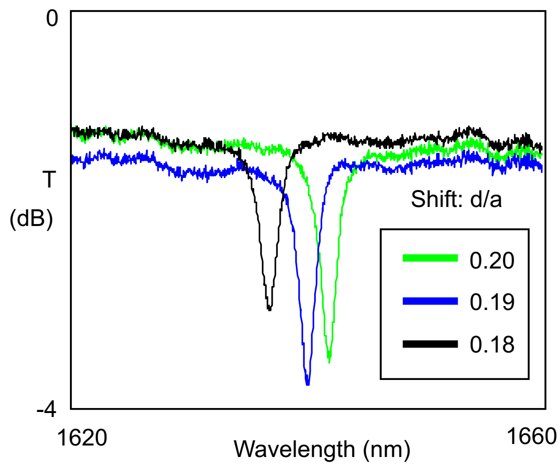


Figure 3: Transmission spectra for coupling to three different PhCNC, each trace corresponding to a varying displacement of adjacent holes to the cavity.

These data indicate that simple chalcogenide PhC resonators can exhibit sufficiently high Q to make all-optical switching feasible. The required Q is determined by the maximum fractional change in the refractive index, $\Delta\,\mathrm{n/n}$, achievable using a Kerr nonlinear optical response of a chalcogenide which is likely to be limited to $\Delta\,\mathrm{n/n} \approx \mathrm{few}\,10^{-4}$. Hence for the Kerr response to pull the resonator frequency by an amount equal to the linewidth of the resonance, Q values in the range of 5000 are required. The main requirement for effective switching will be to increase the inherent resonator Q (before loading via fibre coupling) sufficiently so that coupling losses dominate the effective Q that will result in much larger modulation of the transmission. Experiments on all-optical switching are now underway.

ACKNOWLEDGMENT

The support of the Australian Research Council through its Federation Fellow and Centres of Excellence programs is gratefully acknowledged. The assistance of Myung-Ki Kim and Yong-Hee Lee from KAIST in the production of PhC samples via e-beam lithography and CAIBE is gratefully acknowledged.

REFERENCES

1. Yablonovitch, E., "Inhibited spontaneous emission in solid-state physics and electronics," *Phys. Rev. Lett.*, Vol. 58, 2059, 1987.
2. John, S., "Strong localization of photons in certain disordered dielectric superlattices," *Phys. Rev. Lett.*, Vol. 58, 2486, 1987.
3. Loncar, M., et al., "Design and fabrication of silicon photonic crystal optical waveguides," *J. Lightwave Techn.*, Vol. 18, 1402, 2000.
4. Desieres, Y., et al., "Propagation losses of the fundamental mode in a single line-defect photonic crystal waveguide on an InP membrane," *J. Appl. Phys.*, Vol. 92, 2227, 2002.
5. Akahane, Y., et al., "High-Q photonic nanocavity in a two-dimensional photonic crystal," *Nature*, Vol. 425, 944, 2003.
6. Srinivasan, K., et al., "Optical-fiber-based measurement of an ultra-small volume high-Q photonic crystal microcavity," *Phys. Rev. B*, Vol. 70, 081306, 2004.
7. Tanabe, T., et al., "All-optical switches on a silicon chip realized using photonic crystal nanocavities," *Appl. Phys. Lett.*, Vol. 87, 151112, 2005.

8. Barclay, P. E., et al., "Nonlinear response of silicon photonic crystal microresonators excited via an integrated waveguide and fiber taper," *Opt. Express*, Vol. 13, 801, 2005.
9. Freeman, D., et al., "Fabrication of planar photonic crystals in a chalcogenide glass using focused ion beam," *Opt. Express*, Vol. 13, 3079, 2005.
10. Grillet, C., et al., "Characterization and modeling of Fano resonances in chalcogenide photonic crystal membranes," *Optics Express*, Vol. 14, No. 1, 369–376, 2006.
11. Ruan, Y., et al., "Fabrication of high-Q chalcogenide photonic crystal resonators by e-beam lithography," Submitted to *Applied Physics Letters*.
12. Hwang, I.-K., et al., "Optimization of coupling between photonic crystal resonator and curved microfiber," *IEEE Journal of Quantum Electronics*, Vol. 42, 131, 2006.
13. Grillet, C., et al., "Efficient coupling to chalcogenide glass photonic crystal waveguides via silica optical fiber nanowires," *Optics Express*, Vol. 14, 1071, 2006.

Photonic Crystals in Diamond for Quantum Information Technology

Joseph Salzman
Department of Electrical Engineering and Microelectronics Research Center
Technion — The Israel Institute of Technology, Haifa 32000, Israel

Abstract— One of the approaches to realize entanglement for quantum information processing, is *cavity quantum electrodynamics*. In this approach, the interaction of material qubits (atomic or electronic quantum states) with a high finesse optical resonator is used for atom-atom, atom-photon, or photon-photon entanglement. The main challenge in this approach is to avoid de-coherence induced by the cavity modes that leak to the environment. An attractive architecture to overcome this obstacle is formed by high-Q, coupled nano-cavities in a *Photonic Crystal*. The coupling of the nano-cavities can be gated to avoid de-coherence. The nitrogen-vacancy (NV) defect center in diamond, was recently shown to allow successful measurement of its spin quantum state, thus becoming a strong candidate for solid state quantum information processing. Our initial findings on diamond-based, photonic crystal coupled cavity studies, and preliminary fabrication experiments are encouraging, leading to the possibility to demonstrate a gated qubit in a diamond-based photonic crystal cavity.

1. INTRODUCTION

Quantum Information Technology (QIT), and in particular Quantum Computation (QC), show great potential for revolutionizing the methods by which one can collect and distribute information. Over the past 5 years, the explosive growth of schemes for quantum computation triggered a revolutionary reappraisal of the role of quantum mechanics in technology. One of the key issues for practical realization of QIT is the implementation of quantum bits, or qubits, that form the basis of useful quantum devices. While methods based on extensions of established technology (such as: silicon, NMR, linear optics and ion traps) have shown promise, experimental achievements in QC are only at the proof-of-principle stage in terms of their abilities to perform basic QC tasks [1]. Recently, it has been recognized that due to some properties of diamond, ***diamond based technologies appear to offer a promising route for a practical, scalable, and controllable implementation of qubits for QC*** [2–5].

Applications of Quantum Electrodynamics (QED) systems to quantum information derive mostly from the ability to coherently interchange quantum states between material (atomic) qubits and photonic qubits. Proof-of-principle experiments have been performed to demonstrate atom-atom, atom-photon, and photon-photon entanglement as the basis for two-qubit gate protocols. However, to achieve coherent dynamics with just a single atom and a single photon, a small, extremely low loss optical cavity is required, to work in the "strong coupling regime". The implementation of such cavities is extremely challenging.

Recently, we proposed a novel and original architecture for the control and manipulation of atom-photon entangled states (single qubit) in a nano-cavity, and cavity-cavity coupling for two qubit operartion, using **single crystal Diamond** as the base material, and a **Photonic Crystal** as the optical medium [6].

2. QUANTUM PROPERTIES OF DIAMOND — THE N-V CENTER

Diamond is an extraordinary material. It is well known as the hardest of all materials, but is also the most thermally conducting and displays a range of other highly desirable properties such as optical transparency from the ultra-violet into the infra-red, high electron and hole mobilities, and resistance to both chemically and radiation harsh environments. Less well known are the extraordinary quantum properties of the optical centers in diamond. In fact, the only photo-stable single-photon sources at room temperature reported so far **in any material** are nitrogen-vacancy (N-V) centers in diamond.

The N-V center in diamond consists of a nitrogen substituting for a carbon plus a nearest neighbor vacancy (Figure 1) [2]. It has a ^{3}A ground state and a ^{3}E excited state. The optical transition between these two states has an energy of 1.945 eV (637.3 nm). Most importantly, the transition has extraordinarily high quantum efficiency, such that single defect spectroscopy has been demonstrated very convincingly over the past few years. The transition is spin selective and under white light illumination, it is possible to rapidly spin polarize the system into the ground state [3]. Other crucial properties are: 1. long decoherence times, of the order 1 sec or more in dilute samples. Only few qubit candidates other than the N-V center in diamond that are so long lived exist in nature; 2. Rabi oscillations [4], simple 2 qubit (using a neighbouring ^{13}C atom) gate operations [5], and NMR spectroscopy have all been demonstrated at room temperature on single N-V centers; 3. the diamond matrix can withstand extremely high electric fields, thus opening the possibility to affect the optical transition of the N-V center by Stark shift, and tune the nano-cavity eigenstates; 4. the host is non-magnetic, thus providing long spin lifetimes; and 5. the optical center is especially stable, no photo bleaching is observed.

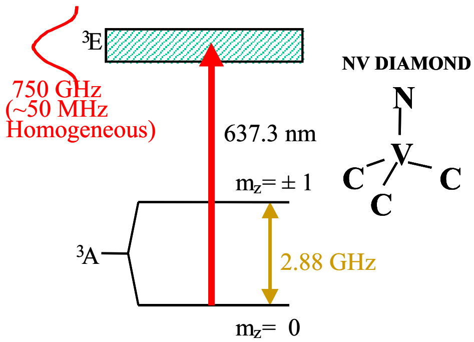


Figure 1: The physical and electronic structure of the [N-V] optical centre in diamond [2].

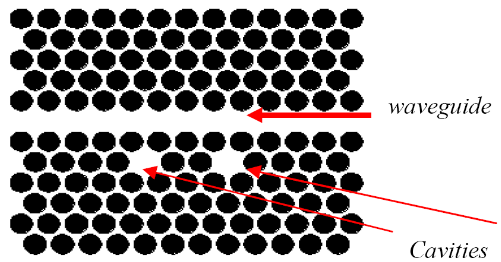


Figure 2: Photonic Crystal structure with 2 cavities coupled to a waveguide.

These remarkable and unique properties, coupled with the diamond specific fabrication tools which we recently developed [7], form the basis for the proposed quantum devices.

3. PHOTONIC CRYSTAL ARCHITECTURE

The input-output, or the "bus" for an operational set of qubits is an optical waveguide. The waveguide is surrounded by a 2D Photonic Crystal in which light propagation is forbidden. Individual defects in the photonic crystal region (missing holes) are formed, to behave as optical nano-cavities. In each optical cavity, an N-V center is inserted. When excited, the NV center relaxation dynamics is spin-dependent. Thus, the atomic transition and the electromagnetic field form a "dressed state" composed by the photon-atom interaction.

When two cavities are placed in close proximity, the photonic component of the dressed state can leak from one cavity to the other (photon hopping) to form two coupled cavities (two coupled qubits). More cavities can be added to form a chain of coupled qubits.

Coherence in a single qubit or in a chain of coupled qubits is maintained as far as these cavities are isolated from the environment (no coupling to the waveguide). Read-out of the state can be accomplished by using a low Q tunable cavity as a quantum Q-switch [8].

4. MODELLING

The first calculation of the proposed structure was a model of a 2D Photonic Crystal composed of a hexagonal array of cylindrical air holes in a dielectric slab [9]. Diamond has smaller refractive index ($n = 2.4$) than other semiconductors, leading to a weaker guidance in the vertical direction and a cavity mode that is less confined vertically than in a silicon based slab. The plane wave expansion method (PWE) was used to calculate the photonic band-structure, a 3D finite-difference time-domain method (FDTD) was used to determine the quality factor ($\boldsymbol{Q}$) of the cavity, with several geometric designs. Details of the bulk crystal parameters (hole radii, slabthickness, cavity geometry, etc.) are presented in a separate paper in this volume. It was found that for small volume nano-cavities, a quality factor of $\boldsymbol{Q} \approx 3.0 \times 10^4$ can be obtained.

In order to obtain higher values of $\boldsymbol{Q}$, the double heterostructure (DH) design, based on a Photonic Crystal waveguide, was considered [10]. The waveguide is defined by a linear defect into the PC slab in certain direction. In our case, one row of holes in the ΓJ direction was filled. The modes were indexed by the symmetry of the B_y field component in the xz plane, where sub-index "e" ("o") stands for the even (odd) parity, for the x and z directions. There are two waveguide modes inside the band gap: B_{oe} and B_{eo}. Each of these modes can serve as a basis for the cavity. Moreover, by decreasing the holes radii in the rows adjacent to the filled row of holes, the effective index is increased resulting in the entrance of the next mode B_{ee} into the band-gap region.

The PC waveguide provides confinement in the lateral direction. For confinement in *longitudinal* direction the lattice constant, a, of a small number of periods along the waveguide axis direction is increased. The region of increased a has a higher refractive index, leading to the formation of a quantized, x-localized mode. The frequency of this mode is lower than the lowest frequency of the surrounding waveguide, but still higher than that of the dielectric band, hence, providing field confinement to the cavity. For the DH cavity with $\Delta a = 0.05a$, and a slab height of $h = 0.8a$ addressing the 1st waveguide mode (B_{oe}), the obtained result is $Q_v = 67,000$. This result seems to be significant for further consideration of photonic crystals in diamond for quantum electrodynamic applications [11].

5. MICROMACHINING OF DIAMOND

The demonstration of the capability to form three-dimensional microstructures in bulk single-crystal diamond was recently presented by Olivero et al., [7] by means of a Focused Ion Beam (FIB) assisted lift-off technique. The lift-off technique consists in the removal of a surface sheet from the bulk sample through the selective chemical etching of a sacrificial graphitic layer created at a well-defined depth by means of MeV ion implantation followed by thermal annealing.

Single-crystal High-Pressure-High-Temperature (HPHT) artificial diamond samples were implanted with 2 MeV He ions in $100 \times 100\,\mu m^2$ square regions, at a fluence of $1 \cdot 10^{17}$ ions·cm^{-2}; the implantation was performed by raster scanning the ion beam focused to a micrometric spot.

MeV ion implantation in a crystalline material induces the creation of lattice defects (i.e., vacancies and interstitials) resulting from the collisions with atomic nuclei. The maximum rate of nuclear energy loss occurs when the ion energy decreases to $\sim 10^2$ eV, i.e., at the end of ion range. Figure 3 shows the density profile of the vacancies induced in diamond by 2 MeV Helium ions, resulting from Monte Carlo simulations; the graph also reports the damage threshold above which graphitization occurs upon thermal annealing. A buried heavily damaged layer is created at a depth of $\sim 3.5\,\mu m$ below the surface, while the region between the surface and the end of range exhibit a much lwer damage density.

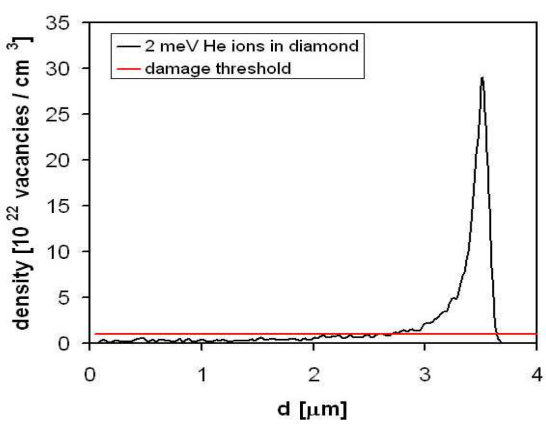


Figure 3: TRIM Monte Carlo simulation of the density profile of vacancies induced in diamond by 2 MeV Helium ions, together with the damage threshold above which graphitization occurs in diamond upon annealing.

The annealed samples were then processed with a wet chemical etching in boiling acid (1:1:1 H_2SO_4 : $HClO_4$: HNO_3). The chemical attack selectively etches the graphitic phase that was exposed to the surface through the drilled trenches, while leaving intact the chemically inert diamond phase. A suspended bridge, fabricated in order to perform preliminary waveguiding experiments was formed, as shown in Figure 4. Subsequently, input and output 45° mirrors were produced at each side of the bridge by FIB.

By using membranes fabricated by this method as a starting material, a 2D array of holes was etched into the diamond membrane, to form a 2D-lab photonic crystal. An example a such a 2D array is presented in Figure 5. Optical experiments are being conducted to study the electromagnetic properties of the produced samples.

In conclusion, some of the basic components for the implementation of basic QIT operations, based on the properties of NV centers in diamond were demonstrated. These include photonic

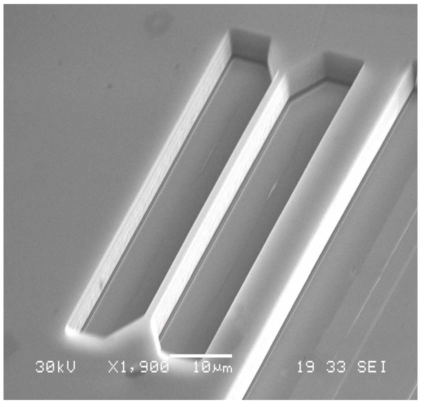


Figure 4: Suspended waveguide in diamond.

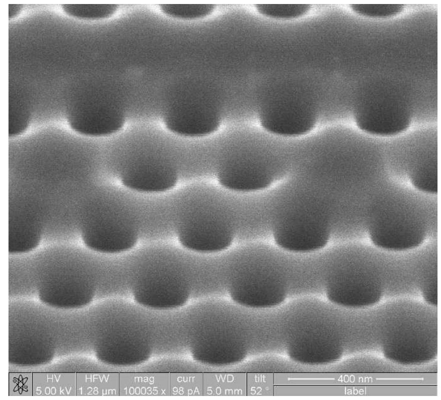


Figure 5: Photonic crystal cavities in diamond.

crystal waveguides, cavities, and coupling elements. Intensive optical testing of these preliminary devices are being performed in our laboratories. Much work is needed to assess the potential of diamond as the base material for QIT.

ACKNOWLEDGMENT

The author is indebted to I. Beyn, B. Meyler, A. Lahav, P. Olivero, A. Greentree and S. Tomljenovic-Hanic who contributed in the experimental and modeling work.

REFERENCES

1. "A quantum information science and technology roadmap, Part 1: Quantum computation," http://qist.lanl.gov.
2. Van Oort, E., N. B. Manson, and M. Glasbeek, "Optically detected spin coherence of the N-V diamond centre in its triplet ground state," *J. Phys. C*, Vol. 21, 4385, 1988.
3. Manson, N. and J. P. D. Martin, "Optically induced spin orientation of the N-V centre in diamond," *Proc. Of the Int. Conf. on Quantum Information*, Rochester, 2001.
4. Jelezko, F., T. Gaebel, I. Popa, A. Gruber, and J. Wrachtrup, *Phys. Rev. Lett*, Vol. 92, 076401, 2004.
5. Jelezko, F., T. Gaebel, I. Popa, M. Domhan, A. Gruber, and J. Wrachtrup, *Phys. Rev. Lett.*, Vol. 93, 130501, 2004.
6. Salzman, J., S. Prawer, and D. Jamieson, *Photonic Crystal Devices and Systems in Diamond*, provisional patent, CCID 131000480, 2004.
7. Olivero, P., J. Salzman, et al., "Ion-beam assisted liftoff technique for 3D micromachinning of freestanding single crystal diamond," *Advanced Materials*, Vol. 17, 2427, 2005.
8. Greentree, A., S. Prawer, and J. Salzman, "Quantum gate for Q switching in monolithic photonic bandgap cavities containing two level atoms," *Physical Review A*, Vol. 73, 013818, 2006.
9. Tomljenovic-Hanic, S., M. J. Steel, C. M. de Sterke, and J. Salzman, "Diamond based photonic crystal microcavities," *Opt. Express*, Vol. 14, 3556, 2006.
10. Song, B., S. Noda, T. Asano, and Y. Akahane, "Ultra-high-Q photonic double heterostructure nanocavity," *Nature*, Vol. 425, 944–947, 2003.
11. Beyn, I. and J. Salzman, "High Q photonic crystal nanocavities on diamond for QED," *J. Europ. Phys. Appl. Phys.*, 2006, to be published.

Broadband Slow Light and Nonlinear Switching Devices

G. Böttger[1], **J.-M. Brosi**[1], **A. Maitra**[1], **J. Wang**[1], **A. Y. Petrov**[2], **M. Eich**[2]
J. Leuthold[1], **and W. Freude**[1]

[1]Institute of High-Frequency and Quantum Electronics
University of Karlsruhe, Engesserstr. 5, 76131 Karlsruhe, Germany
[2]Materials in Electrical Engineering and Optics
University of Technology Hamburg-Harburg, Eissendorfer Str. 38, 21073 Hamburg, Germany

Abstract— In this presentation two classes of compact waveguide structures with functionalities important in high-speed telecommunication networks are reviewed: The first one demonstrates slowing-down of light to 4% of the vacuum speed of light in a photonic crystal line-defect waveguide (PC-LDWG). A fabrication-tolerant bandwidth of 1300 GHz for carriers in the 200 THz range is shown both in 3D-simulations and upscaled microwave experiments. The second class introduces optically bistable and direction-dependant transmission in nonlinear materials for stopband-tapered waveguide Bragg gratings (ST-WBG). Optical switching powers of less than 2.7 mW are required to achieve an optical isolation ratio of 11 within 300 μm propagation through an active material.

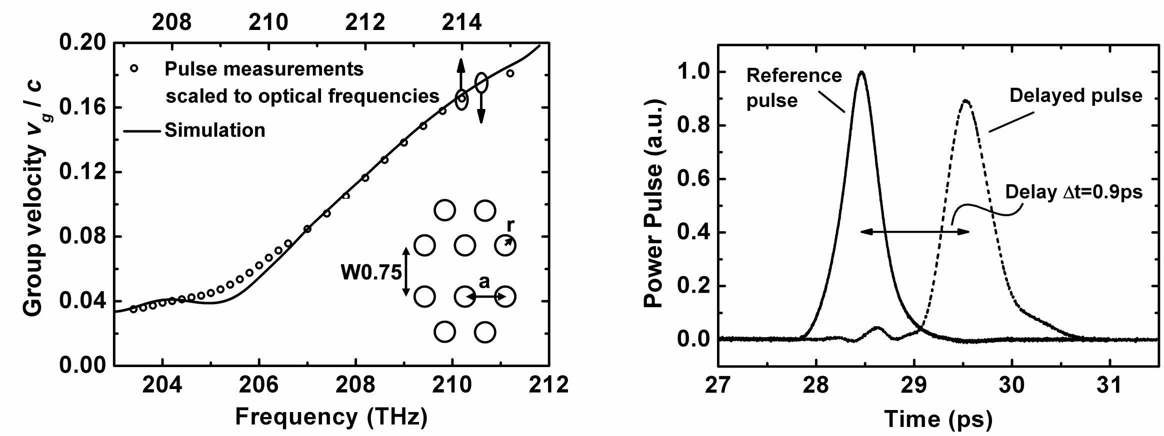
1. INTRODUCTION

Photonic crystals (PCs) offer unique possibilities to control the flow of light. Nanocavities with extremely high-Q factors can be realized [1], dispersion properties of photonic crystal waveguides can be engineered, light pulses can be slowed down or even stopped [2, 3]. Slow-light structures significantly reduce the size and power needed for optical modulators and nonlinear elements [4]. Optical bandwidths greater than 20 GHz and low dispersion needed for optical communication systems have however not been demonstrated experimentally so far. Furthermore, the effect of unavoidable fabrication-induced disorder on the group velocity in photonic crystal line-defect waveguides (PC-LDWG) has not yet been studied. We investigate a PC-LDWG designed for broadband slow-light transmission according to [3]. Experiments were conducted in an upscaled microwave model [5] and can be directly translated to optical frequencies because of the scaling laws of Maxwell's equations. Waveguide structures for the experiments are upscaled by a factor of 20000, while the frequency is downscaled by the same amount. This way it is possible to study, e.g., the effect of disorder in PC air-hole radii on the group velocity systematically. Measured results were transformed back into the optical regime and agreed well with 3D finite integration simulations performed with CST Microwave Studio (MWS).

We also show a second class of nonlinear material waveguide devices that are asymmetric concerning forward and backward propagation of light. The asymmetry is caused by Bragg gratings on the sidewalls of strip waveguides. The grating amplitude is varied linearly along the waveguide, consequently field confinements change along the waveguide. In conjunction with nonlinear materials a directionality of the waveguide device is established. Furthermore, transmission not only depends on the propagation direction, but also on optical input powers. Effectively a hysteresis loop can be opened, which is usable for bistable switching devices or isolators. Modified gratings on the waveguide sidewalls change the position and extension of a photonic stopband along the waveguide, justifying the term "stopband-tapered waveguide Bragg grating" (ST-WBG) [6, 7]. Equivalently structured active materials with nonlinear gain, e. g., semiconductor optical amplifiers (SOAs), enhance the device performance by reducing required switching powers.

2. BROADBAND SLOW SLIGHT WAVEGUIDE STRUCTURE

A slow light PC-LDWG was designed in a membrane structure with relative permittivity $\varepsilon_r = 10$, which is close to the value for semiconductor materials in optical devices. The specific parameters of the PC-LDWG are a lattice constant $a = 0.45$ μm, membrane height $h = 0.27$ μm, hole radius $r/a = 0.25$ and a central linear defect width $0.75\,a$, see inset schematic of Figure 1(a). This structure confines light on a specific mode of the central defect, and may be probed by measuring the propagation of light pulses. Figure 1(a) characterizes the frequency dependant group velocity of such pulses in simulations and experiments. The chosen design slows down light pulses to group velocities as small as 4% of the speed of light. This is achieved over a very wide range of 1300 GHz on

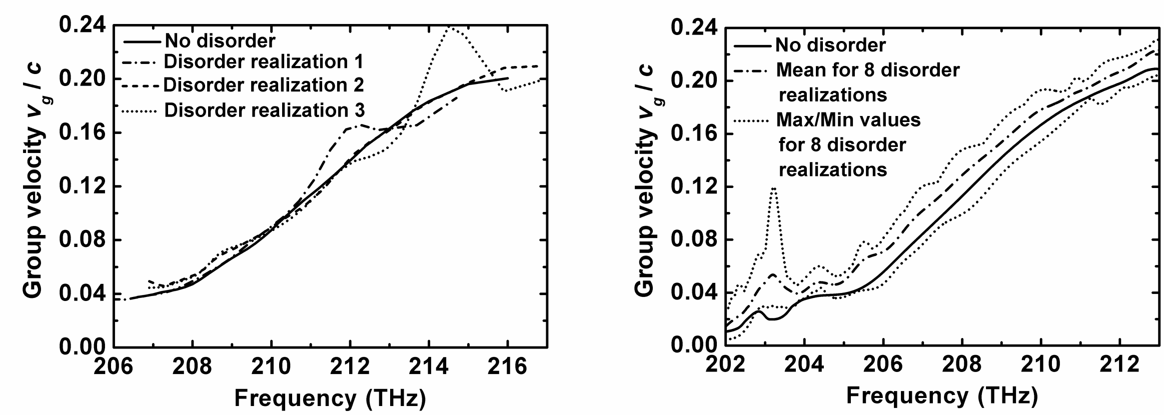


(a) (b)

Figure 1: (a) Group velocity plotted as a function of optical frequency for simulated and scaled-back measurement. The inset shows the structure parameters of the disorder-free PC-LDWG, (b) Pulse measurements. The delayed pulse is transmitted through 30 periods a of the slow-light device. The time axis is scaled to the optical regime, and the pulse energy normalized.

the left side of the plot in Figure 1(a). The experimentally measured curve deviates only slightly from the 3D-simulation. The upper frequency axis has been scaled from the microwave regime experiment to the optical domain. It is shifted with respect to the lower, simulation-frequency axis to fit curves. This mismatch is attributed to an uncertainty in the effective relative permittivity of the dielectric material used. Furthermore, simulations assume ideal, lossless materials.

For the microwave experiment, a slot antenna was designed which excites the TE-polarized mode in a strip waveguide. To better couple to the low group velocity mode of the PC-LDWG, taper sections were introduced, slightly increasing the width of the PC-LDWG toward the input and output strips. Pulse transmission experiments were carried out with a modulated microwave source and a detector. Pulses were nearly Gaussian with an equivalent optical duration of 0.5 ps and a bandwidth of 1.3 THz. Figure 1(b) shows the actual delay of a pulse transmitted through a perfectly structured W0.75 waveguide with a length of 30 lattice constants. The reference pulse passes through the tapered regions, only, with the slow-light portion of the device taken out.



(a) (b)

Figure 2: (a) Influence of disorder on group velocity. MW measurements scaled to optical frequencies, (b) numerical study of disorder influence on group velocity for fixed material constants.

3. INFLUENCE OF DISORDER ON GROUP VELOCITY

To study the effect of disorder on the group velocity characteristics of the device, we introduced normally distributed variations of the radii of the air cylinders. The standard deviation was set to 5% of the nominal radius. We produced three different realizations, each with a length of 15 periods. In Figure 2(a) the group velocity measurements from the microwave pulse experiments are displayed. The samples had slightly different material properties which introduced frequency offsets between the curves. These have been leveled out to compare the effect of the geometric disorders between the different realizations. The low group velocity propagation of the probing pulse does not deteriorate. We also conducted 3D-simulations for 8 different realizations of radially disordered PC-LDWGs with constant material properties, see Figure 2(b). The mean group velocity of the disordered systems (dash-dot) in comparison with an ideal structure (solid line) is higher. Single system performance varies over the indicated range. This behavior will be subject of future investigations.

The PC linear defect waveguides described up to this point were highly symmetric, apart from unwanted disorder from fabrication inaccuracies. In the following two sections we drop the symmetry requirement with regard to forward and backward direction of the waveguide. This second class of waveguides also involves periodic structuring. Stopbands, however, are modified with the side wall corrugations of strip waveguides, resulting in stopband-tapered waveguide Bragg gratings (ST-WBG). In combination with nonlinear materials, either passive or active, the formed strip waveguides have propagation characteristics which depend on the direction and level of the optical input powers.

4. MODELING OF NONLINEAR SWITCHING DEVICES

Passive and active nonlinear Bragg gratings of spatial period Λ are modeled using coupled mode theory [8]. We formulate a system of coupled partial differential equations [6, 7] for wavelengths close to the Bragg wavelength in either passive or active material systems, and we neglect dispersion:

$$\begin{aligned} &\pm\frac{\partial A_{f,b}}{\partial z}+\frac{\partial A_{f,b}}{\partial t}=i\Delta\beta_{f,b}A_{f,b}+i\kappa A_{b,f},\\ &\tau\frac{\mathrm{d}g(z,t)}{\mathrm{d}t}=g_0(z)-\left[1+\left(|A_f|^2+|A_b|^2\right)/P_{\mathrm{sat}}\right]g(z,t),\\ &\Delta\beta_{f,b}=\begin{cases}\delta(z)+\Gamma_{NL}(z)\left(|A_{f,b}|^2+2|A_{b,f}|^2\right), & \text{(passive)}\\ \delta(z)-i\frac{1}{2}(1-i\alpha)\,g(z,t) & \text{(active)}\end{cases}\end{aligned} \tag{1}$$

The envelopes of the forward and backward traveling A_f and A_b are assumed to vary slowly with time t and position z. The carrier angular frequency is ω_0, k is the propagation constant, $k_B=\pi/\Lambda$ the Bragg wavenumber, $\Gamma_{NL}\propto n_2$ the contribution from the Kerr type nonlinearity, α the SOA linewidth-enhancement factor, g is the modal gain, g_0 is the unsaturated gain, P_{sat} the saturation power, τ the carrier lifetime, and κ is the grating's linear coupling coefficient. The detuning parameter is defined by $\delta_0 := \delta(\omega_0) = k(\omega_0) - k_B$.

Structural non-uniformity can be introduced by e.g., a chirped or stopband-tapered grating, by an asymmetric phase shift, or by a combination of these. If nonlinearities are involved, such devices show a nonreciprocal transmission characteristic. By way of illustration, we discuss a grating with length L, where the effective refractive index $n(z)$ changes both with z and the electric field magnitude E according to $n(z) = n_0 + \Delta n_G(z)\cos(2\pi z/\Lambda) + n_2|E|^2$. The structure's bandgap is tapered by varying the grating amplitude $\Delta n_G(z)$ linearly along z. This changes the coupling coefficient $\kappa(z) = \pi\Delta n_G(z)/(2n_0\Lambda) = \kappa(0)[1 + \Delta\kappa z/L]$, with $\Delta\kappa$ as tapering coefficient. For the same operating point the penetration depth of the input field is different if the light is incident from different directions, which in turn leads to a nonreciprocal behaviour, as shown in the following.

5. PERFORMANCE OF NONLINEAR SWITCHING DEVICES

We consider an InP-based ST-WBG with $n_0 = 3.4$, an effective cross section area $0.36\,\mu\mathrm{m}^2$, a length $L = 300\,\mu\mathrm{m}$, and a Bragg wavelength $\lambda_B = 1.55\,\mu\mathrm{m}$ ($k_B = 2\pi n_0/\lambda_B = \pi/\Lambda$, $\Lambda = 228\,\mathrm{nm}$). For the active grating $\kappa L = 3$, $\Delta\kappa = -30\%$, $g_0L = 1.2$, $\delta_0L = 5$ and $\alpha = 5$ hold, and for the passive grating $\kappa L = 5$, $\Delta\kappa = -15\%$, $\delta_0L = 4.75$ are valid. From a numerical evaluation of Eq. (1) we find the hysteresis curves for the active and the passive grating with and without ($\Delta\kappa = 0\%$) stopband-tapering. The input is a CW signal with a detuning δ_0 inside and near the upper edge of the shifted stopbands, compare Figure 3, either incident on the ST-WBG from *left-to-right* (LTR), in which direction the gratings are *negatively* tapered, or, with the same intensity, from *right-to-left* (RTL), where the tapering is *positive*.

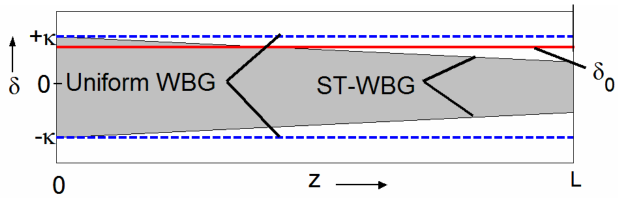


Figure 3: Stopband-tapered waveguide Bragg grating (ST-WBG) of length L. Schematic stopband dependence on z with band edges for uniform grating (− − −), and for ST-WBG (upper and lower boundaries of shaded region), and detuning parameter δ_0 of CW input signal (—).

For the active (passive) grating and an RTL input power of 2.7 mW (130 mW) the output power is 6.8 mW (120 mW) according to Figure 4(a) (Figure 4(b)), while the same input power of 2.7 mW (130 mW) for LTR operation leads to an output power of 0.62 mW (15 mW) with an associated isolation ratio of 11 (8). The up- and down-switching thresholds scale linearly with the grating length [10], so doubling L would reduce the relevant powers by a factor of two. Figure 4(c) displays the impulse response of the passive grating with the same detuning $\delta_0 L = 4.75$ as before (within the stopband near its upper edge at z=0) and for an input peak power of 170 mW, either for LTR or RTL operation. The average input power of 85 mW is close to the RTL up-switching threshold in Figure 4(b). For RTL propagation the transmitted impulse is larger (and more compressed because of the high dispersion near the band edge) than for the LTR direction (where the carrier frequency is farther away from the band edge and dispersion is lower).

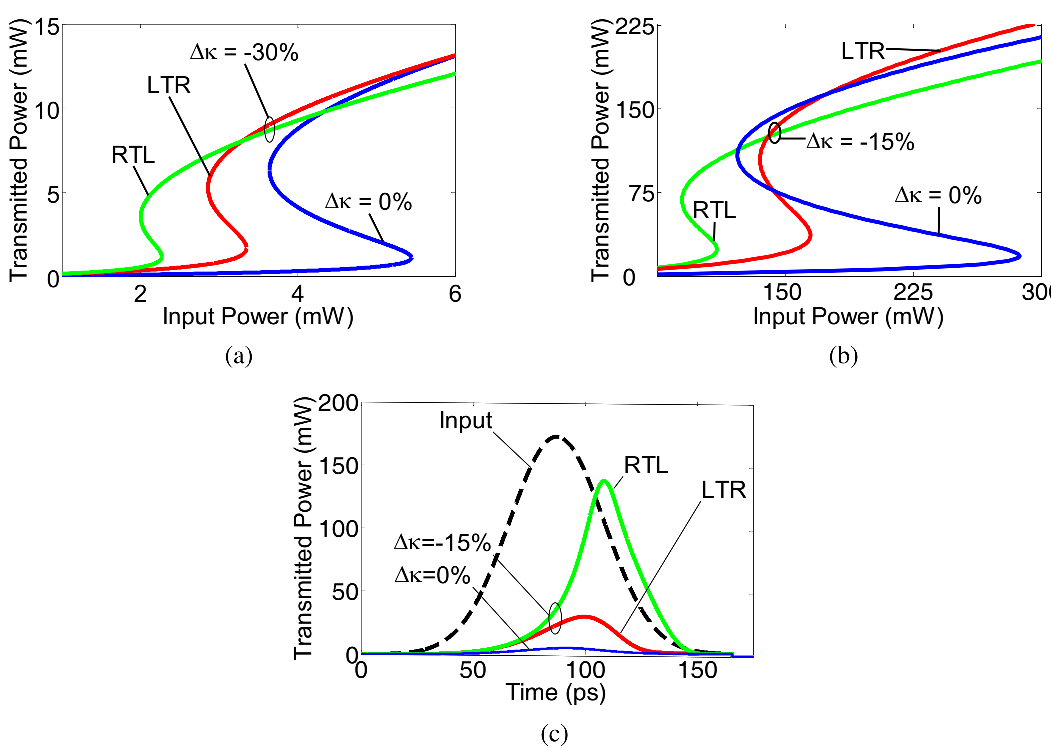


Figure 4: Nonreciprocal behaviour in active (a) and passive (b), (c) ST-WBG with different stopband tapering. The fraction of power which is not transmitted is reflected. Hysteresis curves in (a), (b), impulse responses in (c).

6. CONCLUSIONS

We have discussed two classes of waveguide devices. The first one consists of photonic crystal line-defect waveguides realizing slow modes propagating with group velocities of 0.04 c over an unprecedented bandwidth of 1300 GHz. Measurements in PC-LDWGs geometrically scaled up for operation at microwave frequencies confirmed the performance of the 3D-simulated designs. A variation of the radii of air cylinders in a range to be expected when producing respective PC-LDWGs at optical frequencies did not degrade the slowing down of pulses, which is prospective for telecommunication applications.

Breaking symmetries in a second class of structured devices, stopband-tapered waveguide Bragg gratings, enabled all-optically switching or isolating devices, which are of high interest in the conditioning of light in optical networks. It should be pointed out that elements of the latter class may be incorporated into PC-LDWGs of the former class, be it the non-uniformity of the waveguide itself, the explicit use of active materials, or the operation at different power levels. Methods to enable and investigate such novel nanostructured devices have been put forward.

REFERENCES

1. Song, B.-S. et al., "Ultra-high-Q photonic double-heterostructure nanocavity," *Nature Materials*, Vol. 4, 207–210, 2005.
2. Mori, D. and T. Baba, "Wideband and low dispersion slow light by chirped photonic crystal coupled waveguide," *Opt. Express*, Vol. 13, 9398–9408, 2005.
3. Petrov, A. Y. and M. Eich, "Zero dispersion at small group velocities in photonic crystal waveguides," *Appl. Phys. Lett.*, Vol. 85, 4866–4868, 2004.
4. Soljačić, M. and J. D. Joannopoulos, "Enhancement of nonlinear effects using photonic crystals," *Nature Materials*, Vol. 3, 211–219, 2004.
5. Freude, W., G. A. Chakam, J.-M. Brosi, and C. Koos, "Microwave modelling of photonic crystals," *Photonic Crystals-advances in Design, Fabrication, and Characterization*, K. Busch, S. Lölkes, R. Wehrspohn, and H. Föll, Eds., Wiley-VCH, Weinheim, 2004.
6. Maitra, A., C. G. Poulton, J. Wang, J. Leuthold, and W. Freude, "Low switching threshold using nonlinearities in stopband-tapered waveguide Bragg gratings," *IEEE J. Quant. Electron.*, Vol. 41, 1303–1308, Oct. 2005
7. Freude, W. A. Maitra, J. Wang, C. Koos, C. Poulton, M. Fujii, and J. Leuthold, "All-optical signal processing with nonlinear resonant devices," *Proc. 8th Intern. Conf. on Transparent Optical Networks (ICTON'06)*, Nottingham, UK, paper We.D2.1, June 2006.
8. De Sterke, C. M. and J. E. Sipe, "Switching dynamics of finite periodic nonlinear media: A numerical study," *Phys. Rev. A*, Vol. 42, 2858–2869, Sept. 1990.
9. Winful, H. G., J. H. Marburger, and E. Garmire, "Theory of bistability in nonlinear distributed feedback structures," *Appl. Phys. Lett.*, Vol. 35, 379–381, Sept. 1979.
10. De Sterke, C. M. and J. E. Sipe, "Gap solitons," *Progress in Optics*, Vol. XXXIII, 203–260, North-Holland Amsterdam, 1994.

Making Quarter Wavelength Notch Antennas Wideband

P. J. Massey[1], **K. R. Boyle**[1], **A. J. M. de Graauw**[2], **M. Udink**[2], **and D. L. Raynes**[3]

[1]NXP Semiconductors, UK
[2]NXP Semiconductors, The Netherlands
[3]Ansoft Corporation, UK

Abstract— Very compact yet efficient UWB antennas are created by utilizing the ground plane as the antenna. There are notches in the ground plane, some to act as the coupling element(s), and some for optimizing the antenna bandwidth. The performance of the antennas compares favorably with wideband dipole designs occupying the same area.

1. INTRODUCTION

Notch antennas have a long history of being used as compact designs, having originally been integrated into missiles and aircraft flying surfaces [1, 2]. More recently notch antennas have been proposed as compact structures for UWB (Ultra WideBand) applications [3–5]. Some designs achieve high bandwidth using flared notches [3, 6], and others use relatively wide notches [4]. The very compact structures discussed in [5] conserve the ground plane area that can be shared with circuitry by using relatively narrow notches. This paper re-examines the structures discussed in [5], describes the geometrical and electromagnetic features used their design, and discusses the reasons for the exceptionally wide bandwidth achieved in their limited size.

2. FEATURES USED IN NARROW WIDTH NOTCH UWB ANTENNAS

Figure 1 illustrates how the impedance of a notch antenna is affected by ground plane dimensions, and the principles of how a very wideband notch antenna is designed. The notch configuration is shown on the left of the figure. The feed is 5 mm away from the short end of the notch, and for discussion purposes, there is no dielectric substrate. The 18 mm notch length is a quarter wavelength long at about 4 GHz.

The first Smith chart shows the response when the ground plane is very large. Increasing the notch width from 1 mm to 2 mm would result in the loop reducing in diameter. Introducing a dielectric substrate would both change the centre frequency of the loop, increase the loop diameter, and move the loop downwards and to the left of the chart. When the feed is moved further away from the notch's shorted end, the loop in the Smith chart grows and moves to the downwards and to right. If the feed is far enough away from the shorted end, the loop crosses the real impedance axis, and therefore for a relatively narrow bandwidth, the response is well matched to a real impedance. In the past notch antennas have been matched using these dielectric substrate and feed location mechanisms.

As the Smith chart to the top right of the figure shows, providing the ground plane depth is significantly greater than the notch length, reducing the depth of the ground plane has relatively little effect upon the antenna response. However, as shown on the bottom right of the figure, adjusting the width can make a huge difference to the response. Finite ground plane widths introduce a second loop into the response curve. As this loop tends to be centered on a frequency at which the ground plane width is close to half wavelength across, it is believed to be associated with a dipole like response in the ground plane (the notch acting as the separation between the two dipole elements. When the ground plane width is about twice the electrical length of the notch, the two loops are centered on similar frequencies. This results in the impedance varying little over a wide bandwidth, with slight reduction in reactance with increasing frequency over this bandwidth. Fortunately, this reactance slope can easily be compensated for using a series capacitor.

The bottom left Smith chart shows the response when the antenna has been matched with a series capacitor. This preserves the relative constancy of the impedance between 2.5 and 5.5 GHz, and achieves a return loss of better than 10 dB between these frequencies. Beyond this frequency range, the impedance diverges to the edges of the Smith chart.

Note that the constructive combination of the dipole and notch resonance mechanisms relies upon the notch's open end lying close to halfway across the width of the dipole. Offsets of around 1/6th of the ground plane width have been shown to detrimentally widen the response shown in the figure's lower right Smith chart.

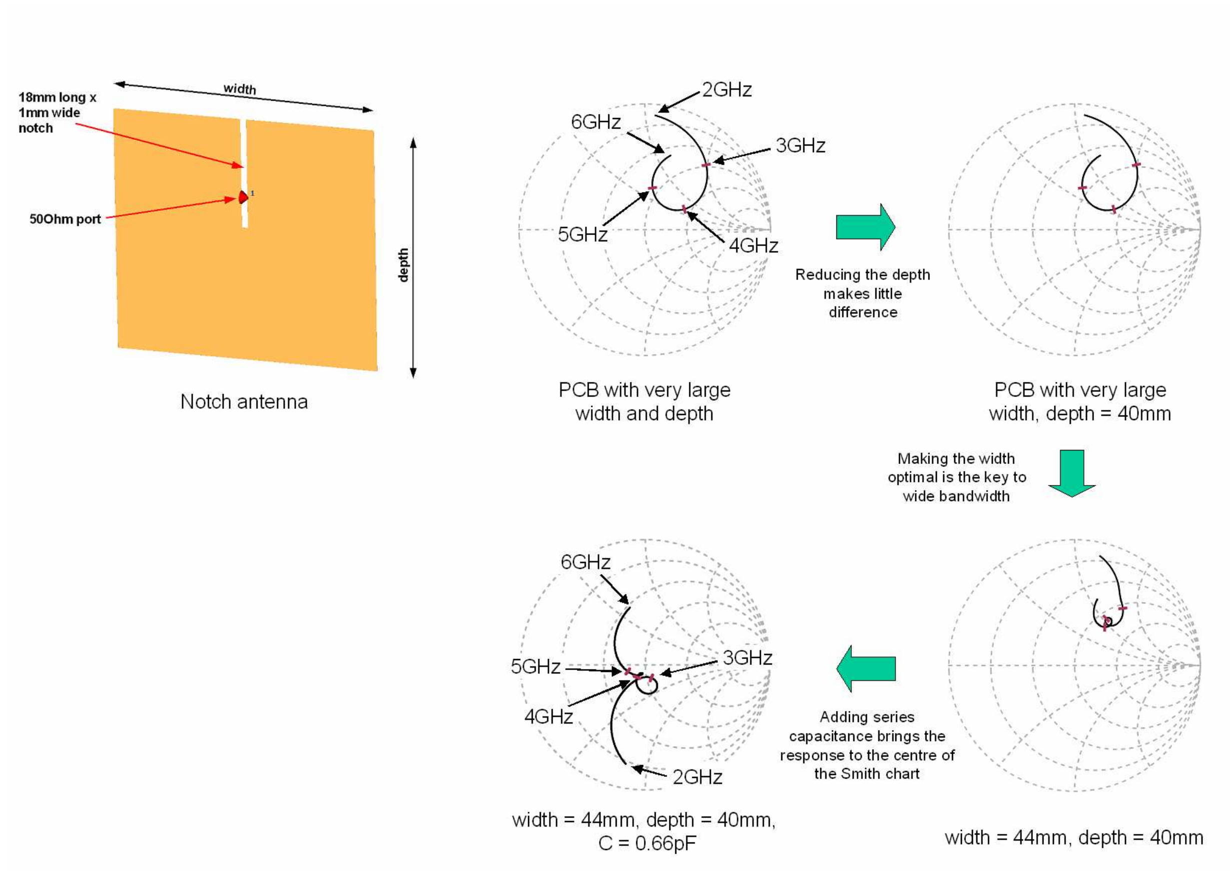


Figure 1: Variation in impedance with ground plane dimensions of a notch antenna.

The conference presentation includes charge and field plots that identify and illustrate the contributions of the dipole and notch resonance mechanisms to the antenna's behavior.

3. PRACTICAL IMPLEMENTATIONS

For clarity of presentation, the above description has used a straight notch with no substrate. In practice, to save space, the notch can be folded into an L shape, and the ground plane's depth can be reduced. Also the ground plane can be a PCB that includes conventional dielectric substrates. Reference [5] describes some examples of practical designs, including the use of two fed notches and choke notches in achieving coverage of the FCC UWB 3.1 to 10.6 GHz band, while suppressing coupling to the 5.5 GHz WLAN band. The choke notches act to restrict the effective width of the ground plane in situations where the overall width of the ground plane is larger than that required for wide bandwidth antenna operation.

ACKNOWLEDGMENT

The original work on notch antennas for consumer devices was undertaken as a joint development between Philips Semiconductors Nijmegen and the antennas team in Redhill at Philips Research UK. Later work specifically on UWB notch antennas was undertaken at Redhill with the support of Philips Semiconductors' San Jose (USA) branch. All these organizations are now part of NXP Semiconductors.

REFERENCES

1. Johnson, W. A., "The notch aerial and some applications to aircraft radio installations," *Proc. IEEE (London)*, Vol. 102, Part B, 211–218, 1955.
2. Johnson, R. C. and H. Jasik, *Antenna Engineering Handbook*, McGraw-Hill, New York, 1961.
3. Foster, P., "UWB antennas issues," *Proceedings of IEE Seminar on "Ultra Wideband Communications Technologies and System Design"*, 69–83, London, UK, July 2004.

4. Latif, S., L. Shafai, and S. K. Sharma, "Bandwidth enhancement and size reduction of microstrip slot Antennas," *IEEE Trans. AP*, Vol. 53, No. 3, 994–1003, 2005.
5. Massey, P. J., K. R. Boyle, A. J. M. de Graauw, M. Udink, and D. L. Raynes, "Optimised UWB notch antennas for miniaturized consumer electronics applications," *Proceedings of IET Seminar on "Ultrawideband Systems, Technologies and Applications"*, 263–267, London, UK, April 2006.
6. Burberry, R. A., *VHF and UHF Antennas*, Peter Peregrinus, London, 1992.

A Frequency Notched Inverted-trapezoid UWB Antenna

Min Ding, Ronghong Jin, and Junping Geng
Department of Electronic Engineering, Shanghai Jiao Tong University, 200240, China

Abstract— A new ultra wideband monopole antenna is presented in this paper. An inverted-trapezoid structure is constructed to obtain ultra wideband (UWB) performance. The parameters and characteristics of the antenna and the simulation results show that the ultra wideband performance is achieved with the inverted-trapezoid structure. An inverted-trapezoid slot is used to be exempt from interference with existing wireless systems.

1. INTRODUCTION

The UWB antenna has become an intensive topic of current antenna research due to some of its unique features such as transmitting and/or receiving very short time durations of electromagnetic energy and avoiding frequency dispersion and space dispersion. Recently, many methods are developed to realize conventional UWB antennas [1–3] in the commercial domain, the frequency range of which is between 3.1 GHz and 10.6 GHz meeting with FCC standard. However, they cause interference for existing wireless communications, for example, wireless LAN with IEEE 802.11a, and so on. Consequently, a good antenna for UWB applications should provide low voltage standing wave ratio (VSWR) and be exempt from interference with existing wireless systems. Recently, UWB antennas with band-notched characteristic were presented in [4, 5].

In this paper, an inverted-trapezoid structure is used to construct a UWB antenna, on which an inverted-trapezoid slot is embedded to reject the existing wireless local area network (WLAN) band, such as the 5.25-GHz band (5.15–5.35 GHz) and 5.8-GHz band (5.725–5.875 GHz). This kind of band-notched UWB antenna requires no external filters and thus greatly simplifies the circuit design of the communication system.

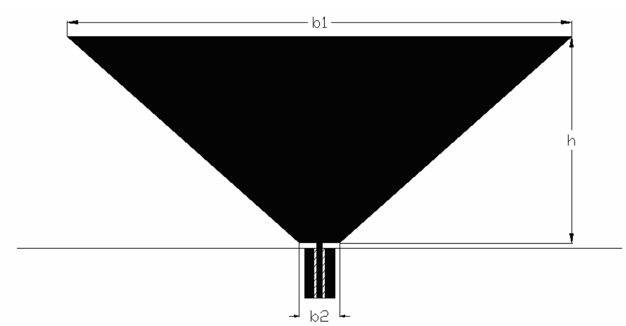


Figure 1: The structure of the inverted-trapezoid antenna without slots.

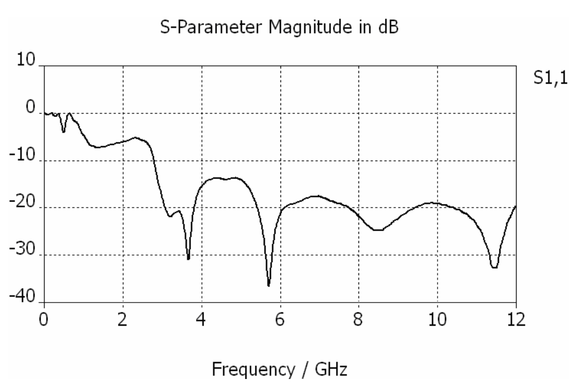
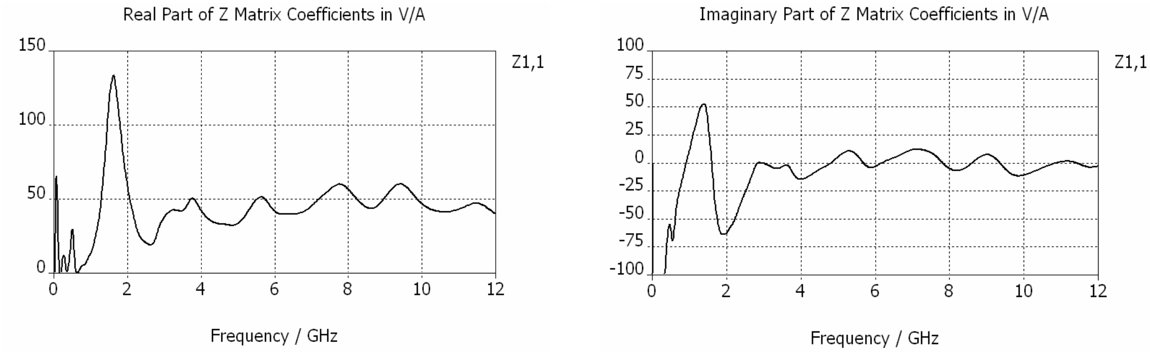


Figure 2: S_{11} of the inverted-trapezoid antenna without slots.

2. ANTENNA CONFIGURATIONS AND SIMULATION RESULTS

The inverted-trapezoid antenna structure without slots is shown in Figure 1. In particular, the inverted-trapezoid antenna element is made by a 0.18-mm copper sheet, which is located vertically



(a) (b)

Figure 3: The input impedance of the inverted-trapezoid antenna without slots. (a) the real part, (b) the imaginary part.

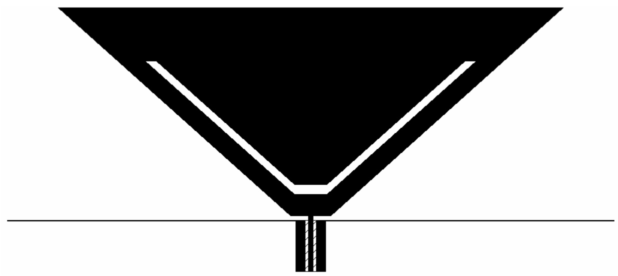


Figure 4: The structure of the antenna with the trapezoid-shape slot.

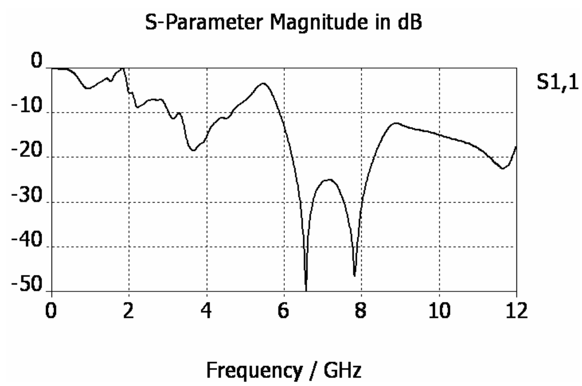
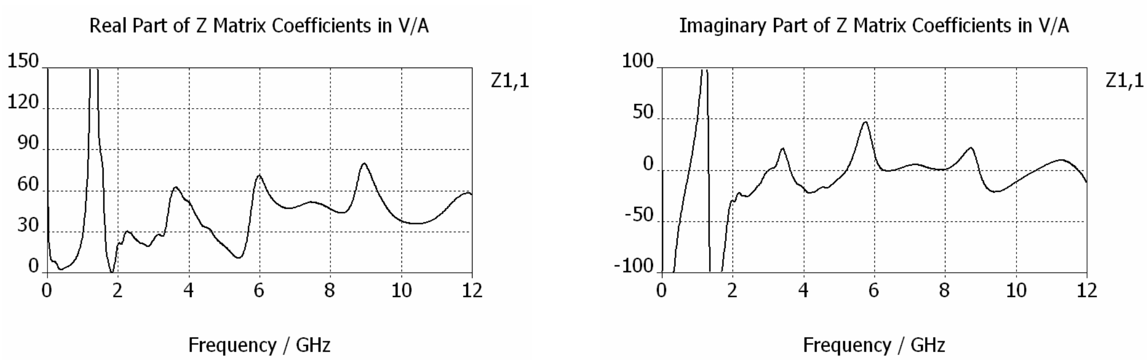


Figure 5: S_{11} of the antenna with the trapezoid-shape slot.



(a) (b)

Figure 6: The input impedance of the antenna with the trapezoid-shape slot. (a) the real part, (b) the imaginary part.

above a finite-size ground plane with an SMA connector. The rectangular ground plane has dimensions of $120 \times 50 \times 1$ mm, with a thickness of 1 mm. Return losses are simulated for the prototype antenna without slots as shown in Figure 2. It is known from Figure 2 that this prototype antenna

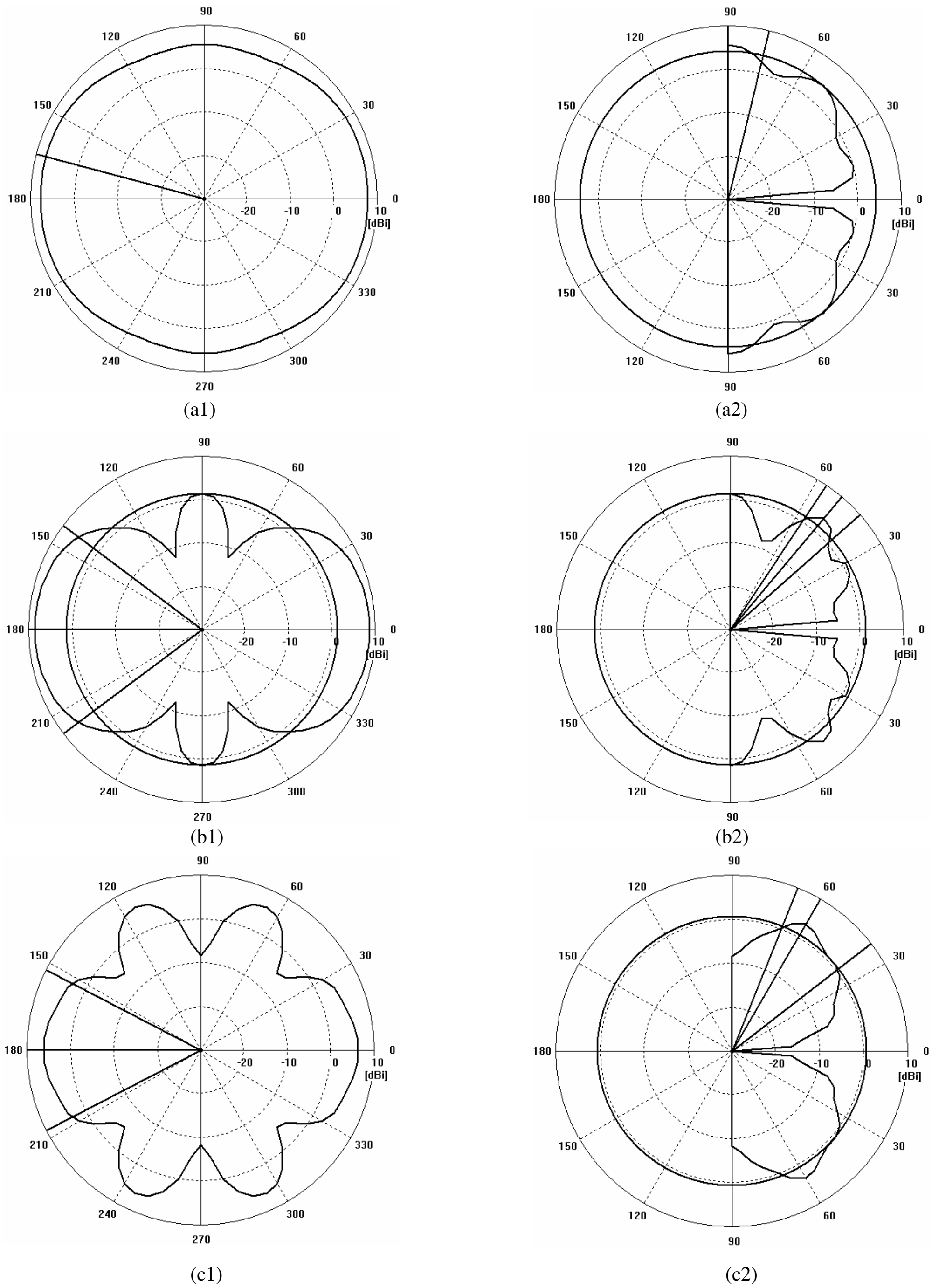


Figure 7: Radiation patterns of the second order iteration structure. (a1) E-plane in 4 GHz, (a2) H-plane in 4 GHz, (b1) E-plane in 7 GHz, (b2) H-plane in 7 GHz, (c1) E-plane in 10 GHz, and (c2) H-plane in 10 GHz

covers frequency range from 2.8 GHz to above 12 GHz. The simulated input impedance of the prototype antenna in Figure 3 is affected by changing parameters of the top side width $b1$ of the inverted-trapezoid antennathe bottom side width $b2$ of the inverted-trapezoid antenna, the height h of the inverted-trapezoid antenna, and the distance between the bottom side of the inverted-trapezoid antenna and the ground plane. Figure 3 demonstrates that the prototype antenna yields good UWB impedance performance.

The band-notched antenna is composed of the inverted-trapezoid prototype antenna structure with the inverted-trapezoid slot is shown in Figure 4. The return losses are simulated for the band-notched antenna with the inverted-trapezoid slot as shown in Figure 5. It is known from Figure 5 that the band-notched antenna with the inverted-trapezoid slot covers frequency range from 3 GHz to above 12 GHz, except from 4.7–5.9 GHz. As a result, the notched bandwidth of the antenna with the inverted-trapezoid slot satisfies the requirement for exempt from interference with existing wireless systems. It is shown in Figure 6 that input impedance of the band notched antenna is matched for 3 GHz to above 12 GHz, except from 4.7–5.9 GHz, which satisfies the requirement for exempt from interference with existing wireless systems, too. Figure 7 plots the simulated E-plane and H-plane radiation patterns at 4 GHz, 7 GHz and 10 GHz.

3. CONCLUSIONS

A new inverted-trapezoid ultra wideband antenna has been presented in this paper, and the interference with existing wireless systems is prevented successfully by the inverted-trapezoid slot. The simulated return losses and input impedance of the inverted-trapezoid prototype antenna demonstrate that the prototype antenna provides good UWB frequency performance. The simulated return losses, input impedance and radiation pattern of the band notched antenna demonstrate that an impedance bandwidth of 3–12 GHz, with VSWR$<$2, while avoiding interference with the 5.15–5.35-GHz and 5.15–5.825-GHz bands, which are occupied by existing wireless systems.

REFERENCES

1. Yang, T. and S.-Y. Suh, "Compact antennas for UWB applications," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 15, No. 5, 16–20, 2004.
2. Qiu, X. N., H. M. Chiu, and A. S. Mohan, "Techniques to improve ultra wide band performance of planar monopole antenna," *IEEE/ACES International Conference on Wireless Communications and Applied Computational Electromagnetics*, 186–190, 3-7 April, 2005.
3. Qing, X. and Z. N. Chen, "UWB characteristics of disc cone antenna," *IEEE International Workshop on Antenna Technology: Small Antennas and Novel Metamaterials*, 97–100, 7–9 March, 2005.
4. Lee, W. S., D. Z. Kim, and K. J. Kim, "Multiple frequency notched planar monopole antenna for multi-band wireless systems," *The European Conference on Wireless Technology*, 535–537, 3–4 Oct., 2005,
5. Qiu, J., Z. Du, and J. Lu, "A band-notched UWB antenna," *Microwave and Optical Technology Letters*, Vol. 45, No. 2, 152–154, 2005.

The Design of an UWB Antenna with Notch Characteristic

Jaehoon Choi, **Seokjin Hong**, **and Uisheon Kim**
Department of Electrical and Computer Engineering, Hanyang University
17 Haengdang-dong, Seongdong-gu, Seoul 133-791, Korea

Abstract— Four types of novel and compact ultra-wideband (UWB) microstrip-fed monopole antennas having frequency band-stop function are presented. To generate the band-stop performance, modified planar monopoles with inverted U-slot, U-slot, small strip bar and $\lambda/4$ C-shaped slot are used. The designed antennas satisfy the -10 dB return loss requirement in the frequency band between 3 and 11 GHz, while showing the band-stop characteristic in the frequency band of 5.0 to 5.9 GHz.

1. INTRODUCTION

Since the Federal Communications Commission (FCC) released the commercial use of UWB radio system, many researchers have been paying much attention to modern in door radar applications. For the reliable use of these radar services, the development of UWB antennas operating from 3.1 to 10.6 GHz is inevitable. To satisfy such a requirement, various wide band antennas have been studied [1–8]. Among the many possible alternatives, printed monopole antennas have been extensively investigated because of their attractive features such as light weight, simple structure, and ease of mass production. Since the existing wireless LAN and Hiper-LAN service bands (5.15–5.825 GHz) overlap the UWB service band, UWB radio signal can interfere with those of WLAN and Hiper-LAN services. To overcome this problem, UWB antennas with band-stop performance are desirable. In this paper, wideband antennas with band-stop characteristic are proposed.

2. ANTENNA DESIGN

2.1. An Inverted U-slot Method

The inverted U-slot is inserted in the middle of the radiating patch as shown in Fig. 1. The inverted U-slot has a length and the slot width 0.5 mm. The tilt angle (α) of the U-slot is about 16 and other dimensions of the antenna are shown in Fig. 1. The length equals about 0.29λ at the desired notch frequency of 5.45 GHz.

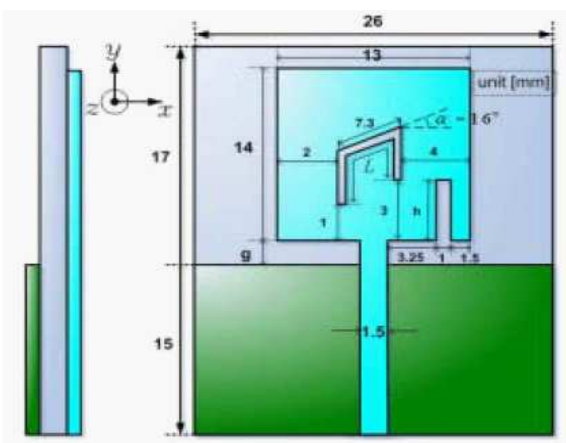


Figure 1: The geometry of proposed antenna.

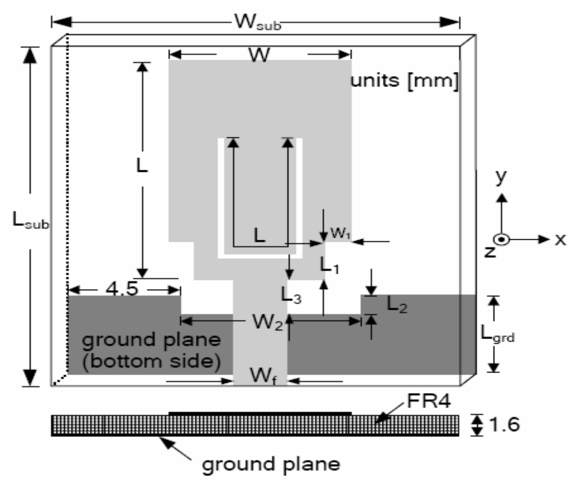


Figure 2: The geometry of proposed antenna.

2.2. An U-slot Method

The geometry of the proposed compact microstrip-fed monopole antenna with band-stop characteristic for ultra-wideband applications is illustrated in Fig. 2. the antenna structure is a rectangular patch of $7 \times 11.5\,\text{mm}^2$ with the two notches of dimensions $1 \times 2.5\,\text{mm}^2$ at the two lower corners of the patch. In addition, $\lambda/4$ U-shaped thin slot is inserted into the radiating patch. The use of a U-shaped slot inserted into the patch yields the frequency band-stop characteristic.

2.3. Method Using a Small Strip Bar

The geometry of the proposed microstrip-fed monopole antenna is illustrated in Fig. 3. The proposed antenna consists of two monopoles with the same size ($w \times l$) and a small strip bar ($l_s \times w_s$), and is printed on the FR4 substrate with height (h) of 1.5 mm and relative permittivity of 4.4. A small strip bar is located between the two monopoles and connected at the end of microstrip feed line. The insert of strip bar into patch yields the frequency band-stop characteristic.

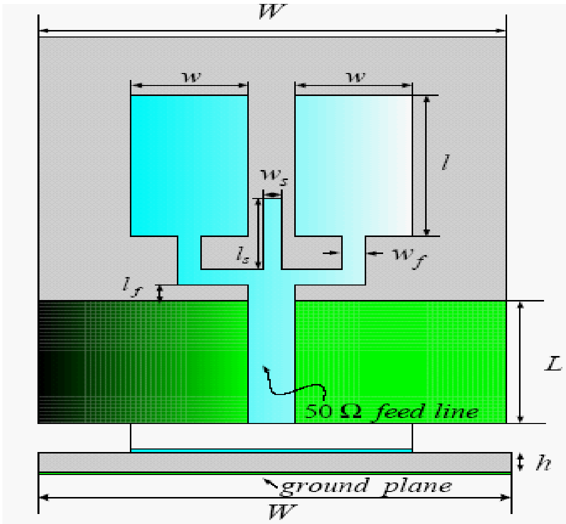


Figure 3: The geometry of proposed antenna.

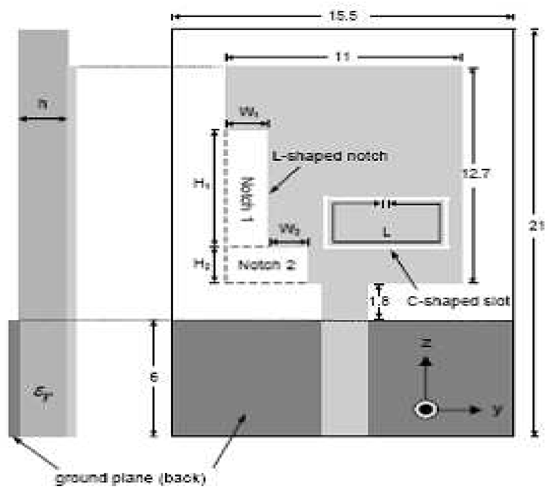


Figure 4: The geometry of proposed antenna.

2.4. Method Using a $\lambda/4$ C-shaped Slot

The configuration of the proposed antenna design is shown in Fig. 4. The antenna is printed on FR4 substrate ($\varepsilon_r = 4.4$, $h = 1.6$ mm), and fed by a microstrip line of width 2.2 mm and has compact size of $15.5 \times 21\,\text{mm}^2$. The use of a C-shaped slot inserted into the patch yields band-stop characteristic and the slot length (L) of about 16 mm is approximately $\lambda/4$ at the center frequency of the rejected frequency band.

3. EXPERMENTAL RESULTS

3.1. Result A

Figure 5 shows the measured VSWR performance of the proposed monopole antenna with and without inverted U-slot. For the purpose of comparison, the VSWR of the simple monopole antenna is also shown in Fig. 5. As shown in Fig. 5, simple monopole antenna has poor VSWR characteristics at the frequency band over 7.1 GHz. After adding a narrow slit, however, the measured bandwidth of the antenna ranges from 3.1 GHz to more than 11 GHz for VSWR less than 2.0. It is also observed that the sharp frequency band-stop characteristic is obtained very close to the desired frequency of 5.45 GHz when an inverted U-slot with a length of 16 mm is inserted into the antenna. This length of the slot corresponds to 0.29λ at the frequency of about 5.45 GHz.

3.2. Result B

Figure 6 shows the measured return losses for the simple monopole antenna and proposed antenna with a notch in the ground plane only. For the purpose of comparison, the result of the proposed antenna with a notch in the ground plane and an U-slot inserted into the radiating patch is also shown in Fig. 6. The notch structure in the ground plane improves the impedance bandwidth especially over 7.9 to 10.5 GHz. It is also observed that the stop band of 4.92 to 5.86 GHz is created by inserting the U-slot while the antenna has an operating frequency range from 3.1 to more than 11 GHz for the return loss of < -10 dB.

3.3. Result C

Figure 7 shows the measured return loss for the optimized antenna. It is clearly observed that the addition of a small strip bar achieves the band-stop operation and improves input impedance at higher frequency band at the same time. The designed antenna satisfies the 10 dB return loss

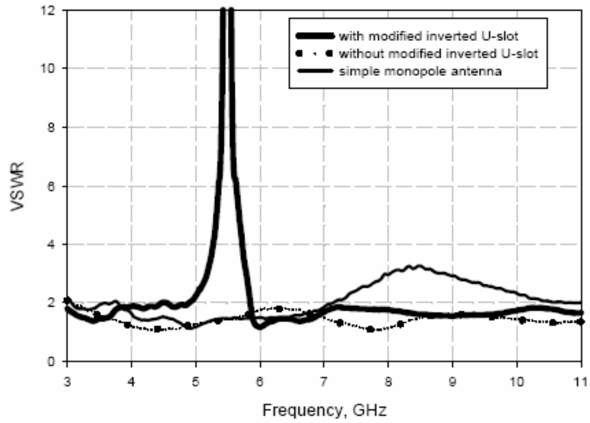


Figure 5: Measured return losses for the proposed antenna with inverted U-slot.

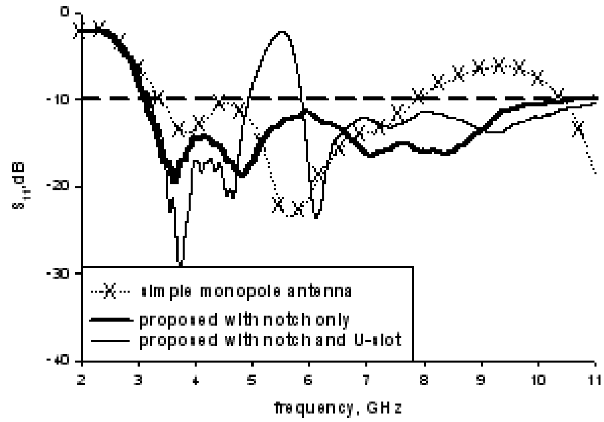


Figure 6: Measured return losses of the proposed internal antenna with U-slot.

requirement in the frequency band between 3.1 GHz and 13 GHz while showing the band rejection performance in the frequency band of 4.9 to 6.0 GHz. Also, as the length of strip bar increases, the center frequency of the lower band decreases in experiment.

3.4. Result DThe

measured and simulated return losses for the proposed antenna are shown in Fig. 8, together with the measured one for the antenna with L-notch only. The measured 10 dB impedance bandwidth for the antenna is from 3.08 to 10.97 GHz while a stop frequency band from 5.03 to 5.91 GHz was created.

Good radiation patterns are obtained for all four type antennas throughout the frequency band

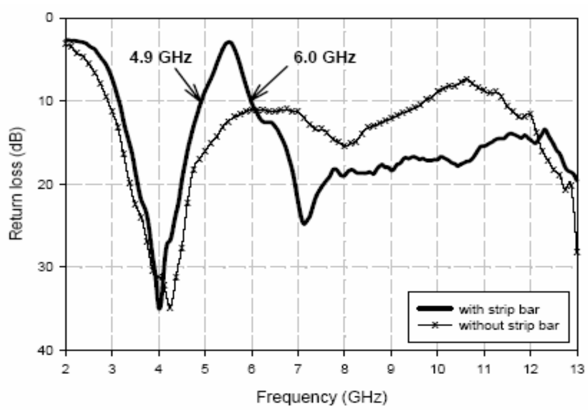


Figure 7: Measured return losses for the proposed antenna with strip bar.

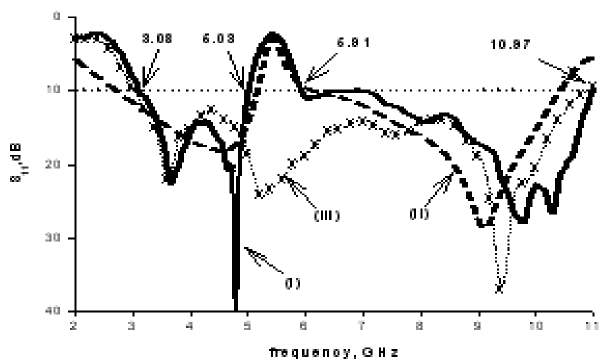


Figure 8: Measured return losses of the proposed internal antenna with C-shaped slot.

Table 1: Measured gains.

Antenna 1	Frequency (MHz)	3	5.5	7	9
	Gain (dBi)	3.2	7.2	4.6	4.3
Antenna 2	Frequency (MHz)	3	5.5	7	9
	Gain (dBi)	2.8	7.75	3	2.45
Antenna 3	Frequency (MHz)	3	5.5	7	9
	Gain (dBi)	4	7	4	5.5
Antenna 4	Frequency (MHz)	3	5.5	7	9
	Gain (dBi)	3	2.8	3.4	3.78

of interest. The measured maximum gains of four types of antenna are listed in Table 1.

4. CONCLUSION

Antenna design techniques for microstrip-fed monopole UWB antenna with frequency band-stop characteristic have been proposed and implemented. Modified planar monopoles with inverted U-slot, U-slot, small strip bar and $\lambda/4$ C-shaped slot to create the frequency stop band of 4.92 to 5.86 GHz are used while maintaining wide bandwidth performance. The proposed antenna could be a good candidate for handheld UWB applications.

ACKNOWLEDGMENT

This research was supported by the MIC (Ministry of Information and Communication), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment)" (IITA-2005-(C1090-0603-0019).

REFERENCES

1. Kim, J., T. Yoon, J. Kim, and J. Choi, "Design of an ultra wideband printed monopole antenna using FDTD and genetic algorithm," *IEEE Microwave and Wireless Components Letters*, Vol. 15, No. 6, 395–397, June 2005.
2. Choi, W., J. Jung, K. Chung, and J. Choi, "Compact microstrip-fed antenna with band-stop characteristic for ultra-wideband applications," *Microwave and Optical Technology Letters*, Vol. 47, No. 1, 89–92, October 2005.
3. Jung, J., W. Choi, and J. Choi, "A small wideband microstrip-fed monopole antenna," *IEEE Microwave and Wireless Components Letters*, Vol. 15, No. 10, 703–705, October 2005.
4. Choi, J., K. Chung, and Y. Roh, "A band rejection antenna with parametric analysis for UWB application," *Microwave and Optical Technology Letters*, Vol. 47, No. 3, 287–290, November 2005.
5. Chung, K., J. Kim, and J. Choi, "Wideband microstrip-fed monopole antenna having frequency band-notch function," *IEEE Microwave and Wireless Components Letters*, Vol. 15, No. 11, 766–768, November 2005.
6. Choi, W., K. Chung, J. Jung, and J. Choi, "Compact ultra-wideband printed antenna with band-rejection characteristic," *IEE Electronics Letters*, Vol. 41, No. 18, 990–991, September 2005.
7. Jung, J., K. Seol, W. Choi, and J. Choi, "Wideband monopole antenna for various mobile commucation applications," *IEE Electronics Letters*, Vol. 41, No. 24, 1313–1314, November 2005.
8. Jung, J., W. Choi, and J. Choi, "A compact broadband antenna with an L-shaped notch," *IEICE Transaction on communication*, Vol. E89-E, No. 6, 1968–1971, June 2006.

Adaptive Wideband Beamforming with Combined Spatial/Temporal Subband Decomposition

Wei Liu and Richard Langley

Communications Research Group
Department of Electronic & Electrical Engineering, University of Sheffield, UK

Abstract— An adaptive wideband beamforming structure with a lower computational complexity and higher convergence speed is proposed, which is achieved by combined spatial/temporal subband decomposition. First, the received array signals are processed by a transformation matrix, and then split into subbands by a series of analysis filter banks. At each subband, an independent beamformer can be operated, where some of the lower subbands can be discarded without affecting its performance due to the highpass filtering effect of the transformation matrix.

1. INTRODUCTION

Beamforming has found many applications in various areas ranging from sonar and radar to wireless communications [1]. It is a signal processing technique to form beams in order to receive signals illuminating a sensor array from specific directions, whilst attenuating signals from other directions. In a statistically optimum beamformer [2], since the statistics of the array data are often not known or may change over time, adaptive algorithms may be used to determine the beamformer's coefficients. A beamformer structure for wideband array signals is shown in Fig. 1, where each of the M received signals is processed by an FIR filter with a length of J and the outputs of these FIR filters are then sumed up to form the final output.

To perform adaptive wideband beamforming with high interference rejection and angular resolution, arrays with a large number of sensors and filter coefficients have to be employed. Reducing the resultant high computational complexity and increasing its convergence speed has motivated many different solutions, including partially adaptive beamforming [3], transform-domain or frequency-domain beamforming [4, 5], and subband beamforming [6, 7]. Recently, a generalised sidelobe canceller (GSC) with combined subband decomposition in both the temporal and spatial domains was proposed in [8], where the columns of the blocking matrix are judiciously designed to generate a highpass filtering effect to the array signals. The bandlimited spectra of the resultant outputs are then exploited by subband decomposition and appropriately discarding the low-pass subbands, where there is no signal existing.

In this paper, we generalize the idea in [8] and propose a subband-selective transformation matrix, which is applied to the received array signals as a preprocessing step. Each of the outputs

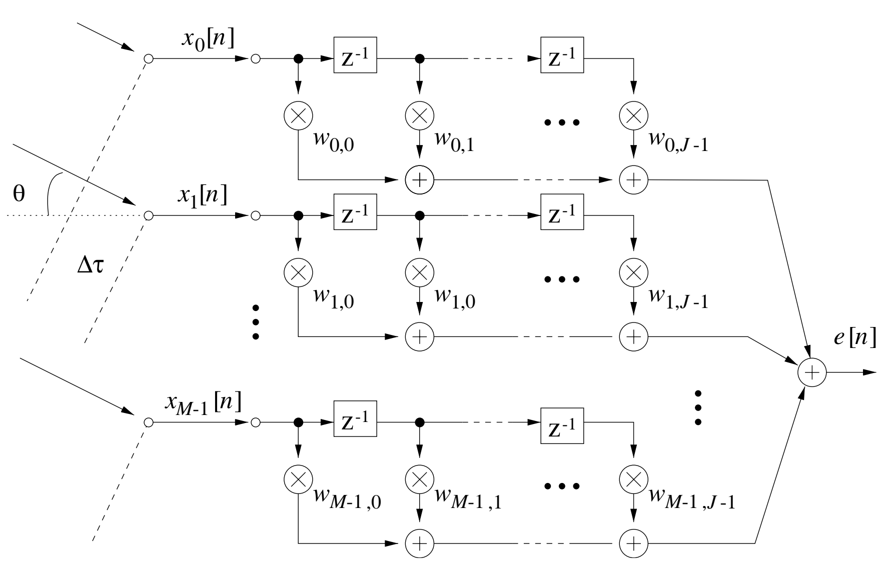


Figure 1: A general structure for broadband beamforming.

of the transformation matrix are then split into K decimated frequency bands ("subbands") by an analysis filter bank and the corresponding subband signals form K sets of subband arrays and an independent beamformer can be operated at each of the subband arrays. Because of the highpass filtering effect of the transformation, some of the lower subbands can be discarded without affecting the beamformer's performance. As a result, the number of inputs to each of the subband beamformers is different from one another. Due to the combined spatial/temporal decorrelation, a faster convergence speed is also achieved in addition to a lower computational complexity.

2. SPATIALLY/TEMPORALLY SUBBAND DECOMPOSITION

As shown in Fig. 1, the received M array signals at time index n are denoted by $x_m[n]$, $m = 0, \ldots, M-1$. If we apply a full-rank $M \times M$ transformation matrix $\mathbf{A}$ to them, then its output $\mathbf{y}[n]$ can be expressed as

$$\mathbf{y}[n] = \mathbf{A}\mathbf{x}[n] \quad \text{with} \tag{1}$$

$$\begin{aligned} \mathbf{x}[n] &= [x_0[n]x_1[n]\ldots x_{M-1}[n]]^T \\ \mathbf{y}[n] &= [y_0[n]y_1[n]\ldots y_{M-1}[n]]^T \\ [\mathbf{A}]_{m,l} &= a_{m,l}. \end{aligned} \tag{2}$$

The m-th output $y_m[n]$ is given by

$$y_m[n] = \sum_{l=0}^{M-1} a_{m,l} \cdot x_l[n]. \tag{3}$$

Suppose a signal with an angular frequency ω and an angle of arrival θ impinges on the uniformly spaced linear array as shown in Fig. 1 and the sampling period is T_s, then the response of the m-th row vector of $\mathbf{A}$ can be expressed as

$$R_m(\omega, \theta) = \sum_{l=0}^{M-1} a_{m,l} e^{-jl\omega\Delta\tau} \tag{4}$$

with $\Delta\tau = \frac{d}{c}\sin\theta$, where d is the array spacing, and c is the wave propagation speed. With the normalised angular frequency $\Omega = \omega T_s$, we obtain the response as a function of Ω and θ

$$R_m(\Omega, \theta) = \sum_{l=0}^{M-1} a_{m,l} e^{-jm\mu\Omega\sin\theta} \quad \text{with} \quad \mu = \frac{d}{cT_s}. \tag{5}$$

As the sampling frequency is in general twice the highest frequency component of the signal and array spacing is half the wavelength of the highest frequency component, we have $d = \frac{1}{2} \cdot c \cdot (2T_s) = cT_s$ and $\mu = 1$. In the following, without loss of generality, we will always assume $\mu = 1$.

With $A_m(\hat{\Omega}) = \sum_{l=0}^{M-1} a_{m,l} e^{-jm\hat{\Omega}}$, i. e., $A_m(\hat{\Omega})$ is the frequency response of the filter with impulse responses $a_{m,0}, a_{m,1}, \ldots, a_{m,M-1}$, we have

$$R_m(\Omega, \theta) = A_m(\Omega\sin\theta). \tag{6}$$

Now, suppose the row vectors of $\mathbf{A}$ are designed to form a series of bandpass frequency responses $A_m(\hat{\Omega})$, each with a bandwidth of $2\pi/M$ and together they cover the whole bandwidth of $[-\pi; \pi]$. Fig. 2 shows such an arrangement with M being an odd number.

As discussed in [8], the bandpass characteristic of the row vectors will have a highpass filtering effect on the received array signals. As an example, let us consider the l-th row vector with a frequency response given in Fig. 2, i. e.,

$$\left|A_l(\hat{\Omega})\right| = \begin{cases} 1 & \text{for } \hat{\Omega} \in [\hat{\Omega}_{l,L}; \hat{\Omega}_{l,U}] > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{7}$$

With such a response, all of the received array signal components with a frequency $\Omega \in (-\hat{\Omega}_{l,L}; \hat{\Omega}_{l,L})$ will not be able to pass this filter, since, no matter which direction it comes from, $\hat{\Omega} = \Omega\sin\theta$ will

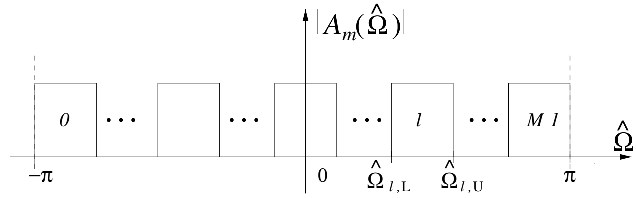


Figure 2: Characteristics of the $M-1$ row vectors contained in $\mathbf{A}$.

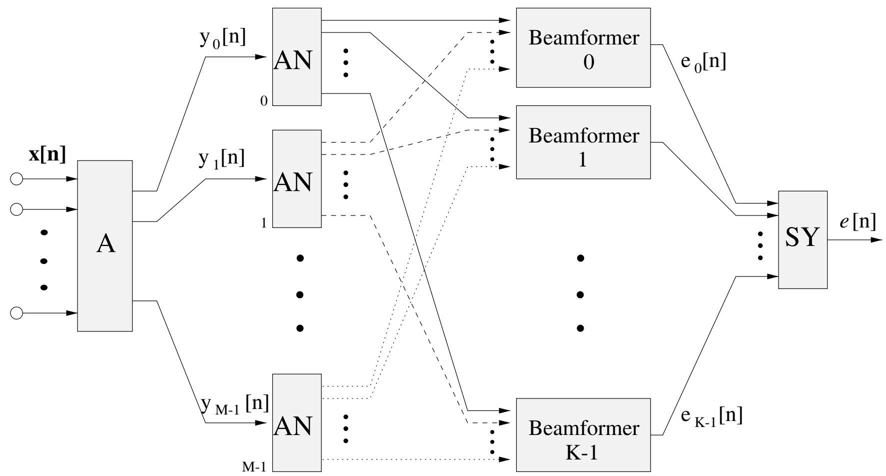


Figure 3: The proposed beamforming structure with combined spatial/temporal subband decomposition.

never fall into the passband of $[\hat{\Omega}_{l,L}; \hat{\Omega}_{l,U}]$. As a result, the frequency range of its output will be $|\Omega| \geq \hat{\Omega}_{l,L}$ and its lower bound is defined by the lower cutoff frequency $\hat{\Omega}_{l,L}$ when $\hat{\Omega}_{l,L} > 0$. On the contrary, when $\hat{\Omega}_{l,L} < \hat{\Omega}_{l,U} < 0$, the lower band will be decided by $|\hat{\Omega}_{l,U}|$.

Now consider the transformed array signals $\mathbf{y}$. We can apply the same subband technique as in [6] and split each of the signal $y_m[n]$ by a K-channel analysis filter bank [9] and an independent adaptive beamformer can then be set up at each set of corresponding subband array signals. The output of these K subband beamformers are then combined by a synthesis filter bank to form the fullband beamforming output. Fig. 3 shows this structure, where the blocks labeled "AN" are the analysis filter banks and the block labeled "SY" is the synthesis filter bank. Note, due to the highpass filtering effect, some of the lower subbands will not carry any signal, and therefore can be discarded during the beamforming process, in a similar way as in [8]. For the example discussed before, any subbands with a frequency range between 0 and $\hat{\Omega}_{l,L} > 0$ for the l-th row vector output can be discarded without affecting the beamforming performance.

3. IMPLEMENTATIONS OF THE PROPOSED BEAMFORMING STRUCTURE

3.1. Transformation Matrix

The key to the proposed method is the design of the transformation matrix with required bandpass characteristics. Since each of the row vector can be considered as an FIR filter with a length M, we can design them one by one using the existing FIR filter design methods, with specified magnitude responses given in Fig. 2. Another method is to design a prototype lowpass filter and then modulate it to the required positions and one of the most commonly used is the discrete Fourier transform (DFT) matrix given by

$$\mathbf{A} = \begin{bmatrix} w^{0\cdot 0} & w^{0\cdot 1} & \cdots & w^{0\cdot(M-1)} \\ w^{1\cdot 0} & w^{1\cdot 1} & \cdots & w^{1\cdot(M-1)} \\ \vdots & \vdots & \ddots & \ddots \\ w^{(M-1)\cdot 0} & w^{(M-1)\cdot 1} & \cdots & w^{(M-1)\cdot(M-1)} \end{bmatrix}, \tag{8}$$

where $\omega = e^{-j\frac{2\pi}{M}}$. However, we can find that the sidelobe attenuation of the DFT matrix is only about 15 dB. To increase it further, as an example, we can use a hamming window function [10] as the prototype filter and then modulate it using a DFT to the corresponding positions. The magnitude responses of a 15×15 transformation matrix obtained by modulating a hamming window with a length 15 are provided in Fig. 4, where the stopband attenuation is almost 40 dB, although with an increased overlap between adjacent passbands.

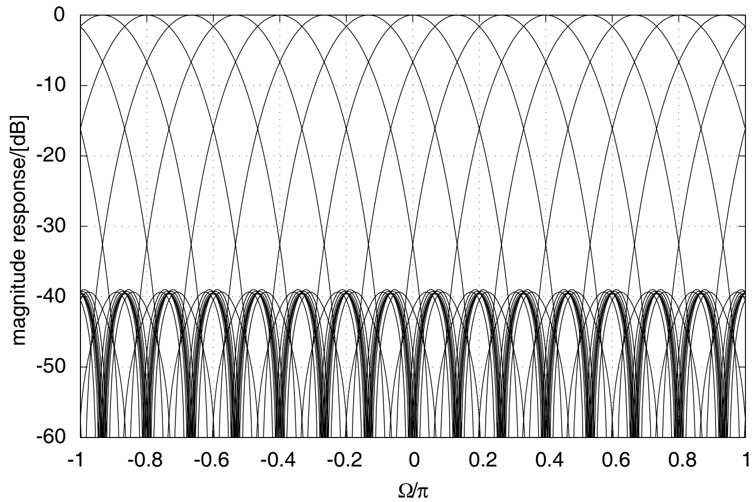


Figure 4: The magnitude responses of a 15×15 transformation matrix obtained by modulating a hamming window.

3.2. Beamformer Implementations

Another issue is the choice of the subband beamformer in the proposed structure given by Fig. 3. If we know the direction of arrival of the signal of interest, we can use a linearly constrained minimum variance beamformer or alternatively a generalised sidelobe canceller [11] and in this case the fullband constraints should be transformed into different subbands so that each subband beamformer can have its own constraints [6], which in general are different from one another. If there is a reference signal available, then fullband beamforming can be performed using a multi-channel adaptive filter (MCAF), as shown in Fig. 5(a), where the M received sensor signals are fed into the MCAF and its coefficients are adjusted by minimizing the difference signal $e[n]$ between the reference signal and the MCAF output. The adaptive filter length for each channel is J. Such a reference signal based beamformer can be implemented in our proposed structure as shown in Fig. 5(b), where the reference signal is split into K decimated subbands by the same analysis filter bank and each of the subband reference signals works as the reference signal for the corresponding subband array signals.

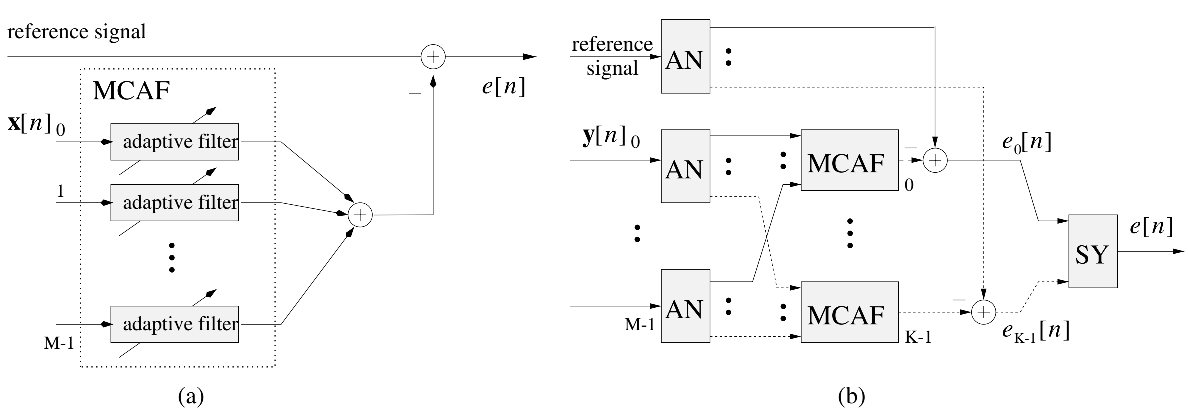


Figure 5: (a) A fullband beamformer with a reference signal, (b) the corresponding subband beamformer structure, where "MCAF" represents "multi-channel adaptive filter".

4. SIMULATIONS

Here we compare the performance of the proposed beamformer as in Fig. 5(b) with that of Fig. 5(a), where $M = 15$ and the fullband adaptive filter length $J = 140$. The responses of the transformation matrix for our proposed method have already been given in Fig. 4 and each of its outputs is divided into $K = 16$ subband channels by oversampled GDFT filter banks [9] with decimation ratio $N = 14$. The prototype lowpass filter employed for the GDFT filter banks has a length of 448 with a bandwidth of $\pi/8$ and a stopband cutoff frequency $\Omega = \pi/14$. The length of the subband adaptive channel for each MCAF is $J/N = 10$. The signal of interest illuminates the array from $\theta = 40°$. and four interfering signals impinge from $\theta = -80°, -60°, -20°, -80°$, respectively, each of them with a signal to interference ratio (SIR) of -30 dB. All the signals have a bandwidth of $\Omega \in [0.3\pi; 0.95\pi]$. Additionally, all sensors receive spectrally and spatially uncorrelated noise at a 20 dB SNR (signal to noise ratio).

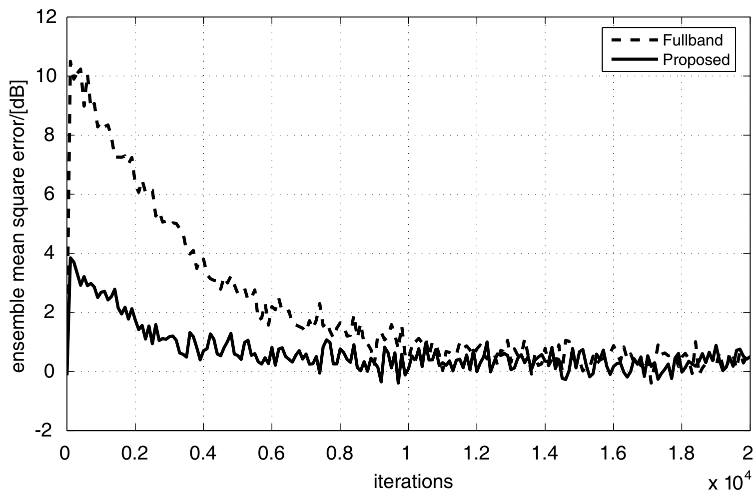


Figure 6: Learning curves of the fullband adaptive beamformer and the proposed subband-selective beamformer.

For our proposed method, in each MCAF block, we have discarded channels with very low signal power according to the responses of the filter banks and the transformation matrix. In this case, there are in total $K = 16$ subband beamformers and 15 channels for each subband beamformer. The number of channels discarded was 66 so that only $15 \times 16 - 66 = 174$ channels were processed for each update.

Figure 6 shows the learning curves of a normalised LMS (least mean square) algorithm with a step size of 0.06. The depicted performance measure is the squared output error, i. e., the difference between the beamformer output and the reference signal, averaged over 150 independent simulations. The convergence rate of the proposed method is much higher than the adaptive fullband beamformer due to the combined decorrelation effect of both the transformation matrix and the filter banks. In terms of computational complexity, the proposed method only needs 21% multiplications of its fullband counterpart.

5. CONCLUSIONS

An adaptive wideband beamforming structure with combined spatial/temporal subband decomposition has been proposed. In this structure, the received array signals are first processed by a transformation matrix, then decomposed into decimated subbands and at each subband, an independent beamformer can be operated. Due to the highpass filtering effect of the transformation matrix, some of the lower subbands can be discarded during beamforming without affecting its performance. An example with a reference signal based beamformer has been provided and simulation results show that the proposed method can achieve a much faster convergence speed and a lower computational complexity.

ACKNOWLEDGMENT

This work was partially supported by the Nuffield Foundation award.

REFERENCES

1. Van Trees, H. L., *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*, John Wiley & Sons, Inc., New York, U.S.A., 2002.
2. Van Veen, B. D. and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoustics, Speech, and Signal Processing Magazine*, Vol. 5, No. 2, 4–24, April 1988.
3. Wang, Y. Y. and W. H. Fang, "Wavelet-based broadband beamformers with dynamic subband selection," *IEICE Trans. Communications*, Vol. E83B, No. 4, 819–826, April 2000.
4. Chen, Y. H. and H. D. Fang, "Frequency-domain implementation of Griffiths-Jim adaptive beamformer," *J. Acoust. Soc. Am.*, Vol. 91, 3354–3366, 1992.
5. An, J. and B. Champagne, "GSC realisations using the two-dimensional transform LMS algorithm," *IEE Proc. Radar, Sonar and Navig.*, Vol. 141, 270–278, October 1994.
6. Weiss, S., R. W. Stewart, M. Schabert, I. K. Proudler, and M. W. Hoffman, "An efficient scheme for broadband adaptive beamforming," *Proc. Asilomar Conference on Signals, Systems, and Computers*, Vol. I, 496–500, Monterey, CA, November 1999.
7. Liu, W., S.Weiss, and L. Hanzo, "Subband adaptive generalized sidelobe canceller for broadband beamforming," *Proc. IEEE Workshop on Statistical Signal Processing*, 591–594, Singapore, August 2001.
8. Liu, W., S. Weiss, and L. Hanzo, "A subband-selective broadband GSC with cosine-modulated blocking matrix," *IEEE Transactions on Antennas and Propagation*, Vol. 52, 813–820, March 2004.
9. Weiss, S. and R. W. Stewart, *On Adaptive Filtering in Oversampled Subbands*, Shaker Verlag, Aachen, Germany, 1998.
10. Oppenheim, A. V. and R. W. Schafer, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, 1975.
11. Griffiths, L. J. and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Transactions on Antennas and Propagation*, Vol. 30, No. 1, 27–34, January 1982.

Magnetic Plasmon Propagation along a Chain of Connected Subwavelength Resonators at Infrared Frequencies

H. Liu[1], **D. A. Genov**[1], **D. M. Wu**[1], **Y. M. Liu**[1]
J. M. Steele[1], **C. Sun**[1], **S. N. Zhu**[2], **and X. Zhang**[1]

[1]5130 Etcheverry Hall, Nanoscale Science and Engineering Center
University of California, Berkeley, California 94720-1740, USA
[2]Department of Physics, Nanjing University, Nanjing 210093, China

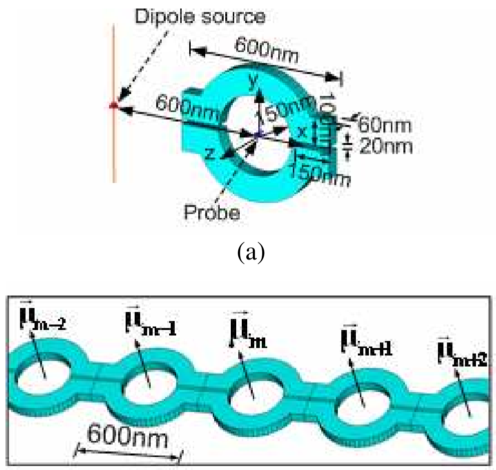
Abstract— A one-dimensional magnetic plasmon propagating in a linear chain of single split ring resonators is proposed. The subwavelength size resonators interact mainly through exchange of conduction current, resulting in stronger coupling as compared to the corresponding magnetoinductive interaction. Finite-difference time-domain simulations in conjunction with a developed analytical theory show that efficient energy transfer with wave attenuation of less then 0.57 dB/μm and group velocity higher than 1/4c can be achieved. The proposed novel mechanism of energy transport in the nanoscale has potential applications in subwavelength transmission lines for a wide range of integrated optical devices.

A fundamental problem of integrated optics is how to transport electromagnetic (EM) energy in structures with transverse dimensions that are considerably smaller than the corresponding wavelength of illumination. The main reason to study light-guiding in the nanoscale has to do with the size of transmission lines being a limiting factor for substantial miniaturization of integrated optical devices. Planar waveguides and photonic crystals are currently key technologies enabling a revolution in integrated optical components. [1, 2] However, the overall size and density of the optical devises based on these technologies is limited by the diffraction of light, which sets the spatial extend of the lowest guided electromagnetic mode at about half wavelength.

In 1999, [3] Pendry reported that nonmagnetic metallic element, double split ring resonator (DSRR), with size below the wavelength of incident radiation, exhibits strong magnetic response and behaves like an effective negative permeability material. In such systems, there are no free magnetic poles. However, the excitation of displacement currents in the DSRR results in induction of a magnetic dipole moment that is somehow similar to a bar magnet. In analogy to the electrostatic SP resonances in metal nanoparticles, an effective media made of DSRRs could support resonant magnetic plasmon (MP) oscillations at GHz [3–5] and THz frequencies [6–9]. Combined with an electric response, characterized by negative permittivity, such systems could lead to development of meta-materials with effective negative indexes of refraction [4–5].

In this letter we propose a sub-wavelength size metal structure, referred to as a single split ring resonator (SSRR) which, (a) demonstrate magnetic resonance in the THz range, and (b) could be used to support propagation of long range MP polaritons. It is well known that radiation loss of a magnetic dipole is substantially lower as compared to the radiation of an electric dipole of similar size [10]. Thus, application of MP for guiding EM energy at long distances has great potential for direct application in novel sub-diffraction sized transmission lines. Indeed, magnetic plasmons have been already shown to play an important role in the excitation of magneto-inductive [11], and electro-inductive waves [12] in the microwave range. In this work we show that at high frequencies a coupling mechanism based on exchange of conduction current between specially designed resonators, may be utilized to efficiently transfer energy along a sub-wavelength sized metal nanostructures. The interaction due to the conduction current is found to be much stronger than the corresponding magneto-inductive coupling, which leads to significant improvement in the properties of the guided MP wave.

Figure 1(a) presents a novel design of a SSRR characterized by two half-space metal loops with tails adjacent to their ends. Pendry's double split ring structure has no tail, nevertheless the space between the rings acts as a capacitor allowing the flow of a displacement current. [3] In the present design it is the gap in the tails that plays the role of a capacitor. Excitation of magnetic response in a system of SSRRs fabricated on a planar substrate, results in induction of magnetic dipoles moments that are perpendicular to the substrate plane (see Fig. 1). Parallel dipoles are characterized with small spatial field overlap and consequently the magneto-inductive interactions between them are expected to be rather weak. To substantially increase coupling between the dipoles, we physically connect the SSRRs, as shown in Fig. 1(b). The contact between the rings serves as a "bond" for conduction current that flow from one SSRR to its neighbor. Thus, in addition to the magneto-inductive coupling, our system interacts directly by exchange of conduction current. This type of coupling is somewhat similar to the electron exchange interaction between two magnetic atoms in a ferromagnetic material. [13] As shown below, the direct physical link between the resonators leads to stronger interaction between the SSRRs and an efficient EM energy transport taking place along a chain of SSRRs.



(a)

(b)

Figure 1: (a) Structure of single split ring resonator; (b) structure of infinite SSRR chain.

To study the electromagnetic response of the proposed SSRR we perform a set of finite-difference time-domain (FDTD) calculations using a commercial software package CST Microwave Studio (Computer Simulation Technology GmbH, Darmstadt, Germany). In the calculations we rely on the Drude model to characterize the bulk metal properties. Namely, the metal permittivity in the infrared spectral range is given by $\varepsilon(\omega) = 1-\omega_s^2/(\omega^2+i\omega\omega_\tau)$, where ω_s is the bulk plasma frequency and ω_τ is the relaxation rate. For gold, the characteristic frequencies fitted to experimental data, are $\hbar\omega_s = 9.02\,\text{eV}$ and $\hbar\omega_\tau = 0.027\,\text{eV}$ [14].

To better understand magnetic plasmon in the infinite chain of connected SSRRs, we develop a comprehensive semi-analytic theory based on the attenuated Lagrangian formalism. If q_m ($m = 0, \pm1, \pm2, \pm3, \ldots$) is the total oscillation charge in the m-th SSRR, L is the induction of the ring and C is the capacitance of the gap, then we can write the Largangian of the coupled system as

$$\Im = \sum_m \left(\frac{1}{2} L \dot{q}_m^2 - \frac{1}{2C}(q_m - q_{m+1})^2 + M\dot{q}_m\dot{q}_{m+1} \right) \tag{1}$$

where the first two terms correspond to the kinetic energy stored in the inductors and the static energy stored in the capacitors, the third interaction term $M\dot{q}_m\dot{q}_{m+1}$ is due to magneto-inductive coupling. The dissipation function of the system can be written as

$$\Re = \sum_m \frac{1}{2}\gamma\dot{q}_m^2 \tag{2}$$

From the damping Euler-Lagrangian equation

$$\frac{d}{dt}\left(\frac{\partial\Im}{\partial\dot{q}_m}\right) - \frac{\partial\Im}{\partial q_m} = -\frac{\partial\Re}{\partial\dot{q}_m} \tag{3}$$

we can obtain the decay oscillation equation of the m-th SSRR as

$$\ddot{Q}_m = -\omega_0^2 Q_m - \Gamma\dot{Q}_m + \frac{\omega_0^2}{4}(Q_{m-1} + 2Q_m + Q_{m+1}) - \frac{M}{L}\left(\ddot{Q}_{m-1} + \ddot{Q}_{m+1}\right) \tag{4}$$

If we define the m-th SSRR as a single magnetic dipole $\mu_m = Q_m/S$ (S is the area of SSRR), then

$$\ddot{\mu}_m + \omega_0^2\mu_m + \Gamma\dot{\mu}_m = \frac{\kappa_1}{2}\omega_0^2(\mu_{m-1} + 2\mu_m + \mu_{m+1}) - \kappa_2\left(\ddot{\mu}_{m-1} + \ddot{\mu}_{m+1}\right) \tag{5}$$

Here, the interaction term $\kappa_1\omega_0^2(\mu_{m-1} + 2\mu_m + \mu_{m+1})/2$ comes from exchange current between two SSRRs and $-\kappa_2\left(\ddot{\mu}_{m-1} + \ddot{\mu}_{m+1}\right)$ comes from magnetic induction. The coupling coefficients $\kappa_1 = 0.5$ and $\kappa_2 = M/L$. The general solution of Eq. (5) corresponds to an attenuated MP wave: $\mu_m = \mu_0 \exp(-m\alpha d)\exp(i\omega t - imkd)$, where ω and k are the angular frequency and wave vector respectively, α is the attenuation per unit length and d is the SSRR's size. By substituting $\mu_m(t)$ into Eq. (5), and working in a small damping approximation ($\alpha d \ll 1$), simplified relationships for the MP dispersion and attenuation are obtained

$$\omega^2(k) = \omega_0^2\frac{1 - \kappa_1[1 + \cos(kd)]}{1 + 2\kappa_2\cos(kd)} \quad \text{(6a)} \qquad \alpha(\omega) = \frac{\omega\Gamma}{\kappa_1\omega_0^2 + 2\kappa_2\omega^2}\frac{1}{\sin(kd)\cdot d} \quad \text{(6b)}$$

The range of applicability and overall accuracy of the predicted relationships are compared in Fig. 2 to FDTD results for a finite chain of SSRRs. The chain size is restricted to 50 resonators lengths, which assure reliable estimates of the system properties without imposing overwhelming computational constrains. The MP polariton is excited by a dipole source placed at a distance of 600 nm from the center of the leading SSRR element, and the H field along the chain is analyzed to determine the wave vector k of the propagating mode. Numerically and analytically estimated MP dispersion and attenuation curves are depicted in Figs. 2(b) and (c), respectively.

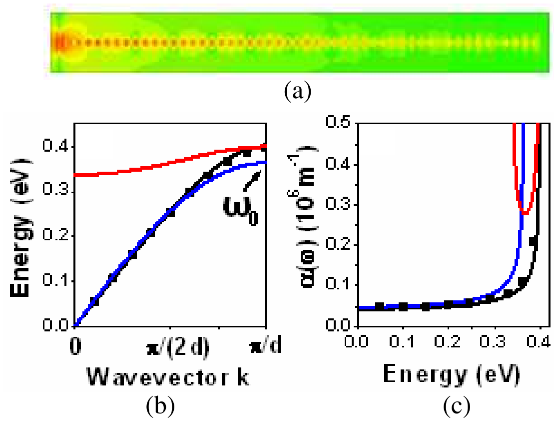


Figure 2: (a) FDTD simulation of a MP propagation along a connected chain of 50-SSRRs at $\eta\omega = 0.3$ eV; (b) dispersion $\omega(k)$, and (c) attenuation coefficient $\alpha(\omega)$. The analytical result (Eq. (6)) including conduction current and magneto-inductive interactions, black solid curve, matches well with the FDTD numerical data (square dots) The predicted MP characteristics based singularly on exchange current interactions ($\kappa_2 = 0$), or magneto-inductive interactions ($\kappa_1 = 0$) are presented with blue and red curves, respectively.

The magnetic plasmon is exclusively a transversal wave. It is described by a single dispersion curve (black solid line in Fig. 2(b)) which cover a broad frequency range $\omega \in (0, \omega_c)$, with a cutoff frequency $\hbar\omega_c = \hbar\omega_0/\sqrt{1 - 2\kappa_2} \approx 0.4$ eV. The precise contribution of each coupling mechanisms in the MP dispersion can be studies readily using Eq. (6a). Exclusion of the magneto-inductive term, results in slight decrease in the cutoff frequency $\omega_c \to \omega_0$ (blue solid curve in Fig. 2(b)). On the other hand, if the SSRRs interact only through the magneto-inductive force, a dramatic change in the MP dispersion is observed (red solid curve in Fig. 2(b)). Namely, the propagating band shrinks to a very narrow range of frequencies $\Delta\omega \cong 2\omega_0\kappa_2$ centered around ω_0.

Strong wave dissipation has been one of the major obstacles for utilization of surface plasmons in optical transmission lines. The sub-diffraction sized MP transmission line, proposed in this work, promises a considerable improvement in the electromagnetic transmission. Fig. 2(c) presents both, analytical and numerical results for the MP wave attenuation along the SSRR's chain. For most of the propagation band, $\alpha(\omega)$ stays constant and have relatively low value. For instance, at an incident frequency $\hbar\omega = 0.3$ eV (wavelength 4.1 μm), the MP attenuation coefficient is $\alpha = 0.65 \times 10^5 m^{-1}$, which corresponds to a field decay length of approximately 15.4 μm (26 unit cells).

The attenuation of a signal that travels along the SSRR chain is thus in the order of $0.57\,\mathrm{dB}/\mu\mathrm{m}$. The reason behind the efficient energy transfer is easy understood by looking at the expected attenuation when one of the coupling mechanisms is artificially impeded. Clearly, MP polaritons that are excited entirely by inductive coupling, similarly to the EP polaritons, exhibit strong attenuation (red curve in Fig. 2(c)), while introduction of direct physical link between the resonators dramatically improves transmission (blue curve in Fig. 2(c)). This effect is also manifested through the MP's group velocity $V_g = \partial\omega/\partial k \approx \frac{1}{2}\sqrt{\kappa_1}\omega_0 d\sin(kd)/\sqrt{1-\cos(kd)}$, which reaches values up to $0.25c$ close to the center of the propagation band.

In conclusion, we have proposed and studied a one-dimensional magnetic plasmon propagating in a linear chain of novel single split ring resonators. We show that at high frequencies a coupling mechanism based on exchange of conduction current, could be used to improve energy transmission along the chain. The current exchange interaction is found to be much stronger than the corresponding inductive coupling. A comprehensive analytical model is developed for calculation of MP dispersion, attenuation coefficient and group velocity. The theory is consistent with the performed FDTD simulations, representing a direct evidence of effective energy transfer below the diffraction limit. Excitation of magnetic plasmon waves could be a promising candidate for the development of a wide range of fast nanoscale optical devices. This includes in-plane, CMOS compatible, subwavelengh optical waveguides, fast optoelectronic switches and transducers. The localized MP modes could also play a substantial role in surface enhanced Raman spectroscopy for molecule detection and bio-sensing.

ACKNOWLEDGMENT

This work was supported by AFOSR MURI plasmonics program (Grant No. FA9550-04-1-0434) and the Center for Scalable and Integrated Nanomanufacturing (SINAM), an NSF Nanoscale Science and Engineering Center (NSEC) under award number DMI-0327077.

REFERENCES

1. Saleh, B. E. A. and M. C. Teich, *Fundamentals of Photonics*, Wiley, New work, 1991.
2. Mekis, A., J. C. Chen, I. Kurland, S. Fan, P. R. Villeneuve, and J. D. Joannopoulos, "High transmission through sharp bends in photonic crystal waveguides," *Phys. Rev. Lett.*, Vol. 77, 3787–3790, 1996.
3. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, 2075–2084, 1999.
4. Smith, D. R., W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, "Composite medium with simultaneously negative permeability and permittivity," *Phys. Rev. Lett.*, Vol. 84, 4184–4187, 2000.
5. Shelby, R. A., D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science*, Vol. 292, 77–79, 2001.
6. Yen, T. J., W. J. Padila, N. Fang, D. C. Vier, D. R. Smith, J. B. Pendry, D. N. Basov, and X. Zhang, "Terahertz magnetic response from artificial materials," *Science*, Vol. 303, 1494–1496, 2004.
7. Linden, S., et al., "Magnetic response of metamaterials at 100 terahertz," *Science*, Vol. 306, 1351–1353, 2004.
8. Enkrich, C., M. Wegener, S. Linden, S. Burger, L. Zschiedrich, F. Schmidt, J. F. Zhou, T. Koschny, and C. M. Soukoulis, "Magnetic metamaterials at telecommunication and visible frequencies," *Phys. Rev. Lett.*, Vol. 95, 203901, 2005.
9. Shalaev, V. M., W. Cai, U. K. Chettiar, H. Yuan, A. K. Sarychev, V. P. Drachev, and A. V. Kildishev, "Negative index of refraction in optical metamaterials," *Opt. Lett.*, Vol. 30, 3356–3358, 2005.
10. Jackson, J. D., *Classical Electrodynamic*, John Wiley & Sons, Inc., New York, 1999.
11. Sydoruk, O., O. Zhuromskyy, E. Shamonina, and L. Solymara, "Phonon-like dispersion curves of magnetoinductive waves," *Appl. Phys. Lett.*, Vol. 87, 072501, 2005.
12. Beruete, M., F. Falcone, M. J. Freire, R. Marqués, and J. D. Baena, "Electroinductive waves in chains of complementary metamaterial elements," *Appl. Phys. Lett.*, Vol. 88, 083503, 2006.
13. Majlis, N., *The Quantum Theory of Magnetism*, World Scientific, Singapore, 2000.
14. Ordal, M. A., et al., "Optical properties of the metals Al, Co, Cu, Au, Fe, Pb, Ni, Pd, Pt, Ag, Ti, and W in the infrared and far infrared," *Appl. Opt.*, Vol. 22, 1099–1120, 1983.

Plasmonics — The Missing Link between Nanoelectronics and Microphotonics

Mark L. Brongersma, Rashid Zia, and Jon Schuler
Geballe Laboratory for Advanced Materials, Stanford University, Stanford, CA 94305, USA

Abstract— Plasmonics is an exciting new device technology that has recently emerged. It exploits the unique optical properties of metallic nanostructures to enable routing and manipulation of light at the nanoscale. A tremendous synergy can be attained by integrating plasmonic, electronic, and conventional dielectric photonic devices on the same chip and taking advantage of the strengths of each technology. We will provide a perspective on future directions and possibilities for integrating plasmonic devices on a Si chip.

1. INTRODUCTION

The development of integrated electronic and photonic circuits has led to remarkable data processing and transport capabilities that permeate almost every facet of our daily lives. Scaling these devices to smaller and smaller dimensions has enabled faster, more efficient, and less expensive components but has also brought about a myriad of challenges. Currently, two of the most daunting problems preventing significant increases in processor speed above 10 GHz are thermal and RC delay time issues with electronic interconnection [1, 2]. Optical interconnects do not exhibit such problems and posses an almost unimaginably large data carrying capacity. Unfortunately, the reduction in size of dielectric waveguides is fundamentally limited by the diffraction limit of light, imposing a lower size limit on a guided light mode of about $\lambda/2n$ (about 0.5 μm). Indeed, photonic structures tend to still be at least 1 or 2 orders of magnitude larger than their electronic counterparts. This obvious size mismatch between electronic and photonic components has presented major problems in interfacing these technologies, creating a bottleneck that prevents higher data processing speeds. It thus appears that further progress will require the development of a radically new device technology that can facilitate information transport between nanoscale devices at optical frequencies and form a bridge between world of electronic and photonics.

Metals may posses exactly the right combination of properties to tackle the issues outlined above and realize the dream of even faster chips. The metals commonly used in electrical interconnection such as copper (Cu) and aluminum (Al) allow for the excitation of surface plasmon-polaritons (SPPs) [3]. SPPs are surface electromagnetic ("light") waves that propagate at a metal-dielectric interface and exhibit a strong coupling to the free electrons in the metal, as pictorially shown in Fig. 1.

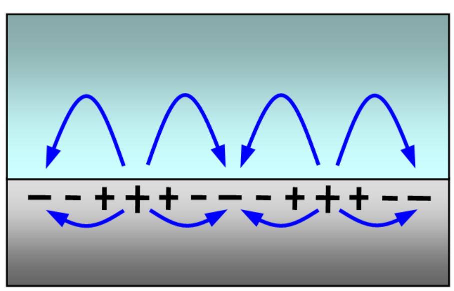


Figure 1: Surface plasmon-polariton (SPP) propagating along a metal-dielectric interface. These waves are transverse magnetic (TM) in nature. Their electromagnetic field intensity is highest at the surface and decays exponentially away from the interface.

Metallic stripes that support such waves can thus serve as tiny optical waveguides termed *plasmonic waveguides*. In contrast to dielectric waveguides, plasmonic waveguides can simultaneously carry optical and electrical signals, giving rise to new capabilities. A systematic study will be presented on the transport properties of such metallic stripes\interconnects and their potential for routing information around on future chips in the form of SPPs ("light") will be discussed. If this could be made a reality, a tremendous synergy could be attained by integrating plasmonic, electronic, and conventional photonic devices on the same chip and taking advantage of the strengths of each technology.

The latest advances in electromagnetic simulations and nanofabrication techniques allow for the design and use of plasmonic structures in real devices. Current silicon-based integrated circuit technology is already used to making nanoscale metallic structures such as copper and aluminum interconnects to route electronic signals between the transistors on a chip. This mature processing technology can be used to our advantage to enable low-cost fabrication of plasmonic devices and integration with electronic and dielectric photonic components on a silicon chip.

In order to study the study the propagation of SPPs along metallic stripe waveguides and plasmonic nanocircuits, we have constructed a Photon Scanning Tunneling Microscope (PSTM). Fig. 2(a) shows schematically how a microscope objective can be used to focus a light beam onto a metallic launchpad and thereby launch a SPP into the metallic stripe waveguide. This method of exciting SPPs makes use of the well-known Kretchmann geometry that enables phase matching of the free space excitation beam to the SPP on the top metal surface [3]. A sharp metal-coated tip with a nanoscale aperture (about 50 nm diameter) is used to locally *tap into* the guided SPP wave and scatter light into the far-field. The scattered light is then detected with a photomultiplier tube. The detected signal provides a measure of the local light intensity in the waveguide and by scanning the tip over the surface of the waveguide with piezo-motors maps of the light intensity anywhere in the waveguide can be produced.

Using electron beam lithography, gold (Au) stripes with widths ranging from 5 μm to 50 nm were generated on SiO_2 glass slides. Au is easy to deposit, does not oxidize, and exhibits a qualitatively similar plasmonic response as Cu and Al. Fig. 2(b) shows some of these stripes attached to a Au launchpad from which SPP are launched. Fig. 2(c) shows a PSTM image of a surface plasmon-polariton propagating along a 4.5 μm wide stripe, and a propagation distance of several tens of micrometers is observed.

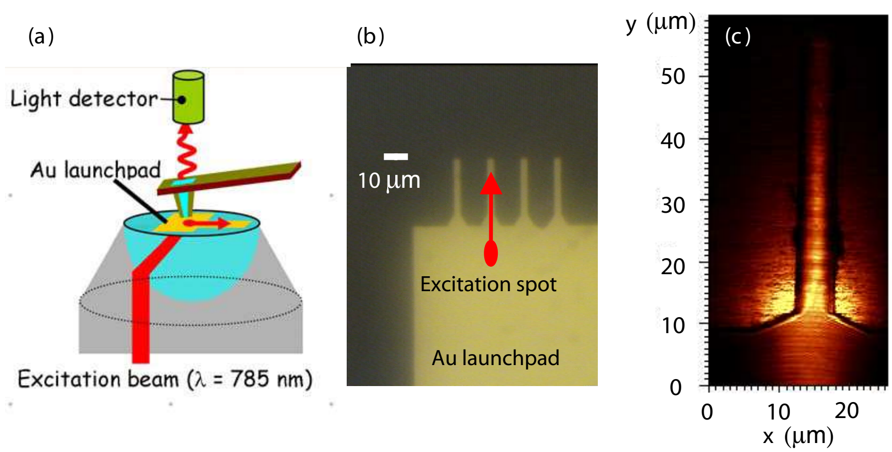


Figure 2: (a) Schematic of the operation of a photon scanning tunneling microscope (PSTM) that enables the study of surface plasmon-polariton ("light") propagation in nanoscale metallic stripe waveguides and plasmonic nanocircuits. A red arrow indicates how a SPP is launched from an excitation spot into a gold (Au) stripe waveguide using a high numerical aperture microscope objective. (b) Optical microscopy image of a SiO_2 substrate with an array of Au stripes attached to a launchpad generated by electron beam lithography. The red arrow again indicates the launching of surface plasmon-polaritons. (c) A PSTM image of a surface plasmon-polariton propagating along a 4.5 μm wide and 48 nm thick Au stripe at $\lambda = 785$ nm.

Measurements on stripes of different widths convincingly show that the propagation distance of SPPs decreases with decreasing stripe width. This is also predicted by full-vectorial electromagnetic simulation of the attenuation coefficient, α_{spp}, versus waveguide width for the lowest order quasi-TM SPP modes of a Au stripe ($t = 55$ nm, $\lambda_{\text{excitation}} = 800$ nm), as shown in Fig. 3 [4, 5]. It is seen that the attenuation significantly increases as soon as he waveguide width drops below a few times the free space excitation wavelength.

These results suggest that the possibilities for using single metallic stripes for intra-chip communication may be limited to short distances. Before continuing with our discussion on using

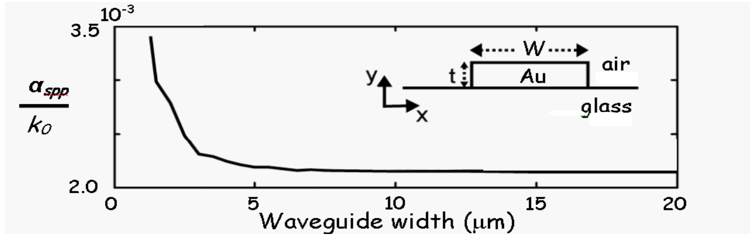


Figure 3: The normalized attenuation constant, α_{spp}, for the lowest order leaky, quasi-TM surface plasmon modes on a Au stripe waveguide versus waveguide width, W. The stripe thickness is, t, is 55 nm and the excitation wavelength is $\lambda = 800$ nm. The inset shows the simulated geometry.

stripe waveguides for interconnections, it is worth pointing out that active plasmonic devices or SPP-based sensor technologies may be realized that are significantly smaller than the SPP decay length.

Other waveguide geometries may offer interesting capabilities for short distance transport as well. Metal-insulator-metal (MIM) waveguides support deep sub-wavelength SPP modes that can transport information over short distances [6]. In combination with optical antenna structures, such MIM waveguides may serve as a bridge between micronscale dielectric photonic devices and nanoscale electronic devices, as shown schematically in Fig. 4.

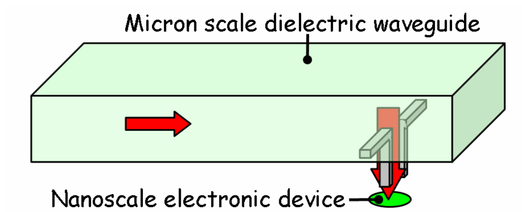


Figure 4: Schematic illustration how a nanoscale antenna structure can serve as a bridge between micronscale dielectric components and nanoscale electronic devices.

Recent studies on optical antennas have demonstrated their effectiveness in concentrating electromagnetic waves into nanoscale volumes. By using metallic nanostructures as a bridge between photonics and electronics, one plays to the strengths of the metallic nanostructures (concentrating fields and subwavelength guiding), the dielectric components (long distance information transport) and the nanoscale electronic ones (high speed information processing). It is envisioned that solutions of this type may be able to unleash the full potential of nanoscale functionality for chip-scale information transport and processing.

REFERENCES

1. Bohr, M., "Interconnect scaling — The real limiter to high performance ULSI," *Tech. Dig. IEDM*, 241–244, 1995.
2. Banerjee, K., A. Amerasekera, G. Dixit, and C. Hu, "The effect of interconnect scaling and low-k dielectric on the thermals characteristics of the IC metal," *Tech. Dig. IEDM*, 65–68, 1996.
3. Raether, H., *Surface Plasmons*, Vol. 111, Springer-Verlag, New York, 1988.
4. Zia, R., M. D. Selker, and M. L. Brongersma, "Leaky and bound modes of plasmonic waveguides," *Phys. Rev. B*, Vol. 71, 165431-1, 2004.
5. Zia, R. and M. L. Brongersma, "Near-field characterization of guided polariton propagation and cutoff in surface plasmon waveguides," Accept in *Phys. Rev. B*, 2006.
6. Zia, R., M. D. Selker, P. B. Catrysse, and M. L. Brongersma, "Geometries and materials for sub-wavelength surface plasmon modes," *J. Opt. Soc. Am*, Vol. A21, 2442–2446, 2004.

Linear, Nonlinear and Ultrafast Behavior of Surface Plasmon Polaritons in Nanostructures

M. Sandtke[1], **R. J. P. Engelen**[1], **J. Prangsma**[1], **J. A. H. van Nieuwstadt**[1]
R. H. Harmsen[1], **S. Enoch**[2], **and L. Kuipers**[1]

[1]Center for Nanophotonics, FOM Institute AMOLF, Amsterdam, The Netherlands
[2]Institut Fresnel, Marseille, France

Abstract— Surface plasmon polaritons (SPPs) are able to concentrate electromagnetic radiation in subwavelength volumes and are potentially useful for miniaturized all-optical chips. These chips will require SPP guiding and nonlinear interactions. Here, we will present phase-sensitive, timeresolved visualization of propagating SPP wavepackets in SPP waveguides. We will also show that the nonlinear optical response of subwavelength holes arrays can be greatly affected by hole shape.

The use of SPPs [1, 2] in combination with (nano) structured metal films, allows for the manipulation of optical fields on a scale (much) smaller than the wavelength of light. SPPs are a good candidate for signal processing in an all-optical chip the basis for SPPs on a chip, the SPP guide, has been studied extensively, both experimentally and theoretically. The large field concentrations associated with SPPs in nanostructures play a key role in surface enhanced Raman scattering (SERS) [3], where the use of noble metal clusters van even enhance the signal of single molecules [4]. Combining SPP guiding structures with structures in which their propagation is influenced by nonlinear interactions are two crucial ingredients for future all-optical plasmonic chips. In order to make progress towards this goal investigations of SPP propagation and nonlinear optics in model metallic structures are of key importance.

Conventional far-field studies are able to answer some of the questions rising about SPP behavior, but for a detailed understanding on a subwavelength scale, near-field microscopy a necessary instrument. Using this technique mode profiles of guided SPPs in straight SPP guides have been studied experimentally [5]. Also the influence of the width of the metal guide on the propagation length is studied [6].

In this contribution we present two ingredients necessary for all-optical SPP circuits: local, timeresolved investigation of SPP wavepackets and a strong modification of the nonlinear response of a plasmonic nanostructure. SPP wavepackets are tracked as they propagate along a metallic waveguide. We determine their phase- and group velocity. By investigating second harmonic generation (SHG) in sub-wavelength hole arrays we gain insight in their nonlinear response. We show that the arrays exhibits a “hot” holes shape for which effective nonlinear susceptibility of the structure is enhanced by an order of magnitude. This enhancement yields an increase in SHG on top of the increase caused in a more trivial way by a change in the linear properties of the array. The increase in effective nonlinear susceptibility is attributed to a low group velocity of the fundamental as it is transmitted through the holes.

In the past few years we have successfully used time- and phase-resolved photon scanning tunneling microscopy (PSTM) for investigations of optical phenomena inside photonic crystal structures [7–12]. This technique uses ultrafast pulses and combines conventional near-field microscopy in collection mode with a Mach-Zehnder type interferometer. The light is split in two branches. In the first, called the signal branch, the light is incident on the sample and is subsequently collected by the near-field probe. The size of the aperture of this probe is the determining factor for the spatial resolution of the microscope. The second branch, called the reference branch contains an optical delay line and an acousto-optic modulator (AOM), which shifts the frequency of the light in the reference branch by 40 kHz. The AOM enables heterodyne detection. The branches are brought together and the interference is detected. For a fixed length of the reference branch, this allows the phase evolution of the light propagating in nanostructures to be determined by scanning the near-field probe across the sample. When ultrafast pulses, with a small coherence length, are used, interference is only detected if there is time overlap between the pulses from the signal and reference branch. As a result, scanning the probe for a fixed length of the reference branch yields a “snapshot” of the pulse inside the structure. By changing the length of the reference branch with the delay line and therefore the time needed for the reference pulse to reach the mixing point, the "snapshot" is obtained for a different time.

Figure 1(a) shows a topography image of a typical SPP waveguide obtained by means of the shear-force height feedback of the near-field microscope. The Au waveguides are fabricated by means of e-beam lithography and followed by a lift off step. The dimensions of this particular guide are 6 μm wide, 80 μm long and 55 nm thick. A femtosecond Optical Parametric Oscillator, at a wavelength (in air) of 1500 nm, is used to excite femtosecond SPP wavepackets at the entrance of the guide, using a Kretschmann-Raether configuration [13]. Figure 1(b) depicts "snapshots" at

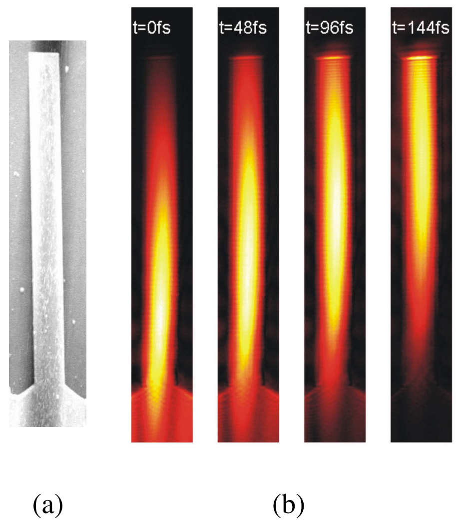


Figure 1: (a) Greyscale image of the topography of the SPP waveguide obtained from the shear-force height feedback of the near-field microscope. The Au waveguide is 6 μm wide, 80 μm long and 55 nm thick. In the bottom of the image the top of the so-called launch pad is visible, where the SPPs are excited via a Kretschmann-Raether configuration. (b) False color image of four snapshots of the amplitude of the SPP waveguide is it propagates inside the waveguide. The time between each frame is 48 fs.

different times of the SPP wavepacket as it propagates along the waveguide. The propagation of the wavepacket is clearly resolved. From the position of the center-of-mass of the wavepacket as a function of time, we can unambiguously obtain the SPP group velocity. Figure 2 shows the position of the SPP center-of-mass versus time. We find a group velocity of 0.8c for the group velocity of the SPP wavepacket. When compared to the expected 0.98c for a plain Au-air interface, it is clear that the guide influences the dispersion of the SPP.

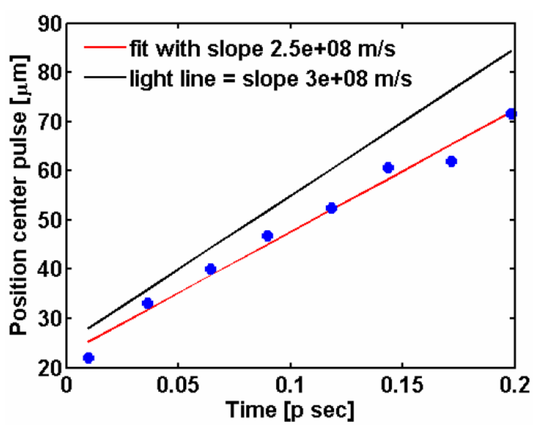


Figure 2: Graph of the position of the center-of-mass of the SPP wavepacket versus time. A linear fit to the data points yields the SPP group velocity without any need for modeling.

Figure 3 depicts the complete information provided by our measurement of the SPP wavepacket. The false color image shows the product of the amplitude times the cosine of the phase (difference between the branches). All the individual phase fringes are resolved. The spacing of the phase fringes yields the phase velocity of the wavepacket as its optical frequency is given by the wavelength of the incident laser and is therefore known. Within the experimental accuracy the phase velocity is the same as that expected for a plain Au-air interface: 0.99c. Clearly, the measurement of the group velocity is a more accurate means to determine the influence of the geometrical confinement on the dispersion of the SPPs.

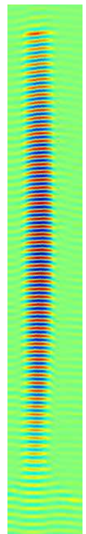


Figure 3: False color representation of the raw measurement data as obtained with the phase-sensitive nearfield microscope. The data is proportional to the amplitude of the SPP time the cosine of the phase difference between the signal and the reference branch. The fringes indicate the actual SPP wave fronts. The spacing between the wave fronts yields the phase velocity.

An excellent model system to investigate nonlinear effects in plasmonic nanostructures is a subwavelength hole array. These arrays are known to exhibit a phenomenon that has been named both enhanced, but also extraordinary transmission [14]. In this phenomenon the fraction of the incident light that is transmitted by the array can exceed the geometrical open fraction of the film due to the presence of the holes. Associated with this extraordinary transmission are large local field enhancements inside the holes. It was found that changing the hole shape and size, changes the linear transmission properties of the array [15–17]. These arrays also exhibit nonlinear effects such as second-harmonic generation (SHG) [18].

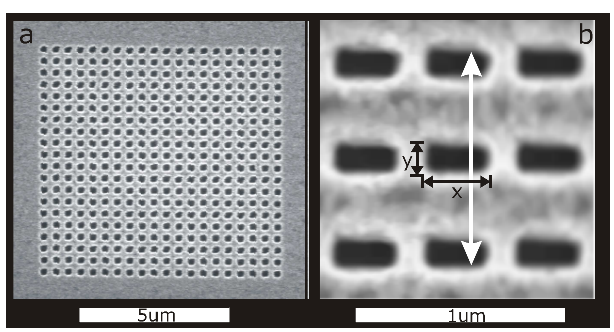
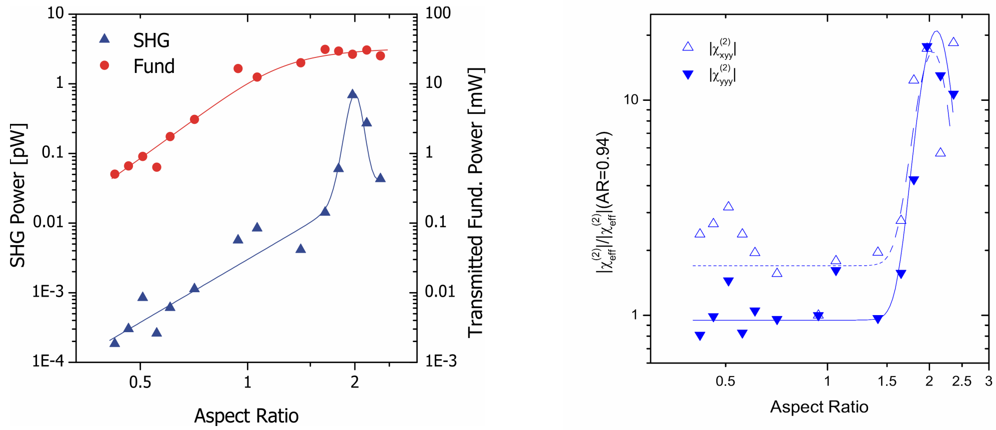


Figure 4: (a) SEM micrograph of an array of square holes. (b) SEM micrograph of an array with AR = x/y = 2.0, where x and y (black arrows) are defined as the hole dimensions perpendicular and parallel to the polarization of the incident light (white arrow), respectively.

We investigated SHG in more detail by measuring multiple arrays containing 20×20 rectangular holes in a 160 nm thick Au film (see Figure 4(a)). The hole area are fixed at $3.4 \times 10^4\,\mathrm{nm}^2$ and 410 nm, respectively. The aspect ratio of the holes is defined as: AR = x/y, with x and y defined in Figure 5(b). The white arrow in Figure 4(b) denotes the polarization of the fundamental light. The sample was illuminated using a Ti:Sapphire laser (830 nm, 80 fs pulses).



(a) (b)

Figure 5: (a) The experimentally determined SHG power and fundamental power measured in transmission as a function of the aspect ratio (AR) of the subwavelength holes in the array. Lines are guides to the eye. (b) The normalized effective nonlinear susceptibility for different AR. Lines are guides to the eye. It is clear that around AR = 2 the effective nonlinear susceptibility is enhanced by an order of magnitude.

Figure 5(a) shows the amount of SHG and fundamental light transmitted by the hole array. We observe a large increase in both signals for increasing AR. The SHG signal even increases by 4 orders of magnitude. The increase in the fundamental transmission is consistent with previous findings as the (1, 0) shifts as the aspect ratio is increased [16].

It is obvious that the SHG will increase as a result of an increase in the transmission of the fundamental, which will result in higher fields inside the holes and at the sample surface. For a better understanding of the actual nonlinear response of the hole arrays, we eliminate this trivial increase by approximating the hole arrays as a slab of effective nonlinear medium that generates SHG. By taking into account the linear transmission coefficients of the array for both the fundamental and the second harmonic we can determine the effective n nonlinear susceptibility of the array. The effective second-order nonlinear susceptibility $\chi^{(2)}_{eff}$ for each array is presented in Figure 5(b). We observe that between AR = 0.4 and AR = 1.5 no clear trend is discernable in the effective nonlinear susceptibility. This indicates that the gradual increase of the SHG signal over two orders of magnitude, is primarily explained by linear transmission effects only. However, around AR = 2 we observe an increase of the $\chi^{(2)}_{eff}$, by an order of magnitude.

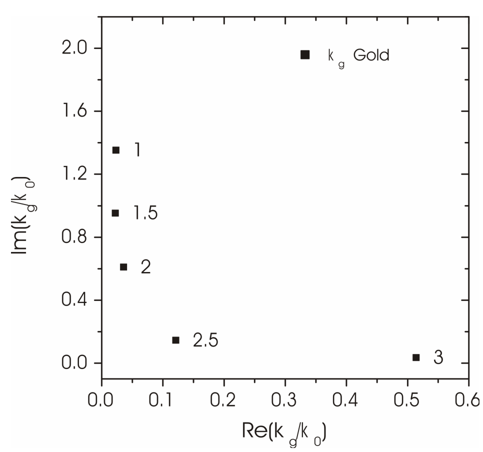


Figure 6: Calculated imaginary vs. real part of the propagation constant of the fundamental modes inside the subwavelength holes for an incident wavelength of 830 nm. Roughly where the imaginary part of the propagation constant is equal to the real part is the cut off aspect ratio for the incident wavelength.

To explain the enhancement of the effective nonlinear susceptibility around AR = 2, we consider the wavevectors of the guided modes for the fundamental light inside the different hole shapes, calculated by Fourier modal method [19]. These wavevectors are depicted in Figure 6. We can distinguish two regimes. The first regime starts at $\mathrm{AR} = 1$ and runs roughly until $\mathrm{AR} = 2$; here the imaginary part of k_g is dominant, i.e., modes excited with a free space wavelength of 830 nm are mainly evanescent. On the other hand, the second regime, starting at $\mathrm{AR} = 2.5$ and extending to higher AR, shows modes with a dominance of the real part of k_g. Here, the modes are propagating. In between the evanescent and the propagating regime will be the cutoff region for a single hole. When a waveguide is close to cutoff, the modes in that guide will have a low group velocity. It has been shown that a reduction in group velocity leads to enhanced nonlinear responses due to an increase of lightmatter interactions. We suggest that these slow modes for the fundamental wavelength cause the sharp increase in the $\chi^{(2)}_{eff}$ [20].

In conclusion we have presented the first investigations in which SPPs are tracked on a femtosecond time scale as they propagate along an SPP waveguide. Both the phase- and group velocity of the wavepackets are unambiguously determined. We have also shown that the effective nonlinear susceptibility of metallic nanostructures can be enhanced by changing their shape. The resulting enhancement in nonlinear effects occurs on top the enhancements produced by local field enhancements. In the case of subwavelength hole arrays we find that the effective nonlinear susceptibility can be enhanced by an order of magnitude by changing their shape. The susceptibility increases as a result of slow light propagation inside the holes.

ACKNOWLEDGMENT

This work was made possible by the facilities of the Amsterdam nanoCenter. The work is part of the research program of the Stichting for Fundamenteel Onderzoek der Materie (FOM), which is financially supported by the Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO). Support of the EC-funded project PHOREMOST (FP6/2003/IST/2-511616) is gratefully acknowledged.

REFERENCES

1. Raether, H., *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, Springer-Verlag, Berlin, Germany.
2. Barnes, W. L., A. Dereux and T. W. Ebbesen, *Nature*, Vol. 424, 824–830, 2003.
3. Chang, R. and T. Furtak, *Surface enhanced Raman scattering*, Plenum, New York, 1981.
4. Nie, S. and S. R. Emory, *Science*, Vol. 275, 1102, 1997.
5. Weeber, J-C., et al., *Phys. Rev. B*, Vol. 64, 045411, 2001.
6. Zia, R., M. D. Selker, and M. L. Brongersma, *Phys. Rev. B*, Vol. 71, 165431, 2005.
7. Balistreri, M. L. M., et al., *Science*, Vol. 294, 1080–1082, 2001.
8. Bozhevolnyi, S. I. and L. Kuipers, *Semiconductor Science and Technology*, Vol. 26, R1–R16, 2006.
9. Gersen, H., T. J. Karle, R. J. P. Engelen, W. Bogaerts, J. P. Korterik, N. F. van Hulst, T. F. Krauss, and L. Kuipers, *Phys. Rev. Lett.*, Vol. 94, 123901, 2005.
10. Engelen, R. J. P., T. J. Karle, H. Gersen, J. P. Korterik, T. F. Krauss, L. Kuipers, and N. F. van Hulst, *Optics Express*, Vol. 13, 4457–4464, 2005.
11. Gersen, H., T. J. Karle, R. J. P. Engelen, W. Bogaerts, J. P. Korterik, N. F. van Hulst, T. F. Krauss, and L. Kuipers, *Phys. Rev. Lett.*, Vol. 94, 073903, 2005.
12. Engelen, R. J. P., Y. Sugimoto, Y. Watanabe, J. P. Korterik, N. Ikeda, N. F. van Hulst, K. Asakawa, and L. Kuipers, *Optics Express*, Vol. 14, 1658–1672, 2006.
13. Kretschmann, E. and H. Raether, *Z. Naturforsch. A*, Vol. 23, 2135, 1968.
14. Ebbesen, T. W., et al., *Nature*, Vol. 391, 667, 1998.
15. Klein Koerkamp, K. J., et al., *Phys. Rev. Lett.*, Vol. 92, 183901, 2004.
16. Van der Molen, K. L., K. J. Klein Koerkamp, S. Enoch, F. B. Segerink, N. F. van Hulst, and L. Kuipers, *Phys. Rev. B*, Vol. 72, 045421, 2005.
17. Van der Molen, K. L., F. B. Segerink, N. F. van Hulst, and L. Kuipers, *Appl. Phys. Lett.*, Vol. 85, 2004, Lastname, F. M., *Book Title*, Wiley-Interscience, New York, 1986.
18. Airola, M., Y. Liu, and S. Blair, *J. Opt. A*, Vol. 7, 118, 2005.
19. Li, L. F., *J. Opt. Soc. Am. A*, Vol. 13, 1024, 1996.
20. Van Nieuwstadt, J. A. H., M. Sandtke, R. H. Harmsen, F. B. Segerink, J. C. Prangsma, S. Enoch, and L. Kuipers, accepted for publication in *Phys. Rev. Lett.*.

Modeling Lightning Attachment to Tall Towers

Rajeev Thottappillil and Nelson Theethayi
Uppsala University, Uppsala, Sweden

Abstract— Lightning return stroke resulting from the downward leader attachment to tall towers is modeled as a point current source in series with the lightning channel and tower. Assuming a quasi-transverse electromagnetic field structure, expressions for the current at the top of the tower, at the base of the tower and along the tower are presented.

1. INTRODUCTION

Modeling the lightning strike to tall towers is an interesting problem, especially due to the fast expansion of the mobile telecommunication network and the resulting increase in the number of towers and associated lightning events to the towers. Measurements show that return stroke peak currents at the top and bottom of a tall tower are different. For instance, Gorin and Shkilev [1] reports median return stroke peak currents of 9 kA at 533 m level of the 540 m tall Ostankino TV tower in Moscow, whereas they have measured 18 kA at the bottom at 47 m level. The increased value of the return stroke peak current at the bottom of the tower is due to the transient processes in the tower during the lightning return stroke. Also, it has been observed that the average value of the peak electric and magnetic fields from the from tower lightning return strokes are 2 to 3 times larger than that would be expected for a return stroke of similar peak current striking level ground [2]. Analysis of current waveforms at the top and bottom of the tower gives indications of current reflections at the tower bottom and tower top [3], which could be approximately explained by modeling the tower as a transmission line with reflection coefficients at the tower ground interface and at the tower-lightning channel interface and the lightning as a current or voltage source at tower top. However, there are differences among researchers on the nature of the source to be used at the tower top for modeling the lightning interaction [4–7]. This issue is addressed in this paper and a new model for lightning interaction with the tower is proposed.

2. THE MODEL

In normal negative ground flashes, a stepped-leader, usually negatively charged, descends towards the ground from the thundercloud. When the leader approaches the ground, electric fields at the tips of grounded objects are enhanced. When the field exceeds the break down value of the air an upward leader is initiated from the object towards the descending leader. When the upward and downward moving leaders meet, a lightning return stroke happens, associated with large currents and large field changes. The upward wave effectively neutralizes the negative charge and the downward wave neutralizes the positive charge of the upward leader from the tower, charge on the tower, and charge from the grounding network. The upward current can be treated as due to positive charge flow and the downward current can be treated as due to negative charge flow, giving rise to current vectors in the same direction in the channel and the tower initially.

According to the above picture, the undisturbed current source, representative of the return stroke initiation at the attachment point between the two oppositely charged leaders, is in SERIES with the lightning channel and can not be in parallel with remote ground as reference as suggested in [4]. The series point current source is very similar to the source in the Transmission Line (TL) model of the return stroke to ground proposed by [8]. In that model, a point current source at the base of the lightning channel launches a current wave upward and this current is assumed to be the same as the current that would be measured at the base of the channel. The current wave is assumed to be traveling without any attenuation and distortion as if in a lossless transmission line, and hence the name of the model, TL model. The actual geometry of the twoconductor TL or the associated field structure was not discussed until quite recently. Thottappillil et al., [9–11] showed that the exact field solution of the problem of a vanishingly thin perfectly conducting vertical wire above a perfectly conducting ground plane, excited by a point source at the junction of the wire and ground plane, is a spherical transverse electromagnetic field structure (spherical TEM) centered at the point source. This field structure is associated with an unattenuating current traveling at speed of light and the problem can be modeled as a TL with lightning channel as one conductor and the ground plane as the second conductor. In the TL model of return stroke, the speed is assumed to be the optically observed speed of the return-stroke front, typically one-third to two-third the speed of light, and hence one can think of an imperfect spherical TEM, approaching to spherical TEM as return stroke speed approaches light speed. The concept of series point source model for tower lightning return stroke described above is illustrated in Fig. 1.

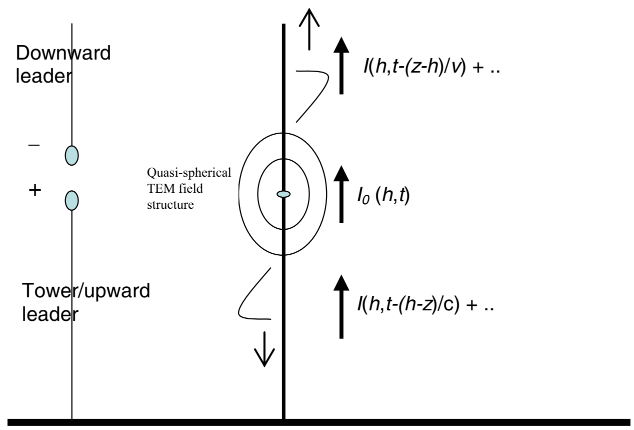


Figure 1: Illustration of the concept of series current source model for tower lightning return stroke.

The wave impedance associated with pure spherical TEM wave is the free space impedance $377\,\Omega$. However, due to the quasi-spherical TEM field structure, effective wave impedance different from $377\,\Omega$ for the tower and the channel can be considered. The concept illustrated in Fig. 1 can also be viewed as a two conductor transmission line problem, with one conductor being the tower and the channel and the other conductor being the ground. In terms of distributed TL parameters, the surge impedance of a two-conductor TL supporting quasi-TEM field structure is given by,

$$Z = \sqrt{\frac{R + j\omega L}{G + j\omega C}} \tag{1}$$

For nearly spherical waves between the vertical conductor and ground plane, G, L, and C are similar for the tower and the channel, but due to the high resistance of the lightning plasma channel the surge impedance of the channel is higher than that of the tower. The finite conductivity of ground and associated losses during wave reflection can be modeled as lumped ground impedance. From these considerations, the schematic diagram of the series point current source model can be drawn as given in Fig. 2.

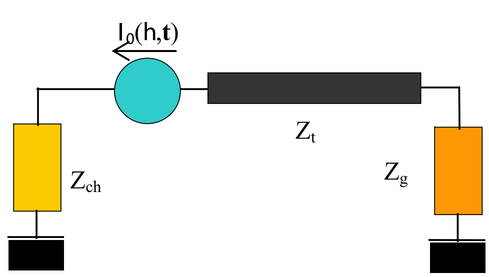


Figure 2: Schematic diagram of series point current source representation of tower lightning return stroke.

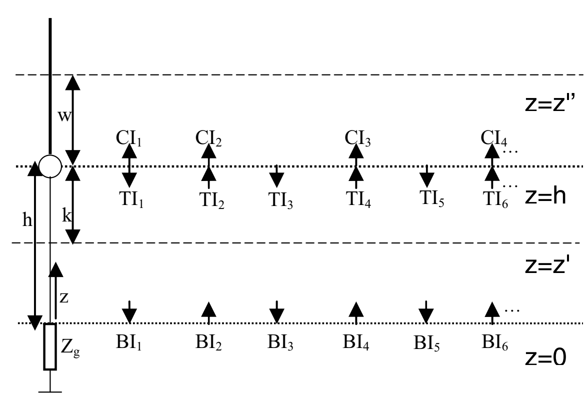


Figure 3: Schematic diagram showing the direction of current waves at the tower-ground interface and tower-lightning channel interface.

In Fig. 2, Z_{ch} is the lightning channel surge impedance, Z_t is the tower surge impedance, and Z_g is the ground surge impedance. Also, $Z_{ch} > Z_t > Z_g$. The point current source $I_0(h, t)$ is assumed to be the same as the current that would be measured at ground in the absence of the tower and is the same as the input to the classical TL model of the return stroke starting from ground. The current launched initially into the channel and the tower is the same, that is,

$$i(h^-,\, t) = i(h^+,\, t) = I_0(h,\, t) \tag{2}$$

Let us now derive the expressions for the tower current and the channel current at a given time t considering the reflections at various interfaces. Fig. 3 shows the direction of the current waves and the points at which total currents are determined.

In Fig. 3 the BI_i are the current waves incident and reflected at the tower base, TI_i are the corresponding current waves in the tower incident or reflected at the tower-channel junction and CI_i are the current waves leaving the tower into the channel. The reflection coefficient for the current wave incident at the towerground interface is ρ_g, given by

$$\rho_g = \frac{Z_t - Z_g}{Z_t + Z_g} \tag{3}$$

and the reflection coefficient at the tower-channel interface for the current wave incident from the tower is ρ_t, given by

$$\rho_t = \frac{Z_t - Z_{ch}}{Z_t + Z_{ch}}. \tag{4}$$

For most tower structures and ground, $Z_{ch} > \underline{Z_t} > Z_g$. Therefore, ρ_t is negative and ρ_g is positive. The presence of any upward leader is neglected, but could be easily included into the model if required. As mentioned before, speed of light is assumed for the waves in the tower, and speed v is assumed for the upward wave in the lightning channel. The total current received at the bottom of the tower at any instant t is

$$\begin{aligned} i(0,\, t) &= BI_1 + BI_2 + BI_3 + BI_4 + BI_5 + BI_6 + \ldots \\ &= I_0\left(h,\, t-\frac{h}{c}\right) + \rho_g I_0\left(h,\, t-\frac{h}{c}\right) + \rho_t\rho_g I_0\left(h,\, t-\frac{3h}{c}\right) + \rho_t\rho_g^2 I_0\left(h,\, t-\frac{3h}{c}\right) \\ &\quad + \rho_t^2\rho_g^2 I_0\left(h,\, t-\frac{5h}{c}\right) + \rho_t^2\rho_g^3 I_0\left(h,\, t-\frac{5h}{c}\right) + \rho_t^3\rho_g^3 I_0\left(h,\, t-\frac{7h}{c}\right) + \rho_t^3\rho_g^4 I_0\left(h,\, t-\frac{7h}{c}\right) + \ldots \\ &= (1+\rho_g)\sum_{n=0}^{\infty} \rho_t^n\rho_g^n \cdot I_0\left(h,\, t-\frac{(2n+1)h}{c}\right) \end{aligned} \tag{5}$$

The magnitude of total current in the tower at the tower-channel interface at any instant t is

$$\begin{aligned} i(h,\, t) &= TI_1 + TI_2 + TI_3 + TI_4 + TI_5 + TI_6 + \ldots \\ &= I_0\,(h,\, t) + \rho_g I_0\left(h,\, t-\frac{2h}{c}\right) + \rho_t\rho_g I_0\left(h,\, t-\frac{2h}{c}\right) + \rho_t\rho_g^2 I_0\left(h,\, t-\frac{4h}{c}\right) \\ &\quad + \rho_t^2\rho_g^2 I_0\left(h,\, t-\frac{4h}{c}\right) + \rho_t^2\rho_g^3 I_0\left(h,\, t-\frac{6h}{c}\right) + \rho_t^3\rho_g^3 I_0\left(h,\, t-\frac{6h}{c}\right) + \rho_t^3\rho_g^4 I_0\left(h,\, t-\frac{8h}{c}\right) + \ldots \\ &= I_0(h,\, t) + (1+\rho_t)\sum_{n=0}^{\infty} \rho_t^n\rho_g^{n+1} \cdot I_0\left(h,\, t-\frac{2(n+1)h}{c}\right) \end{aligned} \tag{6}$$

The magnitude of currents launched into the channel from the tower-channel interface at any instant t is

$$\begin{aligned} i(h,\, t) &= CI_1 + CI_2 + CI_3 + CI_4 + CI_5 + CI_6 + \ldots \\ &= I_0\,(h,\, t) + (1+\rho_t)\rho_g I_0\left(h,\, t-\frac{2h}{c}\right) + (1+\rho_t)\rho_t\rho_g^2 I_0\left(h,\, t-\frac{4h}{c}\right) \\ &\quad + (1+\rho_t)\rho_t^2\rho_g^3 I_0\left(h,\, t-\frac{6h}{c}\right) + (1+\rho_t)\rho_t^3\rho_g^4 I_0\left(h,\, t-\frac{8h}{c}\right) + \ldots \\ &= I_0(h,\, t) + (1+\rho_t)\sum_{n=0}^{\infty} \rho_t^n\rho_g^{n+1} \cdot I_0\left(h,\, t-\frac{2(n+1)h}{c}\right) \end{aligned} \tag{7}$$

The above expressions in (6) and (7) are identical and show that the currents in the channel just above the tower-channel interface and the currents in the tower just below the tower-channel interface are the same at any instant, and it should be so as per the continuity equation of charge and current.

The total current at any point $z = z'$ at any instant in the tower is the sum of currents incident at that point from the tower tip and the tower bottom with appropriate delays and is given by

$$\begin{aligned} i(z',t) &= I_0\left(h,\, t-\frac{k}{c}\right)+\rho_g I_0(h,\, t-\frac{h}{c}-\frac{z'}{c})+\rho_t\rho_g I_0\left(h,\, t-\frac{2h}{c}-\frac{k}{c}\right)+\rho_t\rho_g^2 I_0\left(h,\, t-\frac{3h}{c}-\frac{z'}{c}\right) \\ &\quad +\rho_t^2\rho_g^2 I_0\left(h,\, t-\frac{4h}{c}-\frac{k}{c}\right)+\rho_t^2\rho_g^3 I_0\left(h,\, t-\frac{5h}{c}-\frac{z'}{c}\right)+\ldots \\ &= \sum_{n=0}^{\infty}\rho_t^n\rho_g^n\cdot I_0\left(h,\, t-\frac{2nh}{c}-\frac{k}{c}\right)+\sum_{n=0}^{\infty}\rho_t^n\rho_g^{n+1}\cdot I_0\left(h,\, t-\frac{(2n+1)h}{c}-\frac{z'}{c}\right) \end{aligned} \tag{8}$$

where $k = h - z'$. The current at any point $z = z''$ at any instant in the channel is the sum of currents incident at that point which are launched into the channel from the tower tip with appropriate delays and is given by

$$\begin{aligned} i(z'',t) &= I_0\left(h,\, t-\frac{w}{v}\right)+(1+\rho_t)\rho_g I_0(h,\, t-\frac{2h}{c}-\frac{w}{v})+(1+\rho_t)\rho_t\rho_g^2 I_0\left(h,\, t-\frac{4h}{c}-\frac{w}{v}\right) \\ &\quad +(1+\rho_t)\rho_t^2\rho_g^3 I_0\left(h,\, t-\frac{6h}{c}-\frac{w}{v}\right)+(1+\rho_t)\rho_t^3\rho_g^4 I_0\left(h,\, t-\frac{8h}{c}-\frac{w}{v}\right)+\ldots \\ &= I_0\left(h,\, t-\frac{w}{v}\right)+(1+\rho_t)\sum_{n=0}^{\infty}\rho_t^n\rho_g^{n+1}\cdot I_0\left(h,\, t-\frac{2(n+1)h}{c}-\frac{w}{v}\right) \end{aligned} \tag{9}$$

where $w = z'' - h$.

As a demonstration, the equations presented above are used to evaluate the current distributions on a tower of height 168 m representative of Peissenberg tower in Munich [3, 12]. The point current source is representative of typical subsequent return strokes as shown in Fig. 6 in [13]. In the simulations the current waveforms at tower top (source point), tower bottom, middle of tower and also the current injected into the channel is shown in Fig. 4. The adopted reflection coefficients are at tower top $\rho_t = -0.53$ and $\rho_g = 0.7$, and these values are estimated from measured current waveforms [5].

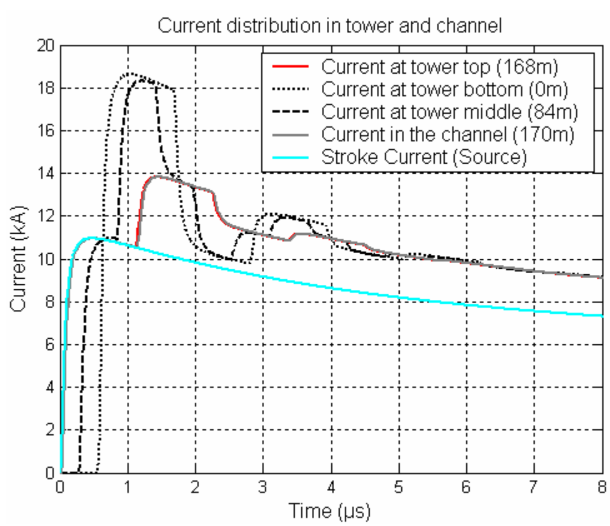


Figure 4: Current distributions in a 168 m tall tower and the channel using the series point source current injection model.

It is seen from the simulations that the current at the tower top is the same as that of the injected into the channel. The current in the channel is calculated at 2 m above the tower tip.

The reflections are dominant at the middle of the tower and the bottom of the tower due to larger value of the ground reflection coefficient i.e., the current peak at the tower bottom is $(1+\rho_g)$ times point source current. The current in the middle of the tower is affected by the reflections from the tower top and tower bottom, hence we see half the width for the peaks for currents in the middle of the tower as shown in Fig. 4. The overall features of the waveform at the tower top in Fig. 4 are consistent with the features of measured currents in actual tower.

3. DISCUSSION

At the beginning of this paper, it has been argued that series point current source model is perhaps closer to reality as far as lightning attachment process to tower is concerned. This point is discussed further here. There are several return stroke models having distributed current sources along the lightning channel. These are referred to as the so called current discharge type models [14]. These distributed sources are due to the charge deposited by the leader on the channel (including corona) being neutralized by the upward propagating return stroke wavefront. Therefore the distributed sources are in fact distributed current sources in series with the lightning channel and not in parallel to it. In pictorial representation of the process, corona sources are shown as distributed current sources connected perpendicular to the vertical channel. The other end of the corona current source is not often well-defined, but is most likely the outer sheath of corona envelope, which is part of the vertical conductor itself. If the other end of corona source is interpreted as the remote reference or ground (the second conductor), as done in [4], such a current source would be in parallel connection across the two conductors of the transmission line, between vertical conductor and the ground plane. Another interpretation for the parallel current source would be to consider the outer sheath of corona to be the reference or ground conductor and also consider a cylindrical TEM wave.

The series voltage source representation of lightning attachment to tower [6] is very similar to the series point current source representation of this paper. However, the voltage source is unknown and therefore first the current source is to be specified and the voltage source determined from the knowledge of channel and tower surge impedances. We believe that the series point current source representation as presented in this paper is a simpler and straight forward approach to modeling the lightning interaction with the tower.

It is also possible to extend the series point current source model to include upward leaders from the tower. In that case the point current source is assumed at the junction between the upward leader and downward leader. An additional reflection coefficient at the tower tip for current waves traveling down has to be defined.

ACKNOWLEDGMENT

This work is carried out with a grant from Swedish Research Council.

REFERENCES

1. Gorin, B. N. and A. V. Shkilev, "Measuements of lightning currents at the Ostankino tower," *Elektrichestvo*, Vol. 8, 64–65, 1984.
2. Rachidi, F., W. Janischewskyj, A. M. Hussein, C. A. Nucci, S. Guerrieri, B. Kordi, and J. S. Chang, "Current and electromagnetic field associated with lightning return strokes to tall towers," *IEEE Trans. on Electromagnetic Compatibility*, Vol. 43, No. 3, 356–366, 2001.
3. Fuchs, F., "On the transient behaviour of the telecommunication tower at the mountain Hoher Peissenberg," *Proc. 24th Int. Conf. Lightning Protection*, 36–41, Birmingham, UK, 1998.
4. Rachidi, F., V. Rakov, C. A. Nucci, and J. L. Bermudez, "The effect of vertically-extended strike object on the distribution of current along the lightning channel," *Journal of Geophysical Research*, Vol. 107, No. D23, 4699, 2002.
5. Bermudez, J. L., M. Rubinstein, F. Rachidi, F. Heidler, and M. Paolone, "Determination of reflection coefficients at the top and bottom of elevated strike objects struck by lightning," *Journal of Geophysical Research*, Vol. 108, No. D14, 4413, doi:10.1029/2002JD002973, 2003.
6. Baba, Y. and V. Rakov, "On the use of lumped sources in lightning return stroke models," *Journal of Geophysical Research*, Vol. 110, D03101, doi:10.1029/2004JD005202, 2005a.
7. Guerrieri, S., C. A. Nucci, F. Rachidi, and M. Rubinstein, "On the influence of elevated strike objects on directly measured and indirectly estimated lightning currents," *IEEE Trans. on Power Delivery*, Vol. 13, No. 4, 1998.
8. Uman, M. A. and D. K. McLain, "Magnetic field of the lightning return stroke," *Journal of Geophysical Research*, Vol. 74, No. 28, 6899–6910, 1969.

9. Thottappillil, R. and M. A. Uman, "Reply to the 'Comment on "Return stroke transmission line model for stroke speed near and equal that of light by R. Thottappillil, J. Schoene, and M.A. Uman"' by G. Kordi, R. Moini, and V. A. Rakov," *Geophysical Research Letters*, Vol. 29, May 2002.
10. Thottappillil, R. and M. A. Uman, "Electric and magnetic fields from a semi-infinite vertical thin-wire antenna above a conducting plane," *Progress in Electromagnetic Research Symposium*, Pisa, Italy, March 28–31, 2004.
11. Thottappillil, R., M. A. Uman, and N. Theethayi, "Electric and magnetic fields from a semi-infinite antenna above a conducting plane," *J. Electrostatics*, Vol. 61, 209–221, 2004.
12. Heidler, F., J. Wiesinger, and W. Zischank, "Lightning currents measured at a telecommunication tower from 1992 to 1998," paper presented at *14th International Zurich Symposium on Electromagnetic Compatibility*, Zurich, Switzerland, 2001.
13. Nucci, C. A., G. Deindorfer, M. A. Uman, F. Rachidi, M. Ianoz, and C. Mazzetti, "Lightning return stroke current model with specified channel-base current: A review and comparison," *J. Geophys. Res.*, Vol. 95, No. 20, 395–20,408, 1990.
14. Rakov, V. A. and M. A. Uman, "Review and evaluation of lightning return stroke models including some aspects of their application," *IEEE Trans. EMC*, Vol. 40, 403–426, 1998.

Electromagnetic Scattering from an Arbitrarily Shaped Three-dimensional Inhomogeneous Bianisotropic Body

C. Mei, M. Hasanovic, J. K. Lee, and E. Arvas

Department of Electrical Engineering and Computer Science
Syracuse University, Syracuse, NY 13244, USA

Abstract— This paper presents a numerical solution to scattering problems that involve 3D inhomogeneous bianisotropic scatterers of an arbitrary shape. The constitutive relations are assumed to be of the most general form and composed of four 3×3 matrices or tensors. The problem is described through a mixed potential formulation. The electric and magnetic potentials are related to electric and magnetic bound charges and polarization currents and then to the electric and magnetic polarizations. The electric field and magnetic field integral equations are constructed. The method of moments technique is then applied to obtain a numerical solution to the problem. The volume of the scatterer is meshed by tetrahedral cells and face-based functions are applied to expand unknown quantities. The proposed formulation has been evaluated and verified through examples of scattering by various chiral and gyrotropic scatterers illuminated by a plane electromagnetic wave. Numerical results of scattering from a chiroferrite sphere are also presented showing the flexibility of the proposed method.

1. INTRODUCTION

In recent years, there has been a growing interest in the study of interaction between electromagnetic fields and bianisotropic materials. Various complex media such as uniaxial materials, gyrotropic materials, chiral materials, and combinations of chiral materials and ferrites (chiroferrites) have found their place in circuit substrates, waveguides, antennas, and many other electronic and optical devices. A study of inhomogeneous bianisotropic scattering problems is also important for applications such as radar detection, medical diagnostics, and power absorption in biological bodies.

With a development of material techniques, there is an urgent need for a fast and accurate general-purpose electromagnetic field solver that could handle all kinds of inhomogeneities, dispersion, anisotropy, chirality, and even bianisotropy. There exist several methodologies to analyze electromagnetic field interaction with complex media. For example analytical methods such as Mie series expansion [1] and exponential matrix technique [2] are successfully used to study source radiation and wave propagation problems in simple structures such as biisotropic spheres or bianisotropic multilayer substrates. 3D full-wave numerical algorithms in time domain such as finite-difference time-domain (FDTD) method, the method of line (MoL), and transmission line modeling method (TLM) have been extended to model electromagnetic interaction with various complex media such as chiral materials and gyrotropic materials [3, 4]. Latest research efforts are related to scattering from homogenous dispersive anisotropic or biisotropic materials. However each of these methods has certain limitations. Analytical methods are difficult to apply to complex structures that involve 3D arbitrary shaped geometries. In solving problems that involve dispersive materials, time domain methods rely on the Z-transform of analytical expressions that describe dispersion properties of a material [5]. These analytical expressions are in many cases very difficult to obtain. Corresponding computer programs also need to be adapted for different dispersion properties.

Current frequency domain methods such as finite-element method (FEM), FEM-boundaryelement method (FEM-BEM), and method of moments (MoM) are also able to treat wave propagation and radiation problems related to chiral media or anisotropic materials [6–9]. Most of the time problems are solved in frequency domain, so there is no need to obtain analytic expressions of the material dispersion in advance. In general, there is little research work done in solving problems that involve general bianisotropic materials. Main purpose of this paper is to offer a MoM solution to problems that involve 3D inhomogeneous bianisotropic scatterers of an arbitrary shape. The developed formulation does not put any limits on the geometrical assignment of material properties to the scatterer. It can be applied to multilayered scatterers, scatterers with materials assigned to different regions of the scatterer in a linear, exponential or any other fashion, etc.

2. FORMULATION

The proposed problem can be formulated in the following way. Let us consider a general inhomogeneous bianisotropic body characterized by the following constitutive relations in a region free of sources,

$$\overline{D} = \overline{\overline{\epsilon}} \cdot \overline{E} + \overline{\overline{\xi}} \cdot \overline{H}, \quad \overline{B} = \overline{\overline{\zeta}} \cdot \overline{E} + \overline{\overline{\mu}} \cdot \overline{H} \tag{1}$$

where $\overline{\overline{\epsilon}}$ is the permittivity tensor, $\overline{\overline{\mu}}$ is the permeability tensor, and $\overline{\overline{\xi}}$ and $\overline{\overline{\zeta}}$ are the magnetoelectric tensors. They are all assumed to be generalized 3×3 tensors.

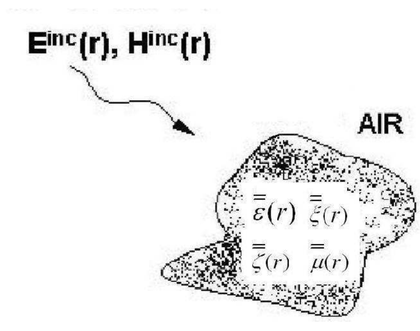


Figure 1: Inhomogeneous bianisotropic body illuminated by an electromagnetic wave.

As shown in Figure 1, when the body is illuminated by a time-harmonic electromagnetic wave, the total field can be expressed through the following mixed potential formulation:

$$\overline{E} = \overline{E}^{inc} - j\omega\overline{A} - \nabla V - \frac{1}{\epsilon_0}\nabla \times \overline{F}, \quad \overline{H} = \overline{H}^{inc} - j\omega\overline{F} - \nabla U + \frac{1}{\mu_0}\nabla \times \overline{A} \tag{2}$$

where $\overline{A}$ and $\overline{F}$ and are electric and magnetic vector potentials and V and U are electric and magnetic scalar potentials respectively.

The electric and magnetic vector potentials $\overline{A}$ and $\overline{F}$ and electric and magnetic scalar potentials V and U are related to the electric and magnetic polarization currents $\overline{J}_{ep}(r)$ and $\overline{J}_{mp}(r)$, electric and magnetic bound charges $\overline{\rho}_{ep}(\overline{r})$ and $\overline{\rho}_{mp}(\overline{r})$, and electric and magnetic surface bound charges $\overline{\sigma}_{ep}(\overline{r})$ and $\overline{\sigma}_{mp}(\overline{r})$ in the following way:

$$\begin{aligned}
\overline{A} &= \mu_0 \int_V dv' G\left(\overline{r}, \overline{r}'\right) \overline{J}_{ep}\left(\overline{r}\right), \quad V = \frac{1}{\epsilon_0}\int_V dv' G\left(\overline{r}, \overline{r}'\right) \rho_{eb}\left(\overline{r}\right) + \frac{1}{\epsilon_0}\int_S ds' G\left(\overline{r}, \overline{r}'\right) \sigma_{eb}\left(\overline{r}\right) \\
\overline{F} &= \epsilon_0 \int_V dv' G\left(\overline{r}, \overline{r}'\right) \overline{J}_{mp}\left(\overline{r}\right), \quad U = \frac{1}{\mu_0}\int_V dv' G\left(\overline{r}, \overline{r}'\right) \rho_{mb}\left(\overline{r}\right) + \frac{1}{\mu_0}\int_S ds' G\left(\overline{r}, \overline{r}'\right) \sigma_{mb}\left(\overline{r}\right)
\end{aligned} \tag{3}$$

where $G(\overline{r}, \overline{r}')$ is the free space Green's function.

Electric and magnetic polarization currents and electric and magnetic bound charges are further related to electric and magnetic polarizations $\overline{P}$ and $\overline{M}$, as

$$\begin{aligned}
\rho_{eb} &= -\nabla \cdot \overline{P}, \qquad \sigma_{eb} = -\left(\overline{P}_1 - \overline{P}_2\right) \cdot \hat{n}, \qquad \overline{J}_{ep} = j\omega\overline{P} \\
\rho_{mb} &= -\mu_0\nabla \cdot \overline{M}, \quad \sigma_{mb} = -\mu_0\left(\overline{M}_1 - \overline{M}_2\right) \cdot \hat{n}, \quad \overline{J}_{mp} = j\omega\mu_0\overline{M}
\end{aligned} \tag{4}$$

where $\hat{n}$ is the normal direction of the boundary pointing from medium 2 to medium 1.

By definition electric and magnetic polarizations are related to electric and magnetic fields by constitutive relations,

$$\overline{P} = \overline{D} - \epsilon_0\overline{E} = \left(\overline{\overline{\epsilon}} - \epsilon_0\overline{\overline{I}}\right) \cdot \overline{E} + \overline{\overline{\xi}} \cdot \overline{H}, \quad \overline{M} = \frac{1}{\mu_0}\overline{B} - \overline{H} = \frac{1}{\mu_0}\left[\overline{\overline{\zeta}} \cdot \overline{E} + \left(\overline{\overline{\mu}} - \mu_0\overline{\overline{I}}\right) \cdot \overline{H}\right]. \tag{5}$$

Substituting (3) into (2), and replacing bound charges and polarization currents by the electric polarization and magnetic polarization and then by electric and magnetic fields using (4) and (5), the only unknowns left in (2) are the total electric and magnetic fields. This way we constructed the electric integral equation and magnetic integral equation.

3. APPLYING THE METHOD OF MOMENTS TECHNIQUE

Equations (2) cannot be solved analytically. Expanding the unknown quantities $\overline{D}(\overline{r})$ and $\overline{B}(\overline{r})$ in terms of a set of face-based functions $\overline{f}_n(\overline{r})$ on a tetrahedral mesh as described in [10], where

$$\overline{D}\left(\overline{r}\right)=\sum_{n=1}^{N} D_n \overline{f}_n\left(\overline{r}\right), \quad \overline{B}\left(\overline{r}\right)=\sum_{n=1}^{N} B_n \overline{f}_n\left(\overline{r}\right), \tag{6}$$

then testing with $\overline{f}_m(\overline{r})$ as in Galerkin's method, (2) are transformed into the following set of equations,

$$\begin{aligned}
&j\omega\left\langle \overline{f}_m(\overline{r}), \overline{A}\left(\overline{r}\right)\right\rangle+\left\langle \overline{f}_m(\overline{r}), \nabla V\left(\overline{r}\right)\right\rangle+\left\langle \overline{f}_m(\overline{r}), \nabla\times\frac{\overline{F}\left(\overline{r}\right)}{\epsilon_0}\right\rangle+\left\langle \overline{f}_m(\overline{r}), \overline{E}(\overline{r})\right\rangle=\left\langle \overline{f}_m(\overline{r}), \overline{E}^{inc}\left(\overline{r}\right)\right\rangle\\
&j\omega\left\langle \overline{f}_m(\overline{r}), \overline{F}\left(\overline{r}\right)\right\rangle+\left\langle \overline{f}_m(\overline{r}), \nabla U\left(\overline{r}\right)\right\rangle-\left\langle \overline{f}_m(\overline{r}), \nabla\times\frac{\overline{A}\left(\overline{r}\right)}{\mu_0}\right\rangle+\left\langle \overline{f}_m(\overline{r}), \overline{H}(\overline{r})\right\rangle=\left\langle \overline{f}_m(\overline{r}), \overline{H}^{inc}\left(\overline{r}\right)\right\rangle
\end{aligned} \tag{7}$$

where for arbitrary $\overline{X}$ and $\overline{Y}$, $\left\langle\overline{X},\overline{Y}\right\rangle$ is the symmetric product of $\overline{X}$ and $\overline{Y}$ defined to be the integral of their dot product over the volume of the scatterer. Unknown vectors $\overline{E}$ and $\overline{H}$ in (7) are replaced by $\overline{D}$ and $\overline{B}$ using

$$\overline{E}=\overline{\overline{\alpha}}_1\cdot\overline{D}+\overline{\overline{\alpha}}_2\cdot\overline{B}, \quad \overline{H}=\overline{\overline{\alpha}}_3\cdot\overline{D}+\overline{\overline{\alpha}}_4\cdot\overline{B} \tag{8}$$

where

$$\overline{\overline{\alpha}}=\begin{bmatrix}\overline{\overline{\alpha}}_1 & \overline{\overline{\alpha}}_2\\ \overline{\overline{\alpha}}_3 & \overline{\overline{\alpha}}_4\end{bmatrix}=\begin{bmatrix}\overline{\overline{\epsilon}} & \overline{\overline{\xi}}\\ \overline{\overline{\zeta}} & \overline{\overline{\mu}}\end{bmatrix}^{-1}, \tag{9}$$

and then by the expansion of basis functions $\overline{f}_n(\overline{r})$ using (6). Equations (7) can then be solved numerically for D_n and B_n.

The process described above involves evaluating several different types of volume integrals over tetrahedra and surface integrals over triangles:

$$\begin{aligned}
I_{mn}&=\int_{T_n}\overline{f}_m\left(\overline{r}\right)\cdot\left[\overline{\overline{\alpha}}_n\overline{f}_n\left(\overline{r}\right)\right]d\tau', \quad \overline{I}_{1n}\left(\overline{r}\right)=\int_{T_n}\overline{f}_n\left(\overline{r}\right)G\left(\overline{r},\overline{r}'\right)d\tau', \quad I_{2n}\left(\overline{r}\right)=\int_{T_n}G\left(\overline{r},\overline{r}'\right)d\tau'\\
I_{3mn}&=\int_{T_n}\overline{f}_m\left(\overline{r}\right)\cdot\nabla\times\overline{\overline{\beta}}_n\overline{I}_{1n}\left(\overline{r}\right)d\tau', \quad I_{Sn}=\int_{S_n}\overline{\overline{\beta}}_n\left[\overline{f}_n\left(\overline{r}\right)\cdot\hat{n}_n\right]G\left(\overline{r},\overline{r}'\right)d\tau'.
\end{aligned} \tag{10}$$

These integrals are treated in the same way as in [8–14].

4. NUMERICAL RESULTS

Numerical data obtained through the MATLAB implementation of the proposed formulation is given here. We considered three types of scatterers: a two-layered chiral sphere, a gyromagnetic sphere, and a two-layered chiroferrite sphere. These bodies are illuminated by a θ-polarized plane electromagnetic wave incident from the direction where $\theta=180^\circ$ and $\phi=0^\circ$ ($\overline{E}^{inc}=\hat{a}_x E^{inc}e^{jkz}$).

We first investigated a two-layered chiral sphere of radius R with $k_0R=k_0r_2=1.5$ where k_0 is the free-space wave number. The radius of the core is half the radius of the whole sphere and has relative chirality $\xi_{r1}=0.3$ while the outside layer has relative chirality $\xi_{r1}=0.5$. Both layers have the same relative permittivity $\epsilon_r=2$ and relative permeability $\mu_r=1$. Results presented in Figure 2 are copolarized and cross-polarized bistatic radar cross sections, respectively. Obtained numerical results are compared to the results given in [9] and good agreement is observed.

In the second example, the scattering body is a homogeneous gyromagnetic sphere. The sphere is of radius R of $k_0R=0.2\pi$ and filled by a material with relative permittivity $\epsilon_r=1$ and the relative permeability tensor $\overline{\overline{\mu}}_r=\begin{bmatrix}5 & -j & 0\\ j & 5 & 0\\ 0 & 0 & 7\end{bmatrix}$. Total radar cross sections obtained through the implementation of the proposed formulation are given on the left side of Figure 3, and compared to the work of Geng et al. [7]. The results of two methods match well.

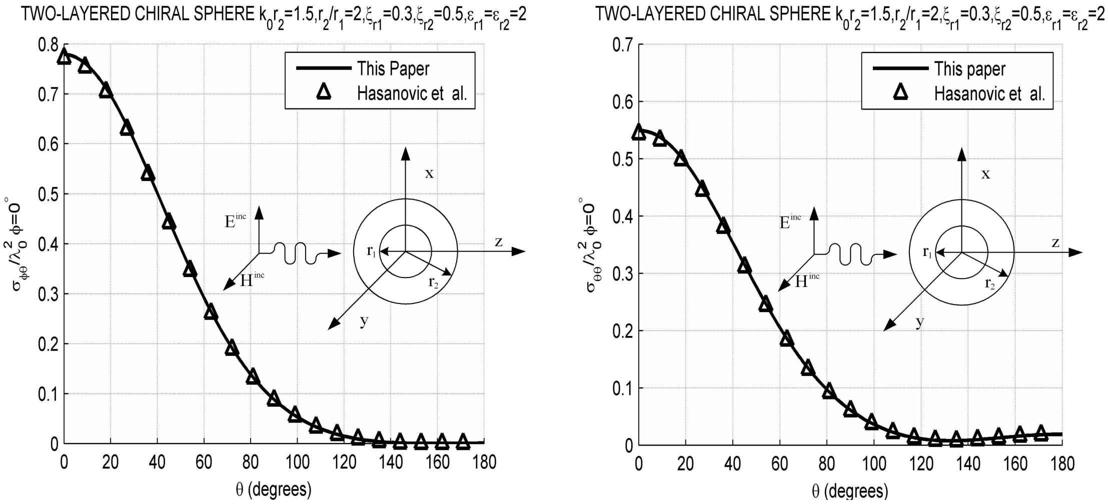


Figure 2: Copolarized (left) and cross-polarized (right) bistatic radar cross sections of a two-layered chiral sphere illuminated by a plane EM wave.

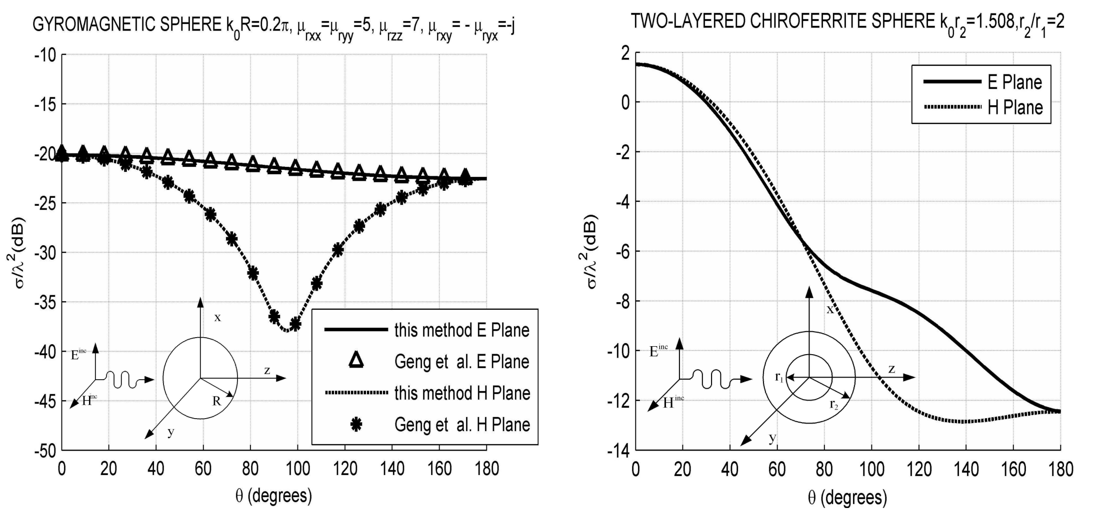


Figure 3: Total radar cross section of a gyromagnetic sphere (left) and a two-layered chiroferrite sphere (right). Both scatterers are illuminated by a plane EM wave.

The last example we show here is a two-layered chiroferrite sphere of radius R with $k_0R = k_0r_2 = 1.508$. The radius of the core is half the radius of the whole sphere and the material is gyromagnetic material with the relative permeability tensor $\overline{\overline{\mu}}_r = \begin{bmatrix} 2.3231 - j0.1101 & 0.0879 + j0.6571 & 0 \\ -0.0879 - j0.6571 & 1.7462 - j0.4308 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and relative permittivity $\epsilon_r = 1$. The outside layer has relative chirality $\xi_r = 0.1018 - j0.0076$, relative permittivity $\epsilon_r = 4.7692 - j1.8462$, and relative permeability $\mu_r = 1.7462 - j0.4308$. Results for total radar cross sections are presented on the right side of Figure 3. As the method gave accurate results for the previous cases of chiral and gyromagnetic scatterers, we expect that the numerical results obtained for the chiroferrite scatterer are also valid.

5. CONCLUSION

This paper presents a numerical solution based on the method of moments for electromagnetic scattering from arbitrarily shaped three-dimensional inhomogeneous bianisotropic scatterers. Cases that we studied here are a two-layered chiral sphere, a gyromagnetic sphere, and a two-layered chiroferrite sphere. The proposed solution is applicable to any shape of scatterer and to any kind of spatial and frequency dependence of material properties. However, it may suffer from a rapid growth of computational complexity in the case of electrically large objects with increased mesh resolution. Some acceleration techniques such as those based on the fast Fourier transform [15] can be applied in further research with the objective of increasing the efficiency of the computer program through a reduction of memory and time requirements.

REFERENCES

1. Demir, V., A. Elsherbeni, D. Worasawate and E. Arvas, "A graphical user interface (GUI) for plane wave scattering from a conducting, dielectric or a chiral sphere," Software at ACES web site: http//aces.ee.olemiss.edu, Sept. 2004.
2. Tsalamengas, J. L., "Interaction of electromagnetic waves with general bianisotropic slabs," *IEEE Trans. Microwave Theory Tech.*, Vol. 40, 1870–1878, Oct. 1992.
3. Demir, V., "Electromagnetic scattering from three dimensional chiral objects using the FDTD method," Ph.D. Dissertation, Syracuse University, 2004.
4. Yagli, A. F., "Electromagnetic scattering from three-dimensional gyrotropic objects using the transmission line modeling (TLM) method," Ph.D. Dissertation, Syracuse University, 2006.
5. Demir, V., A. Elsherbeni, and E. Arvas, "FDTD formulation for dispersive chiral media using the Z transform method," *IEEE Trans. Antennas Propagat.*, Vol. 53, 3374–3384, Oct. 2005.
6. Bilotti, F., A. Toscano, and L. Vegni, "FEM-BEM formulation for the analysis of cavity-backed patch antennas on chiral substrates," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, 306–311, 2003.
7. Geng, Y. L., X. B. Wu, and L. W. Li. "A plane electromagnetic wave scattering by a ferrite sphere," *J. of Electromagn. Waves and Appl.*, Vol. 18, No. 2, 161–179, 2004.
8. Kobidze, G. and B. Shanker, "Integral equation based analysis of scattering from 3-D inhomogeneous anisotropic bodies," *IEEE Trans. Antennas Propagat.*, Vol. 52, 2651–2658, Oct. 2004.
9. Hasanovic, M., "Electromagnetic scattering from an arbitrarily shaped three-dimensional inhomogeneous chiral body," Ph.D. Dissertation, Syracuse University, 2006.
10. Schaubert, D. H., D. R. Wilton, and A. W. Glisson, "A tetrahedral modeling method for electromagnetic scattering by arbitrarily shaped inhomogeneous dielectric bodies," *IEEE Trans. Antennas Propagat.*, Vol. AP-32, 77–85, Jan. 1984.
11. Wilton, D. R., S. M. Rao, A. W. Glisson, D. H. Schaubert, O. M. Al-Bundak, and C. M. Butler, "Potential integrals for uniform and linear source distributions on polygonal and polyhedral domains," *IEEE Trans. Antennas Propagat.*, Vol. AP-32, 276–281, Mar. 1984.
12. Graglia, R. D., "Static and dynamic potential integrals for linearly varying source distributions in two- and three-dimensional problems," *IEEE Trans. Antennas Propagat.*, Vol. AP-35, 662–669, June 1987.
13. Hasanovic, M., C. Mei, J. R. Mautz, and E. Arvas, "Scattering from an arbitrarily shaped three-dimensional inhomogeneous magnetic and dielectric scatterer," *22nd Annual Review of Progress in Applied Computational Electromagnetics (ACES 2006)*, Miami, USA, Mar. 2006.
14. Hasanovic, M., C. Mei, J. R. Mautz, and E. Arvas, "Electromagnetic scattering from an arbitrarily shaped three-dimensional inhomogeneous chiral body," *IEEE AP-S International Symposium, USNC/URSI National Radio Science Meeting*, Albuquerque, USA, July 2006.
15. Nie, X. C., N. Yuan, L. W. Li, Y. B. Gan, and T. S. Yeo, "A fast combined field volume integral equation solution to EM scattering by 3-D dielectric objects of arbitrary permittivity and permeability," *IEEE Trans. Antennas Propagat.*, Vol. 54, 961–969, Mar. 2006.

Numerical Analysis of a Virtual Optical Probe Based on Surface Plasmon Polariton

Jiying Xu, Jia Wang, and Boxiong Wang

State Key Laboratory of Precision Measurement Technology and Instruments
Department of Precision Instruments, Tsinghua University
Beijing 100084, China

Abstract— Surface plasmon polariton (SPP) based on metal-dielectric interface can greatly enhance the evanescent field, and so its propagation control is very important. SPP on the flat and structured silver film is numerically studied by 2D finite-difference time-domain method. Confined SPP distributions of two kinds of structured silver films demonstrate a virtual probe with a constant size of 140 nm. The influencing factors, modulation depth and incident spot size, are discussed.

Near-field nanometric light source is of great interest for its super-resolution ability in imaging, data storage, and so on. The concept of a virtual probe is proposed by T. Grosjean [1, 2] to overcome the extreme difficulty of separation regulation between probe and sample. The principle is based on the destructive combination of TM polarized evanescent light beams, and confined intensity distribution can be generated and keep a constant size in a certain range. T. Hong [3] obtained a simulation result of virtual probe by finite-difference time-domain (FDTD) method. E. Descrovi [4] proposed to realize a virtual probe based on the excitation of surface plasmon polariton (SPP) on structured thin metal film. In this paper, by 2D-FDTD method, the excitation of SPP on flat and structured metal film is numerically analyzed, and some influencing factors of virtual probe are studied.

For a flat metal film on a glass substrate, the incident light will be totally reflected at the metal-glass interface when the incident angle is larger than the critical angle. When TM polarized light with an appropriate incident angle matches SPP wavevector, SPP can be resonantly excitated as shown in Fig. 1(a). A collimated Gaussian beam with wavelength 470 nm is incident at 43.2°, and the illuminant spot size is 1 μm. The silver film is 25 nm in thickness and denoted by the green dashed line. The resonant SPP propagates along the silver-air interface and its amplitude decreases in exponential form along vertical direction. However, for TE polarized light under the same conditions, SPP cannot be excitated as shown in Fig. 1(b). Some light penetrates the silver film for it is thin.

When the silver film is modulated by certain structures, SPP and localized SPP can be resonantly excitated. By design of appropriate structure, propagation and combination of SPPs can be controlled to meet some needs. In Fig. 2(a), the modulated structure is a flat surface region with length of 110 nm surrounded by approximate sinusoidal modulation with period of 220 nm and modulation depth of 10 nm (similar to the structure of Ref. 4). When TM polarized Gaussian

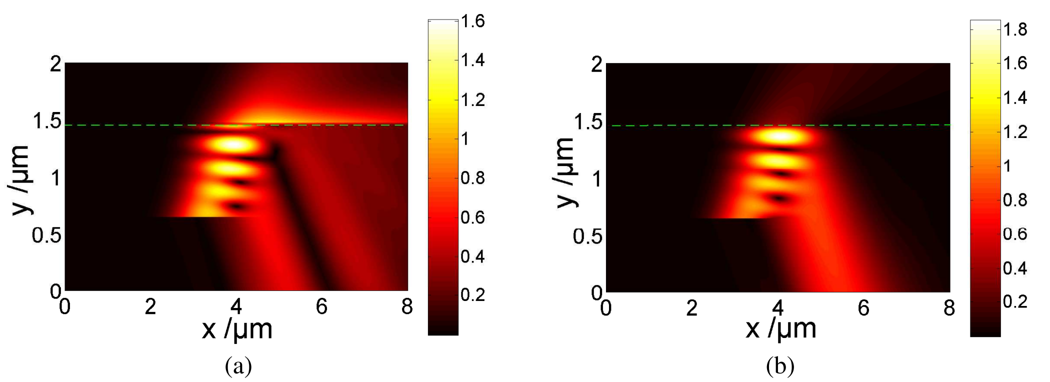
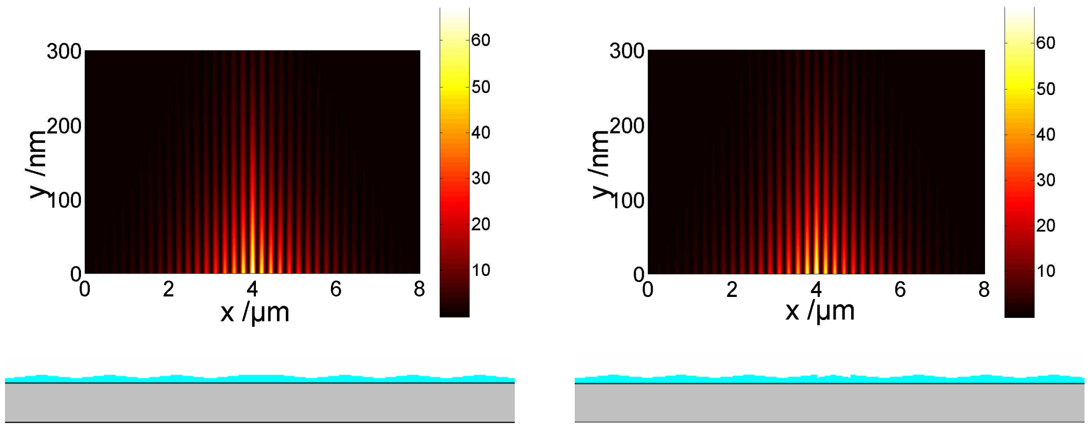


Figure 1: Excitation of SPP by (a) total reflection of TM polarized light, and (b) total reflection of TE polarized light.

wave with 3.5 µm spot size is incident at 45°, the intensity distribution of magnetic field in air is confined to form a central peak with full width at half maximum (FWHM) about 140 nm. The size of the central peak does not increase with distance. In Fig. 2(b), another modulated structure is presented. The central flat surface region of Fig. 2(a) is changed into a tapered tip and the tip is in the same plane of the silver film. It also produces a confined intensity distribution of magnetic field with central peak intensity a little higher than Fig. 2(a).



(a) (b)

Figure 2: Intensity distribution of magnetic field of two different structured silver films: (a) central flat surface region, (b) central tapered tip.

Although the central peak has a maximum intensity, but the intensity of the side lobes cannot be omitted. By increasing modulation depth of surrounded structures, the effect of side lobes can be controlled to a certain extent. As shown in Fig. 3, the field confinement factor of central peak increases with modulation depth, and there is an optimum modulation depth for obtaining maximum magnetic amplitude.

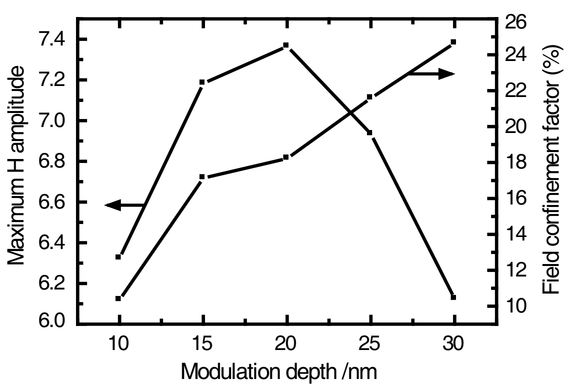


Figure 3: Influence of modulation depth.

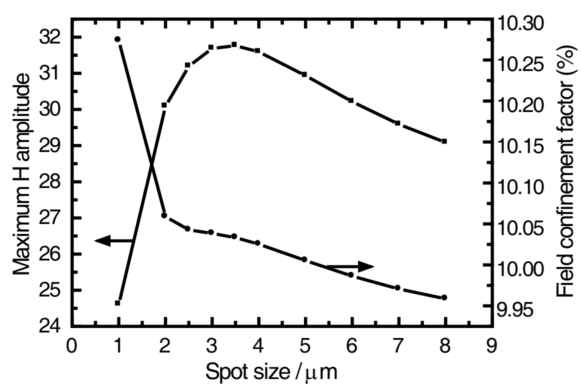


Figure 4: Influence of incident spot size.

The incident spot size also influences the maximum magnetic amplitude and field confinement factor. As shown in Fig. 4, when the spot size of incident collimated Gaussian beam with constant power increases, the field confinement factor slowly decreases; however, the maximum magnetic amplitude varies a lot, and there is an optimum size of 3.5 µm.

In conclusion, the structured metal film can modulate the excitation and propagation of SPP. By an appropriate structure design, a confined evanescent field distribution can be obtained and used as a near-field optical virtual probe. It can be potentially applied in the field of super-resolution optical imaging, near-field high-density optical storage, nano-lithography and so on.

REFERENCES

1. Grosjean, T. and D. Courjon, "Immaterial tip concept by light confinement," *J. Microsc.*, Vol. 202, 273–278, 2000.

2. Grosjean, T., D. Courjon, and D. Van Labeke, "Bessel beams as virtual tips for near-field optics," *J. Microsc.*, Vol. 210, 319–323, 2003.
3. Hong, T., J. Wang, L. Q. Sun, and D. C. Li, "Numerical simulation analysis of a near-field optical virtual probe," *Appl. Phys. Lett.*, Vol. 81, 3452–3454, 2002.
4. Descrovi, E., V. Paeder, L. Vaccaro, and H. P. Herzig, "A virtual optical probe based on localized surface plasmon polaritons," *Opt. Express*, Vol. 13, 7017–7027, 2005.

Directivity of Light Transmission through a Single Subwavelength Aperture without Plasmon Resonance

Evgeni Popov, **Nicolas Bonod**, and **Michel Nevière**
Institut Fresnel, Domaine universitaire de Saint Jérôme, Université d'Aix Marseille III
CNRS UMR 6133, 13397 Marseille Cedex 20, France

Abstract— We present at first a rigorous differential method to resolve Maxwell equations in cylindrical coordinates. Such a method permits the calculation of field amplitudes in both spatial and spectral space in case of subwavelength hole pierced in metallic films with finite conductivity. From the rigorous method, we present a perturbation theory based on the single scattering approximation. It provides analytical formulas in the inverse space of the field amplitude in the spectral space. They lead to an equivalence with the well known model of single magnetic dipole excited inside a pierced layer. Light emitted by this dipole is then radiated to the cladding following the Fresnel's coefficients of light transmission from the layer to the cladding. This approach is in concordance with numerical results obtained from the rigorous method. The theory predicts that the directivity of the radiated pattern increases for smaller values of the layer permittivity modulus, as well as for a metallic layer than for a dielectric one, independently from surface plasmon excitation.

1. INTRODUCTION

Light transmission through a subwavelength hole is of practical importance since subwavelength hole are used to enhance the sensitivity of single molecules detection techniques, with very fruitful applications in chemistry and biology [1]. The theory of Bethe [2] predicts that a small aperture diffracts as if the screen and the aperture were replaced by two emitting electric and magnetic dipoles. In the case of normally incident linearly polarized plane wave, it remains a single magnetic dipole, lying in the plane of the screen and perpendicular to the incident polarization vector. Thus, the diffraction in the plane of polarization (i.e., in the plane perpendicular to the dipole axis) is uniform angularly presenting no directivity, while in the perpendicular plane in follows a simple cos [2] law. Jackson [3] predicts a stronger angular dependence for larger apertures, but his predictions for small holes coincide with Bethe's theory, at least concerning the directivity. However, recent theoretical and experimental works [4–7] indicate that the directivity of the radiation pattern of a single aperture in real metals is larger than the theoretical predictions valid for perfectly conducting screens. These facts indicate the necessity to revisit Bethe's theory in order to obtain better understanding of the process of light diffraction by small apertures in real-metal screens. Based on the rigorous differential method in cylindrical geometry [8], we propose an approach valid in the first-order approximation with respect to R^2. It is used in the study of the directivity of the radiation pattern when varying the screen permittivity. The analytical equations provide a simple physical interpretation, enabling us to demonstrate that for sufficiently large in modulus permittivities ε_2, the diffraction pattern is equivalent to the diffraction pattern of a magnetic dipole excited by the incident field and emitting inside the plane screen. Its radiation is transferred into the cladding by the refraction law guided by Fresnel coefficients. The stronger the contrast between the screen and the cladding, the smaller the angular variation of these coefficients and thus of the diffracted field. In the limit $|\varepsilon_2| \to \infty$, it is demonstrated that the diffraction pattern corresponds to the radiation of a magnetic dipole without the screen confirming Bethe's interpretation [9].

2. DIFFERENTIAL METHOD IN CYLINDRICAL COORDINATES

In cylindrical coordinates (see Fig. 1), any object and field is periodic with respect to the polar angle θ, with period 2π, so that the field components can be represented on the $\exp(in\,\theta)$ basis. The device under study being independent with respect to the variable θ, the Fourier components are mutually independent.

Moreover, it has been previously established on waveguide theory [10] that each Fourier component of the electric field can be expressed as an integral involving Bessel function $J_n(k_r r)$, where k_r is the r-component of the wavevector of each elementary spectral component of the field varying continuously from 0 to ∞. The numerical treatment of the problem requires the discretization and

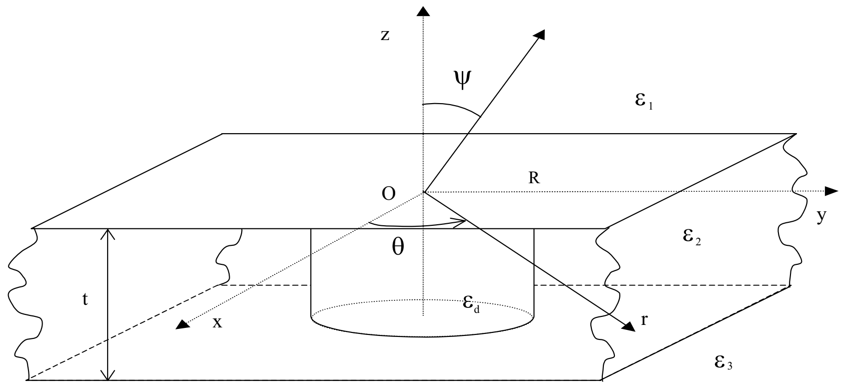


Figure 1.

truncation of the basis so that the field components finally read:

$$\begin{aligned}
E_r(r,\theta,z) &= i\sum_{n=-N}^{N}\sum_{m=1}^{\mathrm{Max}} k_m\Delta k_m\left[b_{n,m}^{E}(z)J_{n+1}(k_m r)-c_{n,m}^{E}(z)J_{n-1}(k_m r)\right]\exp(in\,\theta)\\
E_\theta(r,\theta,z) &= \sum_{n=-N}^{N}\sum_{m=1}^{\mathrm{Max}} k_m\Delta k_m\left[b_{n,m}^{E}(z)J_{n+1}(k_m r)+c_{n,m}^{E}(z)J_{n-1}(k_m r)\right]\exp(in\,\theta)\quad,\\
E_z(r,\theta,z) &= \sum_{n=-N}^{N}\sum_{m=1}^{\mathrm{Max}} k_m\Delta k_m E_{z,n,m}(z)J_n(k_m r)\exp(in\,\theta)
\end{aligned}\tag{1}$$

Similar equations apply to the cylindrical components of $\mathbf{H}$, with $b_{n,m}^{E}$ changed into $b_{n,m}^{H}$, $c_{n,m}^{E}$ into $c_{n,m}^{H}$ and $E_{z,n,m}$ into $H_{z,n,m}$. Using Eqs. (1), Maxwell equations can be reduces into a set of $2N+1$ differential equations:

$$\begin{aligned}
\frac{d}{dz}b_{n,m}^{E} &= b_{n,m}^{H}-\frac{k_m^2}{k_0^2\varepsilon_2}\left(b_{n,m}^{H}-c_{n,m}^{H}\right)-\sum_{m'\neq m}\left(\varepsilon^{-1}\right)_{m,m'}^{n,n}\frac{k_m k_{m'}}{2k_0^2}\left(b_{n,m'}^{H}-c_{n,m'}^{H}\right)\\
\frac{d}{dz}c_{n,m}^{E} &= -c_{n,m}^{H}-\frac{k_m^2}{k_0^2\varepsilon_2}\left(b_{n,m}^{H}-c_{n,m}^{H}\right)-\sum_{m'\neq m}\left(\varepsilon^{-1}\right)_{m,m'}^{n,n}\frac{k_m k_{m'}}{2k_0^2}\left(b_{n,m'}^{H}-c_{n,m'}^{H}\right)\\
\frac{d}{dz}b_{n,m}^{H} &= \frac{k_m^2}{2}\left(b_{n,m}^{E}-c_{n,m}^{E}\right)-k_0^2\varepsilon_2 b_{n,m}^{E}-k_0^2\sum_{m'\neq m}\left(\varepsilon\right)_{m,m'}^{n+1,n+1}b_{n,m'}^{E}\\
\frac{d}{dz}c_{n,m}^{H} &= \frac{k_m^2}{2}\left(b_{n,m}^{E}-c_{n,m}^{E}\right)+k_0^2\varepsilon_2 c_{n,m}^{E}+k_0^2\sum_{m'\neq m}\left(\varepsilon\right)_{m,m'}^{n-1,n-1}c_{n,m'}^{E}
\end{aligned}\tag{2}$$

where the matrix elements of ε are given by:

$$(\varepsilon)_{m,m'}^{n,n}=k_m\Delta_{m'}\int_0^\infty \varepsilon(r)J_n(k_m r)J_n(k_{m'}r)r\,dr.\tag{3}$$

We consider devices being z-independent in a piecewise manner so that the integration is made with the use of the eigenvalue/eigenvector technique. With the knowledge of the incident plane wave, it is then able to calculate each spectral component $b_{n,m}^{E}$, $c_{n,m}^{E}$, $b_{n,m}^{H}$, $c_{n,m}^{H}$ and to reconstruct the electromagnetic field with Eqs. (1).

3. SINGLE-SCATTERING THEORY OF LIGHT DIFFRACTION

Formula (3) can be written as:

$$\varepsilon_2\delta_{m,m'}+(\varepsilon_d-\varepsilon_2)k_m\Delta_{m'}\frac{R}{k_m^2-k_{m'}^2}\left[J_{n+1}(k_m R)J_n(k_{m'}R)k_m-J_n(k_m R)J_{n+1}(k_{m'}R)k_{m'}\right].\tag{4}$$

Taking into account that for small arguments $J_n(k_mR)J_n(k_mR)\sim(k_mR)^{|n|11}$, the off-diagonal terms of $(\varepsilon)^{n,n}$ are then proportional to: $(\varepsilon)_{m,m'}^{n,n}\sim k_m\Delta_{m'}R^{2(|n|+1)}$, $m'\neq m$. In normal incidence, Eq. (4) is simplified into $(\varepsilon_d - \varepsilon_2)\Delta_i R J_1(k_m R)$ with $\Delta_i \equiv \Delta_0$. Working with the sum and the difference of the field components:

$$\begin{aligned} M_m^E &= b_{1,m}^E - c_{1,m}^E, \qquad & P_m^E &= b_{1,m}^E + c_{1,m}^E, \\ M_m^H &= b_{1,m}^H - c_{1,m}^H, \qquad & P_m^H &= b_{1,m}^H + c_{1,m}^H \end{aligned} \tag{5}$$

the set of differential equations are further simplified to take the form corresponding to the unperturbed system:

$$\begin{aligned} \left(M_i^E\right)' &= P_i^H, \qquad & \left(P_i^H\right)' &= -k_{i_z}^2 M_i^E, \\ \left(M_i^H\right)' &= -k_0^2 \varepsilon_2 P_i^E, \qquad & \left(P_i^E\right)' &= \tfrac{k_{i_z}^2}{k_0^2 \varepsilon_2} M_i^H \end{aligned} \tag{6}$$

Eqs. (6) lead to a second-order inhomogeneous equation for M_m^E which has a solution in the form:

$$M_m^E = \mathfrak{M}_m^{E\pm} \exp(\pm i k_{m_z} z) - R J_1(k_m R) \frac{\Delta_\varepsilon}{k_{m_z}^2 - \gamma_i^2} \hat{C}_i^{\pm} \exp(\pm i \gamma_i z) \tag{7}$$

where $\mathfrak{M}_m^{E\pm}$ are two unknown amplitudes, subjected to the boundary conditions between the different domains (cladding/pierced layer/substrate) and the "sources" are given by:

$$\hat{C}_i^{\pm} = C_i^{\pm} \Delta_i = \Delta_i (\mathfrak{P}_i^{E\pm} - \mathrm{M}_i^{E\pm})/2 \tag{8}$$

It is possible to obtain similar solutions for M_m^H while the form of P_m^H and P_m^E is obtained from the expressions for M_m^E and M_m^H. Considering a single interface, the "source" amplitude is obtained with $\mathfrak{P}_i^{E-} = i k_{i_z} \mathfrak{M}_i^{H-}/k_0^2 \varepsilon_2$. Taking into account that there is no incident wave from above when $m \neq i$, the continuity of M_m^E and P_m^H leads to:

$$\alpha_{m_z} \mathfrak{M}_m^{E,R} = -k_{m_z} \mathfrak{M}_m^{E-} + R J_1(k_m R) \frac{\Delta_\varepsilon \gamma_i}{k_{m_z}^2 - \gamma_i^2} \hat{C}_i^{-} \tag{9}$$

which gives for the amplitudes scattered in the upper medium:

$$\mathfrak{M}_m^{E,R} = -\frac{R J_1(k_m R) \Delta_\varepsilon \hat{C}_i^{-}}{(\alpha_{m_z} + k_{m_z})(k_{m_z} + \gamma_i)}. \tag{10}$$

In a similar manner, it is able to obtain:

$$\mathfrak{M}_m^{H,R} = i \frac{k_{m_z}}{\varepsilon_2} \frac{R J_1(k_m R) \Delta_\varepsilon \hat{C}_i^{-}}{(\alpha_{m_z}/\varepsilon_1 + k_{m_z}/\varepsilon_2)(k_{m_z} + \gamma_i)}. \tag{11}$$

4. FAR FIELD DIRECTIVITY

In light diffraction by small apertures, directivity is important to increase the efficiency of detecting in biophysical and physicochemical experiments on laser-induced fluorescence and Raman scattering inside nanovolumes, aiming to study single molecules. The far-field increase of directivity has been largely observed experimentally when using highly conducting metals and excitation of a surface plasmon wave is usually invoked in this phenomenon. However, it is not clear how this plasmon can be radiated from the surface apart from the aperture boards in order to increase the angular directivity of the radiation pattern, if the surface is without defects.

To physically explain the angular dependence of light emmission, we use both the numerical and the analytical method in the case of highly conducting screens. Following figures represent the radial component of the angular density of the Poynting vector, defined as:

$$P_\rho = \frac{1}{2\rho^2} \left(\mathbf{E} \times \bar{\mathbf{H}}\right) . \hat{\rho} \tag{12}$$

where $\rho = \sqrt{r^2 + z^2}$ is the distance between the centre of the aperture and the observation point, and $\hat{\rho}$ is the corresponding unit vector.

Two different metals are considered, real Al and an artificial one having permittivity multiplied by 100. In the plane $\theta = 0$ (see Fig. 1), the higher conductivity presents an angular dependence weaker than for Al screen (hollow squares). Both curves are much wider than in the perpendicular plane, $\theta = 90°$, which has values varying insignificantly with the conductivity, as observed further on. The theory of Bethe predicts this behavior, but more correct formulas are available for greater radii by Jackson (p.492 of ref.):

$$P_\rho \sim \left| \frac{J_1\left(\sin \frac{2\pi}{\lambda} R \sin\psi\right)}{\sin \frac{2\pi}{\lambda} R \sin\psi} \right|^2 \left(\cos^2\psi + \sin^2\psi\cos^2\theta\right) \tag{13}$$

While for the highly conducting material the three methods (Eq. (13)), the analytical and the rigorous method) predict variation of the radiation pattern in the two planes $\theta = 0,\ 90°$, one can observe that the directivity of $P_\rho(\psi)$ increases when $|\varepsilon_2|$ decreases and this fact cannot be explained by surface plasmon whose field decreases away from the aperture. To confirm this hypothesis, Fig. 2 shows results with a dielectric instead of a metallic material. Similar behavior in the plane $\theta = 0$ is obtained, the dependence becomes flatter with the increase of $|\varepsilon_2|$, i.e., for materials with smaller $|\varepsilon_2|$, the directivity is higher. And once again, this fact cannot be explained by a surface wave which is not supported by a single dielectric-dielectric interface.

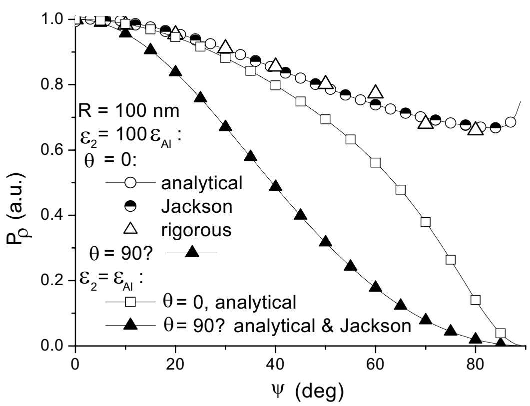


Figure 2.

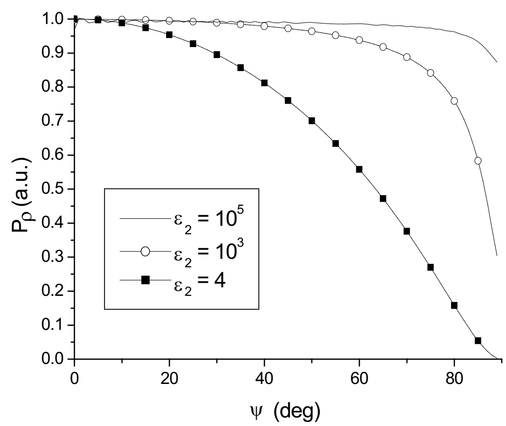


Figure 3.

5. EQUIVALENCE WITH A DIPOLE RADIATION

Let us now explain the physical reasons of the stronger directivity observed when $|\varepsilon_2|$ becomes much larger than ε_d. In the far-field zone $|\varepsilon_2| \gg k_m^2$, so that $k_{m_z} + \gamma_i \approx 2\sqrt{\varepsilon_2}$, and with $J_1(k_m R) \approx k_m R/2$, expressions of $\mathfrak{M}_m^E$ and $\mathfrak{M}_m^H$ in Eqs. (10) and (11) can be simplified into:

$$\mathfrak{M}_m^{E,D_M} \approx k_m \frac{\varepsilon_2}{8|\varepsilon_2|} R^2 \hat{C}_i^- = k_m k_2 \frac{Z_2 D_M}{8\pi} \tag{14}$$

which represents the $\mathfrak{M}_m^{E,D_M}$-component of the field radiated by a magnetic dipole orientated in the y-direction with a dipole moment equal to $D_M = \pi R^2 \frac{\varepsilon_2}{|\varepsilon_2|} \frac{\hat{C}_i^-}{k_2 Z_2}$. It can be demonstrated in a similar manner that:

$$\mathfrak{M}_m^{H,D_M} \approx -ik_m \frac{k_0^2 |\varepsilon_2|}{k_{m_z}} k_2 \frac{Z_2 D_M}{8\pi} \tag{15}$$

represents the $\mathfrak{M}_m^{H,D_M}$-component of the same dipole. The incident field induces a scattered field inside the layer pierced by the aperture. For high values of $|\varepsilon_2|$, the scattered field can be considered as the field of a magnetic dipole having a dipole moment given proportional to the incident field inside the layer equal to $\hat{C}_i^- = \Delta_i(\mathfrak{P}_i^{E-} - \mathfrak{M}_i^{E-})/2$, as following Eq. (8). The emission of the dipole in the plane perpendicular to the dipole direction (i.e., the plane containing the incident electric field vector) is uniform angularly. The field emitted by this dipole is transmitted through the layer surface through the Fresnel transmission coefficients:

$$T_m^{TE,inv} = \frac{2k_{m_z}}{\alpha_{m_z} + k_{m_z}},\ T_m^{TM,inv} = \frac{2k_{m_z}/\varepsilon_2}{\alpha_{m_z}/\varepsilon_1 + k_{m_z}/\varepsilon_2} \tag{16}$$

As they are angularly dependent, the emission in the cladding will depend on the angle of transmission. As a consequence, even in the plane $\theta = 0$ one can expect an angular variation of the emission. But this dependence decreases with increasing the contrast between the cladding and the layer with the aperture, so that one can expect that with $|\varepsilon_2|$ growing up, the radiation patter in the plane $\theta = 0$ will become less depending on the polar angle ψ. And indeed, larger values of $|\varepsilon_2|$ lead to a weaker angular dependence of $\mathfrak{P}_m^{E,R}/k_m$, for example. By taking the asymptotic expressions of the transmission coefficients and multiplying them by the dipole field components, it is easily found that the components of the field radiated in the cladding are equal to:

$$\mathfrak{M}_m^{E,R} \simeq \frac{k_m}{4} R^2 \frac{\varepsilon_2 \hat{C}_i^-}{|\varepsilon_2|}, \qquad \mathfrak{P}_m^{H,R} \simeq i\alpha_{m_z} \frac{k_m}{4} R^2 \frac{\varepsilon_2 \hat{C}_i^-}{|\varepsilon_2|} \tag{17}$$

$$\mathfrak{M}_m^{H,R} \approx -i \frac{k_0^2 \varepsilon_1}{\alpha_{m_z}} \frac{k_m}{4} R^2 \frac{\varepsilon_2 \hat{C}_i^-}{|\varepsilon_2|}, \qquad \mathfrak{P}_m^{E,R} \simeq -\frac{k_m}{4} R^2 \frac{\varepsilon_2 \hat{C}_i^-}{|\varepsilon_2|} \tag{18}$$

which are nothing but the field components of a magnetic dipole emitting in the cladding as if the substrate were absent. These expressions represent the equivalence between the scattering by an aperture in an infinitely conducting screen in normal incidence and a radiation pattern of a magnetic dipole.

REFERENCES

1. Levene, M. J., J. Kotach, S. Turner, M. Foquet, H. G. Craighead, and W. W. Webb, “Zero-mode waveguide for single-molecule analysis at high concentrations,” *Science*, Vol. 209, 682–686, 2003.
2. Bethe, H. A., “Theory of diffraction by small holes,” *Phys. Rev.*, Vol. 66, 163–182, 1944.
3. Jackson, J. D., *Classical Electrodynamics*, 3rd Ed., John Wiley, New York, 1998.
4. Degiron, A., H. J. Lezec, N. Yamamoto, and T. W. Ebbesen, “Optical transmission properties of a single subwavelength aperture in a real metal,” *Opt. Commun.*, Vol. 239, 61–66, 2004.
5. Yin, L., V. K. Vlasko-Vlasov, A. Rydh, J. Pearson, U. Welp, S. H. Chang, S. K. Gray, G. C. Schatz, D. B. Brown, and C. W. Kimball, “Surface plasmons at single nanoholes in Au films,” *App. Phys. Lett.*, Vol. 85, 467–469, 2004.
6. Zakharian, R., M. Mansuripur, and J. V. Moloney, “Transmission of ligth through small elliptical apertures,” *Opt. Express*, Vol. 12, 2631–2648, 2004.
7. Wanneracher, R., “Plasmon-supported transmission of light through nanometric holes in metallic thin films,” *Opt. Commun.*, Vol. 195, 107–118, 2001.
8. Bonod, N., E. Popov, and M. Nevière, “Differential theory of diffraction by finite cylindrical objects,” *J. Opt. Soc. Am. A*, Vol. 22, 481–490, 2005.
9. Popov, E., M. Nevière, A. Sentenac, N. Bonod, A-L. Fehrembach, J. Wenger, P-F. Lenne, and H. Rigneault, “Single-scattering theory of light diffraction by a circular subwavelength aperture in a finitely conducting screen,” *J. Opt. Soc. Am. A*, in press.
10. Snyder, A. and J. D. Love, *Optical Waveguide Theory*, 250, Chapman et Hall, London, New York, Tokyo, Melbourne, Madras, 1983.
11. Abramovitz, M. and I. A. Stegun, (eds.), *Handbook of Mathematical Functions [9th printing (1970)] National Bureau of Standards Applied Mathematics Series 55*, US Gov't Printing Office, Washington, DC, 1964.

Scattering by a Finite Strip Under Complex Beam Incidence—Asymptotic Evaluation in the Complex Space Domain

M. J. González-Morales[1] and E. Gago-Ribas[2]

[1]University of Valladoid, Spain
[2]University of Oviedo, Spain

Abstract— This communication deals with the resolution of electromagnetic scattering problems by applying complex spaces. The particular problem presented is the scattering by a finite strip under complex beam incidence. It is resolved by the asymptotic evaluation of the radiation integral in the complex space domain. In addition, the outlook for the possibilities of this technique is presented: a summary of proposed problems so far, the potential and limitations.

1. INTRODUCTION

Some classic techniques for EM problems resolution are based on complex variable. For instance, conformal mapping between complex planes to solve problems involving the Laplace equation, or complex contour path integration to obtain Green's function for the wave equation. The authors have been concentrated on understanding the meaning of the technique of *analytical continuation* of wave solutions, the complex quantities associated to it and its application to EM problems, [1]. The starting point is the analytical continuation of the Greens function for the 2D complex point source, which was proposed to obtain a Gaussian beam mathematical formulation, [2, 3]. This technique has been use by decades and it is still being used, [4].

Some steps have been already made satisfactorily by the authors in this matter: The rigorous characterization of (i) the *complex distances* appearing when the source is located at a complex position, [5] (ii) the *complex angles* which are defined when any observation point is related to that complex source, [6]. This constitutes a duo distances-angles which allows the real observation space characterization in terms of complex spaces.

This mathematical formulation, apparently artificial, makes sense when it is applied to the 2D radiation problem of the complex point source, that is, the analysis of the *complex beams*, [7]. One of the most relevant aspect of this technique, which makes it so powerful, is its capability to manage exact solutions and to parameterize the different approximations and their validity ranges. Specifically, complex beams include plane waves, cylindrical waves and Gaussian beams as particular cases contained in its formulation.

In addition, another success achieved is the spectral characterization of the complex beams, [8]. This opens a whole world of possibilities, inasmuch all the resolution techniques in the spectral domain maybe applied to complex beams. Even when finding the complex space domain possibilities has meant a great newness, concerning to the way of thinking and working, we do not expect, in the spectral domain, a so revolutionary results, as long as the spectral domain is complex itself, and a lot of problems resolved some time ago, [9], did already operate with the complex spectral domain means.

Coming to the scattering problems, we have already posed and solved, (wholly or partially), scattering problems evaluated by asymptotic techniques in the complex space domain. The starting point was similar problems formulated in the real space domain, [10, 11]. Specifically, starting from the scattering by a strip or an array of strips under a different homogeneous field illuminations. We have proposed the scattering by a strip or an array of strips under some non-homogeneous fields illuminations, [12]. The first incident fields we considered were non-homogeneous plane waves and Gaussian beams. Our contribution to these problems is no relevant, as long as they may be resolved by using the classical techniques, ([9, 11]). The reason considering these scenarios is to calibrate the method and the results. The main subject of this communication is the scattering by a finite strip under complex beam incidence which will be analyzed in next sections.

In the light of the scenarios considered, [12], we can assert that the complex space formulations are very appropriate to analyze problems concerning to non-homogeneous fields, because they are included in a natural way inside the formulation. The disadvantage when combining with the asymptotic evaluation of integrals is that the results are approximated at the end. The accuracy is not lost but the attractive characteristic of exact solution is lost. We wonder if we will be able to apply some exact procedure (a Sommerfeld-style) to non trivial scattering problems involving complex beams.

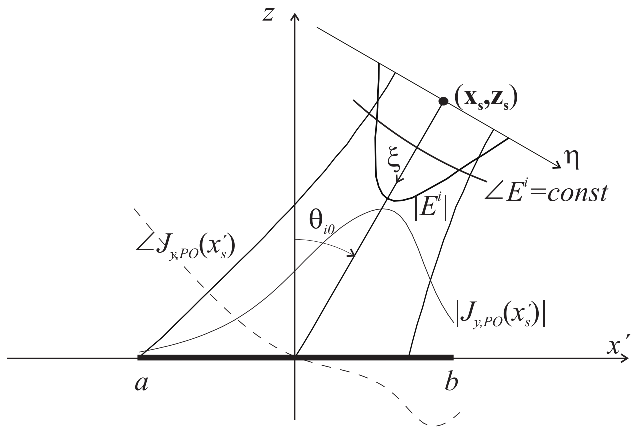


Figure 1: Scattering by a finite perfect conductor strip under complex beam incidence. Incident field and induced current profiles.

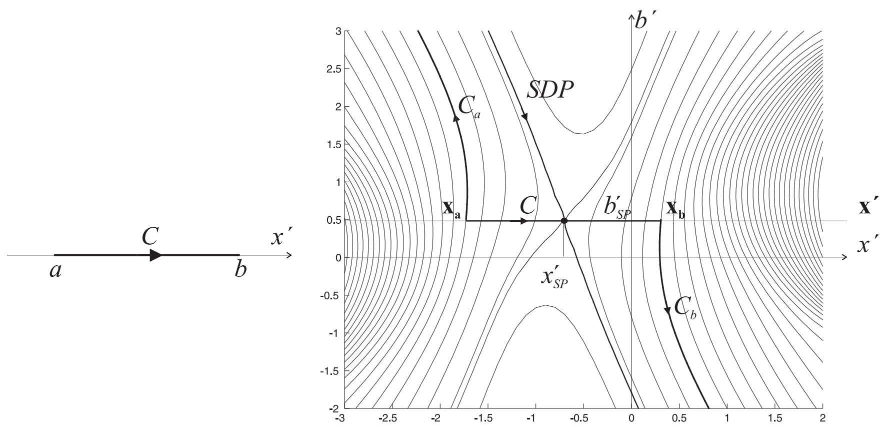


Figure 2: Analytical continuation of the radiation integral into the complex space domain. The integration C path is deformed into the SDP.

Disclosing one of the main conclusions of this paper, we can say that the asymptotic techniques together with the complex space analysis opens the possibility to extend the results, from 2D scattering problems involving cylindrical waves to new problems involving complex beams.

2. SCATTERING BY A FINITE STRIP UNDER COMPLEX BEAM INCIDENCE

The scenario is described in Fig. 1. The scatterer is a perfect electric conductor strip located on x axis and infinite on y direction. The incident field is radiated by a 2D point source located at (x_s, z_s). The analytical continuation from the real coordinates of the source into the complex space ($\mathbf{x}_s = x_s + ib\sin\theta_{i0}$, $\mathbf{z}_s = z_s + ib\cos\theta_{i0}$) gives the representation for the incident complex beam, (θ_{i0} is related to the beam axis direction and b is related to the waist width),

$$\vec{E}^i = E_y^i \hat{y} = \frac{-\omega\mu_0 I e^{-i\pi/4}}{2\sqrt{2\pi}} \frac{e^{ik_0\mathbf{R}_i}}{\sqrt{k_0\mathbf{R}_0}}\hat{y}; \quad \mathbf{R_i} = \sqrt{(x-\mathbf{x_s})^2 + (z-\mathbf{z_s})^2}. \tag{1}$$

The usual physical optics approximation is applied to obtain the induced currents on the perfect electric conductor strip, $\vec{J}_{PO}(x', z'=0) = 2\hat{z} \times \vec{H}^i = J_{y,PO}\hat{y}$,

$$J_{y,PO}(x', z'=0) \sim \frac{-2}{i\omega\mu_0} \frac{ik_0(-\mathbf{z}_s)}{\sqrt{k_0\mathbf{R_i}^3(x, z=0)}} e^{ik_0\mathbf{R}_i(x,z=0)}. \tag{2}$$

An example of the amplitude and phase profiles of the induced currents on the real strip points is represented in Fig. 1.

The extension from the real coordinates of the strip into the space of complex coordinates is also performed,

$$J_{y,PO}(x', z') = \frac{-2}{i\omega\mu_0} \frac{ik_0(\mathbf{z'} - \mathbf{z_s})}{\sqrt{k_0 \mathbf{R_i}}^3} e^{ik_0 \mathbf{R}_i}. \tag{3}$$

The Green's function of complex arguments continues also being valid solution to the wave equation,

$$G(\vec{r}, \vec{\mathbf{r}}') = \frac{i}{4} H_0^{(1)}(k_0 \mathbf{R_o}) \sim \frac{e^{i\pi/4}}{2\sqrt{2\pi k_0}} \frac{e^{ik_0 \mathbf{R_o}}}{\sqrt{\mathbf{R_o}}}; \quad \mathbf{R_o} = \sqrt{(x - \mathbf{x}')^2 + (z - \mathbf{z}')^2}. \tag{4}$$

The radiation integral, with the extended Green's function and the extended currents on the strip complex locations, is given by,

$$E_y^s(x, z) = i\omega\mu_0 \int_{\mathbf{x_a}}^{\mathbf{x_b}} J_y(x') G(x, z; \mathbf{x}', \mathbf{z}') d\mathbf{x}' = \frac{-e^{i\pi/4}}{\sqrt{2\pi k_0}} \int_{\mathbf{x_a}}^{\mathbf{x_b}} \frac{ik_0 \cos \boldsymbol{\theta_i} e^{ik_0 \mathbf{R_i}} e^{ik_0 \mathbf{R_o}}}{\sqrt{\mathbf{R_i}} \sqrt{\mathbf{R_o}}} d\mathbf{x}', \tag{5}$$

where the trigonometric function of a complex angle has been introduced, $\cos \boldsymbol{\theta_i} = (\mathbf{z'} - \mathbf{z_s})/\mathbf{R_i}$.

3. ASYMPTOTIC ANALYSIS IN THE COMPLEX SPACE DOMAIN

The radiation integral,

$$E_y^s(x, z) = \int_{\mathbf{x_a}}^{\mathbf{x_b}} f(x') e^{k_0 q(x')} dx', \tag{6}$$

may be *asymptotically* evaluated, [13], in the complex space domain, Fig. 2.

$$\int_{\mathbf{x_a}}^{\mathbf{x_b}} = \int_{SDP} + \int_{C_a} - \int_{C_b} = E_{SP} + E_a + E_b. \tag{7}$$

The integral value is made up of three contributions: one coming from the complex saddle point and two coming from the endpoints of the strip.

3.1. Complex Saddle Point

Applying the so called Saddle Point Equation (SPE), $q'(x'_s) = 0$, the saddle point (SP), $(x'_s, \boldsymbol{z'_s})$ is obtained. It is worth mentioning that both $dq(x', \boldsymbol{z'})/dx' = 0$ and $dq(x', \boldsymbol{z'})/d\boldsymbol{z'} = 0$ lead to the same condition,

$$x'_s = \frac{x\mathbf{R_i} + \mathbf{R_o} x_s}{\mathbf{R_i} + \mathbf{R_o}}. \tag{8}$$

The SP is a complex point belonging to the complex straight line which goes from the complex image source $(x_s, -\boldsymbol{z_s})$ to the real observation point (x, z).

$$x'_s = \frac{xz_s + x_s z}{z + z_s}; \quad b' = b \frac{\sin\theta''(x - x'_s) + \cos\theta'' z}{(x - x_s)\sin\theta'' + (z + z_s)\cos\theta''}. \tag{9}$$

In addition, $z'_s = 0$ and $\Im\{\mathbf{z'_s}\}$ is related to $\Im\{\mathbf{x'_s}\}$ by b' parameter.

3.2. Complex Snell Law

Applying the SPE, the Snell law is given by $\sin \boldsymbol{\theta_i} = -\sin \boldsymbol{\theta_o}$. This law is interpreted by means of the mathematical base previously established, [6], in particular, the results concerning to complex angles. Basically, the real part of the complex angle gives information about the propagation direction and the imaginary part is related to the attenuation at right angles to the propagation direction. Once the equation is separated into its real and imaginary parts, the real part gives the usual direction rule for the reflected rays, that is, the usual real Snell law; and the imaginary part relates the attenuation for each saddle point, and accordingly, for each associated observation point.

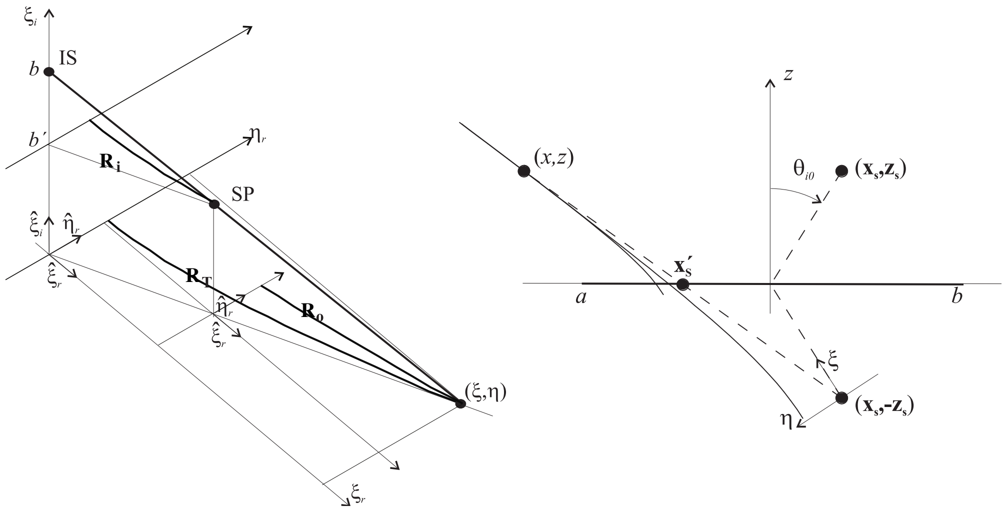


Figure 3: A complex ray (left) is a straight line in the complex space. It goes from the image source (IS), passes through the saddle point (SP) and crosses the real space in the real observation point (ξ, η) (beam adapted coordinates have been used). The complex ray is projected as hyperboles in the real propagation space (right).

3.3. Complex Rays

The complex rays must be 'projected' into the real propagation space leading to some results which may not be found with the conventional analysis with real variable. A complex ray, which follows a straight line in the complex space, becomes a real *curved line* in the real space. This real space interpretation is found out by making hyperbole matching as in the Evanescent Wave Theory, [14]. The comparison between complex rays coming from both the SP and the image source, and reaching a real observation point, are shown in Fig. 3, together with its local matching with hyperbolas.

3.4. Scattered Field

By applying the asymptotic evaluation of integrals results to the radiation integral, [13], the scattered field is obtained. The SP contribution gives the *reflected field* and may be seen as the field radiated by the image point source $(x_s, -\boldsymbol{z_s})$,

$$E_{SP} = \frac{-e^{ik_0(\mathbf{R_i}+\mathbf{R_o})}}{\sqrt{k_0(\mathbf{R_i}+\mathbf{R_o})}} \tag{10}$$

The complex endpoint contributions give the *diffracted field* by the endpoints of the strip,

$$E_{a,b} = \frac{\mp e^{-i\frac{\pi}{4}}}{\sqrt{2\pi}} \frac{\cos\boldsymbol{\theta_i}(a,b)}{\sqrt{\sin\boldsymbol{\theta_i}(a,b)+\sin\boldsymbol{\theta_o}(a,b)}} \frac{e^{k_0\mathbf{R_i}(a,b)}}{\sqrt{\mathbf{R_i}(a,b)}} \frac{e^{k_0\mathbf{R_o}(a,b)}}{\sqrt{\mathbf{R_o}(a,b)}} \tag{11}$$

4. CONCLUDING REMARKS

The scattering by a finite strip under complex beam incidence has been presented. It has been resolved by a new method based on the complex space domain. Both, the problem and the solution presented here, are a generalization of the same scattering problem when the strip is illuminated by some homogeneous waves, as long as the complex beams contain the parametrization of plane waves, cylindrical waves and Gaussian beams, which appear as particular cases included in the formulation.

The complex space analysis assumes to perform the extension of the induced currents on the scatter, from its real position into complex locations. When applying asymptotic techniques, the saddle points are straight complex. Once the mathematical analysis is performed, the most interesting point is the interpretation of the results in the real space. The interpretation of the Snell law requires the results of the complex angles analysis. Basically, the real part of the gives information about the propagation direction and the imaginary part about the attenuation, for each saddle point and its associated observation point. One of the most interesting points is the interpretation of the complex rays, which correspond to curve tracks in the real space.

The formulation provides an analytical expression for the scattered field solution. The interpretation is made in terms of a reflected field and two scattered terms associated to the strip ends. The form of the scattered field fits with the expressions corresponding to a cylindrical incident field, by replacing the real quantities with the complex ones. In spite of this, the content must be carefully evaluated, so long as the diffraction factors and all their terms include complex quantities.

The results lead to think that the solution to some scattering problems in which cylindrical waves are involved, may be generalized by considering an incident complex beam. The difficulty and, at the same time, the interest consist in the interpretation of the quantities which appear in the real propagation space.

ACKNOWLEDGMENT

This work has been supported by the Spanish CICYT under grants TEC2005-04389 and TEC2006-12254-C02-02/TCM, and by de Junta de Castilla y León under grant VA007B05.

REFERENCES

1. González-Morales, M. J., C. Dehesa-Martónez, and E. Gago-Ribas, "About complex extensions and their application in electromagnetics," in I. C. Goknar, L. Sevgi (Eds.), *Complex Computing-networks. A link Between Brain-like and Wave-oriented Electrodynamic Algorithms, Springer Proceedings in Physics*, Vol. 104, 81–86, Springer-Verlag, 2006.
2. Deschamps, G. A., "Gaussian beam as a bundle of complex rays," *Electron. Lett.*, Vol. 7, 684–685, 1971.
3. Felsen, L. B., "Complex Rays," *Philips Res. Repts.*, Vol. 30, 187–195, 1975.
4. Martini, E., G. Pelosi, and S. Selleri, "Line integral representation of physical optics scattering from a perfectly conducting plate illuminated by a gaussian beam modeled as a complex point source," *IEEE Tr. on AP*, Vol. 51, No. 10, 2003.
5. Gago-Ribas, E., M. J. González-Morales, and C. Dehesa-Martínez, "Analytical parametrization of a 2D real propagation space in terms of complex electromagnetic beams," *IEICE Trans. on Electronics: Special Issue on Electromagnetic Theory — Scattering and Diffraction*, Vol. E80-C, No. 11, 1434–1439, Japan, November 1997.
6. Gago-Ribas, E. and M. J. González-Morales, "2D complex point source radiation problem. I. complex distances and complex angles," *Turkish Journal of Electric Engineering and Computer Sciences: Special Issue on Electromagnetic Problems and Numerical Simulation Techniques: Current Status — Future Trends*, Vol. 10, No. 2, 317–343, 2002.
7. González-Morales, M. J. and E. Gago-Ribas, "2D complex point source radiation problem. II. complex beams," *Turkish Journal of Electric Engineering and Computer Sciences: Special Issue on Electromagnetic Problems and Numerical Simulation Techniques: Current Status — Future Trends*, Vol. 10, No. 2, 345–369, 2002.
8. Mahillo-Isla, R., M. J. González-Morales, and C. Dehesa Martánez, "Spectral considerations about 2D paraxial and non-paraxial beam solutions," *Proceedings of PIERS 2006*, Tokio, Japan, August 2006.
9. Lu, I. T., L. B. Felsen, and Y. Z. Ruan, "Spectral aspects of the gaussian beam method: reflection from a homogeneous half-space," *Geophys. J. R. Astr. Soc.*, 915–932, 1987.
10. Gago, E. and L. B. Felsen, *Analitical Study of Finite Strips.(Asymptotic Analysis)*, Weber Research Institute. Polytechnic University of Brooklyn. NY, November 1992.
11. Felsen, L. B. and E. Gago-Ribas, "Ray theory for scattering by two-dimensional quasiperiodic plane finite arrays," *IEEE Trans. on Ant. and Prop.*, Vol. 44, 375–382. 1996.
12. González-Morales, M. J., *Espacios Complejos. Aplicación a Problemas de Radiación y Dispersion Electromagnetica de Haces*, Ph. D. Dissertation, Universidad de Valladolid, Spain, July 2001.
13. Felsen, L. B. and N. Marcuvitz, *Radiation and Scattering of Waves*, Prentice-Hall 1973. Reissued by IEEE Press, 1974.
14. Felsen, L. B., "Evanescent waves," *J. Opt. Soc. Am.*, Vol. 66, No. 8, 751–760, 1976.

Transient Scattering from Arbitrarily Shaped Two-dimensional Objects Located on a Rough Surface

Min Zhang[1], **Hongcheng Yin**[2], **and Zhen Cao**[1]

[1]School of Science, Xidian University, Box 270, Xi'an 710071, China
[2]National Electromagnetic Scattering Laboratory, Beijing 100854, China

Abstract— An electric field time-domain integral equation (TDIE) for arbitrarily two-dimensional bodies located on a time-evolving rough surface is presented. The explicit solution technique of the TDIE is considered, and the time responses of the current distributions and the far-zone field are investigated in detail, meanwhile, the method of moment associated with the inverse discrete Fourier transformation (MoM-IDFT) scheme is also introduced to depict the accuracy of the presented technique for the transient scattering by arbitrarily objects integrated with a rough surface.

Time-domain integral equation (TDIE) techniques possess a number of advantages when used to analyze wideband or nonlinear or time-varying electromagnetic scattering and radiation phenomena [1, 2, 4–6]. During the past decade, many different TDIE approaches have been proposed to model transient scattering by conducting or dielectric bodies, but only few scheme is presented to solve problems of electromagnetic scattering by an object located on a rough surface. Actually, most of previous studies on such a complicated scattering problem have considered the scattering from a target on a single rough surface realization as a function of the aspect angle in term with Monte Carlo methods, or have considered the scattering from target over a rough surface use a four-path model based on the coherent reflection coefficient of the rough surface [3]. It is hard, however, to handle the time-dependence of the scattering by an object located on a realistic time-evolving rough surface (such as a sea surface) in those models. In this paper, the numerical models of transient scattering from one-dimensional (1-D) rough surface and the surface with arbitrarily two-dimensional (2-D) PEC bodies are analyzed. A novel explicit algorithm for the calculations of transient far-fields using TDIE is presented and the numerical results are also compared with MoM-IDFT, to prove the better effectiveness and accuracy of this technique for transient scattering from complex rough surfaces associated with the targets.

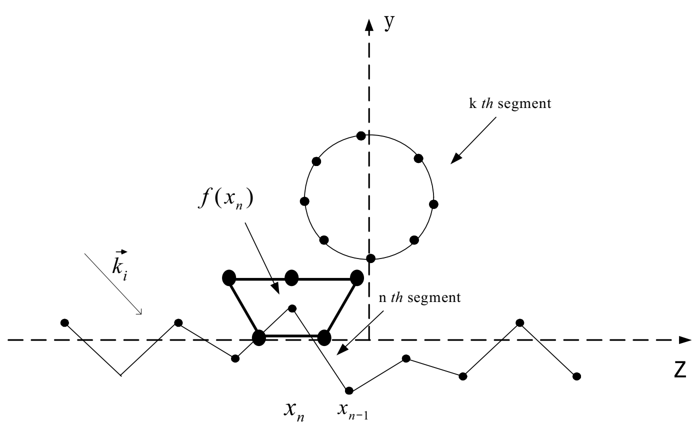


Figure 1: The geometry of a target situated above a rough surface.

We assume that arbitrarily 2-D PEC bodies are located on 1-D rough surface shown in Fig. 1. The boundary condition of zero tangential electrical field is enforced on both target and rough surfaces, so that we have the time-domain electrical field integral equation as

$$\hat{n} \times \left[\vec{E}^{inc}(\vec{r}, t) - \vec{E}^{sca}(\vec{r}, t) \right] = 0 \tag{1}$$

For the TDIE formulation of 2-D object, the scattered field is

$$\vec{E}^{sca}(\vec{r}, t) = -\frac{\partial \vec{A}}{\partial t} - \nabla \varphi \tag{2}$$

where, $\vec{A}(\vec{\rho}, t) = \frac{\mu}{4\pi}\int_S \frac{1}{R}\vec{J}\left(\vec{\rho}', t - \frac{R}{c}\right) ds'$, where $\vec{J}(\vec{\rho}', t)$ denotes the surface current density but the line current density, this is the major difference between the TDIE method with method of moments (MoM). For a TM incident field, the electric field only has a component in the z directed, and $\nabla\varphi = 0$.

We adopt the pulse basis function and the Galerkins method of moments approach, Equation (2) could be integrated as

$$A(\vec{\rho}_m, t) = \int_0^t E^{inc}(\vec{\rho}_m, \tau)\, d\tau = \frac{\mu}{4\pi}\int_C \int_{z'=-\infty}^{\infty} \frac{\sum\limits_{n=1}^{N} I_n(t_j - R_m/c)\, j_n(\vec{\rho}')}{R_m} dz' dc' \tag{3}$$

here I_n is the unknown coeffiecient at the nth zone. The current in a patch does not vary appreciably with time is assumed, and then A could be rewritten as

$$A(\vec{\rho}_m, t_j) = \frac{\mu}{4\pi} F_{mm0} I_m(t_j) + \frac{\mu}{4\pi}\sum_{n=1}^{N}\sum_{k=-\infty}^{\infty} I_n\left(t_j - \frac{R_{mnk}}{c}\right) F_{mnk}\Big|_{m\neq n \text{ or } k\neq 0} \tag{4}$$

Finally, the current I_n is computed by the explicit scheme as

$$I_m(t_j) = \left[4\pi \int_0^{t_j} E^{inc}(\vec{\rho}, \tau)\, d\tau - \tilde{A}(\vec{\rho}_m, t_j)\right] / \mu F_{mm0} \tag{5}$$

where $\tilde{A}$ is the second term of the right-hand side in Eq. (4), $R_{mnk} = \sqrt{|\vec{\rho}_m - \vec{\rho}_n'|^2 + (k\Delta\tau_n)^2}$, $\vec{\rho}_m$ and $\vec{\rho}_n'$ are the center of mth and nth segments, respectively.

In order to keep the continuity of surface current in calculation region, we adopt the taper Gaussian impulse wave with a window function $\exp\left(-\left[\frac{g}{l}(\vec{\rho}\cdot\hat{x} - x_0)\right]^2\right)$,

$$\vec{E}^{inc}(\vec{\rho}, t) = \hat{z}E_0 \frac{4.0}{\sqrt{\pi}T}\exp\left(-\left[\frac{4.0}{T}\left(ct - ct_0 - \vec{\rho}\cdot\hat{k}\right)\right]^2\right)\cdot\exp\left(-\left[\frac{g}{l}(\vec{\rho}\cdot\hat{x} - x_0)\right]^2\right) \tag{6}$$

where g is the attenuation factor, to keep incident field be zero at the edge of the rough surface.

Correspondingly, the far field could be represented as

$$\sqrt{\rho}\vec{E}_f(\vec{\rho}, t) = -\sqrt{\rho}\frac{\partial \vec{A}_f(\vec{\rho}, t)}{\partial t} - \frac{1}{\varepsilon}\sqrt{\rho}\nabla\times\vec{F}_f(\vec{\rho}, t) \tag{7}$$

$$\vec{A}_f(\vec{\rho}, t_j) = \frac{\mu}{2\sqrt{2\rho}\pi}\sum_{n=1}^{N}\Delta l_n \sum_{k=1}^{j-\tau_n} -2\vec{J}_n(t_k)\left(\sqrt{a - ct_k} - \sqrt{a - ct_{k-1}}\right) \tag{8}$$

$$\sqrt{\rho}\nabla\times\vec{F}_f(\vec{\rho}, t_j) = \frac{\varepsilon}{2\sqrt{2}}\sum_{n=1}^{N}\Delta l_n \sum_{k=1}^{j-\tau_n} -\vec{J}_{Mn}(t_k)\times\hat{\rho}\left(\frac{1}{\sqrt{a - ct_k}} - \frac{1}{\sqrt{a - ct_{k-1}}}\right) \tag{9}$$

where $\tau_n = (\rho - \hat{\rho}\cdot\vec{\rho}_n)/c$, $a = ct_j - \rho + \hat{\rho}\cdot\vec{\rho}_n'$, $t_k = \Delta t\cdot k$ is the kth time step, $\vec{J}_n(t_k)$ and $\vec{J}_{Mn}(t_k)$ are the electrical and magnetic currents.

In this study, the surface is meshed by N points, the spatial step size is Δx, the total length of rough surface is $L = N\Delta x$, $x_n = n\Delta x\, (n = 1, 2, \ldots, N)$ and each segment is considered as an infinite conducting strip in three-dimensional space, in order to establish the electrical field integral equations by using TDIE method.

In all our computations, the case of TM polarization is considered, where the parameters $E_0 = 377\,\text{V/m}$, $T = 4\,\text{LM}$, $ct_0 = 6\,\text{LM}$, $L = 25.6\,\text{m}$, $N = 256$, $\Delta t = 0.05\,\text{LM}$ and $g = 5.4$. The unit LM denotes a light meter, so a light meter is the length of time taken by the electromagnetic wave to travel 1 m in the free space.

Figure 2 shows the transient scattering characteristics of a 2-D cylinder located on a gauss rough surface which is shown in Fig. 1, where Fig. 2(a) shows time response of the current at rough surface center point, while Figs. 2(b) and 2(c) dipect the responses from the far fields at forward direction

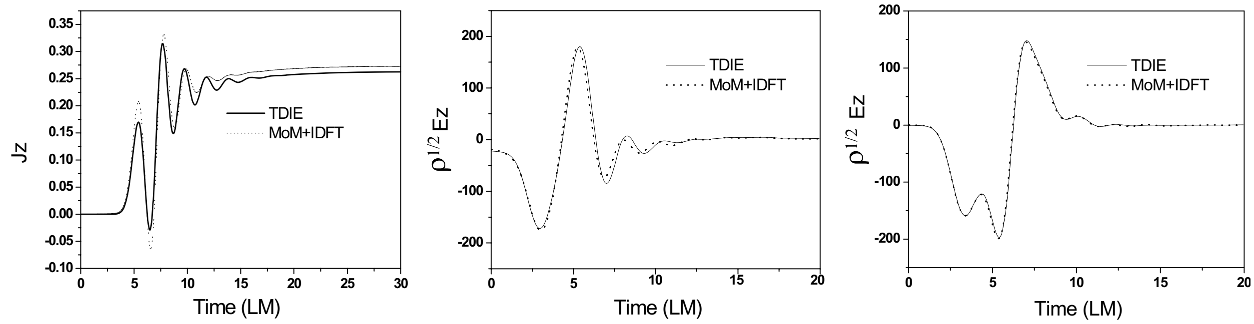


Figure 2: Transient scattering characteristics of a 2-D PEC cylinder with a rough surface.

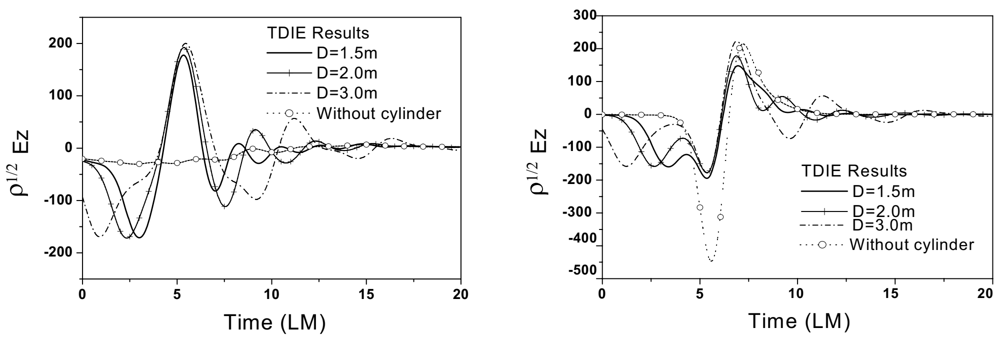


Figure 3: Time response effects of the object located on different height of the surface.

and back direction, respectively. In calculation we choose that correlation length is $l = 0.4\,\mathrm{m}$, rms height is $h = 0.05\,\mathrm{m}$, the radium of the cylinder is 0.5 m, the distance between the cylinder with the rough surface rface is $D = 1.5\,\mathrm{m}$, and the cylinder circle is divided by 30 segments. To test the validity of this approach, we compare the results obtained by using presented method with those using the MoM-IDFT method, in the case of the MoM+IDFT, the frequency ranges are cover 360 MHz with 1000 sample points and the comparison is also depicted in Fig. 2. The TDIE approach works as well as the MoM+IDFT. Actually, it is well known, the MoM+IDFT is a frequency-domain technique with a very time consuming on the wide-band information, but a direct time-domain (TDIE) solution is more efficient for the transient scattering problems. To further characterize the scattering effects of the target located on different position of the rough surface, the time responses of far fields at different height D and even without the cylinder are compared in Fig. 3. It is found that the back scattering has been obviously enhanced when the object is located over the rough surface.

In conclusion, the numerical investigation of electromagnetic scattering by a 2-D target located on a time-evolving rough surface is carried out in this paper. The time responses of the current distribution and far-zone field are computed using the TDIE approach. The solution results indicate the different scattering characteristics between a rough surface with the objects located on the same suface.

REFERENCES

1. Rynne, B. P., "Time domain scattering from arbitrary surfaces using the electric field integral equation," *Electromagnetic Waves Appl.*, Vol. 5, 93–112, 1991.
2. Rao, S. M. and T. K. Sarkar, "Time domain modeling of two dimensional conducting cylinders utilizing an implicit scheme — TM incident," *Microwave Opt. Technol. Lett.*, Vol. 15, No. 6, 342–347, 1997.
3. Burkholder, R. J., M. R. Pino, and F. Obelleiro, "Low angle scattering from 2-D Targets on a time-evolving sea surface," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. GRS-40, 1185–1190, 2002.
4. Sadigh, A. and E. Arvas, "Treating the instabilities in marching-on-in-time method from a different perspective," *IEEE Transactions on Antennas and Propagation*, Vol. AP-41, No. 12, 1695–1702, 1993.

5. Rynne, B. P. and P. D. Smith, "Stability of time marching algorithms for the electric field integral equation," *J. Electromagn. Waves Appl.*, Vol. 4, No. 12, 1181–1205, 1990.
6. Thorsos, E. I., "The validity of the Kirchhoff approximation for rough surface scattering using a Gaussian roughness spectrum," *J. Acoust. Soc. Am.*, Vol. 83, 78–92, 1978.

ISAR Image Modeling of Aircraft with Propeller Blades

Xiaojian Xu and Yao Qin
School of Electronic and Information Engineering, BeiHang University
37 XueYuan Road, Beijing, 100083, China

Abstract— Modeling of the electromagnetic (EM) scattering and inverse synthetic aperture radar (ISAR) images of complex aerial targets with rotating parts is a great challenge due to the extremely complicated scattering mechanisms on the complex target. In this work, we propose a high-frequency approximation procedure for the simulation of wideband signatures of complex targets with rotating parts. Computational examples for period movement part modulation (PMPM) spectrum, radar cross section (RCS) and ISAR images are presented.

1. INTRODUCTION

Modeling of wideband electromagnetic (EM) scattering and inverse synthetic aperture radar (ISAR) images of complex aerial targets with rotating parts is a great challenge due to the fact that the scattering mechanisms are extremely complicated. Great efforts were made in the past two decades [1–5]. To the authors' knowledge, however, references are hardly available to consider the problem of simultaneous EM modeling of period movement part modulation (PMPM) spectrum, radar cross section (RCS) and ISAR images.

This paper presents a high-frequency approximate technique to model the wideband EM scattering and the ISAR images of aircraft with propeller blades. The procedure for time dependent EM scattering calculation of an aircraft is discussed, where the total scattering field is considered to be the vector summation of those from the aircraft hardbody and from the rotating propeller blades. Typical computational examples for PMPM modeling, RCS calculation and ISAR image simulation are presented to demonstrate the feasibility and usefulness of the current work.

2. TIME DEPENDENT EM SCATTERING MODEL OF THE AIRCRAFT

The total scattering field of the aircraft as a function of time t is calculated as the vector summation of the scattering field from the hardbody, $\mathbf{E}_{\text{hardbody}}(t)$, and that from the rotating propeller blades, $\mathbf{E}_{\text{blade}}(t)$, i.e.,

$$\mathbf{E}(t) = \mathbf{E}_{\text{hardbody}}(t) + \mathbf{E}_{\text{blade}}(t) \tag{1}$$

2.1. EM Scattering Calculation of the Hardbody

If the object is geometrically modeled as consisting of K facets plus L wedges, the scattering field of the aircraft hardbody except the rotating parts can be expressed as

$$\mathbf{E}_{\text{hardbody}}(t) = \sum_{k=1}^{K} \mathbf{E}_{\text{facet\_}k}(t) + \sum_{l=1}^{L} \mathbf{E}_{\text{wedge\_}l}(t) \tag{2}$$

where $\mathbf{E}_{\text{facet\_}k}(t)$ and $\mathbf{E}_{\text{wedge\_}l}(t)$ denote the complex physical optic (PO) scattering component from kth facet and the complex diffraction component from nth wedge, respectively. Note that for the scattering from the hardbody, variable t corresponds to the change of target aspect angles.

By PO approximation, the facet scattering is calculated as

$$\mathbf{E}_{\text{facet}} = -E_0 \frac{\mathbf{n} \times \mathbf{H_i}}{2\pi TR} e^{jk\mathbf{r_0}\cdot\mathbf{w}} \sum_{m=1}^{M} (\mathbf{p} \cdot \mathbf{a_m}) e^{jk\mathbf{r}_m\cdot\mathbf{w}} \text{sinc}(\frac{1}{2} k\mathbf{a_m} \cdot \mathbf{w}) \tag{3}$$

And the diffraction component from a wedge cab be calculated using incremental length diffraction coefficient (ILDC) method. Mitzner expressed the scattered field in term of a dyadic diffraction coefficient as [6]

$$\mathbf{E}_{\text{wedge}} = E_0 \frac{\exp[j(kR - \pi/4)]}{\sqrt{2\pi} R} \mathbf{d} \cdot \mathbf{p} dl \tag{4}$$

Detailed parameters and implementation of Eqs. (3) and (4) can be found in [5].

To count for the multiple scattering interactions among facets, by following Blejer [7] and Knott [8], we developed a recursive rapid algorithm for multiple scattering (RRAMS) calculation based on area projection and PO, combining with the reentrant polygon clipping (RPC) method in the context of computer graphics.

2.2. Modeling of the Rotating Propeller Blades

Suppose that the aircraft propeller consists of M rotating blades. Each blade is a perfect conductor which can be divided into N contiguous small plates with different pitch angles, as shown in Fig. 1. The total scattering field of the M blades can be calculated as

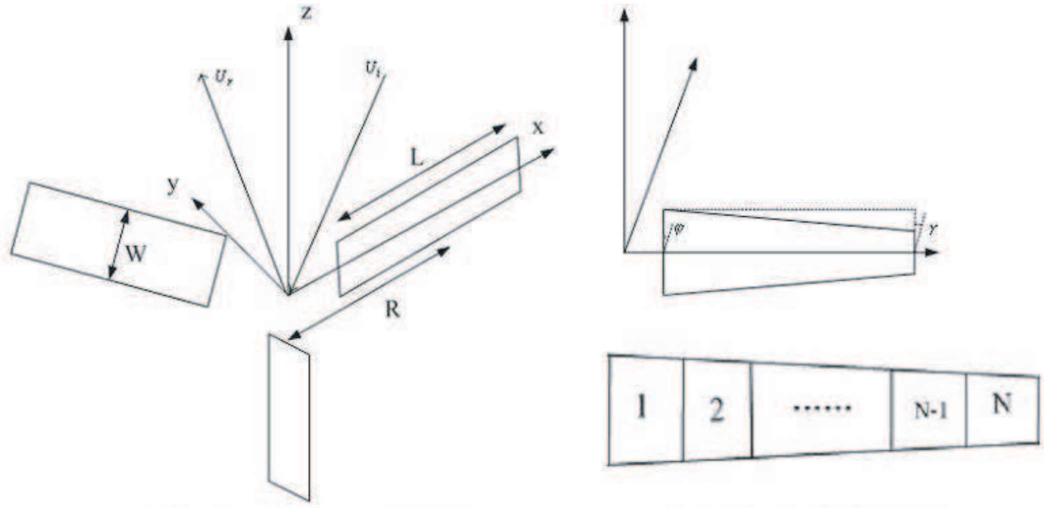


Figure 1: Geometries of the propeller blades. (a) Configuration of the propeller blades. (b) Details for a single blade.

$$\mathbf{E}_{\text{blade}}(t) = \sum_{m=1}^{M}\sum_{n=1}^{N} \mathbf{E}^{s}_{mn}(t) \tag{5}$$

where $\mathbf{E}^{s}_{mn}(t)$ represents the scattering field of the nth small plate of the mth blade, which can be calculated using PO approximation plus equivalent edge current [9], i.e.,

$$\mathbf{E}^{s}_{mn}(t) = \mathbf{E}^{\text{po}}_{mn}(t) + \mathbf{E}^{\text{d}}_{mn}(t) \tag{6}$$

where $\mathbf{E}^{\text{PO}}_{mn}(t)$ and $\mathbf{E}^{\text{d}}_{mn}(t)$ represent the facet and edge scattering fields of the nth small plate of the mth blade, respectively,

$$\mathbf{E}^{\text{PO}}_{mn}(t) = \frac{-jE_0 e^{-jkr_0}}{\lambda r_0} e^{jp_{mn}R_{mn}} L_{mn}W_{mn}\text{sinc}(\frac{p_{mn}L_{mn}}{2})\text{sinc}(\frac{q_{mn}W_{mn}}{2})\overline{\mathbf{u_r}} \times [\overline{\mathbf{u_r}} \times (\overline{\mathbf{z}}_{\mathbf{mn}} \times \overline{\mathbf{h}}_{\mathbf{i}})] \tag{7}$$

and

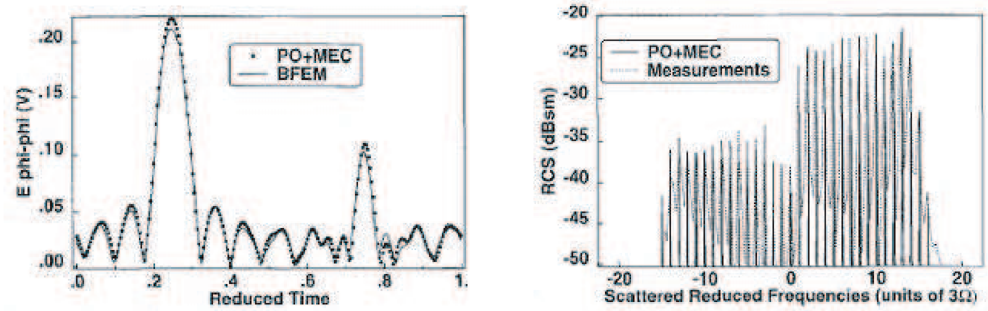
$$\mathbf{E}^{\text{d}}_{\text{mn}}(t) = -\left\{\left[\frac{\mathbf{d}\cdot\mathbf{e_i}}{\sin^2\beta_i}d_e + \frac{\mathbf{d}\cdot\mathbf{h_i}}{\sin\beta_i}d_{em}\right]\mathbf{u_r}\times(\mathbf{u_r}\times\mathbf{d}) - \left[\frac{\mathbf{d}\cdot\mathbf{h_i}}{\sin\beta_i\sin\beta_s}d_m\right]\mathbf{u_r}\times\mathbf{d}\right\}\frac{E_0}{2\pi r(t)}e^{-jkr(t)}I_d \tag{8}$$

Detailed parameters and more calculation equations can be found in [5].

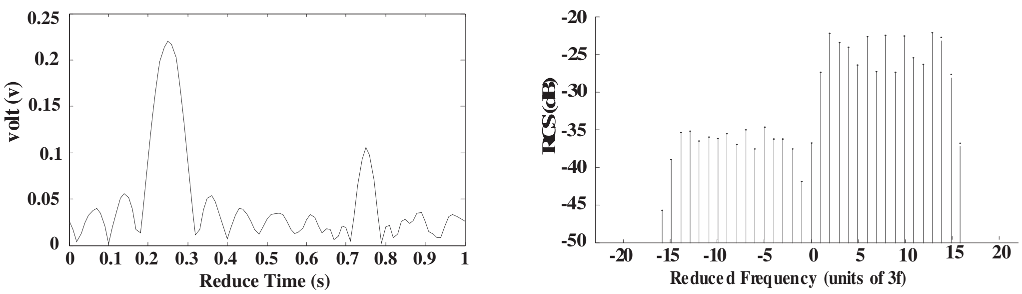
3. COMPUTATIONAL EXAMPLES

In this section we demonstrate a few computational examples to show the usefulness of the current work.

In the first example, we use the same propeller configuration as in [4] and compare the calculation results. The propeller consists of three blades as in Fig. 1. The parameters for the propeller blade are as follows: $R = 0.15$ m, $L = 0.24$ m, $W = 0.06$ m, pitch $\varphi = 15°$, twist angle $\gamma = 0°$, radar wave incidence angle $\alpha_i = 24°$, the angles between radar receiver and the z-axis are $\alpha_r = 24°$, $\theta_r = 0°$, radar wavelength $\lambda = 0.03$ m. The calculated backscattering returns and the corresponding periodic movement part modulation (PMPM) spectrum are shown in Figs. 2(a) and (b), where Fig. 2(a) demonstrates the results obtained in [4]; while Fig. 2(b) illustrates the calculation results in the current work. It is seen that our results are coincident with that in [4].



(a)

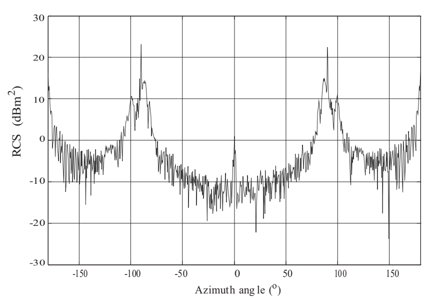


(b)

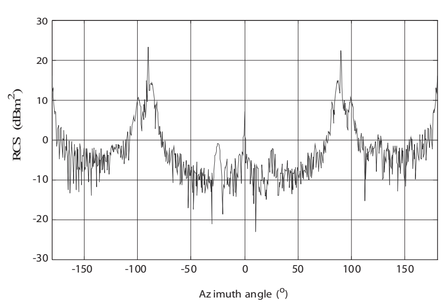
Figure 2: Comparison of our results with that in reference [4]. (a) Backscattering amplitude (left) and PMPM spectrum (right) as in [4]. (b) Backscattering amplitude (left) and PMPM spectrum (right) in this work.

In the second example, we calculate the radar cross section (RCS) and the 2-D ISAR images of a model Spitfire aircraft. The shape of the aircraft is shown in Fig. 3(a), with geometric sizes being as follows: length 9.35 m, wingspan 11.5 m, and height 2.1 m. The rotating blades parameters are: $R = 0.9\,\mathrm{m}$, $L = 1.2\,\mathrm{m}$, $W = 0.15\,\mathrm{m}$, pitch $\varphi = 15°$, twist angle $\gamma = 0°$, the number of blades is $N = 5$ with a rotating rate $f_{rot} = 23\,\mathrm{r/s}$, radar wavelength $\lambda = 0.03\,\mathrm{m}$.

(a)



(b)

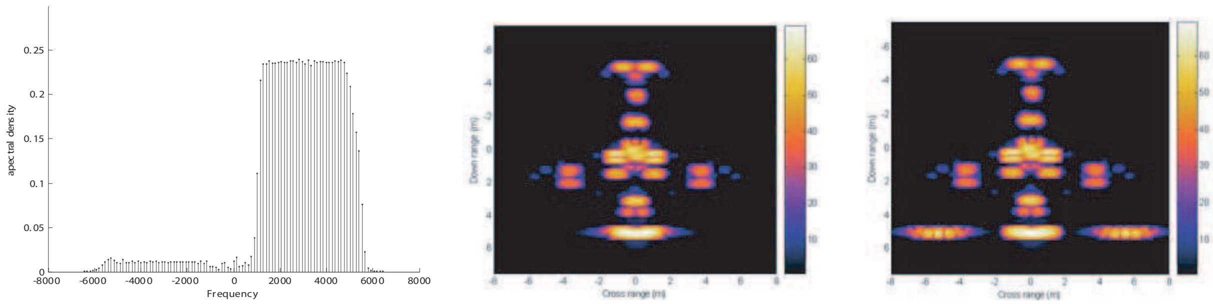


(c)

Figure 3: RCS as a function of azimuth. (a) The aircraft shape. (b) Without blade rotation. (c) Rotation rate of $f_{rot} = 23\,\mathrm{r/s}$.

Figure 3(a) shows the geometric shape of the aircraft model. Figs. 3(b) and (c) demonstrate the aircraft RCS as a function of the azimuth angle at an elevation of 0° for the cases of (b) without rotation and (c) with rotation at a rate of $f_{rot} = 23\,\mathrm{r/s}$, respectively. Around the nose-on azimuth region (especially around azimuth $\phi = 0 \pm 25°$ which corresponds to the pitch angle of the blades), the impact of the PMPM to the total RCS is evident.

Figure 4 illustrates the ISAR images from the nose-on $\pm 2°$ of the aircraft at different radar return sampling frequencies when the propeller is rotating at a rate of $f_{rot} = 23\,\mathrm{r/s}$. The sampling frequencies for Figs. 4(b) and (c) are $f_r = 115\,\mathrm{Hz}$ and $f_r = 140\,\mathrm{Hz}$, respectively. From Figs. 4(b) and (c) it seems that at $f_r = 115\,\mathrm{Hz}$, the ISAR image is well focused without artifacts of the rotating blades; while at $f_r = 140\,\mathrm{Hz}$, artifacts resulted from PMPM are severe. This phenomenon can be explained as follows.



(a) (b) (c)

Figure 4: ISAR images at different sampling frequencies. (a) PMPM spectrum. (b) ISAR image at $f_r = 115\,\text{Hz}$. (c) ISAR image at $f_r = 140\,\text{Hz}$.

Note that the maximum Doppler frequency generated from the blade rotation can be calculated as

$$f_h = \left[2 \cdot 2\pi \cdot f_{rot} \cdot (R + L/2) \sin \alpha_r\right] / \lambda. \tag{9}$$

Putting the simulation parameters into Eq. (9) it is found that at the present case, the maximum Doppler frequency can be as high as 6.1 KHz, as shown in Fig. 4(a), which means that the sampling frequency must be higher than 12.2 KHz for alias-free images. Otherwise, artifacts resulting from the aliasing distortion will appear in the radar images. In the current simulation, the number of blades $N = 5$, rotation rate $f_{rot} = 23\,\text{r/s}$. Therefore, the periodicity of the radar returns must be $f_{rot} \times N = 115\,\text{Hz}$. As a consequence, when the sampling frequency is $f_r = 115\,\text{Hz}$, which is exactly equal to that periodicity, in the ISAR image, only on peak appears at the specified down range cells where the propeller blades locate at. On the other hand, when $f_r = 140\,\text{Hz}$, which is far lower than the required alias-free sampling frequency, multiple peak artifacts arise across the cross range dimension for those down range cells where the propeller corresponds to.

REFERENCES

1. Martin, J. and B. Mulgrew, "Analysis of the effects of blade pitch on the radar return signal from rotating aircraft blades," *Proc. IEE Conference of Radar*, 466–449, 1992.
2. Yang, S. Y. and S. M. Yeh, "Electromagnetic backscattering from aircraft propeller blades," *IEEE Trans. on Magnetics*, Vol. 33, No. 2, 1432–1435, 1997.
3. Tardy, I., G. Piau, and P. Ghabrat, "Computational and experimental analysis of the scattering by rotating fans," *IEEE Trans. on Antennas and Propagation*, Vol. 44, No. 10, 1414–1421, 1996.
4. Chen, V. C., F. Li, and S.-S. Ho, "Analysis of micro-Doppler signatures," *IEE Proc. Radar Sonar Navig.*, Vol. 150, No. 4, 271–276, 2003.
5. Qin, Y., "Radar signature modeling research for aerial targets," MSc. Thesis, Beihang University, 2006.
6. Knott, E. F., "The relationship between Mitzner's ILDC and Michaeli's equivalent currents," *IEEE Trans. on AP*, Vol. 33, No. 1, 112–114, Jan. 1985.
7. Blejer, D. J., "Physical optics polarization scattering matrix for a tophat reflector," *IEEE Trans. on Antennas and Propagation*, Vol. 39, No. 6, 857–859, June 1991.
8. Knott, E. F., "RCS reduction of dihedral corners," *IEEE Trans. on Antennas and Propagation*, Vol. 25, No. 3, 406–409, May 1977.
9. Michaeli, A., "Equivalent edge currents for arbitrary aspects of observation," *IEEE Antennas and Propagation Magazine*, Vol. 36, No. 1, 65–69, 1994.

A High-frequency Hybrid Method to Calculate EM Scattering of a Three-plate Cavity

Kai Cui, Xiaojian Xu, and Shiyi Mao

School of Electronic and Information Engineering, BeiHang University
37 XueYuan Road, Beijing 100083, China

Abstract— This paper focuses on the electromagnetic (EM) scattering modeling of a special kind of cavities consisting of three plates. Physical Optics (PO) approximation and Area Projection (AP) methods are combined to analyze the multiple scattering mechanisms. Computer Graphics (CG) technique is applied to solve the problems when shadowed areas occur. Typical examples demonstrate that the current method has good agreement with published results.

1. INTRODUCTION

Multiple scattering is an important mechanism in electromagnetic (EM) calculation. In many case it makes more scattering contribution than specular reflection. Anderson [1] applied physical optics (PO) method to the backscattering calculation of dihedral corner reflector with various internal angles. Other techniques can also be found in other reference [2]. However, dihedral corner reflector mechanism may not be applied for some kinds of special structures. Assume that adding a plate to dihedral corner reflector to constitute a special kind of three-plate cavities, as shown in Fig. 1. This kind of special structures can be considered as the extension of adihedral corner reflector whose scattering mechanisms are obviously different from those of a normal corner reflector.

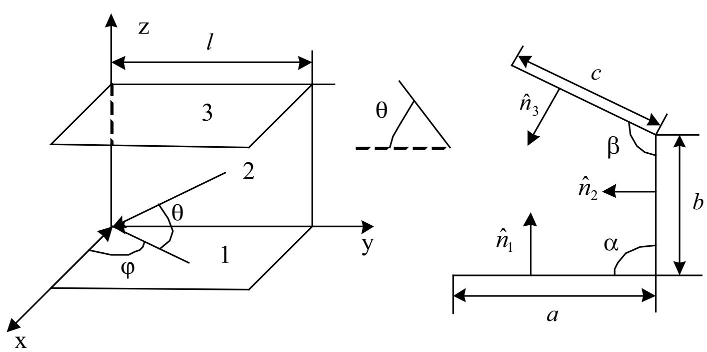


Figure 1: Geometrical structure of the three-plate cavity.

This paper reports a hybrid method that combines PO and Area Projection (AP) to compute the radar cross section (RCS) of three-plate cavities based on Anderson's multiple scattering mechanism calculation for dihedral corner reflector. Different from dihedral corner reflector, the existence of the additional plate causes the phenomenon of shadowed areas, as shown in Fig. 2. The problem can be solved by deploying Computer Graphics (CG) technique. Finally, the method is then validated by comparing the current results with typical examples published previously.

2. THE EM SCATTERING OF A THREE-PLATE CAVITY

The equation is given to compute scattering field of a loaded plate [3]:

$$\vec{\mathbf{E}}^{\mathbf{s}} = -\frac{jk}{4\pi}\frac{e^{-jkr}}{r}\int_s \left(Z_0\overline{M}_1 \cdot \vec{\mathbf{J}}_{\mathbf{s}} + \overline{M}_2 \cdot \vec{\mathbf{J}}_{\mathbf{ms}}\right) e^{j\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}'} dS \tag{1}$$

where $\overline{M}_1 = \begin{bmatrix} 0 & 0 & 0 \\ \cos\varphi\cos\theta & \sin\varphi\cos\theta & -\sin\theta \\ -\sin\varphi & \cos\varphi & 0 \end{bmatrix}$, $\overline{M}_2 = \begin{bmatrix} 0 & 0 & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ -\cos\theta\cos\varphi & -\cos\vartheta\sin\varphi & \sin\theta \end{bmatrix}$, $\vec{\mathbf{J}}_{\mathbf{s}}$ and $\vec{\mathbf{J}}_{\mathbf{ms}}$ denote the surface electric current and magnetic current, respectively:

$$\vec{\mathbf{J}}_{\mathbf{s}} = \hat{\mathbf{n}} \times \vec{\mathbf{H}}, \quad \vec{\mathbf{J}}_{\mathbf{ms}} = \vec{\mathbf{E}} \times \hat{\mathbf{n}} \tag{2}$$

where $\vec{\mathbf{E}}$ and $\vec{\mathbf{H}}$ denote the total electric field and magnetic field on boundary; $\hat{\mathbf{n}}$ is an outward unit vector normal to the surface. For perfect electric conductors (PEC), $\vec{\mathbf{J}}_{\mathbf{ms}} = 0$. Decompose the incidence field and scattering field along the following unit vectors:

$$\begin{aligned}
\vec{\mathbf{E}}^{\mathbf{i}} &= E_\perp \hat{\mathbf{e}}_\perp + E_{//}\hat{\mathbf{e}}^{\mathbf{i}}_{//} \\
\hat{\mathbf{e}}_\perp &= \frac{\hat{\mathbf{i}} \times \hat{\mathbf{n}}}{\left|\hat{\mathbf{i}} \times \hat{\mathbf{n}}\right|}, \quad \hat{\mathbf{e}}^{\mathbf{i}}_{//} = \hat{\mathbf{e}}_\perp \times \hat{\mathbf{i}}, \quad \hat{\mathbf{e}}^{\mathbf{s}}_{//} = \hat{\mathbf{e}}_\perp \times \hat{\mathbf{s}}
\end{aligned} \tag{3}$$

where $\hat{\mathbf{i}}$ and $\hat{\mathbf{s}}$ are unit vectors along incident direction and the direction to the receiver. Then Equation (2) is written as:

$$\begin{aligned}
\vec{\mathbf{J}}_{\mathbf{s}} &= \frac{1}{Z_0}[(1 - R_\perp)E_\perp \cos\theta \cdot \hat{\mathbf{e}}_\perp + (1 + R_{//})E_{//} \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{e}}_\perp)] \\
\vec{\mathbf{J}}_{\mathbf{ms}} &= [(1 - R_{//})E_{//} \cos\theta \cdot \hat{\mathbf{e}}_\perp - (1 + R_\perp)E_\perp \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{e}}_\perp)]
\end{aligned} \tag{4}$$

where θ is the angle between incident direction and the direction normal to the plane; $R_\perp$ and $R_{//}$ are the reflection coefficient of loaded surface for vertical and horizontal polarization, respectively.

Considering the electric polarization vector $\hat{\mathbf{e}}_{\mathbf{r}}$ of the receiver, Equation (1) can be written as:

$$\begin{aligned}
\vec{\mathbf{E}}^{\mathbf{s}} \cdot \hat{\mathbf{e}}_{\mathbf{r}} = &-\frac{jk}{4\pi}\frac{e^{-jkr}}{r}\int_s [(\hat{\mathbf{e}}_{\mathbf{r}} \cdot \hat{e}_\perp)(1 - R_\perp)E_\perp \cos\theta + (1 + R_{//})E_{//}\hat{\mathbf{e}}_{\mathbf{r}} \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{e}}_\perp) \\
&-(\hat{\mathbf{e}}_{\mathbf{r}} \times \hat{\mathbf{s}}) \cdot \hat{\mathbf{e}}_\perp(1 - R_{//})E_{//} \cos\theta + (\hat{\mathbf{e}}_{\mathbf{r}} \times \hat{\mathbf{s}}) \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{e}}_\perp)(1 + R_\perp)E_\perp]e^{j\vec{\mathbf{k}}\cdot\vec{\mathbf{r}}'}dS
\end{aligned} \tag{5}$$

Gordon [4] gave the expression to calculate the PO integration above if the surface is an arbitrary polygon:

$$\begin{aligned}
\sqrt{\sigma} = &-\frac{jk}{\sqrt{\pi}}[(\hat{\mathbf{e}}_{\mathbf{r}} \cdot \hat{\mathbf{e}}_\perp)(1 - R_\perp)E_\perp \cos\theta + (1 + R_{//})E_{//}\hat{\mathbf{e}}_{\mathbf{r}} \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{e}}_\perp) \\
&-(\hat{\mathbf{e}}_{\mathbf{r}} \times \hat{\mathbf{s}}) \cdot \hat{\mathbf{e}}_\perp(1 - R_{//})E_{//} \cos\theta + (\hat{\mathbf{e}}_{\mathbf{r}} \times \hat{\mathbf{s}}) \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{e}}_\perp)(1 + R_\perp)E_\perp]\frac{1}{T}e^{-jk\vec{\mathbf{r}}\cdot\vec{\omega}}\sum_{n=1}^{N} T_n
\end{aligned} \tag{6}$$

where $\vec{\boldsymbol{\omega}} = (\hat{\mathbf{i}}_{\mathbf{n}} - \hat{\mathbf{s}})$, $\hat{\mathbf{i}}_{\mathbf{n}}$ is the nth unit vectors along the incidence direction; T is the length of projection $\vec{\boldsymbol{\omega}}$ onto the surface. T_n is expressed as follows:

$$T_n = \sum_{m=1}^{M} \hat{\mathbf{p}} \cdot \vec{a}_m \cdot e^{ik\vec{r}_m\cdot\vec{\omega}}\mathrm{sinc}(k\vec{a}_m \cdot \vec{\boldsymbol{\omega}}/2) \tag{7}$$

where $\hat{\mathbf{p}} = \hat{\mathbf{n}} \times \vec{\boldsymbol{\omega}} / \left|\hat{\mathbf{n}} \times \vec{\boldsymbol{\omega}}\right|$; $\hat{\mathbf{r}}_{\mathbf{n}}$ is the position vector of the midpoint of the mth edge; $\vec{a}_m$ is the edge vector of the mth side of the polygon; $\mathrm{sinc}(x) = \sin x/x$. If $T = 0$, Equation (6) can be rewritten as

$$\begin{aligned}
\sqrt{\sigma} = &-\frac{jk}{\sqrt{\pi}} \cdot A \cdot [(\hat{\mathbf{e}}_{\mathbf{r}} \cdot \hat{\mathbf{e}}_\perp)(1 - R_\perp)E_\perp \cos\theta + (1 + R_{//})E_{//}\hat{\mathbf{e}}_{\mathbf{r}} \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{e}}_\perp) \\
&-(\hat{\mathbf{e}}_{\mathbf{r}} \times \hat{\mathbf{s}}) \cdot \hat{\mathbf{e}}_\perp(1 - R_{//})E_{//} \cos\theta + (\hat{\mathbf{e}}_{\mathbf{r}} \times \hat{\mathbf{s}}) \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{e}}_\perp)(1 + R_\perp)E_\perp]e^{-jk\vec{\mathbf{r}}\cdot\vec{\omega}}
\end{aligned} \tag{8}$$

where A is the geometrical area of the polygon.

The PO integration of scattering field above can be applied to both loaded and unloaded surfaces. For simplicity we here assume that the three-plate cavity is a PEC. In Fig. 1, the scattering model forms the coherent sum of contributions due to single, double, triple and even high order bounce mechanisms. To use Equations (6) and (8), the key point is how to find the size and shape of the illuminated surface patch in each scattering process so that individual contribution can be computed.

Sutherland [5] used Area Projection to determine the size and shape of individual illuminated surface patch. However, for a three-plate cavity, the phenomenon of shadowed areas occurs with different incident angles, as can be seen in Fig. 2. Because PO method is only effective to compute the contributions of the illuminated area while the existence of shadows changes the total multiple scattering mechanisms. The solution requires two steps: Firstly, determine the illuminated and shadow area in each bounce passes by deploying CG technique. The whole procedure is just inverse to that as in reference [5], seen in Fig. 3. Secondly, only compute the parts that each bounce affects in illuminated area while those in the shadowed area are not considered because some incidence wave will enter into the shadowed area but be reflected to the illuminated area.

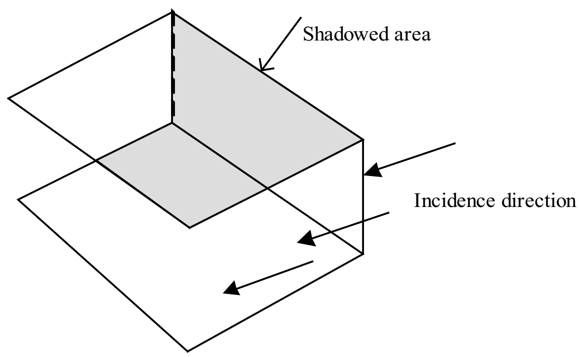


Figure 2: Phenomenon of shadowed area.

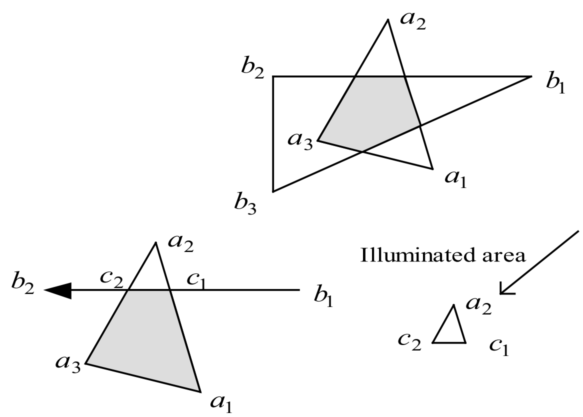


Figure 3: Procedure to determine illuminated area.

3. COMPUTATIONAL EXAMPLES

In order to study the accuracy of the current method, we present a few examples in this section. Our result is compared with that of Burkholder's modal reference solution in reference [6]. In all the cases, we set the radar frequency $f = 10\,\text{GHz}$.

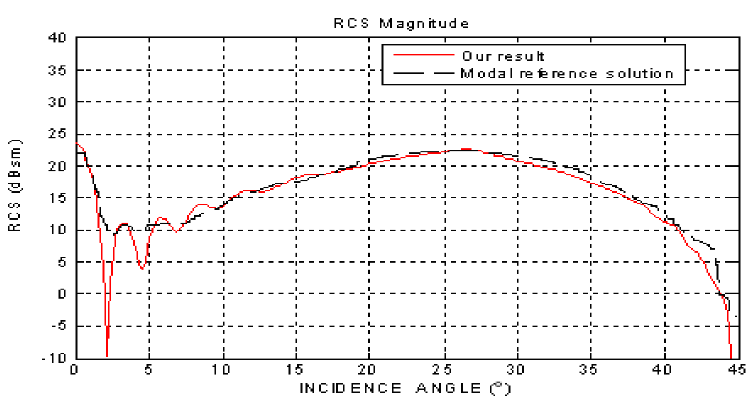


Figure 4: Result comparison of the cavity.

Figure 4 shows the calculated RCS of an electrically large sized cavity as a function of incidence angle. The size parameters for the cavity are $a = b = c = 0.45\,\text{m}$, $l = 0.5\,\text{m}$, $\alpha = \beta = 90°$. For comparison, the result in [6] with the same parameters is also shown in the figure, where solid line shows our computational result while dashed line is from [6]. It is seen that our result has a good agreement with Burkholder's.

Figure 5 presents the computed RCS as a function of the incidence angle for two other different cavities, where $a = b = c = 0.8\,\text{m}$, $l = 2\,\text{m}$. For the first cavity, $\alpha = 90°$ and $\beta = 150°$; while for the second cavity, $\alpha = 75°$ and $\beta = 165°$. In addition, Fig. 6 illustrates the calculated RCS as a function of azimuthal angle at different incidence angles for the second cavity. These two figures demonstrate how the parameters α and β impact on the total RCS of the three-plate cavities.

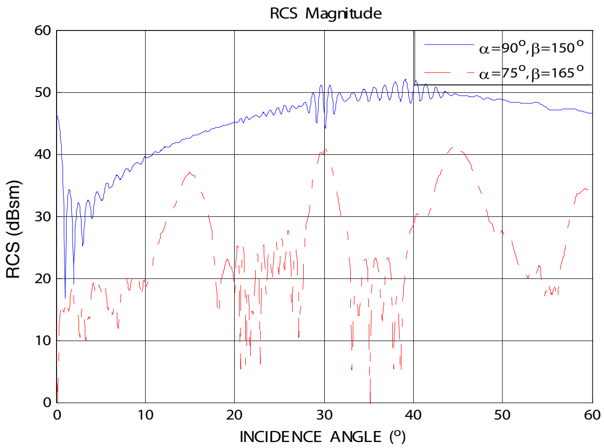


Figure 5: RCS vs. incident angle for different cavities.

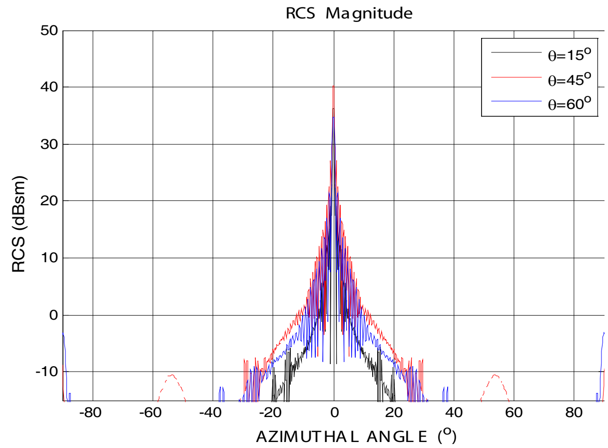


Figure 6: RCS vs. azimuth at different incident angles.

REFERENCES

1. Anderson, W. C., "Consequences of nonorthogonality on the scattering properties of dihedral reflectors," *IEEE Trans. Antennas and Propagation*, Vol. AP-35, No. 10, 1154–1159, Oct. 1987.
2. Knott, E. F., "A tool for predicting the radar cross section of an arbitrary trihedral corner," *IEEE Proc. Southeastcon'81 Conference*, Vol. 137, No. 4, 17–20, Apr. 1981.
3. Zhao, W. J., "Study on prediction techniques of radar cross section of complex targets," PhD dissertation of Xidian University, June 1999.
4. Gordon, W. B., "Far field approximation of the Kirchoff-Helmholtz representation of scattered fields," *IEEE Trans. Antennas and Propagation*, Vol. 23, No. 5, 864–876, July 1975.
5. Sutherland, I. E. and G. W. Hodgman, *CACM*, No. 17, 32–42, 1974.
6. Burkholder, R. J., "High-frequency asymptotic methods for analyzing the EM scattering by open-ended waveguide," NASA-CR-186244, 189–192, Sept. 1989.

The MOM Solution Combining Hybrid Domain Bases and Wire-grid Model for Scattering by Complex Targets

Fangzhi Geng[1,2], **Chunzhu Dong**[2], **Hongcheng Yin**[2], **and Chuangming Tong**[1]

[1]Missile Institute of Air Force Engineering University, Sanyuan 713800, China
[2]National Electromagnetic Scattering Laboratory, Beijing 100854, China

Abstract— An efficient basis function is described in this paper for computation of RCS of PEC wire-grid model of aircraft in the resonance domain based on the method of moment (MoM). Although many achievements on the wire-grid model had been obtained [1], the numerical solutions of this problem are given based on a new hybrid domain basis function and Galerkin method. The aircraft shape is approximated by a grid of thin wires, the length and the radius of this thin wires must be satisfied some special requirements. Computational results for RCS of canonical wire-grid models of targets are enclosed. The final results show that with this hybrid domain basis functions, good agreement with experimental data can be obtained by this technique.

1. INTRODUCTION

The MoM for the wire-grid model was firstly presented by J. H. Richmond in 1966. In [1], the RCS of a sphere and a circular plate were calculated by wire-grid model for continuous conducting surfaces using pulse bases and point matching. The reaction technique subsequently proposed in [2] was also a moment method with sinusoidal bases and Galerkin's method, where several wire-grid models of an airplane with increasing details were used to demonstrate the accuracy of the wire-grid reaction technique by comparing computational results with measured data. So, the technique of wire-grid model combined with subdomain bases was widely used in the computation of the RCS for some complex targets in the past years.

This paper discusses the use of MoM based on the hybrid domain basis functions in wire-grid models, which comes from the thought of almost-entire-domain described in reference [3]. The computational results obtained by this method are compared with experimental data. It is shown that a high efficiency of calculation could be achieved by using wire-grid models to approximate continuous PEC surfaces, and the feasibility is also verified by applying this hybrid domain basis functions to accurately model the thin wire structures' axial current.

2. FORMULATION

2.1. Wire-grid model [1–3]

Consider the wire structure shown in Fig. 1. The structure may be planar or three-dimensional, and it is constructed of straight wire segments.

The current is assumed to vanish at the ends and satisfy Kirchhoff's law at each junction. Considering m wires meeting at a junction, as shown in Fig. 1, to simplify the analysis, let the junction be the starting point of all m wires, i.e., let current intensities from the junction into the wire be I'_k, $k = 1, 2, \ldots, m$. We wish to obtain an expansion which automatically satisfies the current-continuity equation at the junction, that is to say, the Kirchhoff's law at junction must be satisfied by the following equation

$$\sum_{k=1}^{m} I'_k(u^{(k)}) = 0 \tag{1}$$

By Eq. (1), we can obtain

$$I'_m(u^{(m)}) = -\sum_{k=1}^{m-1} I'_k(u^{(k)}) \tag{2}$$

Therefore, for m wires meeting at a junction, there are $m-1$ linear independent unknowns to be solved and the Kirchhoff's law can be satisfied automatically by Eq. (2).

It is frequently of interest to analyze approximately the exact wire model by the right truncated cone, as shown in Fig. 2. We can obtain the parametric equations of the cone axis and its local

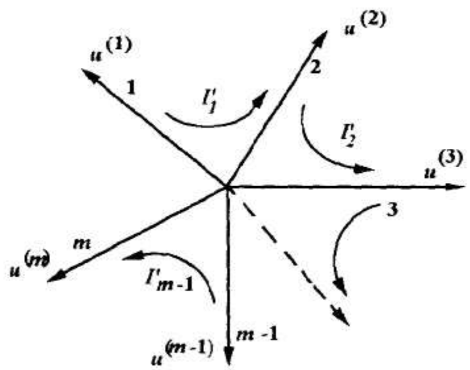


Figure 1: Wire scatterer constructed of straight segments.

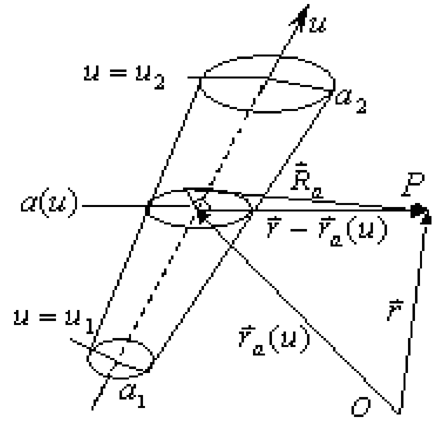


Figure 2: The definition of the distance between source and field points entering the right truncated cone.

radius

$$r_a(u) = r_1 + (u - u_1)\frac{r_2 - r_1}{u_2 - u_1} \tag{3}$$

$$a(u) = a_1 + (u - u_1)\frac{a_2 - a_1}{u_2 - u_1} \tag{4}$$

where $u_1 \leq u \leq u_2$. A right truncated cone is determined by the position vectors and the radii of its beginning and end points, $\boldsymbol{r}_1$ and a_1, $\boldsymbol{r}_2$ and a_2, respectively. The distance between the source point and the field point in this case is defined as (see Fig. 2)

$$R_a = \sqrt{|\boldsymbol{r} - \boldsymbol{r}_a(u)|^2 + a(u)^2} \tag{5}$$

where $\boldsymbol{r}$ and $\boldsymbol{r}_a(u)$ are the position vectors of the field point and source point, respectively. Generally speaking, we define $u_1 = -1$ and $u_2 = 1$ at the beginning and the end of the right truncated cone, for a thin wire with same radius ($a_2 = a_1$).

2.2. Application of Hybrid Domain Bases in MOM [3]

Based on this right truncated cone model, the axis current density function can be expressed as

$$\boldsymbol{I}(u) = \sum_{i=1}^{M} I_i \omega_i(u) \cdot \hat{u} \tag{6}$$

where $\hat{u}$ is the unit vector along the axis, M is the degree of the approximations along the u coordinate, I_i is the unknown coefficient, and $\omega_i(u)$ is arbitrary known basis functions. By setting $f_i(u) = u^{i-1}$, $-1 \leq u \leq 1$, we can adopt hybrid domain basis functions which automatically satisfying Kirchhoff's law, and the basis functions can be written as

$$\omega_i(u) = u^{i-1} + a_i u + b_i \tag{7}$$

The coefficients in Eq. (7) are

$$a_i = \begin{cases} -0.5, (i < 3) \\ 0.5 \times [(-1)^{i-1} - 1], (i \geq 3) \end{cases} \tag{8}$$

$$b_i = \begin{cases} 0.5 \times (-1)^i, (i < 3) \\ 0.5 \times [(-1)^i - 1], (i \geq 3) \end{cases} \tag{9}$$

Inserting Eqs. (7) and (8) in Eq. (6), and expanding the axis direction current function results in

$$\begin{aligned}\mathbf{I}(u) &= \sum_{i=1}^{M_u} I_i\omega_i(u)\cdot\hat{u} = \sum_{i=1}^{M_u} I_i(u^{i-1}+a_iu+b_i)\cdot\hat{u}\\ &= \hat{u}\cdot\left\{\begin{array}{c} I_1\frac{1-u}{2}+I_2\frac{1+u}{2}+\sum\limits_{\substack{i=3\\(2)}}^{M_u} I_i(u^{i-1}-1)\\ +\sum\limits_{\substack{i=4\\(2)}}^{M_u} I_i(u^{i-1}-u)\end{array}\right\}\end{aligned} \tag{10}$$

The electric field integral equation (EFIE) is obtained from the boundary condition for the tangential component of the electric filed vector on the surface of the PEC [3].

$$\hat{n}\times\boldsymbol{E}^s = -\hat{n}\times\boldsymbol{E}^i \tag{11}$$

The scattering electric field vector $\boldsymbol{E}^s$ can be expressed as

$$\boldsymbol{E}^s = -gradV - j\omega\boldsymbol{A} \tag{12}$$

where the electric scalar and magnetic vector potential, V and $\boldsymbol{A}$, are given by

$$V = \frac{-1}{j\omega\varepsilon_0}\int\limits_u \nabla_s\cdot\boldsymbol{I}(u)g(R)du \tag{13}$$

$$A = \mu_0\int\limits_u \boldsymbol{I}(u)g(R)du \tag{14}$$

and the Green function, $g(R)$, is

$$g(R) = \exp(-j\beta R)/4\pi R,\quad \beta = 2\pi/\lambda \tag{15}$$

where $R = R_a$ is given in Eq. (5), β is the propagation constant. Eq. (11) is discretized using Galerkin's method, then a system of linear equations can be expressed in matrix form as

$$[\boldsymbol{Z}_{kl}]\cdot[\boldsymbol{I}_l] = [\boldsymbol{V}_k] \tag{16}$$

where the matrix elements can be obtained by

$$\left\{\begin{array}{l} Z_{kl} =< I_k, L(I_l) >=< I_k, E_l > \\ V_k =< I_k, -E^i > \end{array}\right. \tag{17}$$

So, the elements of the impedance matrix can be got by

$$Z_{kl} = \int\limits_{u_k}\left\{\frac{dI(u_k)}{du_k}V_l - j\omega\left[I(u_k)\frac{dr_a(u_k)}{du_k}\right]\cdot A_l\right\}du_k \tag{18}$$

where $\eta_0 = 120\pi$ is the free space wave impedance.

Aside from what above mentioned, we defined the incident electric field vector as

$$\boldsymbol{E}^i(\hat{r}) = \hat{E}_0\exp\left(-j\beta\hat{k}\cdot\boldsymbol{r}_a(u)\right) \tag{19}$$

where $\hat{k}$ is the incident unit vector, $\boldsymbol{r}_a(u)$ is given in Eq. (3), $\hat{E}_0$ denotes the polarization mode, normalized to $\left|\hat{E}_0\right| = 1$, and the elements of the excitated matrix can be got in the form

$$V_k = -\int\limits_{u_k}\boldsymbol{I}_k\cdot\boldsymbol{E}^i du_k \tag{20}$$

On the base of Eqs. (17)–(20), with the solutions of the matrix equation 16, the current coefficient matrix $[\boldsymbol{I}_l]$ can be obtained naturally, so we can get the parameters used in the computation of scattering field.

2.3. Solutions of RCS

When the scattering field is obtained by Eq. (12), the RCS can be easily obtained in the form

$$\sigma = \lim_{R \to \infty} 4\pi R^2 \left(|\boldsymbol{E}^s|^2 / \left|\boldsymbol{E}^i\right|^2 \right) \tag{21}$$

3. NUMERICAL AND EXPERIMENTAL RESULTS

In order to verify the method presented in this paper, we present some numerical results by means of the hybrid domain bases.

RCS of MIG19 aircraft model (1:200 scale) at two frequencies are presented. The measured data are cited from the reference [2]. All data are presented for monostatic RCS (σ/λ^2) with $\hat{\phi}$ polarization in dB. The coordinate system and the aspect angle ϕ are defined in Fig. 3. And the airplane model is made of 70 segments. We assume the axis length of this model is L, and the radius of thin-wire is a.

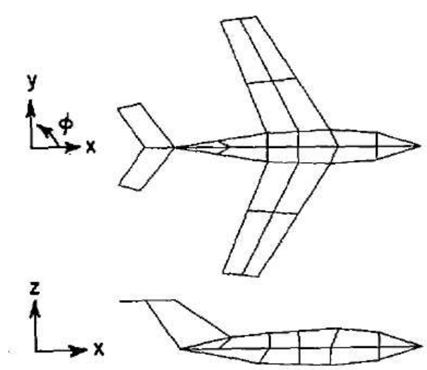


Figure 3: Wire-grid model for MIG19 aircraft.

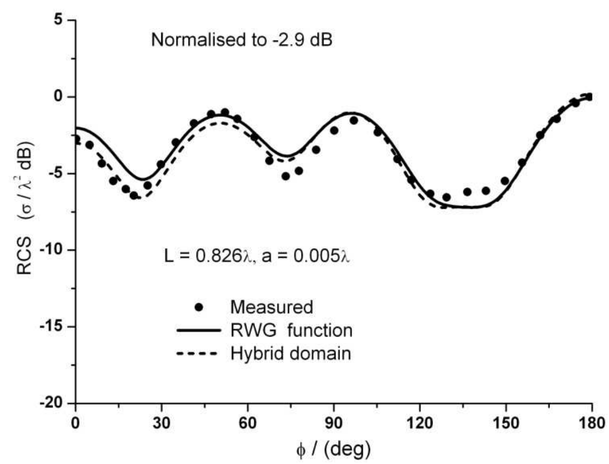


Figure 4: Monostatic RCS of wire-grid model with $L = 0.826\lambda$.

In order to give a better comparison between the computed and measured results, the RWG basis functions based on triangle plate segments [4] are also used to model the continuous conducting surface.

Figure 4 shows that the RCS normalized with λ^2 versus the azimuth angle $\phi(\theta = 90°)$ for the wire-grid model with $\hat{\phi}$ polarization when $L = 0.826\lambda$ and $a = 0.005\lambda$. The black dot indicates the measured data cited from reference [2], the real line result came from the RWG bases results, and

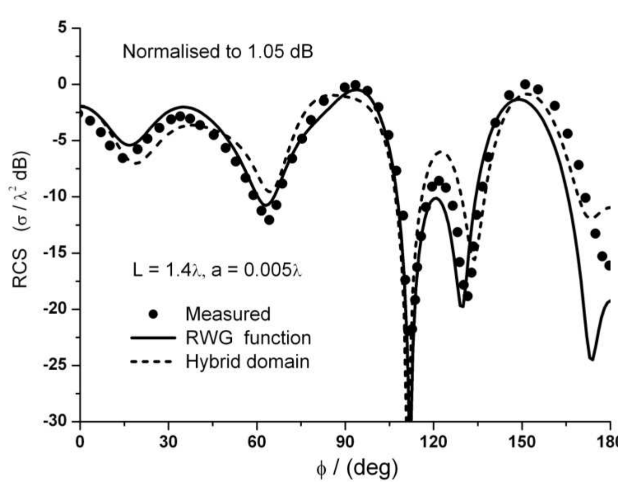


Figure 5: Monostatic RCS of wire-grid model with $L = 1.4\lambda$.

dashed is the result that calculated by this hybrid domain bases. From Fig.4, we can see that the computational results of hybrid domain bases remain in good agreement with the measured data.

Figure 5 shows the RCS results for the model length $L = 1.4\lambda$ and the radius $a = 0.005\lambda$. With the same description as Fig. 4, the black dot indicates the measured data cited from reference [2], the real line result came from the RWG basis functions results, and dashed is the result that calculated by this hybrid domain bases.

4. CONCLUSIONS

The MoM combing the hybrid domain bases with the wire-grid model is presented to calculate the RCS of complex targets. It was found that the apertures in the grid should not exceed $\lambda/4$ in width for wire-grid modeling of a continuous conducting surface. The wire radius is important but not critical within a certain range. A suitable radius is $a = w/25$, where w denotes the width of the apertures. To ensure the validity of the thin-wire model, the radius should not exceed 0.02λ. Good agreements between the theoretical and measured results are achieved to verify the proposed method.

REFERENCES

1. Richmond, J. H., "A wire-grid model for scattering by conducting bodies," *IEEE Trans.*, Vol. AP-14, 782–786, Nov. 1966.
2. Lin, Y. T. and J. H. Richmond, "EM modeling of aircraft at low frequencies," *IEEE Trans.*, Vol. AP-23, No. 1, 53–56, Jan. 1975.
3. Popovic, B. D. and B. M. Kolundzija, *Analysis of Metallic Antennas and Scatters*, The Institution of Electrical Engineers, London, 1994.
4. Rao, S. M., D. R. Wilton, and A. W. Glisson, "Electromagnetic scattering by surfaces of arbitrary shape," *IEEE Trans.*, Vol. AP-30, No. 3, 409–418, May 1982.

EM Scattering from Complex Targets above a Slightly Rough Surface

Chunzhu Dong, **Chao Wang**, **Xiao Wei**, **and Hongcheng Yin**
National Electromagnetic Scattering Laboratory, Beijing 100854, China

Abstract— A hybrid approach which combines the "four-path" model with a quasi-image method is developed in this paper to deal with the high frequency EM scattering problems of complex targets located above a slightly rough surface. The computation process of the multi-path scatterings in the "four-path" model is greatly simplified by using the image targets to treat the shadowing effects and a damped reflection coefficient to correct the scattered field of the rough surface. The effectiveness and efficiency of this method are verified by numerical results, which show that the proposed approach can be used to quickly and accurately evaluate the high frequency scattering from complex targets above a slightly rough surface.

1. INTRODUCTION

When a target is in a definite environment, effects such as multi-path scattering and shadowing will take place, composite EM scattering modeling of the target and the surrounding medium is therefore needed [1]. Unfortunately, actual surrounding surfaces are quite different from their vegetation, roughness and undulation, etc, which makes it rather difficult to model the composite scattering from targets in the environment.

The "four-path" model [2] is a usual approximate method to solve the composite scattering problem of targets located above a rough surface. Johnson [3] confirmed that the "four-path" model is able to consider the primary scattering mechanisms in the composite scattering problem of targets and rough surface.

Slightly rough surface is a kind of very typical surroundings, such as an aerodrome and a flat sea surface, etc. This paper presents a hybrid approach which combines the "four-path" model with a quasi-image method to consider the high frequency EM scattering from complex targets above a slightly rough surface. The computation process of the multi-path scatterings in the "four-path" model is greatly simplified by using the image targets to treat the shadowing effect and a damped reflection coefficient to correct the scattered field of the rough surface. Numerical results quite agree with ones of the Iterative Physical Optics (IPO) in [4], which shows that the proposed approach can be used to quickly and accurately evaluate the high frequency scattering from complex targets above a slightly rough surface.

2. THEORY

2.1. The Reflected Field of Slightly Rough Surface

The height reference plane of the rough surface is defined by xoy plane. The rough surface is classified as slightly rough surface when the surface RMS height $\hat{\sigma}$ and the correlation length l are both smaller than the electromagnetic wavelength λ, and satisfy Eq. 1 as follows,

$$k_0\hat{\sigma} < 0.3, \quad \sqrt{2}\hat{\sigma}/l < 0.3 \tag{1}$$

Scattering returns of the slightly rough surface is dominated by the coherent component, while the incoherent component, i.e., the backscattered field, is negligible in total scattering field [1].

Figure 1 illustrates the geometry of the problem. The rough surface profiles are described by a Gaussian random process with the isotropic Gaussian correlation function defined by surface RMS height $\hat{\sigma}$ and correlation length l.

To calculate the reflected field of slightly rough surface, a damped reflection coefficient Γ_c and an extra random phase shift $\Delta\phi$ are introduced into the rough surface reflection coefficients ($R'_{\perp}$, $R'_{//}$) to correct the Fresnel reflection coefficients ($R_{\perp}$, $R_{//}$) and consider the effect of the surface roughness respectively, i.e.,

$$R'_{\perp} = \Gamma_c R_{\perp} e^{j\Delta\phi}, \quad R'_{//} = \Gamma_c R_{//} e^{j\Delta\phi} \tag{2}$$

where

$$R_{\perp} = \frac{\cos\theta_i - \sqrt{\varepsilon_r - \sin^2\theta_i}}{\cos\theta_i + \sqrt{\varepsilon_r - \sin^2\theta_i}} \quad R_{//} = \frac{\varepsilon_r\cos\theta_i - \sqrt{\varepsilon_r - \sin^2\theta_i}}{\varepsilon_r\cos\theta_i + \sqrt{\varepsilon_r - \sin^2\theta_i}} \tag{3}$$

$$\Gamma_c = \exp(-2(k_0\hat{\sigma}\cos\theta_i)^2) \qquad \Delta\phi = 2k_0 H_z \cos\theta_i \tag{4}$$

In the above two equations, θ_i is the incidence angle, ε_r is the relative permittivity of the rough surface, k_0 is the wave number in free space, H_z that obeys $N(0, \hat{\sigma})$ is the undulating height of the rough surface at the reflection point.

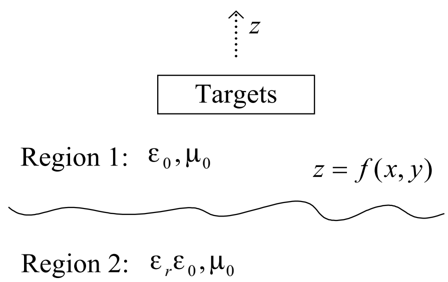


Figure 1: Geometry of the problem.

2.2. The "Four-path" Model

Figure 2 illustrates the backscattering mechanisms of the four-path model. Four multi-path scattering mechanisms are included by the following models: (a) backscattering contributions of the targets only (Path 1); (b) second reflections contributions of target-surface interaction (Path 2); (c) second reflections contributions of surface-target interaction (Path 3); (d) third reflections contributions of target-surface-target interaction (Path 4).

It can be easily seen that Path 1 is the general backscattering path; therefore high frequency asymptotic methods [5, 6], such as Physical Optics (PO), Equivalent electric and magnetic Current Method (ECM) and Shooting Bounce Ray (SBR) method, can be used to calculate the first bounce contribution, edge diffraction and multi-reflection contribution of the targets, respectively. Paths 2 to 4 are all multi-path scattering paths, where both the rough surface reflection coefficients and the shadow effects need to be considered.

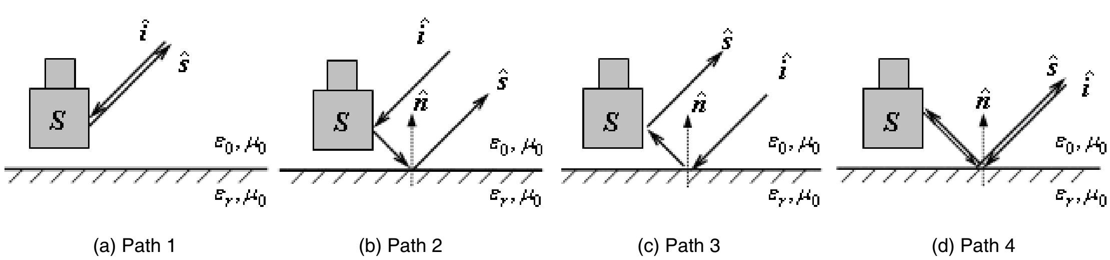


Figure 2: Scattering mechanisms of four-path model.

2.3. The Quasi-image Method

The computation process of the composite scattering by targets above rough surface in the "four-path" model is rather complicated and time-consuming. In order to improve its efficiency, both targets image and radar image can be employed to deal with the shadowing effects and multi-path scattering problem.

Figure 3 illustrates the process using image targets to treat the shadowing effects of targets and surface. For path 2 to path 4, which involves shadowing of the rough surface by the targets S in the incidence direction $\hat{\boldsymbol{i}}$, image target R will be set to eliminate the contribution from the shadowing area.

Figure 4 illustrates the process using image radars to simplify the calculation of the multi-path scattering contribution, where $\hat{\boldsymbol{i}}'$ and $\hat{\boldsymbol{s}}'$ are the mirror image of incidence direction $\hat{\boldsymbol{i}}$ and radar receiving direction $\hat{\boldsymbol{s}}$ through rough surface, respectively. Obviously, path 2 can be equivalent to the "bistatic scattering" problem of the targets with the emitting radar located in the direction of $\hat{\boldsymbol{i}}$ and the receiving radar in the direction of $\hat{\boldsymbol{s}}'$; path 3 can be equivalent to the "bistatic scattering" problem of the targets also with the emitting radar located in the direction of $\hat{\boldsymbol{i}}'$ and the receiving radar in the direction of $\hat{\boldsymbol{s}}'$; while path 4 is a "monostatic scattering" problem with the emitting radar located in the direction of $\hat{\boldsymbol{i}}'$ and the receiving radar in the direction of $\hat{\boldsymbol{s}}'$. Both phase shifts associated with the traveled path lengths and the rough surface reflection coefficient for each path must be considered in the so-called "bistatic scattering" problem and "monostatic scattering" problem.

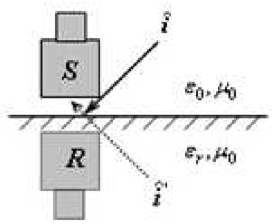


Figure 3: Shadowing effects of the four-path model.

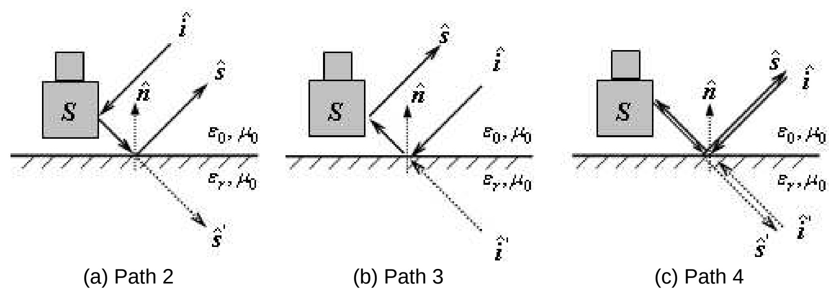


Figure 4: Multi-path scattering problem of the four-path model.

3. NUMERICAL RESULTS

In order to verify the validity of the method, Fig. 5 and Fig. 6 plot the HH-polarized RCS patterns in the $\phi = 0°$ plane of a 1-block target and a 2-block target on a flat sea surface, where the calculation condition is quoted from [4]. Because of the finite flat sea surface area considered in [4], the slight difference is observed between our result and the IPO by [4], but the trend of them agrees well.

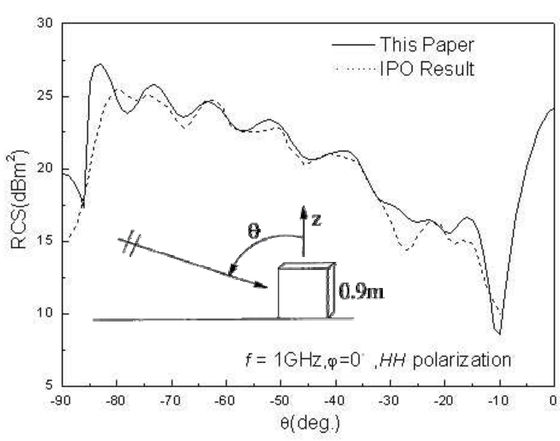


Figure 5: RCS of single PEC cube on the flat sea surface.

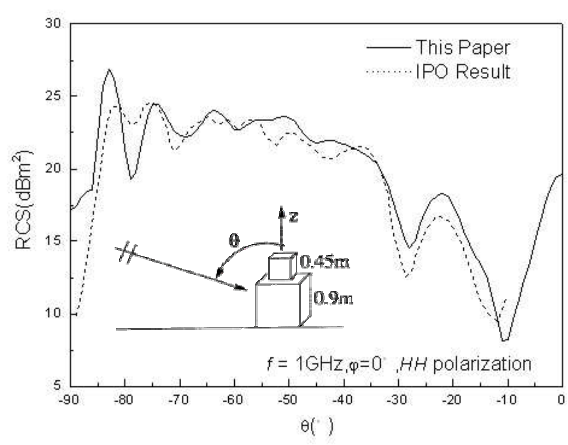


Figure 6: RCS of combined PEC cubes on the flat sea surface.

An aircraft on an airdrome is considered as another example. Assuming that the relative permittivity of the cement ground is $\varepsilon_r = 10 - j2.2$, and a slight roughness is described by $k_0\hat{\sigma} = 0.25$. Fig. 7 depicts the HH-polarized RCS patterns of an aircraft on the airdrome in the $\phi = 0°$ plane. It shows that, as the incidence angle θ_i increases, the interactions between aircraft and airdrome also the multi-path scattering contribution increase.

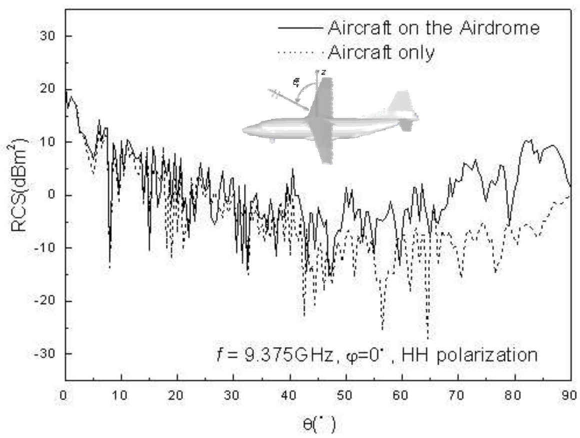


Figure 7: RCS of a aircraft on an airdrome.

4. CONCLUSION

A hybrid approach which combines the "four-path" model with a quasi-image method is developed in this paper to deal with the high frequency EM scattering problems of complex targets above a slightly rough surface. The scattering filed of the slightly rough surface is approximately calculated by modeling the surface as a planar boundary between free space and a homogeneous dielectric medium. And the damped coefficient is introduced to correct the Fresnel reflection coefficients. The computation process of the multi-path scatterings in the "four-path" model is greatly simplified by using the image targets to treat the shadowing effect and the image radar to treat the multi-path effect. The effectiveness and efficiency of this method are verified by numerical results, which shows that the proposed approach can be used for quick and accurate evaluation of the high frequency scattering from complex targets above a slightly rough surface. The EM scattering from complex targets above a rough surface remains as the subject of further study and is ongoing.

REFERENCES

1. Ulaby, F. T., R. K. Moore, and A. K. Dung, *Microwave Remote Sensing, Active and Passive*, Artech House, Norwood, MA, 1986.
2. Johnson, J. T., "A study of the four-path model for scattering from an object above a half space," *Microwave Opt. Technol. Lett.*, Vol. 30, 130–134, July 2001.
3. Johnson, J. T., "A numerical study of scattering from an object above a Rough Surface," *IEEE Transaction on Antennas and Propagation*, Vol. 50, No. 10, 1361–1367, 2002.
4. Burkholder, R. J., P. Janpugdee, and D. Colak, "Development of computational tools for predicting the radar scattering from targets on a rough sea surface," Final Report No. XBONR; ESL-735231-3, ADA388046, Jan. 2001.
5. Michaeli, A., "Equivalent edge currents for arbitrary aspects of observation," *IEEE Transaction on Antennas and Propagation*, Vol. 32, No. 3, 252–258, 1984; Vol. 33, No. 2, 227, 1985.
6. Ling, H., R. Chou, and S. W. Lee, "Shooting and bouncing rays: calculating the RCS of an arbitrarily shaped cavity," *IEEE Transaction on Antennas and Propagation*, Vol. 37, No. 2, 194–205, 1989.

Effects of Different Parameters on Attenuation Rates in Circular and Arch Tunnels

Kamran Arshad, **Ferdinand Katsriku**, and **Aboubaker Lasebae**
School of Computing Science, Middlesex University, London, UK

Abstract— Radio wave propagation in circular and arch shaped cross-section tunnels is analyzed. The electrical and mechanical parameters, such as the shape and transverse dimension of the tunnels, frequency, polarization, electrical parameters like conductivity and permittivity, which have a significant influence on the attenuation inside tunnels are determined. The impact of these parameters on the propagation is analyzed and an optimum set of values is determined. In this paper, taking circular and arch shaped tunnel as an examples, influence of various parameters on UHF radio propagation is analyzed.

1. INTRODUCTION

UHF radio wave propagation in tunnels has been a topic of interest for a long time [1]. Tunnels exists in metropolitan cities and mountainous areas and radio coverage is needed in the tunnels for personal and emergency communications. As a result subject has been very popular, and a few investigations has been carried out in past [5]. Every new city roads includes some tunnel sections or at-least cuttings to reduce noise emission in the environment. In a modern country like Austria, 10% of high priority roads are in tunnels. Common cell planning concepts are not applicable in tunnels due to heavy waveguiding effects, so there are some simplifications introduced in theoretical studies [2].

Traditionally, leaky feeders have been proposed to supply radio services inside tunnels [3]. Leaky feeder cables are expensive, require repeaters at regular distances and maintenance at regular intervals. Their bandwidth is generally equal to one octave and by adjusting the slot configuration, it is even possible to extend this band [4]. However, the attenuation dramatically increases at high frequency, any improvement being only obtained with cables of larger diameters leading to prohibitive cost and weight [4].

Previous experimental studies of radio wave propagation characteristics in tunnel environments were almost entirely concentrated on finding the optimum frequency bands for minimum attenuation [6]. The results showed that the optimum frequency window seems to be between 1–2 GHz. The propagation attenuation in the far region of a straight tunnel is less than that in free space, thus indicating that guided wave phenomena are involved [7]. In general, tunnels may be considered as hollow waveguides surrounded by lossy materials. Radio wave impinges on a wall of the tunnel is partially refracted into the surrounding dielectric and partially reflected back into the tunnel. Characteristics of radio waves in tunnels are influenced by many factors mainly shape of the tunnel (height and width), frequency, polarization, conductivity and permittivity of surrounding material.

Impact of electrical parameters on UHF propagation in straight and circular tunnel has been analyzed in [9]. Usually, cross-section of the tunnel is arched shaped. In arched shaped tunnels, a theoretical treatment of the propagation characteristics is somehow difficult because the boundary does not coincide with a coordinate surface of an orthogonal coordinate system. For this reason, the arched shaped tunnel is modelled in [8] as a circular waveguide having the same cross sectional area. In this paper by considering circular and arched shaped tunnels as examples, we analyze influence of height and width of the tunnel, radio wave frequency, conductivity and permittivity of the tunnel boundary on the attenuation inside the tunnels.

2. ATTENUATION RATE IN TUNNELS

2.1. Circular Tunnel

The attenuation rate, in general, is a function of tunnel dimensions, radio wave frequency, dielectric constant and conductivity of the tunnel surrounding structure. For UHF waves in tunnel environment, the tunnel dimensions are always larger than the free space wavelength. The approximate equations for attenuation rate in circular tunnel for transverse electric (TE) and transverse magnetic mode (TM) is given by [10]:

$$\alpha_{0n} = 8.686\frac{\xi_{1n}^2}{k_0^2a^3}\begin{cases}\mathrm{Re}\frac{1}{\sqrt{\varepsilon_r'-1}} & \mathrm{TE}_{0n}\\ \mathrm{Re}\frac{\varepsilon_r'}{\sqrt{\varepsilon_r'-1}} & \mathrm{TM}_{0n}\end{cases} \tag{1}$$

where ξ_{1n} is the nth nonvanishing root of the first order Bessel function, k_0 is the free-space propagation constant, a is the circular radius in meters, $\varepsilon_r' = \varepsilon_r - j\frac{\sigma}{\omega\varepsilon_0}$ is the complex relative permittivity, ε_r is the relative permittivity of the tunnel wall, ω is the radio wave frequency, ε_0 is the permittivity of free space and σ is the conductivity of tunnel walls.

Approximate solution of the hybrid mode attenuation rate for circular shape tunnel is given as:

$$\alpha_{EH_{mn}} = 8.686\frac{\xi_{m+1,n}^2}{k_0^2a^3}\mathrm{Re}\frac{\varepsilon_r'+1}{2\sqrt{\varepsilon_r'-1}} \qquad \mathrm{EH_{mn}} \tag{2}$$

2.2. Arched Tunnel

Yoshio presented the approximate solution of horizontal and vertical polarized mode for an arched tunnel in [11], for the E_{11}^h and E_{11}^v mode,

$$\alpha_h = K_h\lambda^2\mathrm{Re}\left(\frac{\varepsilon_r'}{d_1^3\sqrt{\varepsilon_r'-1}} + \frac{1}{d_2^3\sqrt{\varepsilon_r'-1}}\right) - K_h\frac{\lambda^3}{2\pi}\mathrm{Im}\left(\frac{\varepsilon_r'^2}{d_1^4(\varepsilon_r'-1)} + \frac{1}{d_2^4(\varepsilon_r'-1)}\right) \tag{3}$$

$$\alpha_v = K_v\lambda^2\mathrm{Re}\left(\frac{1}{d_1^3\sqrt{\varepsilon_r'-1}} + \frac{\varepsilon_r'}{d_2^3\sqrt{\varepsilon_r'-1}}\right) - K_h\frac{\lambda^3}{2\pi}\mathrm{Im}\left(\frac{1}{d_1^4(\varepsilon_r'-1)} + \frac{\varepsilon_r'^2}{d_2^4(\varepsilon_r'-1)}\right) \tag{4}$$

where d_1 and d_2 are the maximum height and width respectively as shown in Figure 1. If (3) and (4) hold true in UHF band, the numerical coefficients K_h and K_v must be constant and independent

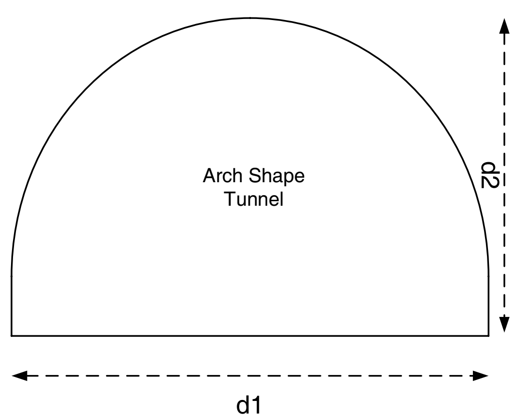


Figure 1: Tunnel dimensions: width d1, height d2.

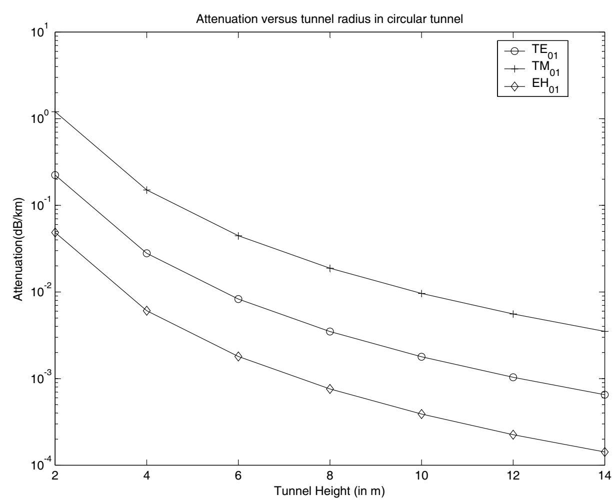


Figure 2: Attenuation versus Radius of a circular tunnel.

of d_1, d_2, λ and ε_r. Numeric values of K_h and K_v have been calculated by using point-marching method as functions of d_1/λ, d_2/λ and ε_r [11]. It has been shown that K_h and K_v approach constant values as d_1/λ and d_2/λ increases, and they are independent of ε_r [11].

3. RESULTS AND DISCUSSION

In this paper, circular and arched tunnels are considered as an example, and influence of various parameters on attenuation rate are analyzed. We consider these two types of tunnels because most practical tunnels are arch shaped tunnels and circular tunnels have close resemblance with arch tunnels. Radius of sample circular tunnel is 4 m and radio wave frequency is suppose to be 1000 MHz.

Figure 2 shows the relation between attenuation rate and radius of the tunnel. Conductivity is assumed to be 0.1 mho/m and relative permittivity is 5. Three modes (TE_{01}, TM_{01} and EH_{11})

are computed and compared. It is clear from Figure 2 that attenuation rate of all modes decreases with increase in tunnel radius, however EH_{11} mode have lowest attenuation rate. Figures 3 and 4 shows the relation between conductivity and relative permittivity of tunnel walls and attenuation in dB/km. It is clear that only in TE_{01} mode, attenuation decreases markedly with increase in conductivity and permittivity, while attenuation in TM_{01} and EH_{11} increases sharply. The certain value of conductivity or permittivity is determined by several factors, such as the radius of the tunnel, the frequency and the permittivity or conductivity etc.

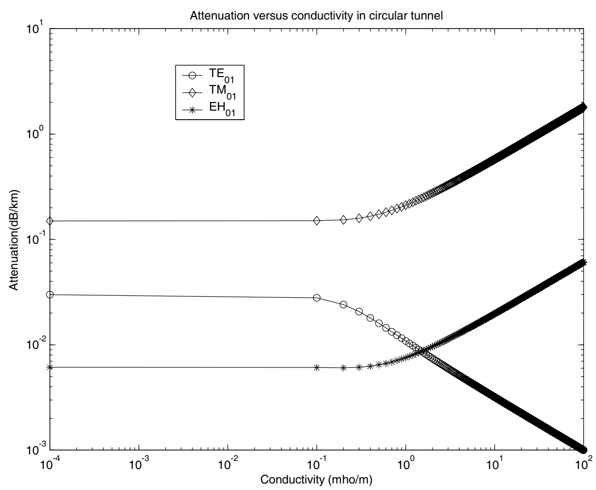


Figure 3: Attenuation versus Conductivity in a circular tunnel.

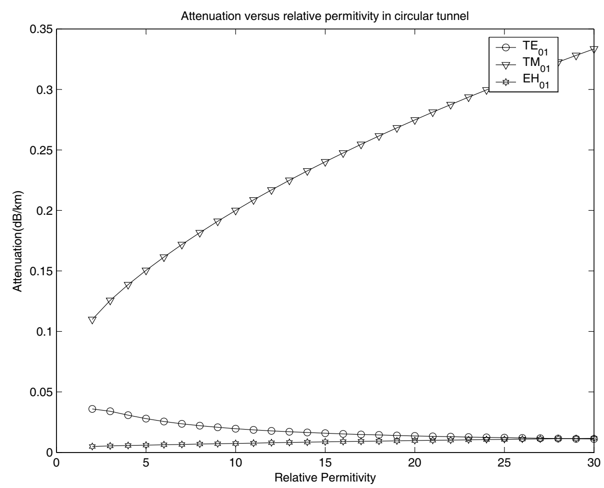


Figure 4: Attenuation versus Relative Permittivity in a circular tunnel.

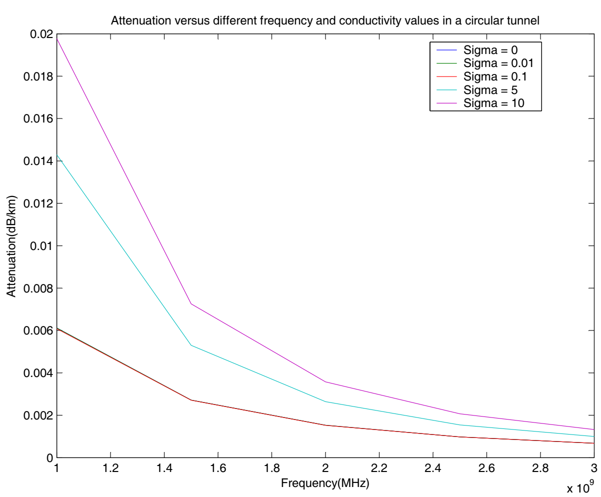


Figure 5: Attenuation versus different Frequencies and Conductivities in a circular tunnel.

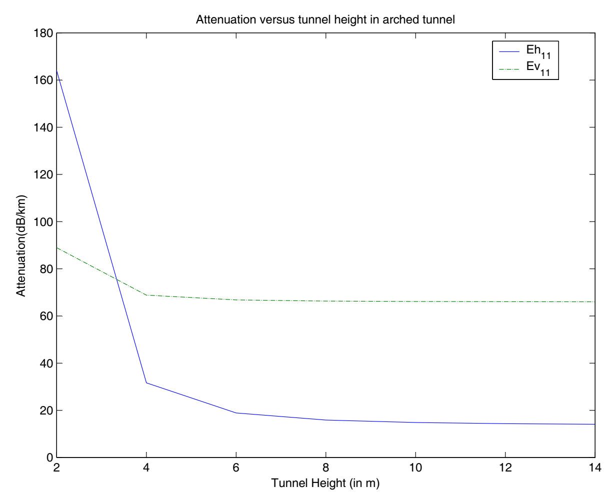


Figure 6: Attenuation versus Tunnel Height in an arched tunnel.

Figure 5 shows the attenuation of EH_{11} mode varying with different frequencies for different values of conductivity σ. We have choose different values of σ, $\sigma = 0, 0.01, 0.1, 5, 10\,\text{mho/m}$ for a wide range of frequencies covering from VHF to UHF. It can be seen that attenuation decreases with increase in frequency, further when $\sigma \leq 1\,\text{mho/m}$, the attenuation is nearly independent of the conductivity, however its different for $\sigma \geq 5\,\text{mho/m}$.

The parameters of the sample arch tunnel are choosen as follows: width 2.5 m, height 4 m, frequency 1000 MHz, conductivity 0.1 mho/m and relative permittivity is 5. Two modes (Eh_{11} and Ev_{11}) are computed and compared. Like circular tunnel, attenuation is plotted versus tunnel height and width in Figures 6 and 7 respectively. From Figure 6, with increase in tunnel height attenuation decreases in both modes, however, Eh_{11} mode decreases very sharply after a certain threshold value of tunnel height. In Ev_{11} mode, change in attenuation with tunnel height is almost constant. This situation is totally different in case of tunnel width as depicted in Figure 7. Ev_{11} mode have high attenuation for lower values of tunnel width and than decreases sharply with increase in tunnel width. For practical values of tunnel width, attenuation in Ev_{11} mode is better than Eh_{11} mode.

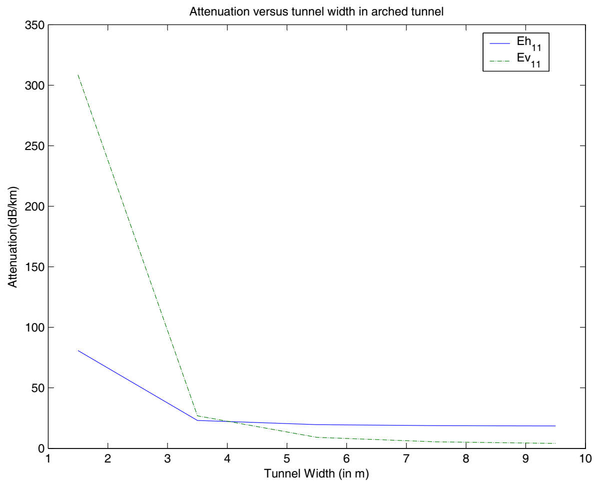


Figure 7: Attenuation versus Tunnel Width in an arched tunnel.

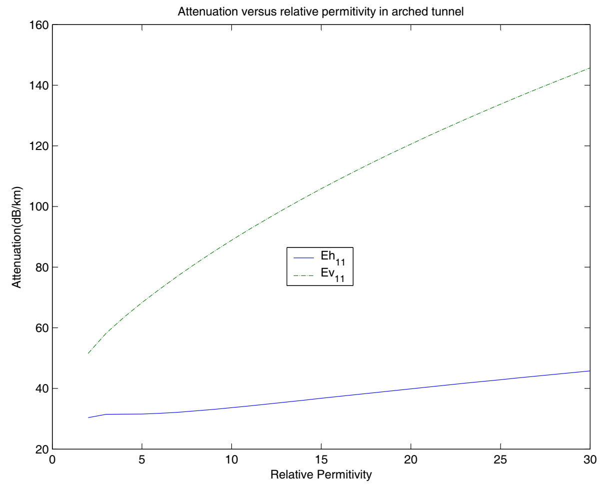


Figure 8: Attenuation versus Relative Permittivity in an arched tunnel.

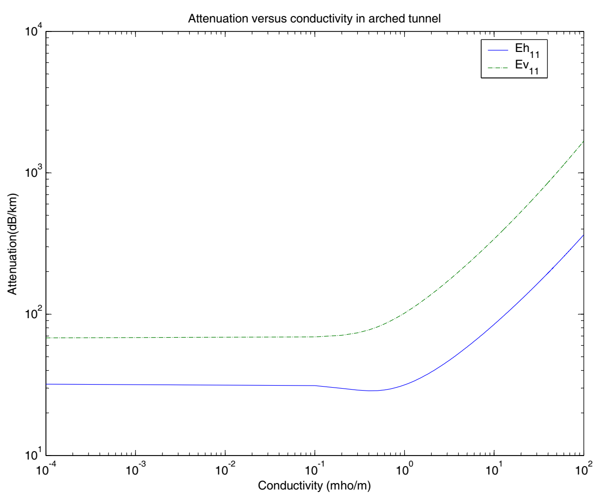


Figure 9: Attenuation versus Conductivity in an arched tunnel.

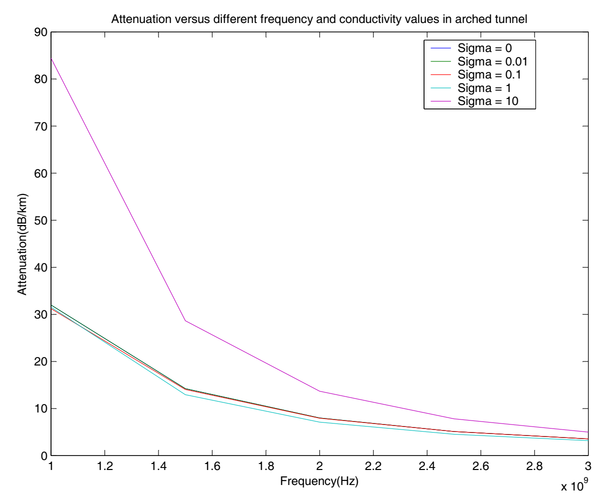


Figure 10: Attenuation versus Frequency in an arched tunnel.

Relation between the relative permittivity and conductivity of arch tunnel walls for two different kinds of mode attenuation are shown in Figure 8 and 9 respectively. Attenuation of Ev_{11} is more than Eh_{11} with relative permittivity and increases sharply. From Figure 9, it is clear that attenuation of two modes are almost constant when conductivity is less than a particular value and than increase sharply, and attenuation of Ev_{11} mode is greater clearly. Same as in circular tunnels, attenuation of Eh_{11} is plotted versus frequency for different values of conductivity in Figure 10. It is seen that like circular tunnel, attenuation is independent of conductivity for smaller values of conductivity however for larger values of conductivity attenuation increases.

4. CONCLUSION

Characteristics of radio waves in tunnels are influenced by many factors which are discussed here in this paper. Conductivity of tunnel walls is very low, so in many tunnels such as underground or mine tunnels, the influence of conductivity can be neglected. However, in some special tunnels like metallic tunnels where conductivity is high, influence of conductivity on attenuation must be considered. Tunnel dimensions also play an important role in attenuation of waves inside tunnels.

Frequency, direction of polarization, permittivity etc. all have influence on the attenuation of signal inside tunnels.

REFERENCES

1. Delogne, P., "EM propagation in tunnels," *IEEE Trans. Antennas Propagat.*, Vol. 39, 401–406, March 1991.
2. Zhang, Y. P. and Y. Hwang, "Theory of the radio-wave propagation in railway tunnels," *IEEE Transactions on Vehicular Technology*, Vol. 47, No. 3, 1027–1036, 1998.
3. Deryck, L., De. R. Keyser, H. Hellin, B. Demoulin, A. Ghetreff, L. Kone, and P. Degauque, "Radio communication in road tunnels — theoretical modles and architecture of leaky feeder system," *Derive Conference*, 1991.
4. Erb, R., "Betrieberfahrungen mit den tunnelfunkanlagen gotthard u. seelisberg," *Technische Mitteilung PTT 4*, 142–148.
5. Mariage, Ph., M. Lienard, and P. Degauque, "Theoretical and experimental approach of the propagation of high frequency waves in road tunnels," *IEEE Trans. Antennas Propagat.*, Vol. 42, No. 1, 75–81, 1998.
6. Parsons, J. D., *The Mobile Radio Propagation Channel*, Pentech, London, U.K., 1992.
7. Zhang, Y. P. and Y. Hwang, "Characterization of UHF radio propagation channels in tunnel environments for microcellular and personal communications," *IEEE Transactions on Vehicular Technology*, Vol. 47, No. 1, 1998.
8. Chiba, J., T. Inaba, Y. Kuwamoto, O. Banno, and R. Sato, "Attenuation constants and phase constants of the tunnels," *Trans. 1987 ISAP Japan*, C-3-4, 389–392, August 1987.
9. Sun, J., L. Cheng, and X. Liu, "Influence of electrical parameters on UHF radio propagation in tunnels," *5th intl Symposium on Multi-Dimensional Mobile Communications*, Vol. 1, 436–438, August 2004.
10. Degauque, P., "Leaky feeders and subsurface radio communications," *IEE Electromagnetic Wave Series*, London, Peregrinus Series, 1982.
11. Yamaguchi, Y., T. Abe, T. Sekiguchi, and J. Chiba, "Attenuation constants of uhf radio waves in arched tunnels," *IEEE Transactions on Microwave Theory and Techniques*, Vol. MTT-33, No. 8, August 1985.

Design and Characterization of Biophotonics Devices

M. Rajarajan, C. Themistos, B. M. A. Rahman, and K. T. V. Grattan
Bio/Photonics Research Centre, School of Engineering and Mathematical Sciences
City University, Northampton Square, London EC1V 0HB, United Kingdom

Abstract— In this paper we study the design and characterization of optical biosensors based on the integrated optics technology. The finite element based perturbation technique has been used to characterize the surface plasmon waveguide. The finite element based beam propagation method has been used to characterize the ring resonators. Some of the novel biosensor structures will be characterized and modal loss and sensitivity and selectivity of these biosensors are studied in detail.

1. INTRODUCTION

There is a significant need today for the better detection and analysis of chemical and biochemical substances, impacting on many important areas of life, including medicine, environmental monitoring, biotechnology, drug discovery and food monitoring and several major systems have been undertaken to study the interactions of bio-molecules. Such analysis methods can provide a clearer understanding of how proteins, encoded by DNA, interact with enzymes, inhibitors, or other proteins. Most of these systems use a reporter method to determine if two or more molecules interact. As a result these conventional methods of pathogen detection require a number of time-consuming steps to arrive at a useable measurement. However, the development and use of more effective biosensor technology could significantly reduce this time, as well as allow the detection of even smaller amounts of pathogens with fewer false positives.

Thus the field of *biosensors* has emerged as a topic of strong scientific interest and economic significance because of this requirement for better measurements in the above areas, added to which is the worldwide concern from the ever-present threats of chemical and bio terrorism. The constant health danger posed by new strands of microbial organisms and the spread of infectious diseases is another major concern requiring effective biosensing for detecting and identifying them to allow rapid remedial action. Biosensors offer several advantages over laboratory based methods, such as *in situ* real-time process monitoring and a high sensitivity to surface modifications, where most of the bioprocess takes place. Although biosensors using surface plasmon resonance (SPR) techniques, ellipsometry, evanescent wave methods, output grating couplers, and reflectance interference spectroscopy are adequately sensitive, they are still quite slow, often bulky, have limited functionality and are usually expensive.

Biosensors represent an application field outside communications technology where integrated optics is expected to play an increasing role and where it is already success commercially. The sensing is performed by the *evanescent tail* of the modal field into the bio-medium. This sensing operation consists of measuring the change of the effective index of a propagating mode when a change of refractive index takes place on the outer surface of the waveguide. The sensitivity of the measurement of the physical or chemical quantity present on the top surface of the waveguide structure depends on the strength and the distribution of the evanescent field in the top surface. The main design task in creating effective biosensors is therefore to find the waveguide structure which maximises the sensitivity on the quantity to be measured. There are several techniques that are currently being used to realise these biosensors. Photonic techniques offer significant possibilities for the development of such devices. One procedure is to make use of spectroscopic techniques for the precise spectrochemical analysis of the chemical species that are present in the biological material. Another is to allow the biological material to fall onto a sensor element, the properties of which are then detected optically. A further, commonly used technique uses the surface plasmon resonance (SPR) principle. It had already been used for several years to investigate organised organic mono- and multilayers on metal surface. A practical and effective method by which to excite the surface plasmon was initially suggested by Kretschmann in 1971. At the beginning of the 1980s Bo Liedberg *et al.*, demonstrated that SPR in the Kretschmann configuration is well suited for both gas and biomolecular sensing purposes. Today SPR-based sensors are manufactured by Biacore, Graffinity Pharmaceuticals, IBIS, Nippon Lasers, and Prolinx for example. There are also resonant mirror based biosensors manufactured by Affinity Sensors commercially available in the market. The waveguide-based biosensors are the attractive candidates for the future and are manufactured and commercialised in the UK by our partner Farfield Sensors, and in Europe by Artificial Sensing Instruments.

However, the analysis of many complex binding events requires multiplexed detection systems that can be used to analyse many binding interactions simultaneously. Commercially available optical biosensors have been limited in their applications due to their small number of surfaces or spots that could be sensed simultaneously. This limitation can be overcome by using integrated optical waveguide biosensors based on the microring resonators with multichannel sensing devices on one chip with the potential for efficient referencing and multicomponent sensor analysis of complex samples. Microring resonators have been proposed as sensitive chemical sensors and biosensors due to their large Q factors. In these devices the high sensitivity is achieved because the light wave interacts many times with each pathogen as a consequence of the resonant circulation of light within the ring structure. The greater the proportion of guided power that passes through the analyte, the greater the sensitivity of the guided modes to changes in the analyte. Specificity of the detected substance can be achieved when a layer of antibodies or other binding materials is deposited onto the active area of the resonator.

This paper aims to develop new biosensor techniques, targeting determining and optimising the *sensitivity* and *specificity* of these biosensors for detecting multiple analytes. Microring resonators are used in communications applications but are new to most areas of sensing. The overall operating performance of the biosensor system will be studied and optimised in the design phase by using powerful and numerically efficient finite element time domain based full vectorial methods.

2. RESULTS

Initially the optical properties of metal-clad optical fiber with a finite cladding thickness, t is studied and effect of the outer cladding materials such as silica, air or acetone on the optical field have been considered. The variation of optical parameters of such structures due to the presence of materials that affect the modal field distribution can be utilized in optical sensor applications. Initially the variation in the effective index of the fundamental RP_{01} SPM, for a fiber diameter of $4\,\mu$m, covered by a metal layer and further covered by silica, air or acetone as the outer cladding, has been studied and presented in Fig. 1. From the above characteristics, it can be seen that for a larger metal thickness the effective index of the RP_{01} optical mode with finite metal thickness is about the same as the effective index of the infinite cladding structure and is not affected by the materials in the outer cladding. For a large metal thickness, t, two SPMs exist at the two dielectric/metal circular boundaries, and when the two dielectric materials are different, their propagation constants are different, and therefore they do not interact to form a supermode. However, as the metal thickness, t, is reduced there is a stronger coupling between the SPMs and two supermodes, with odd- and even-like maximum field intensities at the two interfaces being formed. The first supermode has a higher propagation constant and is confined near the metal interface with the higher refractive index (in this case the inner interface) and has an odd-like field profile. On the other hand the second supermode is confined near the metal interface with the lower refractive index (in this case the outer interface) and has an even-like field profile. The even-like second supermode is dominated by the effective index of the outer metal dielectric interface, and being close to cut-off, as the metal thickness decreases, the even-like mode becomes unbounded. As can be seen from the effective index curves, shown in Fig. 1, when the metal thickness, t, decreases below 50 nm, the effective index of the odd-like supermode increases rapidly. For the waveguides with acetone or air as the outer cladding, the effective index of the even-type SPM is below the cut-off and therefore is not presented here.

Next, a metal-clad fiber with $D = 4\,\mu$m and a finite metal thickness, t, with acetone material around the metal cladding has been considered. The field profile of the fundamental H_x odd-like optical supermode along the y-axis, for different values of the metal thickness, t, have been investigated and are shown in Fig. 2. As can be seen from this figure, for a metal thickness of $0.2\,\mu$m (as shown by a solid line), the field profile has similar properties to those obtained for an infinite cladding fiber, where the field has a maximum intensity at the dielectric/metal interface and decays rapidly in the metal region, except a small negative peak is clearly visible at the copper/acetone interface, due to the weak coupling between the two non-synchronous SPMs. As the metal thickness decreases, the maximum field intensity increases anti-symmetrically at both the interfaces, with the highest maximum being at the silica/copper interface, and exhibiting the properties of an odd-like bounded and strongly coupled SPM in planar dielectric/metal/dielectric

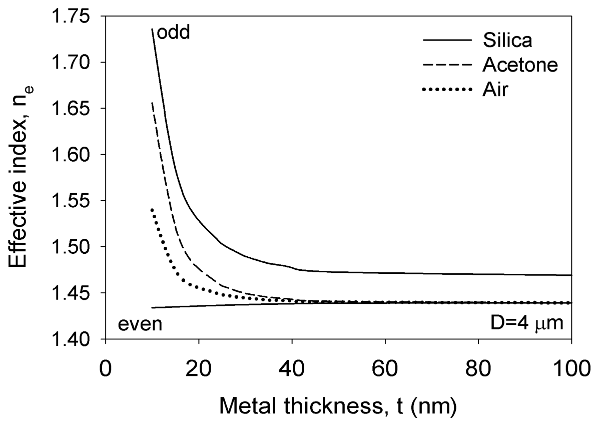


Figure 1: Effective index variation with metal thickness for $D = 4\,\mu$m.

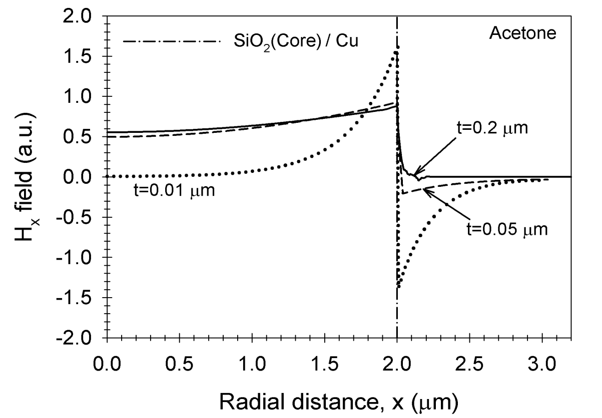


Figure 2: H_x field profiles along the y-axis with metal thickness for acetone in the outer cladding.

waveguide. The field profile is shown by a dotted line for a metal thickness, $t = 0.01\,\mu$m.

Figures 3 and 4 shows the inner and outer radially polarized surface plasmon modes at $D = 4\,\mu$m. As can be seen from these figures the inner mode is more confined and the outer mode is more spread and much more circular. The superimposed field intensity from the H_x and H_y fundamental supermodes form the radially polarized RP01 mode, for a metal-clad silica fibre with a diameter, $D = 4\,\mu$m and a metal thickness, $t = 0.01\,\mu$m and a surrounding index of 1.34 is presented here. From Figure 4 it can be seen that the optical field decays in both the centre of the fibre and the outer cladding, with the anti-symmetric peak field intensities at the inner silica/copper (where the field is positive) and small negative peak (but this is not visible) at the outer copper/air interface.

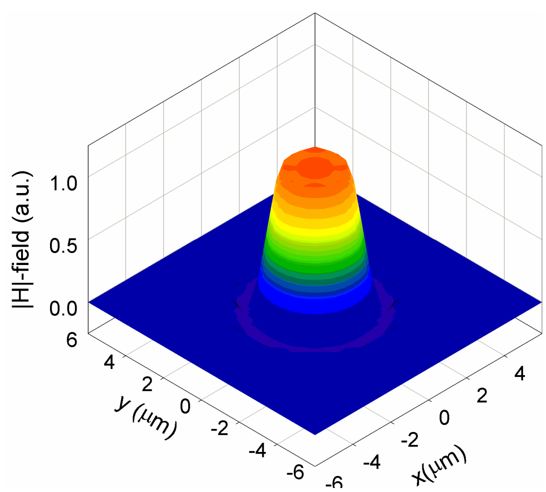


Figure 3: Inner radially polarized surface plasmon mode at $D = 4\,\mu$m.

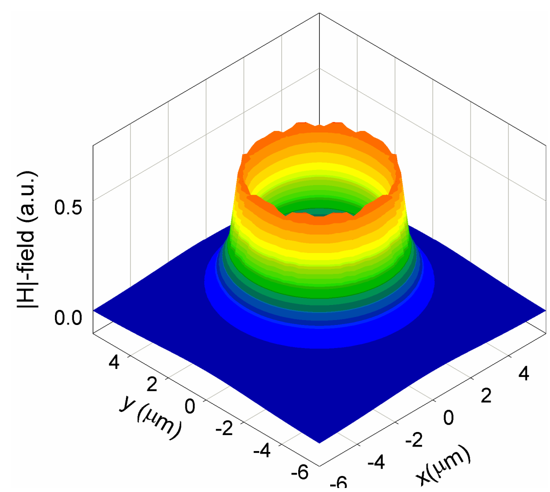


Figure 4: Outer radially polarized surface plasmon mode at $D = 4\mu$m and $n_s = 1.34$.

Figure 5 shows the effective index with the surrounding medium refractive index. In this case the surrounding medium index is varied from 1.30–1.42 to cover most of the materials that can be used in the biosensing applications. As can be seen from Figure 5 the outer surface plasmon effective index changes linearly with the increasing index values. As the metal thickness is small as the index value is increased it almost looks like a fibre with an infinite cladding. However, as the inner mode does not see the surrounding index the refractive index of the inner mode is not affected.

Next the attenuation characteristics of the above modes have been calculated using the vector **H**-field Finite Element Method with perturbation and the variation of the normalized attenuation constant, α/k_0 with the surrounding refractive index is presented for a fibre diameter of $4\,\mu$m in Fig. 6. From this figure it can be seen that as the refractive index of the surrounding medium is increased the modal loss also increases as the mode starts to penetrate more into the surrounding medium. However the modal loss in the inner silica/copper boundary is constant over a range of surrounding index as the field does not see the outside index. The modal loss here is mainly due to the metal loss at the silica/metal inner interface. It can clearly be seen that as the outside index approaches that of the silica index the modal loss is at its maximum. However, to design practical biosensing devices a compromise must be found between the modal loss and the amount of penetration that is needed within the surrounding index medium. This can be attributed to the sensitivity of the biosensor that is needed for different bio medical applications.

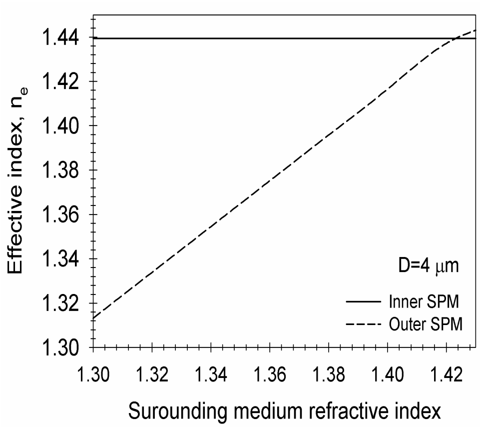


Figure 5: Effective index with surrounding medium (n_s) refractive index.

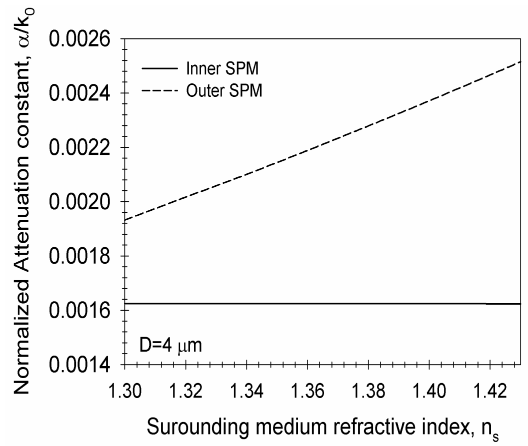


Figure 6: Attenuation with surrounding refractive index.

3. CONCLUSION

A finite-element variational-based vector formulation, in conjunction with the perturbation technique has been used to study detailed optical properties, such as the optical mode field distribution, the effective index and the attenuation constant of a metal-clad silica fiber, with finite metal-clad thickness surrounded by another outer cladding. The variations of the above optical properties with the change of the surrounding materials are very important in several applications, such as optical fiber sensors and biosensors. It has been shown here that a circularly symmetric metal-coated guide supports both vertically and horizontally polarized waves. However, due to the imposition of the electric-wall boundary conditions, at the vertical or horizontal interfaces, the resultant field H_y and H_x field profiles do not have rotational symmetry. However, these modes being degenerate, their superposition produces the typical RP_{01} mode, which has rotational symmetry. By using a segmented metal cladding, the rotational symmetry can be broken, when the vertically and horizontally polarized waves would not degenerate, and a high modal birefringence or polarization-maintaining waveguide thus can be fabricated. For a finite metal thickness, the SPMs exist at both the inner and outer interfaces. By adjusting the metal thickness and refractive index values of the cladding layers, the odd and even-type coupled SPMs can be formed and exploited for various bio sensing appplications.

REFERENCES

1. Homola, J., S. S. Yee, and G. Gauglitz, "Surface plasmon resonance sensors: review," *Sensors and Actuators B*, Vol. 54, 3–15, 1999.
2. Luff, B. F., R. D. Harris, J. S. Wilkinson, R. Wilson, and D. J. Schiffrin, "Integrated optical directional coupler biosensor," *Optics Lett.*, Vol. 21, 618–620, 1996.
3. Luff, B. J., J. S. Wilkinson, J. Piehler, U. Hollenbach, J. Ingenhoff, and N. Fabricius, "Integrated optical Mach-Zehnder biosensor," *J. Lightwave Technol.*, Vol. 16, No. 4, 583–592, 1998.
4. Heideman, R. G., R. P. H. Kooyman, and J. Greve, "Performance of a highly sensitive optical waveguide Mach-Zehnder interferometer immunosensor," *Sensors and Actuators B*, Vol. 10, 209–217, 1993.
5. Kinrot, N., "Analysis of bulk material sensing using a periodically segmented waveguide Mach-Zehnder interferometer for biosensing," Vol. 22, No. 10, 2296–2301, 2004.

6. Mateus, C. F. R., et al., “Compact label-free biosensor using VCSEL-based measurement system,” *Photon Technol. Lett.*, Vol. 16, No. 7, 1712–1714, 2004.
7. Blair, S. and Y. Chen, “Resonant-enhanced evanescent-wave fluorescence biosensing with cylindrical optical cavities,” Vol. 40, No. 4, 570–582, 2001.
8. Boyd, R. W., et al., “Nanofabrication of optical structures and devices for photonics and biophotonics,” *Journal of Modern Optics*, Vol. 50, No. 15–17, 2543–2550, 2003.

Plasmonics in Metal-clad Terahertz Waveguides

Muttukrishnan Rajarajan, Christos Themistos, B. M. A. Rahman, and Kenneth T V Grattan
Bio/Photonics Research Group School of Engineering and Mathematical Sciences
City University, Northampton Square, London EC1V 0HB, UK

Abstract— Finite element analysis, based on the vector **H**-field formulation and incorporating the perturbation technique, is used to calculate the complex propagation characteristics of metal-coated dielectric waveguides at THz frequencies. The propagation and attenuation characteristics of the surface plasmon modes at the metal/dielectric interfaces are presented. The effects on the modal properties of metal-clad dielectric guides with infinite and finite cladding thickness and the formation of the supermodes due to the coupling between the surface plasmon modes in the presence of different surrounding materials are also investigated.

1. INTRODUCTION

The terahertz (THz) region occupies a large portion of the electromagnetic spectrum located between the microwave and optical frequencies and normally is defined as the band from 0.1 to 10 THz. In recent years, this intermediate THz radiation band has attracted a lot of interest, because it offers significant scientific and technological potential for applications in many fields, such as in sensing [1], in imaging [2] and in spectroscopy [3].

Recent progress in sources [4] and receivers of THz waves has generated much interest in studying the waveguiding properties of these waves, both as a part of active or passive components, such as lasers, detectors, or filters, and also to connect various components in a system. Suitable waveguides could be used to direct these beams to the correct locations, for example around bends or corners and such guided-wave technologies offer the possibility of compact sensor systems. However, waveguiding in this intermediate spectral region is a major challenge. Amongst the various THz waveguides that have been suggested, the metal-clad waveguides supporting surface plasmon modes show the greatest promise as low-loss waveguides for use both in active components and as passive waveguides.

2. THEORY

The surface plasmon mode (SPM) is essentially the electromagnetic wave that is located at the metal-dielectric interface because of the interaction with the free electrons of the conductor. Surface plasmon resonances in a metallic layer incorporated inside an optical waveguide structure have been extensively used for various fiber-optic and optoelectronic devices, such as optical polarizers [8], fiber-optic sensors [9], and scanning microscopy [10].

In the design of any waveguide, the key modal parameters are their propagation constants, loss coefficients, the modal field profiles and the dispersion properties. First of all, it is essential to develop a modal solution approach which can provide this information for practical THz waveguides with arbitrary shape, size, and material profiles. In the analysis of THz waveguides incorporating metallic films and the interaction of the metallic films with dielectric materials, in order to accommodate guided waves, is considered to be important for the accurate design of various THz devices which may be used for a range of applications. Practical metallic elements are not perfect conductors, but suffer a small amount of loss and therefore a rigorous model, which accounts for the modal loss is essential [11]. Cao and Jahns [12] have used a field expansion approach, such as the Bessel function expansion and matching the field continuity at the metal dielectric interface of axially symmetrical copper wire. Gallot *et al.* [13] have applied the classical Sommerfeld waveguiding principle to metal waveguides operating in the millimetre wave region, as well as mode expansion and field matching techniques and also a more versatile finite-difference time-domain (FDTD) technique for surface plasmon modes. General field expansion techniques are not sufficiently versatile and cannot be used for waveguides with irregular shapes. The alternative FDTD approach is computationally very expensive. More recently, Deibel *et al.* reported [14] a three-dimensional time-harmonic simulation by using the finite element method, but being a three-dimensional problem this approach requires very large computational resources, such as 20 hours of CPU time on a 64-bit dual processor with 16 GB of RAM.

In the present work, the **H**-field FEM based full-vector formulation, in conjunction with the perturbation technique, has been used for the solution of the metal-clad waveguide modes, where the transverse and longitudinal magnetic field components are analyzed with respect to the rectangular coordinates. Therefore, the waveguide modes are initially presented in terms of the transverse magnetic field components, H^x_{mn} and H^y_{mn}, as commonly used for integrated optical waveguide problems, where the m and n subscripts denote the field maxima along the x- and the y-axes, respectively. In this notation, as an example, for the H^y_{mn} mode (also known as the quasi-TE mode), the H_y (or E_x) field is dominant compared to the non-dominant H_x (or E_y) field component. Furthermore, following the classification of optical modes in a step index fiber for weakly guiding fibers, certain hybrid modes can be grouped together in combinations that are degenerate (have same propagation constant or phase velocity) and a linear combination of these modes can be used to form a new set of modes, called linearly polarized LP_{pq} modes, where the subscripts p and q denote the number of azimuthal pairs of maxima and radial maxima, respectively.

3. RESULTS

Initially, a silicon tube is considered, with a copper coating inside, as shown in Fig. 1. In this case, D is the diameter of the inner air-core and a and t are thicknesses of the outer silicon tube and the thin metal layer, respectively. The refractive index of the air-core and the silicon tube are taken as n_g=1.00 and n_a= 3.4205267 (ε_r = 11.7), respectively, and the complex refractive index of the copper cladding is given by $n_m = 438 - j494$, at an operating frequency of 1 THz.

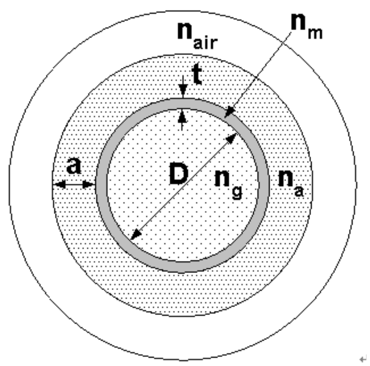


Figure 1: Cross section of a silicon tube, with inner copper coating, at 1 THz.

A two-fold symmetry has been employed in the present analysis, where only a quarter of the waveguide cross-section has been divided into 100 and 120 azimuthal and radial divisions, respectively, thus forming a mesh of 23900 first order triangular elements. The polar coordinate discretization used here matched very accurately the circular cross sectional area of the waveguide core, with the percentage error for the core area being only 0.004%. It takes about 28 seconds of CPU time to obtain a single modal solution on a 3.4 GHz Pentium processor.

There are two metal/dielectric interfaces, one at the outer copper/silicon boundary and the other at the inner copper/air-core boundary. For this guide, the dominant H_x field at the upper and lower metal-dielectric interfaces is tangential to these boundaries, which satisfies the electric-wall boundary condition, $\mathbf{n} \cdot \mathbf{H} = 0$, and supports a surface plasmon mode (SPM) along these metal/dielectric interfaces. The refractive index of the inner and outer cladding materials being very different, the two SPMs have widely different propagation constants, and they do not interact with each other. Variations of the H_x fields along the y-axis for both the SPMs are shown in Fig. 2. In this case, the air-core diameter, $D = 8$ mm, the thickness of the silicon tube, $a = 1$ mm and the thickness of the copper coating, $t = 0.5\,\mu$m. A solid line shows the first SPM at the outer copper/silicon interface, where the field reduces almost linearly inside the 1 mm thick silicon layer and very rapidly inside the very thin 0.5 μm metal layer. The inset on the right shows the rapid field decay near the outer metal/silicon interface. A dashed line shows the variation of the H_x field along the y-axis for the second SPM at the inner copper/air interface. It can be observed that the H_x field decays very slowly in the air-core region and then extremely rapidly inside the metal layer. The rapid decay of the H_x field in the inner metal/air interface is also shown by a dashed line in the inset on the left hand side. It should be noted that in this example, the metal thickness is only 0.5 μm, which is at least three orders of magnitude smaller than the other dimensions. However, at the right and left hand sides of the metal/dielectric interfaces, when the same electric-wall boundary condition is imposed, this forces the H_x field to be zero at the metal boundary and no SPM exists at the vertical walls. The modal field profile along the x-axis (not shown here) has its maximum field intensity at the center of the waveguide and gradually decreases along the radial distance, reaching a zero value at the copper boundary.

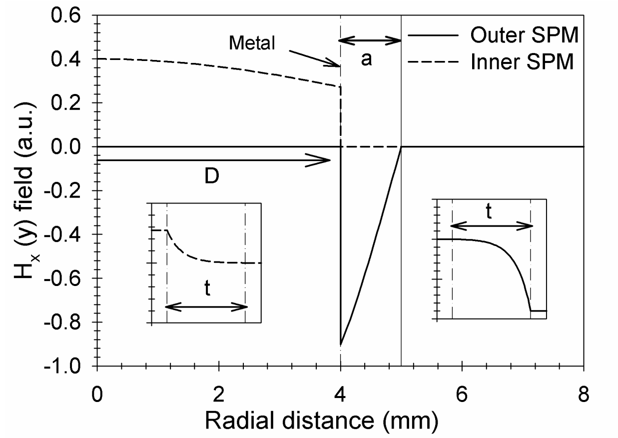


Figure 2: H_x field profile along the y-axis for the even and the odd surface plasmon mode in a hollow silicon tube, with $a = 1$ mm, $D = 8$ mm and inner copper coating, $t = 0.5\,\mu$m.

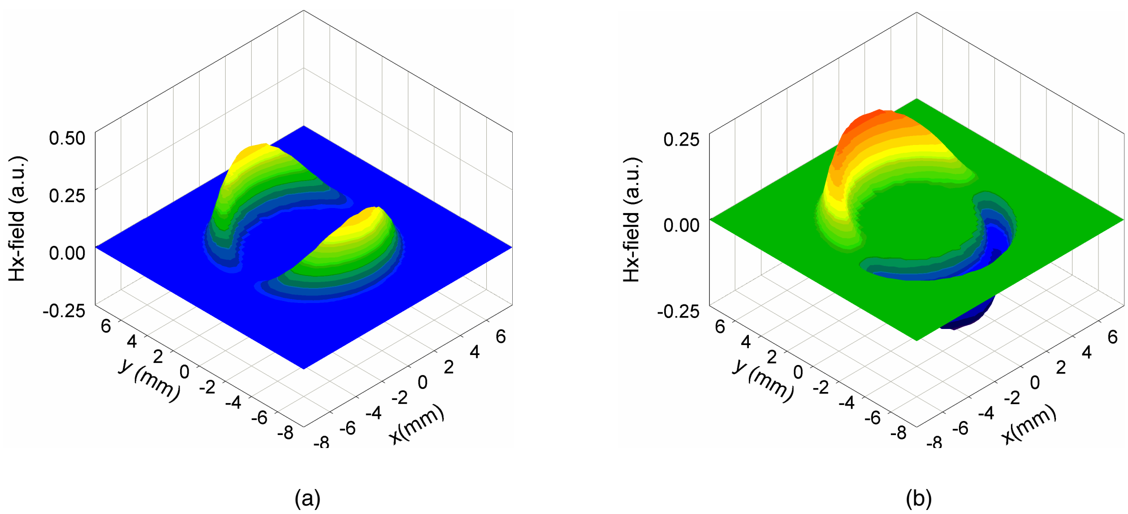


Figure 3: (a) The even-like and (b) the odd-like H_x SPSM at the outer Cu/Si interface, for an inner air-core diameter of $D = 8$ mm.

The 3-D contour profile of the H_x field for the first SPM mode at the outer metal/silicon interface, H_x^{11}(outer) and even-like, is shown in Fig. 3. It can be seen that the field intensity is maximum at the upper and lower interfaces, where the electric-wall boundary condition allows the H_x component to have its maximum value. On the other hand, as mentioned earlier, the same electric-wall boundary condition forces the H_x field to be zero on the left-hand and the right-hand side electric walls. Although for this structure, a rotational symmetry exists, due to the non-identical boundary conditions for a given H_x or H_y field in the vertical and horizontal directions, the modal field profile is no longer circularly symmetric. The SPMs decrease slowly in the inner or outer cladding as shown in Fig. 2, but more rapidly inside the metal layer. For a larger metal thickness, t, the field at one of the interfaces does not extend up to the adjacent interface. So, in this case, although two SPMs exist at the two dielectric/metal circular boundaries, when the two dielectric materials are different, their propagation constants are also different, and therefore they do not easily interact to form a supermode. However, as the metal thickness, t, is reduced the coupling between the SPMs increases and two supermodes, with odd- and even-like maximum field intensities at the inner and outer interfaces may be formed. On the other hand, identical SPMs exist at the upper and lower interfaces and form even and odd-type supermodes with respect to the x-axis. The H_x field profile of this surface plasmon supermode (SPSM), shown in Fig. 3(a), follows an even-function like profile along the y-axis, and can also be termed as an even-like SPSM. Recently, Deibel *et al.* [11] have also reported similar horizontal or vertical 'donut' shaped field electric field profile for the horizontally or vertically polarized wave in a metal wire waveguide. Results presented by Wang and Mittleman [15] also show a similar dipole shaped vertically polarized (E_y or H_x) wave, confined in the upper and lower interfaces of a solid metal wire guide. Further, the H_x field profile of the odd-like SPSM following an odd-like function profile along the y-axis, where the two SPMs are formed at the upper and the lower outer dielectric/metal boundaries, also termed as H_x^{11}(outer) and odd-like, is presented in Fig. 3(b). Propagation constants of these two even and odd-like SPMs are similar.

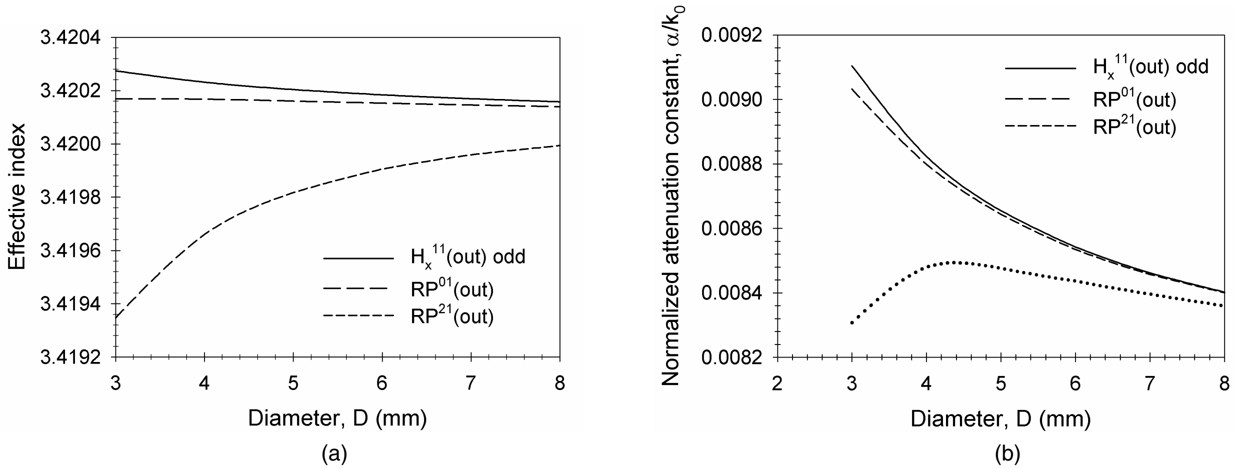


Figure 4: (a) Effective index and (b) attenuation constant variation with air core diameter for the outer SPMs.

Next the attenuation characteristics of the above modes have been calculated using the vector **H**-field FEM with perturbation and the variation of the normalized attenuation constant, α/k_0 with the fiber diameter is presented in Fig. 4(b). From this figure it can be seen that as the fiber diameter increases, the normalized attenuation constants odd-like H_x^{11} and the RP^{01} modes decrease. The attenuation characteristics for the lower order RP^{21} SPM exhibit a maximum attenuation constant for a fiber diameter of about 4.2 mm. The accuracy of the normalized attenuation curves is considered to be within the accuracy limits of the perturbation approach, as discussed by Themistos *et al.* [16].

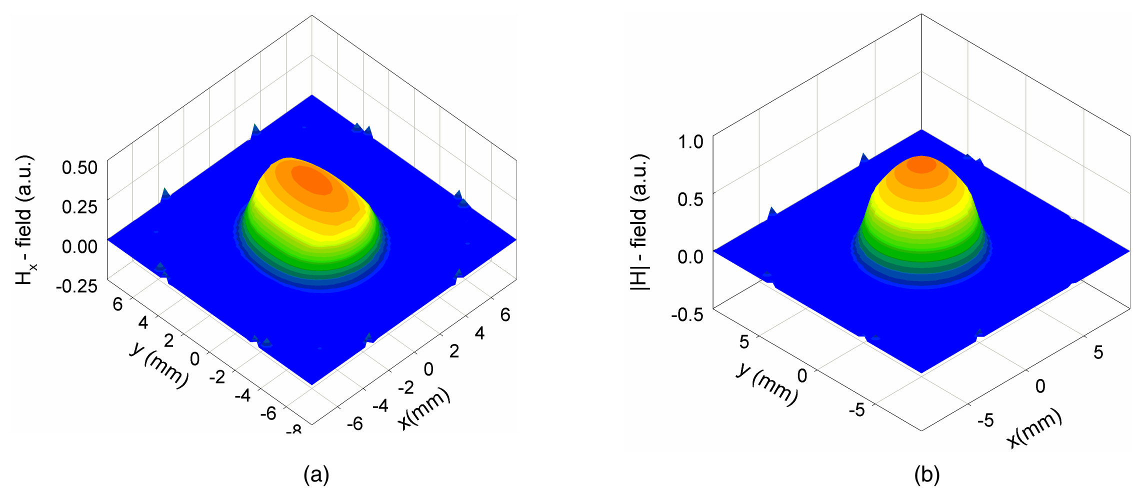


Figure 5: (a) The even-like H_x SPSM and (b) the superimposed RP^{01}-like SPsM at the inner copper/air for an inner air core diameter of $D = 8$ mm.

As mentioned earlier, similar SPMs exist at the inner metal/air interfaces and form an even-type supermodes with respect to the x-axis. The H_x field profile even-like SPSM, exhibiting two field maxima at the upper and at the lower interfaces, is shown in Fig. 5(a). Due to the lack of phase matching it would not form a supermode by interacting with the first SPM at the outer metal/silica interface. However, for this mode the field changes very slowly along the y-axis and has a larger value at the center of the waveguide. The electrical-wall boundary condition forces the H_x field to be zero at the left and right hand side electric walls, so the H_x modal field is also no longer circularly symmetric and appears elleiptical in shape. Similarly another degenerate mode, but with a dominant H_y field, also forms a SPM at the left and right-side inner metal/air interfaces. Its field profile is similar to the H_x field profile, but rotated by 90 degrees. These two modes, being degenerate, can be superimposed to form a circularly symmetric mode. However, vector combination of the H_x and H_y modes gives a resultant field along the radial directions and can be called a radially polarized RP_{01}-like mode, as shown in Fig. 5(b). However, this mode with an almost constant field value inside the air-core region is quite different from the mode shown in Fig. 3(b) at the outer interface. This mode would also be different from the linearly polarized LP_{01} mode in an optical fiber, which can have an almost uniform field profile inside the core region, as shown below.

4. CONCLUSION

A finite-element approach based on a full-vectorial **H**-field formulation, in conjunction with the perturbation technique, has been used to study for the first time the detailed modal properties, such as the mode field distribution, the effective index and the attenuation constant of a metal-clad silicon tube at THz frequency. The origin of the SPM at the inner and outer metal/dielectric interfaces, their phase matching, and the formation of the odd-like and even-like supermodes at these metal/dielectric interfaces are discussed. Subsequently, such odd- or even-like SPSMs at the upper and lower interfaces also couple to forms a more complex coupled supermode. By adjusting the metal thicknesses and the refractive index values of the cladding layers, the odd and even-type coupled SPMs can be formed and potentially exploited for various sensing, imaging and communication systems.

It has been shown here that a circularly symmetric metal-coated guide supports both vertically and horizontally polarized waves. However, due to the imposition of the electric-wall boundary conditions, at the vertical or horizontal interfaces, the resultant H_y and H_x field profiles do not have the rotational symmetry. However, these modes being degenerate, their superposition produces a typical radially polarized RP_{mn} mode, which has rotational symmetry. It is also shown here that for the air-core metal-clad silicon tube, the SPSM at the inner interface may have a uniform modal field profile inside the air-core region. By controlling the waveguide dimensions, a flat field profile with a sharper decay at the inner metal interface can be produced, which could be suitable for near field scanning microscopy. On the other hand, by controlling the metal thickness and adjacent dielectric cladding regions, phase matching can be controlled for various applications, such as THz sensing. By using a segmented metal cladding, the rotational symmetry can be broken, creating a situation where the vertically and horizontally polarized waves would not degenerate, and a high modal birefringence or polarization-maintaining waveguide thus can be fabricated.

REFERENCES

1. Jacobson, R. H., D. M. Mittleman, and M. C. Nuss, "Chemical recognition of gases and gas mixtures with terahertz waves," *Optics Lett.*, Vol. 21, 2011–2013, 1996.
2. Chen, Q., Z. Jiang, G. X. Xu, and X. C. Zhang "Near-field terahertz imaging with a dynamic aperture," *Optics Lett.*, Vol. 25, 1122–1124, 2000.
3. Zhang, J. and D. Grischkowsky, "Waveguide terahertz time-domain spectroscopy of nanometer wave layers," *Optics Lett.*, Vol. 29, 1617–1619, 2004.
4. Kohler, R., A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Richie, R. C. Iotti, and F. Rossi "Terahertz semiconductor-heterostructure laser," *Nature*, Vol. 417, 156–159, 2002.
5. Sun, L. and G. L. Yip, "Analysis of metal-clad optical waveguide polarizers by the vector beam propagation method," *Applied Optics*, Vol. 33, 1047–1050, 1994.
6. Esteban, Ó., R. Alonso, M. C. Navarrete, and A. González-Cano, "Surface plasmon excitation in fiber-optics sensors: A novel theoretical approach," *J. Lightwave Technol.*, Vol. 20, No. 3, 448, 2002.
7. Novotny, L., D. W. Pohl, and P. Regli, "Light propagation through nanometer-sized structures: the two-dimensional aperture scanning near-field optical microscope," *J. Opt. Soc. Am. A*, Vol. 11, 768–1779, 1994.

8. Themistos, C., B. M. A. Rahman, and K. T. V. Grattan, “Analysis of lossy TE/TM modes in metal-clad optical waveguides,” *Applied Optics*, Vol. 37, No. 24, 5747–5754, 1998.
9. Cao, Q and J. Jahns, “Azimuthally polarized surface plasmons as effective terahertz waveguides,” *Opt. Express*, Vol. 13, 511–518, 2005.
10. Gallot, G., S. P. Jamison, R. W. McGowan, and D. Grischkowsky, “Terahertz waveguides,” *J. Opt. Soc. Am, B*, Vol. 17, 851–863, 2000.
11. Deibel, J. A., K. Wang, M. D. Escarra, and D. M. Mittleman, “Enhanced coupling of terahertz radiation to cylindrical wire waveguides,” *Opt. Express*, Vol. 14, 279–290, 2006.
12. Roskas, H, M. C. Nuus, K. W. Goossen, D. W. Kisker, A. E. White, K. T. Short, D. C. Jacobson, and J. M. Poate, “Propagation of picosecond electrical pulses on a silicon-based microstrip line with buried cobalt silicide ground plane,” *Appl. Phys. Lett.*, Vol. 58, 2604–2606, 1991.
13. Ordal, M. A., R. J. Bell, R. W. Alexander, Jr., L. L. Long, and M. R. Querry, “Optical properties of fourteen metals in the infrared and far infrared: Al, Co, Cu, Au, Fe, Pb, Mo, Ni, Pd, Pt, Ag, Ti, V and W,” *Applied Optics*, Vol. 24, 4493–4499, 1985.
14. Prade, B. and J. Y. Vinet, “Guided optical waves in fibers with negative dielectric constant,” *J. Lightwave Technol.*, Vol. 12, 6–12, 1994.
15. Wang, K. and D. M. Mittleman, “Metal wires for terahertz guiding,” *Nature*, Vol. 432, 376–379, 2004.
16. Themistos, C., A. Hadjicharalambous, B. M. A. Rahman, K. T. V. Grattan, and F. A. Fernandez, “Gain/loss characterisation of optical waveguide and semiconductor laser structures,” *IEE Proceedings–Optoelectronics*, Vol. 145, 93–98, 1998

Numerical Simulation Analysis of an Optical Virtual Probe Based on Surface Plasmon Polaritonic Band-gap Structures

Qingyan Wang, Jia Wang, and Shulian Zhang
State Key Laboratory of Precision Measurement Technology and Instruments
Department of Precision Instruments, Tsinghua University, Beijing, 100084, China

Abstract— A 2D optical virtual probe based on surface plasmon polaritons (SPP) in a band-gap structure with a rectangular profile is proposed and investigated numerically by means of finite difference time domain (FDTD) method. The dependence of the distributions of the confined optical fields on different structure parameters is numerically analyzed. The results indicate that with the proper structure design, the full width at half maximal (FWHM) of the confined beam can keep constant in a certain distance range, and the field intensity of the confined virtual probe can be enhanced greatly by SPP. Compared with the optical virtual probe based on interference of evanescent waves, this type of virtual probe has merits of high intensity and high sensitivity, and is likely to promise for the applications in nano-photonics devices based on SPP.

1. INTRODUCTION

Surface plasmon polaritons (SPP) is a two-dimensional electromagnetic excitation existing on the surface of a noble metal, whose electromagnetic field is strongly confined to the vicinity of the surface, and decays exponentially with increasing the distance from the surface. This confinement leads to an enhancement of the electromagnetic field on metal surface, resulting in an extraordinary sensitivity of SPP to surface conditions [1]. Plenty of investigations have revealed that periodically corrugated structures with proper defects in wavelength/sub-wavelength extent on the planar interface between metal and dielectric can control the behavior of SPP on surface. If the period of the corrugated structures equals to half of the effective wavelength of SPP, the periodical structures will result in Bragg scattering of SPP, which produces a gap in the dispersion curve of SPP [2]. When SPPs are propagating in such band-gap structures, the behavior is very similar to the action of photons in photonic crystals. Correspondingly, a line defect in SPP band-gap structures can act as a nano-waveguide [3], and a singular defect as a micro-cavity [4]. Different types of SPP band-gap structures with defects have been proposed and the characteristics of SPP propagating in-plane have been studied widely [5–7].

Emiliano et al. (2005) reported their research on the out-of-plane features of SPP in a band-gap structure consisting of sinusoidal gratings surrounding a flat surface with subwavelength sizes [8]. The out-of-plane distribution of the optical field at the flat region exhibited the features of a non-radiating and highly spatially confined light source, which can be used as a virtual probe in near-field optics [9]. A near-field optical virtual probe is a kind of confined optical field formed by the interferences of two evanescent waves, which is proposed to overcome the drawbacks of the metal probe and fiber probe, such as rigorous distance control and fragility [10, 11]. Nevertheless, the intensity is low (no more than 10% of the maximal intensity of the incident light), and the sensitivity is not enough in detecting metal and dielectric samples. Compared with this, the optical virtual probe based on SPP has merits of high intensity, high sensitivity and potential applications in nanophotonics devices based on SPP.

In this paper, a 2D optical virtual probe based on SPP propagating in a band-gap structure consisting of rectangular grooves with subwavelength period surrounding a singular defect is constructed. The confined field distribution is characterized in detail by the method of Finite-Difference Time-Domain (FDTD). The simulation results show some differences from the sinusoidal grating model, and this simply rectangular profile is promise to reduce the difficulties in fabrication.

2. SIMULATION MODELS

The excitation of SPP by the incident light is implemented under Kretschmann and Raether configuration [12], as shown in Fig. 1. A thin metal film is plated on the bottom of a prism. A beam of p-polarized plane wave with the wavelength of 632.8 nm illuminates the undersurface of the metal film through the prism. If the conditions are satisfied, SPP resonance can be excited on the upper-surface of the film. The relation between the wave vectors of incident light and SPP can be expressed as

$$k_{ix} = \frac{\omega}{c}\sqrt{\varepsilon_{\rm p}}\sin\theta = k_{\rm SP} = \frac{\omega}{c}\left(\frac{\varepsilon_1\varepsilon(\omega)}{\varepsilon_1+\varepsilon(\omega)}\right)^{1/2} \tag{1}$$

where ε_1, $\varepsilon(\omega)$ and $\varepsilon_{\rm p}$ are the permittivities of the air, the metal film and the prism, respectively. θ is the angle of incidence. The reflectivity of the metal film versus the angle of incidence and the thickness of the film $R(\theta, d)$ is shown as Fig. 2. The dark zone denotes small reflectivity where the SPP resonance is excited. If d is 50 nm, the corresponding θ is 43.6°.

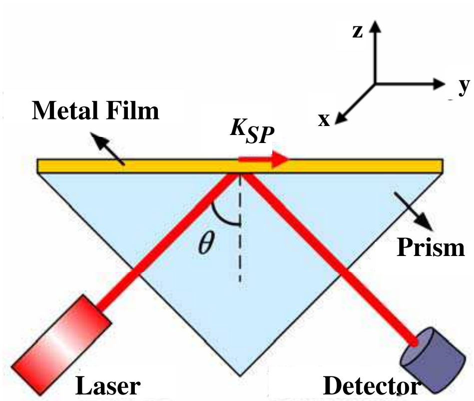


Figure 1: Excitation of SPP resonance in K-R configuration.

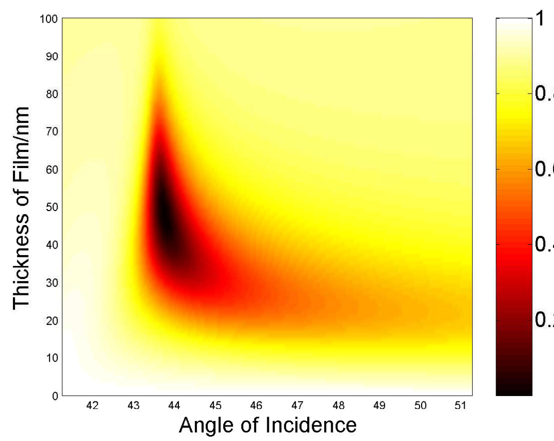


Figure 2: Reflectivity of smooth Au film versus the angle of incidence and the thickness of the film under K-R illumination ($\lambda = 632.8$ nm, $\varepsilon(\omega) = -11.77 + 1.26$i, $\varepsilon_p = 2.3$).

A simple periodical structure consisting of grooves with rectangle profile is constructed on the Au film. With the determinate parameters λ and θ, the effective wavelength of SPP on this Au film is $\lambda_{\rm SP} = 2\pi/{\rm Re}\{k_{\rm SP}\} \approx 608$ nm, thus we choose the period of the grooves is $\Lambda = \lambda_{\rm SP}/2 = 304$ nm, which will match the frequency of SPP in the energy band-gap. The transverse size of a groove is set as about half of the period. The periodical structure is etched on a layer of Au film. We set the depth as $w = 20$ nm. The section of the structured sample and the simulation model are shown as Figs. 3 and 4, respectively. The distance between the observation plane and the sample surface is h, which is 200 nm without further identifications. Calculations were performed by the software package, XFDTD 6.2, from REMCOM.

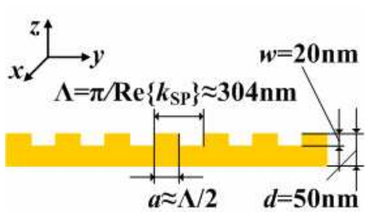


Figure 3: Sketch of the Au film with periodical corrugated structure on surface.

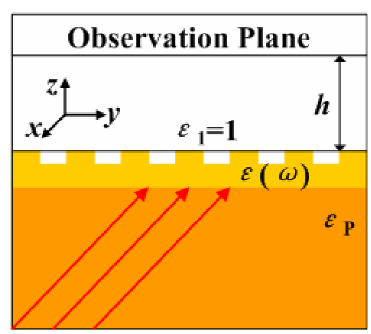


Figure 4: Simulation model by FDTD method.

3. RESULTS AND DISCUSSION

3.1. Realization of 2D Optical Virtual Probe

In order to capture SPP at special location, a singular defect is constructed by changing the transverse size t of a ridge (see Fig. 5). As a result, the rigorous periodical structure is broken, and a defect state is introduced in the energy band-gap.

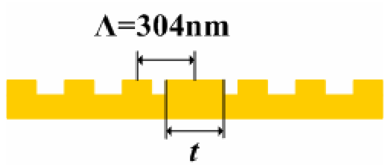


Figure 5: Topographic profile of the periodical corrugated surface with a singular defect structure on surface.

Firstly t is set to 160 nm. SPP is confined at the defect region and the optical field is enhanced here. The field distributions are shown as Fig. 6. In the vertical section along y axis the field distribution shows a central peak and a series of sidelobes. The central peak can be used as a confined virtual probe, the FWHM of which is 224 nm ($\sim 0.35\lambda$), smaller than that of virtual optical probe based on the interference of evanescent waves [10]. The field enhancement factor reaches 150. However, the intensities of the undesirable sidelobes are also high. In order to describe the influence of the sidelobes, we define a parameter F_{sc} as

$$F_{\mathrm{sc}} = (S_{\max}/C_{\max}) \times 100\% \tag{2}$$

where $S_{\max}$ is the maximal intensity of one-order sidelobes and $C_{\max}$ is the maximal intensity of the zero-order central peak. In the structure of $t = 160$ nm, $F_{\mathrm{sc}} = 93\%$ means the effect of sidelobes is very severe, which is harmful for using as a confined nano-source of a virtual probe.

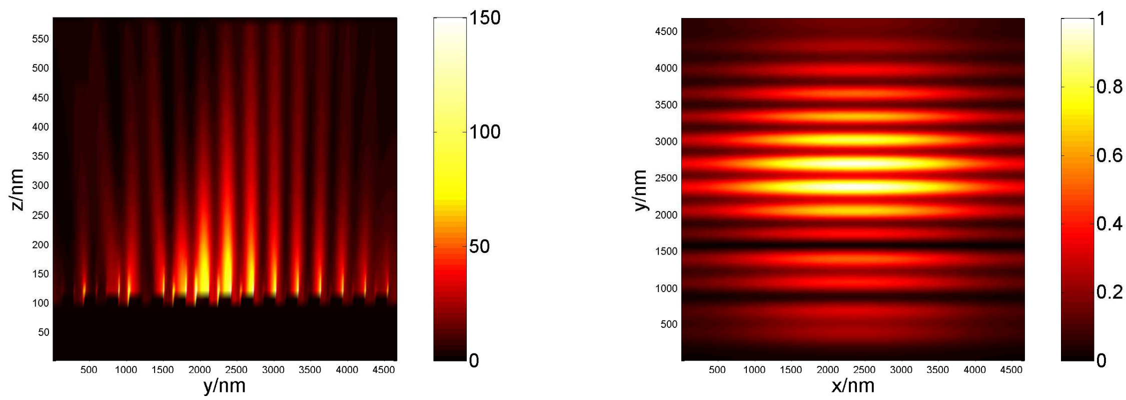
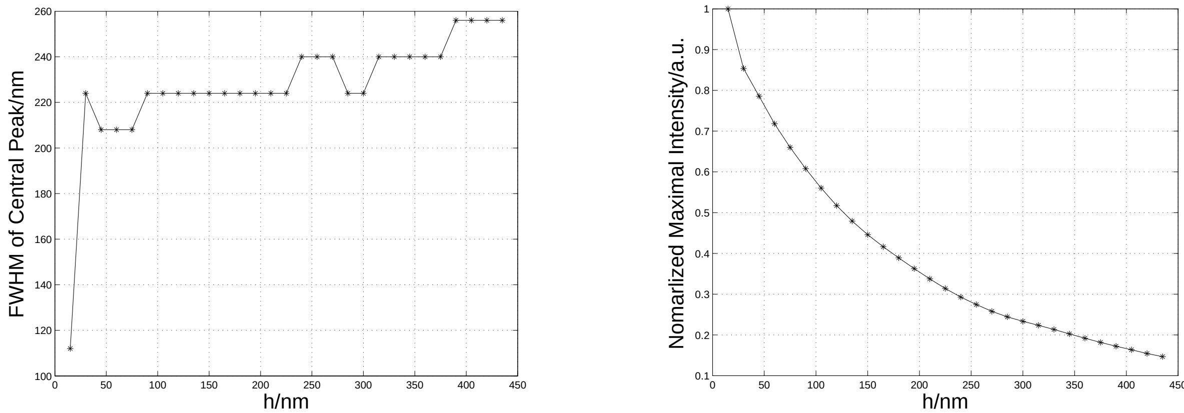


Figure 6: Intensity distribution of the optical field in the structure with $t = 160$ nm in (a) vertical section and (b) transverse section on observation plane where $h = 200$ nm (normalized).

Figure 7 shows the FWHM and the maximal intensity of the central peak versus h. The FWHM retains below 240 nm ($\sim 0.38\lambda$) within 375 nm distance, but the maximal intensity drops fast to 50% at the distance ~120 nm.

3.2. Influence of the Structure Parameters

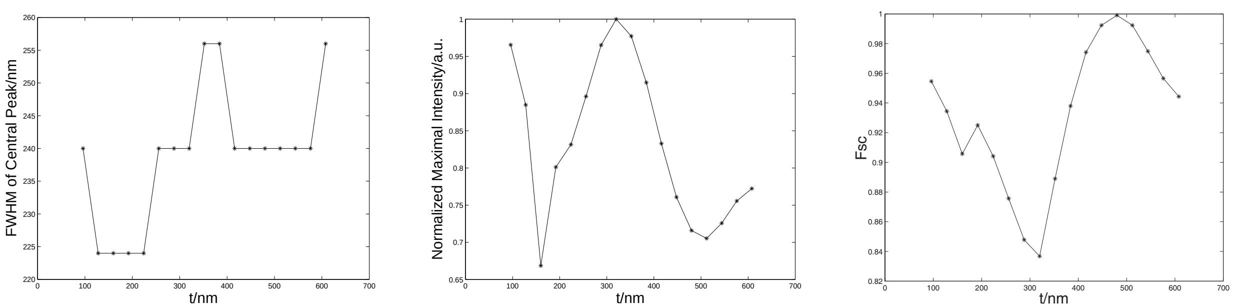
Figure 8 and Fig. 9 show the influences of t and w, respectively. In Fig. 8, F_{sc} curve and the intensity curve both show a typical resonance profile. The two minima of transmitted intensity correspond to $t_0 \approx \Lambda/2$ and $t_1 \approx 3\Lambda/2$, respectively, while in the F_{sc} curve the situation is reverse. High intensity of transmitted field results in the drop of F_{sc}, meaning that the energy in central peak attenuates in a slower manner than that in sidelobes. This result is different from the reports by Emiliano et al. [8]. In their simulation, it is maximal of transmission when the size of the defect equals $\Lambda/2$ or $3\Lambda/2$. The difference may come from the shape of the periodical structure. For a sinusoidal profile, there are a series of dots or lines on the topside, while for a rectangular profile, on the topside are flats. In our models, the transverse size of the ridges is just right the half of the period, which may have cavity-like effect somehow, accordingly lead to the translation of the transmission curve to $\Lambda/2$. Considering the high enhancement and the low sidelobes, we choose $t = \Lambda$. Fig. 9 shows the influences of the etching depth of the corrugated structure w with $t = \Lambda$.



(a) (b)

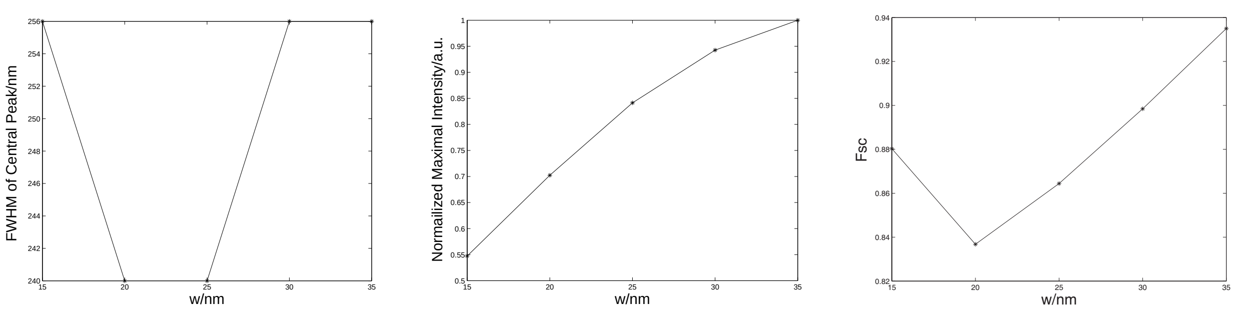
Figure 7: (a) FWHM vs. h, (b) Normalized maximal intensity of the central peak vs. $h(t = 160\,\text{nm})$.

The results indicate that with the rectangle grooves deepening, the transmitted maximal intensity of the central peak increase evidently. We choose $w = 25\,\text{nm}$ as a tradeoff between intensity and F_{sc}, since with this value FWHM is not widened. The field distribution in the vertical section is shown in Fig. 10. The field enhancement factor is 354.



(a) (b) (c)

Figure 8: (a) FWHM of central peak vs. t, (b) Normalized maximal intensity of central peak vs. t, (c) F_{sc} vs. $t(h = 200\,\text{nm})$.



(a) (b) (c)

Figure 9: (a) FWHM of central peak vs. w, (b) Normalized maximal intensity of central peak vs. w, (c) F_{sc} vs. $w(h = 200\,\text{nm}$ and $t = \Lambda)$.

3.3. Interaction with a Nano-particle

Next we try to use the optical virtual probe proposed above to detect a small particle. The simulation model is shown in Fig. 11. The periodically corrugated structure is built with $t = \Lambda$ and $w = 25\,\text{nm}$, and the observation plane is $h = 200\,\text{nm}$, with a small nano-particle placed above it, the radius of which is $100\,\text{nm}$. The distance between the center of the particle and the sample

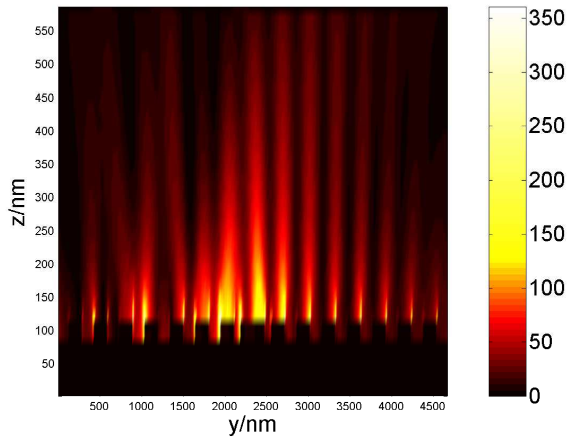


Figure 10: Intensity distribution of the optical field in the structure with $t = \Lambda$ and $w = 25\,\text{nm}$ in vertical section.

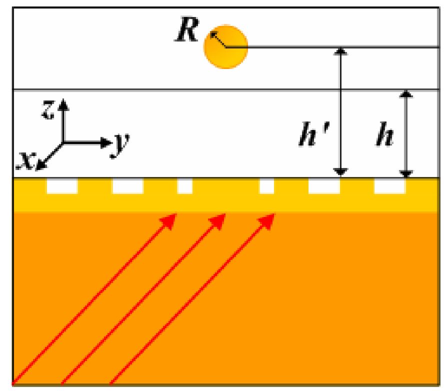


Figure 11: Simulation model of a nano-particle and the sample with $t = \Lambda$ and $w = 25\,\text{nm}$.

surface h' is 350 nm. Two kinds of nanoparticle are used: one is a dielectric particle with $\varepsilon = 2.3$, the other is a golden one with $\varepsilon(\omega) = -11.77 + 1.26\text{i}$.

The simulation results on the observation plane are shown in Fig. 12. The field enhancement factors in golden model and in dielectric model are 420 and 314, respectively. The enhancement in the golden model mainly exists near the particle, while on the observation plane the maximal intensity is decreased instead. Besides, the intensities in the models with particles are both higher than that without particles. This phenomenon may denote that the existence of a particle may extend the effective distance of the virtual probe. It is more obvious in the dielectric model, for the dielectric particle hardly consume the energy of the virtual probe, whereas the golden one can convert the energy into localized SPP (LSPP) on its surface. In the golden model, the FWHM of the central peak is widened to almost twice of that of the model without particles, and F_{sc} is deteriorated, which make the detection of metal sample difficult. However, in the dielectric model, the FWHM is compressed and F_{sc} holds the line. These results show that this type of optical virtual probe can be used to detect dielectric samples, while is not very applicable in detecting metal samples.

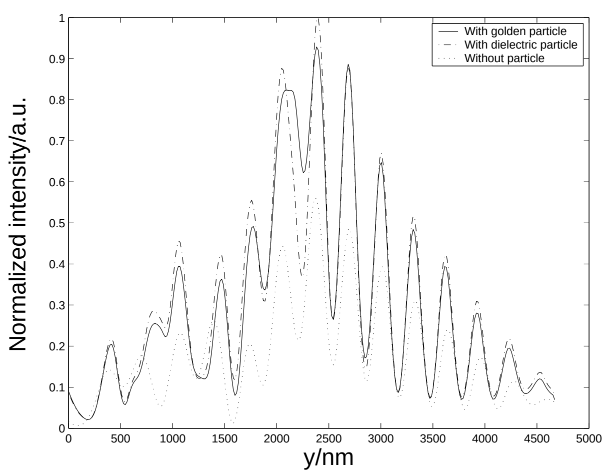


Figure 12: Intensity distribution of the optical field along y-axis on the observation plane ($h = 200\,\text{nm}$)

4. CONCLUSION

A 2D optical virtual probe based on SPP propagating in a band-gap structure consisting of periodically rectangular grooves surrounding a singular defect is proposed, which has merits of high intensity and high sensitivity. By design and optimization, the numerical simulations show that the resolution of the virtual probe (FWHM of the central peak) can reach 240 nm ($\sim 0.38\lambda$) within

200 nm distance from the sample surface, the field enhancement factor is 354, and the ratio of sidelobes suppression is 87%. Further works need to be done to suppress the sidelobes more. This type of optical virtual probe is sensitive to detect dielectric samples, while is not very applicable in detecting metal samples. The characteristics of the virtual probe based on SPP is promise to be applied to optical storage, near-field imaging and optical manipulation, and has potential applications in nanophotonics devices based on SPP.

ACKNOWLEDGMENT

The authors are grateful for the research support by Chinese National Nature Science Foundation Project No. 60278029 and No. 60678028.

REFERENCES

1. Zayatsa, A. V., I. I. Smolyaninov, and A. A. Maradudin., "Nano-optics of surface plasmon polaritons," *Physics Reports*, Vol. 408, 131–314, 2005.
2. Kitson, S. C., W. L. Barnes, J. R. Sambles, "Full photonic band gap for surface modes in the visible," *Phys. Rev. Lett.*, Vol. 77, No. 13, 2670–2673, 1996.
3. Bozhevolnyi, S. I., et al., "Waveguiding in surface plasmon polariton band gap structures," *Phys. Rev. Lett.*, Vol. 86, No. 14, 3008–3011, 2001.
4. Skigin, D. C. and R. A. Depine, "Surface shape resonances and surface plasmon polariton excitations in bottle shaped metallic gratings," *Phys. Rev. E.*, Vol. 63, No. 4, 046608-1–10, 2001.
5. Barnes, W. L., A. Dereux, and T. W. Ebbesen, "Surface plasmon subwavelength optics," *Nature*, Vol. 424, No. 14, 824–830, 2003.
6. Okamoto, T., F. H'Dhili and S. Kawata, "Towards plasmonic band gap laser," *Appl. Phys. Lett.*, Vol. 85, No. 18, 3968–3970, 2004.
7. Krenn, J. R., H. Ditlbacher, G. Schider, et al., "Surface plasmon micro- and nano-optics," *Journal of Microscopy*, Vol. 209, 167–172, 2003.
8. Descrovi, E., V. Paeder, and L. Vaccaro, et al., "A virtual optical probe based on localized surface plasmon polaritons," *Optics Express*, Vol. 13, No. 18, 7017–7027, 2005.
9. Grosjean, T. and D. Courjon, "Bessel beams as virtual tips for near-field optics," *Journal of Microscopy*, Vol. 210, 319–323, 2002.
10. Hong, T., J. Wang, L. Sun, et al., "Numerical simulation analysis of a near-field optical virtual probe," *Appl. Phys. Lett.*, Vol. 81, No. 18, 3452–3454, 2002.
11. Hong, T., J. Wang, L. Sun, et al., "Numerical and experimental research on the near-field optical virtual probe," *Scanning*, Vol. 26, No. 4, 1–6, 2004.
12. Kretschmann, E. and H. Raether, "Radiative decay of non-radiative surface plasmons excited by light," *Zeitschrift fur Naturforschung A*, Vol. 23, No. 12, 2135, 1968.

Hybrid Numerical Simulation of Electrostatic Force Microscopes Considering Charge Distribution

U. B. Bala, M. Greiff, and W. Mathis

Institute of Electromagnetic Theory
Leibniz University of Hannover, Germany

Abstract— The electrostatic force microscope (EFM) is an important tool for imaging and characterizing material surfaces. In this paper a hybrid numerical approach for the simulation of the EFM considering charge distribution inside the sample under investigtion is presented. In the simulation model electrical part is considered. In this paper first a basic knowledge on the EFM and then the numerical model of the EFM considering volume charge distribution are presented. At last several numerical simulation results of the EFM in 3D are pesented.

1. INTRODUCTION

A significant progress in nanotechnology has been observed over the last few years. This progress has also been influenced by the development of new high resolution measurement instruments. Due to the rapid miniaturization of integrated devices into the mesoscopic regime and the increasing interest in very small structures, these instruments have become very important. An interesting example is the atomic force microscope (AFM). Based on the design of the scanning tunneling microscope (STM), the first AFM was developed in 1986 by G. Binnig and his coworkers in collaboration between IBM and Standford university. Since then a new era of topographical imaging, as well as for measuring force-separation interactions between a probe and substrate began. The AFM's ability to scan surfaces with nearly atomic resolution and its versatility make it one of the most important measurement devices in nanotechnics. If the sample under investigation holds a charge distribution and the distance between the AFM tip and the sample is kept large then all other interaction forces except the electrostatic force can be neglected. This special working mode of the AFM is known as electrostatic force microscope (EFM) which can be used for scanning electric field with nearly atomic resolution. The EFM has many materials-related applications including measuring the surface potential or contact potential, detecting charges on surfaces or nanocrystals etc. In this paper a 3D model of the EFM is presented considering charge distribution inside the sample. Several numerical methods are proposed to calculate the electric eld more efficiently.

2. WORKING PRINCIPLE AND MODEL OF THE EFM

For the EFM the interaction force would be the electrostatic force between the biased atomically sharp tip and the sample. In addition the Van der Waals force between the tip and the sample are always present. The Van der Waals force and the electrostatic force have two different dominant regions. The Van der Waals force is proportional to $1/r^6$ where as the elctrostatic force is proportional to $1/r^2$. Thus when the tip is close to the sample the Van der Waals force is dominat and when the tip is moved away from the sample the electrostatic force is dominant. The scanning of the EFM is usually done in two steps. First the topography of the sample is done by tapping scanning mode which is also known as "intermittent-contact" (IC) mode. In this case the Van der Waals force plays a signifiant role. Second using this topgraphical information a constant tip-sample distance is maintained while scanning where the electrostatic force is dominat, a technique which is known as "lift scanning" [1]. In this technique it is assumed that the influence of all short-range forces can be neglected and only the electrostatic force plays the vital role for imaging. To detect the electrostatic force a voltage is applied between the cantilever tip and the sample. The cantilever oscillates near its resonance frequency which changes in response to any additional force gradient. Changes in cantilever resonant frequency can be detected using phase detection, frequency modulation, amplitude modulation etc. A diode laser is focused on the back of the reflective cantilever and the reflected light is collected by a position sensitive detector (PSD). This usually consists of two closely spaced photodiodes. Any angular displacement of the cantilever results in one photodiode collecting more light than the other and therefore in a high output voltage. This voltage then plots the topography of the sample.

Some typical parameters of the EFM i.e., the length of the cantilever is some hundreds of $1\,\mu\text{m}$, the height of the tip is nearly 30 nm and the pick of the tip is usually less than 10 nm. So for modeling and simulating the EFM, multi physics aspects must be taken into consideration. From the numerical point of view additional problems arise since frequently we are confronted with multi-scale problems. Therefore the application of advanced numerical methods is necessary. As the cantilever frequently changes its position during scanning, the coupled mechanical and electrical behavior have to be taken into account. This can be achieved by dividing the model into an electrical part and a mechanical part [4]. The interaction between them can conveniently be realized by using a staggered simulation approach. In the present model the electrical part is considered, i.e., the cantilever deflection is kept fixed.

For developing a model of the EFM different effects have to be considered. For example long distance interaction, charge distribution and possible non-linearity of the material properties, singularity etc. In order to take into consideration these effects the simulation region is divided into three regions as shown in Fig. 1. As high values of the electric field will occur at the pick of the tip, a special numerical method is needed to calculate this electric field more effectively. For this reason an augmented FEM will be applied to region Ω_M. Since charge distribution and nonlinearities of the dielectric properties may have to be considerd, a versatile numerical method such as finite element method (FEM) should be applied to region Ω_F. As boundary element method (BEM) works well when the boundary is infinite or semi-infinite, the large distance interaction between the tip and the cantilever can be conveniently treated by using BEM in region Ω_B. Later all these three numerical methods will be coupled with each other.

3. NUMERICAL FORMULATION OF THE PROBLEM

Consider a bounded domain $\Omega \subset \mathbb{R}$ be bounded with Lipschitz boundary $\Gamma = \partial\Omega$ which is decomposed into three disjoint parts $\Omega = \Omega_M \cup \Omega_F \cup \Omega_B$. At present Poisson's equation in Ω will be applied with mixed boundary conditions. Find $u : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\Delta u &= f \quad \text{in} \quad \Omega \\ u &= u_0 \quad \text{on} \quad \Gamma_D \\ \frac{\partial u}{\partial n} &= t_0 \quad \text{on} \quad \Gamma_N \end{aligned} \tag{1}$$

where f is the charge density, u_0 and t_0 are the Dirichlet and Neumann boundary conditions respectively. The detail Formulation of Laplace problem for FEM-BEM-MFS is shown in [2] by the same authors. Here only some important steps will be presented which is necessary to get the basic understanding of the formulation of Poisson problem.

The energy-related functional in the electrostatic calculation domain Ω (Fig. 1) can be written as

$$W = \int_\Omega (\nabla u)^2 d\Omega \qquad u \in H^1_D(\Omega) : \{u \in H^1 | u_{|\Gamma_D} = u_0\} \tag{2}$$

The electrostatic potential u in the spherical region Ω_M of radius R can be approximated by [13]

$$u(x,y,z) = \sum_{j=1}^{n} u_j \psi_j(x,y,z) + M(r) \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \left(A_{lm} r^l + B_{lm} r^{-(l+1)} \right) Y_{lm}(\theta,\phi), \tag{3}$$

where

$$M(r) = \begin{cases} 1, & 0 \le r \le \frac{R}{2} \\ \frac{1}{2}\left(1 + \cos\left(\frac{2}{R}r - 1\right)\pi\right), & \frac{R}{2} < r < R \\ 0, & R \le r \end{cases} \tag{4}$$

and $Y_{lm}(\theta,\phi)$ can be expressed by the Legendre functions $P_{lm}(\cos\theta)$ [9]. This leads to the stiffness matrix

$$\frac{\partial W}{\partial u_i} = 2\int_\Omega \nabla u \frac{\partial \nabla u}{\partial u_i} d\Omega, \quad \frac{\partial W}{\partial A_{ik}} = 2\int_\Omega \nabla u \frac{\partial \nabla u}{\partial A_{ik}} d\Omega, \quad \frac{\partial W}{\partial B_{ik}} = 2\int_\Omega \nabla u \frac{\partial \nabla u}{\partial B_{ik}} d\Omega. \tag{5}$$

On the FEM-BEM transmission interface [4] $\Gamma_T = \Gamma_B \cap \Gamma_F$, $u_B = u_F$ and $\frac{\partial u_B}{\partial n} + \frac{\partial u_F}{\partial n} = 0$. Using the Gauss theorem on $\Omega_{FM} = \Omega_F \cup \Omega_M$ one obtains [5]

$$\int_{\Gamma_F} \frac{\partial u_{FM}}{\partial n} v \, d\Gamma = \int_{\Omega_{FM}} \mathrm{div}\left(\nabla u_{FM} \cdot v\right) d\Omega = \int_{\Omega_{FM}} \triangle u_{FM} \cdot v \, d\Omega + \int_{\Omega_{FM}} \nabla u_{FM} \cdot \nabla v \, d\Omega \tag{6}$$

i.e., for all $v \in H^1_{D,0}(\Omega_{FM}) := \{v \in H^1(\Omega_{FM}) : v|_{\Gamma_D \cap \Gamma_F} = 0\}$

$$a(u_{FM}, v) := \int_{\Omega_{FM}} \nabla u_{FM} \cdot \nabla v \, d\Omega = \int_{\Omega_{FM}} f \cdot v \, d\Omega + \int_{\Gamma_F} \frac{\partial u_{FM}}{\partial n} v \, d\Gamma =: (f, v)_{\Omega_{FM}} + \left\langle \frac{\partial u_{FM}}{\partial n}, v \right\rangle_{\Gamma_F} \quad (7)$$

where u_{FM} includes u_j, A_{lm} and B_{lm}. The representation formula of the Laplace's equation for the solution of u_B inside Ω_B

$$u_B(x) = \int_{\Gamma_B} \left\{ \frac{\partial}{\partial n(y)} G(x, y) u_B(y) - G(x, y) \frac{\partial u_B}{\partial n(y)} \right\} d\Gamma, \quad x \in \Omega_B \quad (8)$$

with the fundamental solution of the Laplacian in 3D given by

$$G(x, y) = \frac{1}{4\pi} |x - y|^{-1} \quad (9)$$

For Poisson's problem the two boundary integral equations on the BEM region

$$V \frac{\partial u_B}{\partial n} = (I + K) u_B - N_0 f \quad (10)$$

$$W u_B = (I - K') \frac{\partial u_B}{\partial n} - N_1 f \quad (11)$$

where the single layer potential V and the hypersingular operator W are symmetric and the double layer potential K has the dual K' [3]. The integral operators N_0 and N_1 are defined by

$$N_0 f(x) := \int_{\Omega_B} G(x, y) f(y) dy, \quad N_1 f(x) := \frac{\partial}{\partial n_x} N_0 f \quad (12)$$

The saddle point formulation of the problem for all $(w, v, \psi) \in \tilde{H}^{1/2} \times H^1_{D,0}(\Omega_{FM}) \times \tilde{H}^{-1/2}(\Gamma_B)$

$$2a(u_{FM}, v) + \langle W u_B, v \rangle_{\Gamma_T} + \left\langle (I + K') \varphi, v \right\rangle_{\Gamma_T} = 2(f, v)_{\Omega_{FM}} + 2\langle t_0, v \rangle_{\Gamma_n \cap \Gamma_F} - \langle N_1 f, v \rangle_{\Gamma_T} \quad (13)$$

$$\langle W u_B, w \rangle_{\Gamma_B \cap \Gamma_N} + \left\langle (I + K') \varphi, w \right\rangle_{\Gamma_B \cap \Gamma_N} = 2\langle t_0, w \rangle_{\Gamma_B \cap \Gamma_N} - \langle N_1 f, w \rangle_{\Gamma_B \cap \Gamma_N} \quad (14)$$

$$\langle (I + K) u_B, \psi \rangle_{\Gamma_B} - \langle V \varphi, \psi \rangle_{\Gamma_B} = \langle N_0 f, \psi \rangle_{\Gamma_B} \quad (15)$$

If the bases are introduced as $\mathrm{span}\{v_1 \ldots\ldots, v_F\} = X_F$, $\mathrm{span}\{w_1, \ldots\ldots, w_F\} = X_B$ and $\mathrm{span}\{\psi_1, \ldots\ldots, \psi_F\} = Y_B$, the basis functions of X_F and X_B are supposed to be ordered such that

$$\mathrm{span}\{v_1, \ldots\ldots, v_F\} = X_F \cap H^1_{D,0}(\Omega_F)$$

$$\mathrm{span}\{w_1, \ldots\ldots, w_B\} = X_B \cap H^{1/2}(\Gamma_B).$$

If the coefficients of u_{FM} and u_B are denoted by u and the coefficients of φ are denoted by φ again then this system is equivalent to the original differential equation that can be used for descritization. This system corresponds to a matrix formulation which can be written as

$$\begin{pmatrix} M & B^T & 0 & 0 & 0 \\ B & F_{NN} & F_{NC} & 0 & 0 \\ 0 & F_{CN} & F_{CC} + W_{CC} & W_{CN} & \left(K^T + I\right)_C \\ 0 & 0 & W_{NC} & W_{NN} & \left(K^T + I\right)_N \\ 0 & 0 & (K + I)_C & (K + I)_N & -V \end{pmatrix} \begin{pmatrix} u_m \\ u_F \\ u_T \\ u_B \\ \varphi \end{pmatrix} = \begin{pmatrix} b_m \\ b_F \\ b_\Gamma \\ b_B \\ b_\varphi \end{pmatrix} \quad (16)$$

where the subscript C means contribution from the coupling nodes and N means contribution from the non-coupling nodes. Finally the blocks $W, V, K + I$, and $K^T + I$ provide the coupling between the two ansatz spaces X_F and X_B. Here u_m includes the coecients A_{lm} and B_{lm} (3), u_F and u_B are the nodal potentials inside the FE domain and on the boundary of the BE domain respectively, u_T are the nodal potentials on the FE-BE coupling interface and φ are the normal components of the electric field distribution on the boundary of the BE domain. The vector b

includes the corresponding boundary conditions and the charge distribution. As the matrix in (16) is not positive definite, a specific algorithm such as the MINRES algorithm is required for the solution.

A typical simulation of potential distribution and electric field distribution are shown in Figs. 2 and 4 respectively. In the present model some volume charges are injected on the top portion inside the sample. Equal amount of positive volume charge density are present on the left and right volumes where as the negative volume charge density are present in the center volume. As this charges will cause potential inside the volume, they will produce remarkable effects on the simulation result. The result of a typical simulation of potential distribution and electric field distribution considering a moving sample using FEMBEM coupling in the cut plane through the middle of the cantilever is shown in Figs. 3 and 5 respectively. Usually the highest value of electric field appears at the pick of the tip which plays the dominant role for cantilever deflection. But it may be the case due to the different values of volume charge density, the highest value of electric field may not appear on the pick of the tip but near the volume charges. Since the sample is moving there must be a deviation of electric field at the pick of the tip. By utilizing this deviation of electric field the EFM will plot the charge distribution on the moving sample.

Since the scanning process of EFM is dynamic, one has to deal with a moving sample and moving boundaries. As a result the mesh has to be updated at each time step. The approach presented here for mesh updating is based on an arbitrary Lagrangian Eulerian (ALE) algorithm.

4. FIGURES

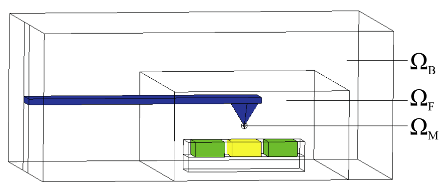


Figure 1: 3D EFM model.

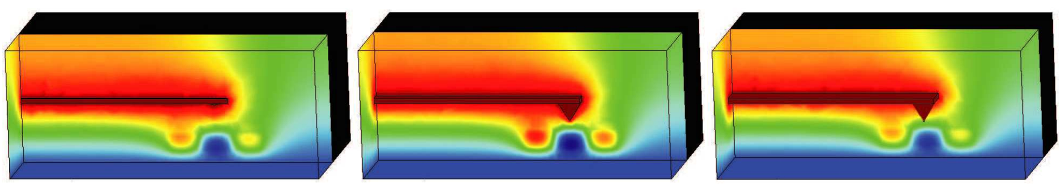


Figure 2: Potential distribution on different places of an EFM.

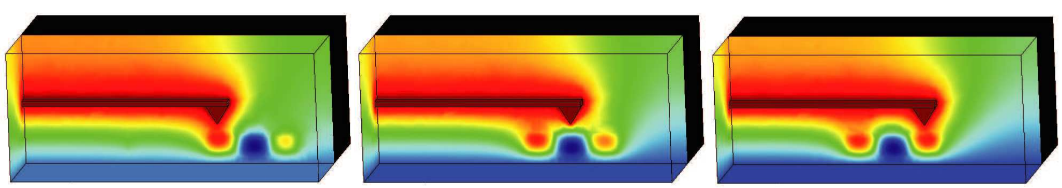


Figure 3: Potential distribution on the middle position of EFM with moving sample.

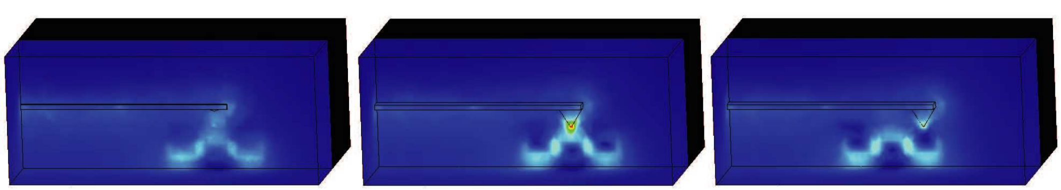


Figure 4: Electric field distribution on different places of an EFM.

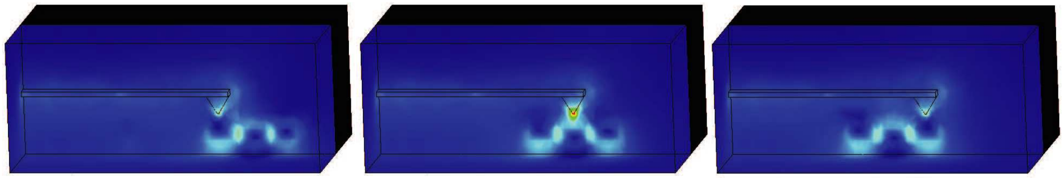


Figure 5: Electric field distribution on the middle position of EFM with moving sample.

5. CONCLUSION

A hybrid numerical approach for the simulation of the EFM considering charge distribution inside the moving sample is presented. In order to fulfill the special requirements different numerical methods are applied to different regions of the EFM. Here the simulation is implemented using FEMBEM coupling. As a high value of electric field is observed near the pick of the tip, implementing augmented FEM near the tip would obviously calculate this field more efficiently. Due to the moving sample and cantilever deflection the mesh needs to be updated at each time step. It can be obtained using ALE.

ACKNOWLEDGMENT

The authors would like to thank PD Dr. rer. nat. M. Maischak (Institute of Applied Mathematics, Hanover University) for his contribution and valuable discussions on numerical coupling. This work has been supported by the German Research Foundation (DFG), GRK 615.

REFERENCES

1. Yan, M. and G. H. Bernstein, "Apparent height in tapping mode of electrostatic force microscopy," *Ultramicroscopy*, Vol. 106, 582–586, 2006.
2. Bala, U. B., M. Greiff, and W. Mathis, "Hybrid numerical simulation of electrostatic force microscopes in 3D," *Proceedings of Progress in Electromagnetic Research Symposium*, Tokyo, Japan, August 2006.
3. Stephan, E. P., "Coupling of boundary element methods and finite element methods," *Encyclopedia of Computational Mechanic*, Vol. 01, Chapter 13, 375–412, 2004.
4. Greiff, M., U. B. Bala, and W. Mathis, "Hybrid numerical simulation of micro electro mechanical systems," *Progress in Electromagnetic Research Symposium (PIERS)*, 270–274, 2006.
5. Reddy, J. N., *Finite Element Method*, McGraw-Hill International Editions, Singapore, 1993.
6. Bhushan, B., *Handbook of Nano-technology*, Springer, Berlin, 2004.
7. Li, Z. C., *Numerical Methods for Elliptic Problems with Singularities*, World Scientific, 1990.
8. Jackson, J. D., *Classical Electrodynamics*, Wiley, 1975.
9. Jacobs, H. O., P. Leuchtmann, O. J. Homan, and A. Stemmer, "Resolution and contrast in Kelvin probe force microscopy," *Journal of Applied Physics*, Vol. 84, No. 3, 1998.
10. Morita, S., R. Wiesendanger, and E. Meyer, *Noncontract Atomic Force Microscopy*, Springer, 2002.
11. Belytschko, T., K. W. Liu, and B. Moran, *Nonlinear Finite Elements for Continua and Structures*, Wiley, Sussex, 2004.
12. Costabel, M., "A symmetric method for the coupling of finite elements and boundary elements," *The Mathematics of Finite Elements and Applications*, Vol. VI, 281–288, 1988.
13. Strang, G., G. and J. Fix, *An Analysis of the Finite Element Method*, Prentice-Hall, 1973.

Characteristics of Scattered Fields from Hydrocarbon Layers in Seabed Logging

F. N. Kong[1], **S. E. Johnstad**[2], **and T. Roesten**[3]

[1] Norwegian Geotechnical Institute, Norway
[2] Hydro Oil & Energy, Norway
[3] Statoil ASA, Norway

Abstract— Analytical results show that the scattered field from a hydrocarbon layer takes the form of an exponentially decay function. Both the real part and the imaginary part of the complex propagation constant are functions of target properties (resistivity and thickness etc.), and can be used as the hydrocarbon indicators.

We have shown the correlation between the position of the pole existing in the TM reflection coefficient and the decay rate of a scattered field. By investigating the variance of the pole position versus the model parameters, we have derived an approximate expression, which relates the measurable complex propagation constant of a scattered field to the hydrocarbon and seabed properties.

1. INTRODUCTION

The principle of seabed logging measurement is based on the fact that the scattered field from a hydrocarbon layer decays slower than the direct wave [1, 3, 6] . However, the questions such as what is the exact decay rate and how the decay rate is related to the model parameters have never been answered clearly before. To investigate quantitatively how the complex propagation constant of a scattered field is related to model parameters is the main purpose of this study.

2. SCATTERED FIELDS FROM HYDROCARBON LAYERS

Figure 1(a) shows the calculated total (direct+scattered) fields along the seabed for different hydrocarbon resistivities for the following model:

- Horizontal electric dipole source, 1 Hz, 50 m above the seabed.
- Hydrocarbon layer at 1200 m depth and 100 m thick.

Figure 1(b) shows the fields for different hydrocarbon thickness for frequency 0.25 Hz and for hydrocarbon resistivity 200 ohmm.

The fields are calculated based on the theory of antenna radiation inside a layering structure [4, 5].

Note that at far offset, the scattered fields dominate. One can see from the figures that the scattered fields are straight lines in a vertical logarithm axis that suggest the scattered field can be written as:

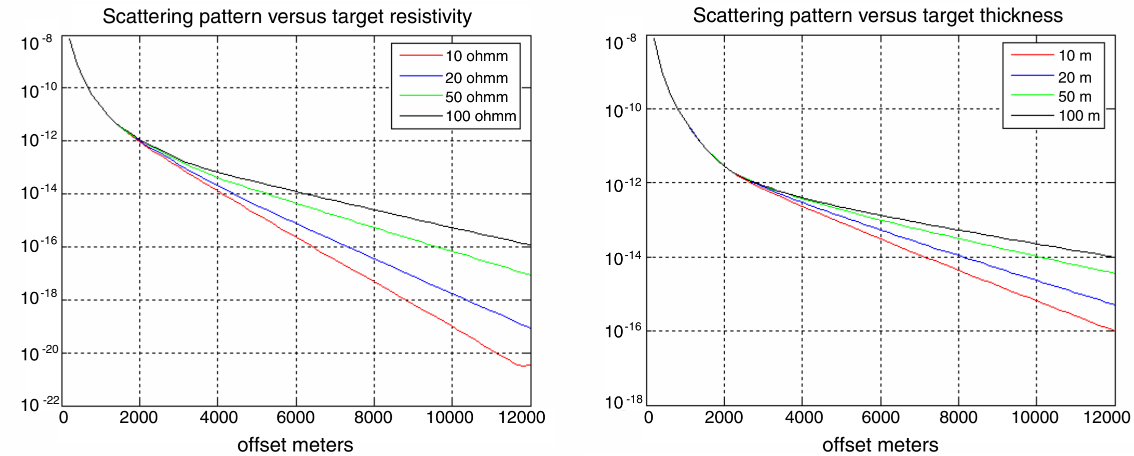
$$s(x) = \text{const.} \exp(\alpha x) = \text{const.} \exp(\alpha_r x + i\alpha_i x) \tag{1}$$

where x is the offset, α is the complex propagation constant, α_r is the decay rate and α_i is the phase constant. In this paper we use the following equation to calculate α_r and α_i values:

$$\alpha = \alpha_r + i\alpha_i = \ln\left(\frac{s(6000)}{s(8000)}\right)/2000 \tag{2}$$

where $s(8000)$ and $s(6000)$ are the scattered field at offset 8000 m and 6000 m respectively. The calculated α values for the curves shown in Figures 1 and 2 are listed in Tables 1 and 2.

From Tables 1 and 2, one can see that both the decay rate and the phase constant are functions of target parameters. Hence they can be used as the hydrocarbon indicators.



(a) (b)

Figure 1: a) Total (scattered and direct) fields for hydrocarbon resistivities 10, 20, 50, 100 ohm-m. b) Total (scattered and direct) fields for hydrocarbon thickness: 10, 20, 50, 100 m.

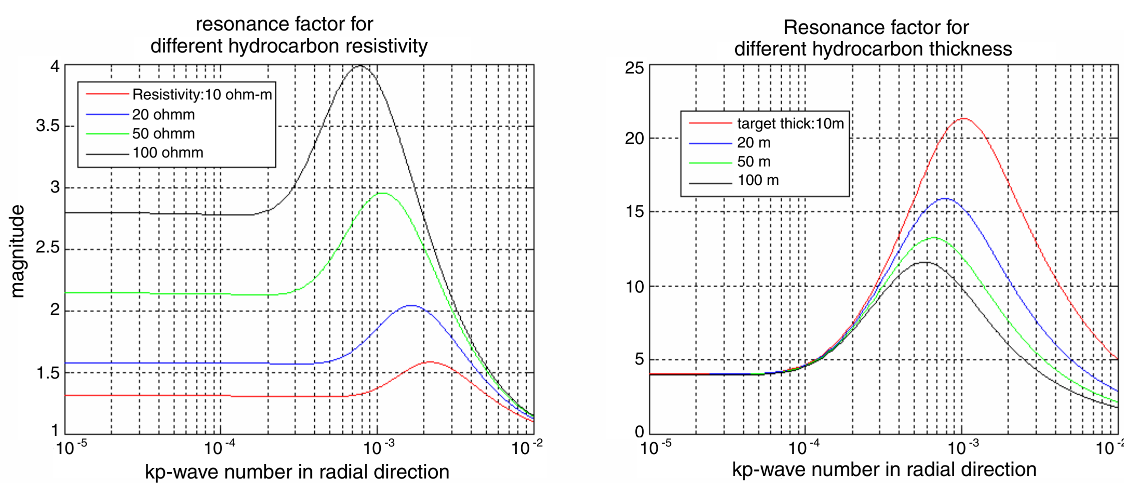
Table 1: Scattered field decay rate and phase constant for different hydrocarbon resistivity.

Resistivity	10 ohm-m	20 ohm-m	50 ohm-m	100 ohm-m
α_r: decay rate (10^{-3})	1.9	1.5	1.1	0.79
α_i: phase constant (10^{-3})	1.2	0.7	0.5	0.33

3. SCATTERED FIELDS IN WAVE NUMBER DOMAIN

According to the AR (auto-regressive) method in spectrum analysis [2], an exponential function in spatial domain will correspond to a pole in its spectrum domain. Hence one may observe a pole in the wave number domain of the scattered field.

It is not difficult to derive that the TM wave reflection is the main contribution to the scattered field from a high resistivity hydrocarbon layer. For a three layer (seabed, hydrocarbon layer and seabed) model, the TM wave reflection coefficient can be written [4], $R = \frac{R_{01}+R_{12}e^{-2k_z d}}{1+R_{01}R_{12}e^{-2k_z d}}$, where R_{01} is the TM reflection coefficients between layer 0 (seabed) and layer 1 (hydrocarbon layer), and R_{12} is the TM reflection coefficients between layer 1 and layer 2 (seabed). The parameter k_z

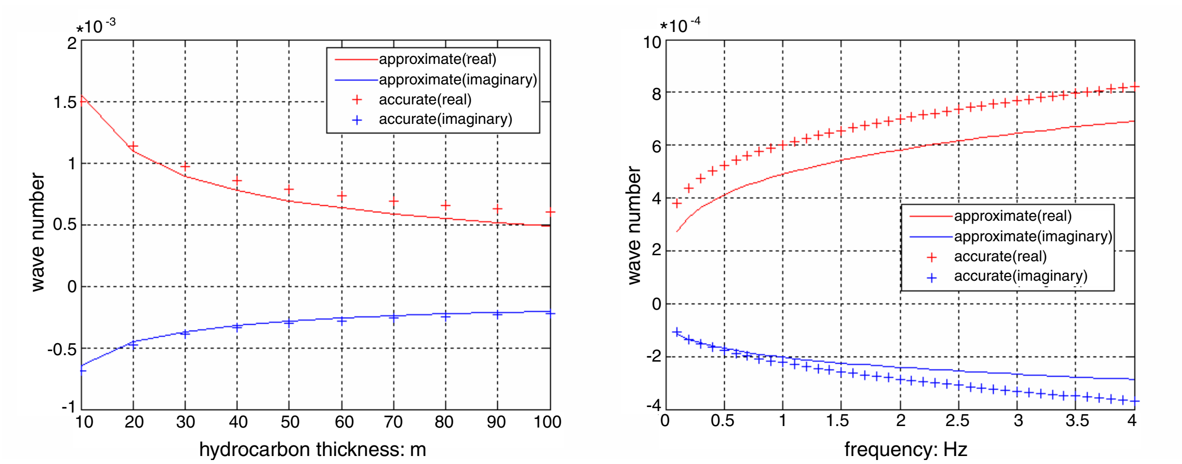


(a) (b)

Figure 2: a) Resonance factor versus k_p for different hydrocarbon resistivity. b) Resonance factor versus k_p for different hydrocarbon thickness.

Table 2: Scattered field decay rate and phase constant for different hydrocarbon thickness.

Thickness	10 m	20 m	50 m	100 m
α_r: decay rate (10^{-3})	0.97	0.8	0.58	0.46
α_i: phase constant (10^{-3})	0.52	0.36	0.20	0.15



(a) (b)

Figure 3: Comparison between analytic results for α_r and α_i using (2) and the approximate results using (4). a) for 1 Hz, 1 ohm-m seabed, 200 ohm-m hydrocarbon, and varying hydrocarbon layer thickness. b) for 1 ohm-m seabed, 200 ohm-m hydrocarbon, 100 m thick, but varying frequency.

is the wave number in z direction, and $k_z = \sqrt{k_p^2 - i\omega\mu\sigma}$ where k_p is the wave number in radial direction, σ is the conductivity of the middle layer, i.e., the hydrocarbon layer in this case. We have plotted in Figure 2 the denominator term of R, which is called later as the resonance factor, and the resonance factor equals:

$$\frac{1}{1 + R_{01}R_{12}e^{-2k_z d}} \tag{3}$$

From Figure 2 one can see that each curve contains a rather sharp peak. They are the poles mentioned before. One can also see that the k_p value at the pole position reduces when the target resistivity increases, and also reduces, when the target thickness increases. These k_p values coincide very well to α_i values listed in the tables.

One can derive a quantitative relation between the k_p values at poles and the model parameters (target resistivity, thickness, etc.) by maximizing the resonance factor is (3). But it seems difficult to obtain a simple relation for general model cases. For the case with high resistivity contrast between the hydrocarbon and seabed, and for the case with model parameters not far from the 'typical' seabed logging parameters, such as 1 Hz frequency, 1 ohm-m seabed, 100 ohm-m and 100 m thick hydrocarbon layer, we have derived the following approximation:

$$k_p = \alpha_r + i\alpha_i = \sqrt{2\frac{\sigma_{\text{oil}}\beta_{\text{seabed}}}{\sigma_{\text{seabed}}\Delta}} \tag{4}$$

where σ_{oil} and σ_{seabed} seabed are respectively the conductivities of hydrocarbon layer and the seabed, and Δ is the thickness of the hydrocarbon layer, and

$$\beta_{\text{seabed}} = \sqrt{-i\omega\mu\sigma_{\text{seabed}}}$$

Equation (4) is only a coarse approximation. Its performance is shown in Figure 3. In Figure 3, the left and right figures show the comparison between analytic results, i.e., α_r and α_i, obtained by using (2), and the approximate results by using (4). We can see the fitting between the analytical results to the approximate results is not exact. Improving the approximation is a future task.

4. CONCLUSION

Analytical results show that the scattered field from a hydrocarbon layer takes the form of an exponentially decay function. Both the real part and the imaginary part of the complex propagation constant are functions of target parameters (resistivity and thickness etc.), and can be used as the hydrocarbon indicators.

We find that there is a pole in the TM reflection coefficient expression. This pole leads to the exponential form for a scattered field. By investigating the variance of the pole position versus the model parameters, we have derived an approximate expression to estimate quantitatively the complex propagation constant of a scattered field from known hydrocarbon and seabed parameters.

ACKNOWLEDGMENT

The authors wish to deliver thanks to Norwegian Research Council, Norsk Hydro and Statoil for finance support through a Petromaks project.

REFERENCES

1. Johnstad, S. E., B. A. Farrelly, and C. Ringstad, "EM seabed logging on the troll field," *EAGE, 67th Conference and Exibition*, Madrid, June 2005.
2. Kay, S. M., *Modern Spectral Estimation*, Englewood Cliffs, Prentice Hall, NJ, 1988.
3. Kong, F. N., H. Westerdahl, S. Ellingsrud, T. Eidesmo, and S. E. Johansen, "Seabed logging: A possible direct hydrocarbon indicator for deepsea prospects using EM energy," *Oil* & *Gas Journal*, May 2002.
4. Kong, J. A., *Electromagnetic Wave Theory*, John Wiley & Sons, Inc., 1990.
5. Kong, J. A., "Electromagnetic fields due to dipole antennas over atratified anisotropic media," *Geophysics*, Vol. 37, No. 6, December 1972.
6. Roesten, T., S. E. Johnstad, S. Ellingsrud, H. E. F. Amundsen, S. Johansen, and I. Brevik, "A SeaBed Logging (SBL) calibration survey over the ormen lange gas field," *EAGE 65th Conference and Exhibition, Extended Abstracts.*, 2003.

Modeling the Response of a Seafloor Antenna in the Limits of Low Frequency and Shallow Water

L.-J. Gelius

Department of Geosciences, University of Oslo
P.O. Box 1047 Blindern, N-0316, Norway

Abstract— In this paper we describe a novel FDFD (finite-difference frequency-domain) method for computing the 2-D earth response of a 3-D subsea horizontal electric dipole (HED) source. It represents scattered field solutions to Maxwell's equations and uses a 1-D three-layer analytical solution for the primary field calculations. Comparison with the 1-D industry type of program CSEM1D gave accurate fits.

1. INTRODUCTION

The basis of the Sea Bed Logging (SBL) approach is the use of a mobile horizontal electric dipole (HED) source and an array of seafloor electric field receivers [3]. The transmitting dipole emits a low frequency electromagnetic signal (fundamental frequency typically ranging from 0.25 to 2 Hz) that diffuses outwards, and the method relies on the large resistivity contrast between hydrocarbon-saturated reservoirs and the surrounding sedimentary layers. The generated electromagnetic soundings can in general be divided into three main contributions: direct EM field, guided modes (associated with high-resistivity zones like hydrocarbons) and air waves (cf. Fig. 2(a)). In case of shallow water depths (e. g., 200 meters or less) the presence of strong airwaves will mask the subsurface signals. In order to develop improved data processing schemes for such cases, modeling tools are needed that can provide controlled test data. In addition, modeling is important when planning new SBL surveys and also when carrying out feasibility analysis of possible new prospects. Since electromagnetic modeling based on time-domain methods give prohibitive large computation times in case of an air layer we concentrate on frequency-domain (methods). Several EM-modeling programs have been developed [1, 5, 7]). However, none of these codes have neither been tailored to the SBL-case nor been documented to work accurately for the case of a low frequency operating seafloor antenna deployed in shallow water.

2. FUNDAMENTALS OF THE MODELING METHOD

We make use of basic ideas employed by Hohmann (1975) [6] to develop a FDFD-method able to compute the earth response of a 3-D sub sea HED-source, assuming a local 2-D earth model (e. g., 2.5-D case). Hohmann introduced scattered field solutions to Maxwell's equations and proposed to use analytical solutions for the primary field calculations (he considered a uniform background only). By adapting the scattered field approach one avoids the source singularity problem. We employed here a reference model consisting of three layers (i. e., 1.5-D modeling) (Fig. 1). This reference model is solved first, employing analytical solution techniques. The full model is then computed using the reference model as a background solution.

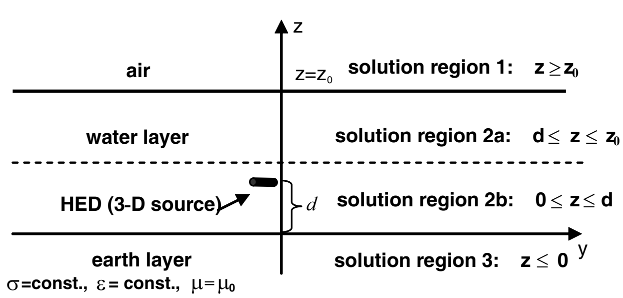


Figure 1: Schematics of the 1.5-D reference model (three layers) including the different solution regions.

2.1. Primary Field (Reference) Solution

We introduce a Cartesian coordinate system (cf. Fig. 1), and let the x-axis represent the cross-line direction, along which the parameters do not vary. Moreover, we assume a cross-line polarized HED-source. The reference model can be divided into four different solution regions (Fig. 1). The Maxwell's equations can now be spatially Fourier transformed along both the x- and y-direction. Based on combinations of these transformed equations, wave-equations for the field components $\hat{E}_{jx}$ and $\hat{H}_{jx}$ can be derived (subscript j representing region j, and the $\{\wedge\}$-symbol implies doubly spatially Fourier transformed fields):

$$\left[\partial_z^2 + k_{jz}^2\right]\hat{H}_{jx}=0 \text{ and } \left[\partial_z^2+k_{jz}^2\right]\hat{E}_{jx} = \frac{\left(k_j^2-k_x^2\right)}{i\omega\varepsilon_j^*}\delta(z-d), \text{ where } k_j^2=\omega^2\mu_0\varepsilon_j^*,\ \varepsilon_j^* = \varepsilon_j+\frac{i\sigma_j}{\omega} \tag{1}$$

where I is the source current and dl is its infinitesimal length. Moreover, the source is placed a distance d above the seafloor. In order to construct feasible solutions a radiation condition must be imposed: only solutions decaying with propagation distance can be allowed within the air and earth half spaces. Hence, a sign-convention for the vertical wavenumber k_{jz} must be set:

$$k_{jz} = \sqrt{k_j^2 - k_x^2 - k_y^2} = \sqrt{i\sigma_j\omega\mu_0 + \varepsilon_j\omega^2\mu_0 - \left(k_x^2 + k_y^2\right)} \tag{2}$$

Inside the water layer (solution regions $2a$ and $2b$ in Fig. 1) the complete and general solutions read

$$\begin{aligned}\hat{E}_{jx} &= \gamma\exp\left[ik_{jz}\cdot z\right] + \eta\exp\left[-ik_{jz}\cdot z\right] - \frac{\left(k_j^2-k_x^2\right)Idl}{2k_{jz}\omega\varepsilon_j^*}\exp\left[ik_{jz}\cdot|z-d|\right],\\ \hat{H}_{jx} &= \alpha\exp\left[ik_{jz}\cdot z\right] + \beta\exp\left[-ik_{jz}\cdot z\right]\end{aligned} \tag{3}$$

where α, β, γ and η are complex constants. For solution regions 1 and 3 (cf. Fig. 1) the solutions must be source-free and outward-propagating only. Next, assuming continuity of the tangential components of the fields across the boundaries gives:

2.2. Solution region 1

$$\hat{H}_{1x} = c_1\exp\left[ik_{2z}\cdot z_0 + ik_{1z}\left(z-z_0\right)\right] - 2ic_2\sin\left(k_{2z}\cdot z_0\right)\exp\left[ik_{1z}\left(z-z_0\right)\right] \tag{4a}$$

$$\begin{aligned}\hat{E}_{1x} &= \omega\mu_0c_3\exp\left[ik_{2z}\cdot z_0 + ik_{1z}\left(z-z_0\right)\right] - 2ic_4\omega\mu_0\sin\left(k_{2z}\cdot z_0\right)\exp\left[ik_{iz}\left(z-z_0\right)\right]\\ &\quad -\frac{\omega\mu_0\left(k_2^2-k_x^2\right)Idl}{ik_2^2k_{2z}}\sin\left(k_{2z}\cdot d\right)\exp\left[ik_{2z}\cdot z_0 + ik_{1z}\left(z-z_0\right)\right]\end{aligned} \tag{4b}$$

2.3. Solution regions $2A$ and $2B$

$$\hat{H}_{2x}=c_1\exp\left[ik_{2z}\cdot z\right] - 2ic_2\sin\left(k_{2z}\cdot z\right) \tag{5a}$$

$$\begin{aligned}\hat{E}_{2z}&=\omega\mu_0c_3\exp[ik_{2z}\cdot z]-2i\omega\mu_0c_4\sin(k_{2z}\cdot z)-\frac{\omega\mu_0\left(k_2^2-k_x^2\right)Idl}{ik_2^2k_{2z}}\sin(k_{2z}\cdot d)\exp[ik_{1z}\cdot z],\ d<z<z_0\\ \hat{E}_{2z}&=\omega\mu_0c_3\exp[ik_{2z}\cdot z]-2i\omega\mu_0c_4\sin(k_{2z}\cdot z)-\frac{\omega\mu_0\left(k_2^2-k_x^2\right)Idl}{ik_2^2k_{2z}}\sin(k_{2z}\cdot z)\exp[ik_{1z}\cdot d],\ 0<z<d\end{aligned} \tag{5b}$$

2.4. Solution region 3

$$\hat{H}_{3x} = c_1\exp\left[-ik_{3z}\cdot z\right],\ \hat{E}_{3x} = \omega\mu_0c_3\exp\left[-ik_{3z}\cdot z\right] \tag{6}$$

Finally, we assume continuity of the y-component of the E- and H-field across the same two interfaces. Hence, we obtain four independent equations giving the solution for the coefficients c_1, c_2, c_3 and c_4. The linear system of equations can be written as:

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4\\ a_3 & a_4 & a_5 & a_6\\ a_7 & a_8 & a_9 & 0\\ a_9 & 0 & a_{10} & a_{11}\end{bmatrix}\begin{bmatrix}c_1\\c_2\\c_3\\c_4\end{bmatrix} = \begin{bmatrix}b_1\\b_2\\0\\b_3\end{bmatrix} \tag{7}$$

where

$$a_1=\left[k_{1z}\left(k_2^2-k_x^2\right)-k_{2z}\left(k_1^2-k_x^2\right)\right]\exp\left[i2k_{2z}\cdot z_0\right],\quad a_2=k_{2z}\left(k_1^2-k_x^2\right)+k_{1z}\left(k_2^2-k_x^2\right)-a_1 \quad (8a)$$

$$a_3=k_xk_y\left(k_2^2-k_1^2\right)\exp\left[i2k_{2z}\cdot z_0\right],\quad a_4=k_xk_y\left(k_2^2-k_1^2\right)-a_3 \quad (8b)$$

$$a_5=\left[k_2^2k_{2z}\left(k_1^2-k_x^2\right)-k_1^2k_{1z}\left(k_2^2-k_x^2\right)\right]\exp\left[i2k_{2z}\cdot z_0\right],\quad a_6=-k_2^2k_{2z}\left(k_1^2-k_x^2\right)-k_1^2k_{1z}\left(k_2^2-k_x^2\right)-a_5 \quad (8c)$$

$$a_7=k_{3z}\left(k_2^2-k_x^2\right)+k_{2z}\left(k_3^2-k_x^2\right),\quad a_8=-2k_{2z}\left(k_3^2-k_x^2\right),\quad a_9=k_xk_y\left(k_3^2-k_2^2\right) \quad (8d)$$

$$a_{10}=-k_3^2k_{3z}\left(k_2^2-k_x^2\right)-k_2^2k_{2z}\left(k_3^2-k_x^2\right),\quad a_{11}=2k_2^2k_{2z}\left(k_3^2-k_x^2\right) \quad (8e)$$

$$b_1=\left[Idlk_xk_y\left(k_2^2-k_x^2\right)\sin\left(k_{2z}\cdot d\right)\left(k_2^2-k_1^2\right)\right]\exp\left[i2k_{2z}\cdot z_0\right]/\left(ik_{2z}k_2\right) \quad (8f)$$

$$b_2=-iIdl\sin\left(k_{2z}\cdot d\right)\left[\left(k_1^2-k_x^2\right)\left(k_2^2-k_x^2\right)-k_{1z}k_1^2\left(k_2^2-k_x^2\right)^2/\left(k_{2z}k_2^2\right)\right]\exp\left[i2k_{2z}\cdot z_0\right] \quad (8g)$$

$$b_3=Idl\left(k_2^2-k_x^2\right)\left(k_3^2-k_x^2\right)\exp\left[i2k_{2z}\cdot d\right] \quad (8h)$$

Having solved for these two components, the remaining field components are given as (follows from combinations of the doubly spatially Fourier transformed Maxwell's equations):

$$\hat{E}_{jy}=\left[-k_xk_y\hat{E}_{jx}+i\omega\mu_0\partial_z\hat{H}_{jx}\right]/\left(k_j^2-k_x^2\right),\quad \hat{E}_{jz}=\left[ik_x\partial_z\hat{E}_{jx}+k_y\omega\mu_0\hat{H}_{jx}\right]/\left(k_j^2-k_x^2\right) \quad (9a)$$

$$\hat{H}_{jy}=\left[-i\omega\varepsilon_j^*\partial_z\hat{E}_{jx}-k_xk_y\hat{H}_{jx}\right]/\left(k_j^2-k_x^2\right),\quad \hat{H}_{jz}=\left[-\omega\varepsilon_j^*k_y\hat{E}_{jx}+ik_x\partial_z\hat{H}_{jx}\right]/\left(k_j^2-k_x^2\right) \quad (9b)$$

Finally, the primary (reference) fields are found after employing an inverse spatial Fourier transform over the y-coordinate. Up till now, we have considered a HED-source polarized in the cross-line direction. The case of an inline polarized source can be solved in a very similar manner using simple coordinate transformations.

2.5. Scattered (Difference) Field Solution

Subtracting the reference model from the general model give the difference fields [7]:

$$\nabla\times\overline{E}^d=i\omega\mu_0\overline{H}^d,\quad \nabla\times\overline{H}^d=-i\omega\varepsilon^*\overline{E}^d-i\omega\Delta\varepsilon^*\overline{E}^0\quad \text{where } \Delta\varepsilon^*=\varepsilon^*-\varepsilon^{0*}\cong i\Delta\sigma/\omega \quad (10)$$

In Eq. (10) the term $-i\omega\Delta\varepsilon^*\overline{E}^0$ can be interpreted as a generalized source where $\overline{E}^0$ represents the reference solution of the electric field. Moreover, for a conductive medium $\Delta\varepsilon^*$ can be interpreted as the scaled difference between the total 2-D conductivity and the 1-D background conductivity. Since we assume a 2.5-D formulation it implies that the electric parameters vary in two dimensions only. But the source is a dipole of finite-length, hence a 3-D one. A spatial Fourier transform of Eq. (10) can be carried out along the invariant x-direction. Combination of these transformed equations (scalar versions) give coupled equations for $\widetilde{E}_x^d$ and $\widetilde{H}_x^d$, the difference, along-strike fields in the spatial FT-domain (with $\gamma^2=k_x^2-k^2$):

$$\partial_y\left[\sigma\partial_y\widetilde{E}_x^d/\gamma^2\right]+\partial_z\left[\sigma\partial_z\widetilde{E}_x^d/\gamma^2\right]-\sigma\widetilde{E}_x^d=\Delta\sigma\widetilde{E}_x^0+ik_x\left[\partial_y\left(\Delta\sigma\widetilde{E}_y^0/\gamma^2\right)-\partial_z\left(\Delta\sigma\widetilde{E}_z^0/\gamma^2\right)\right]$$
$$+ik_x\left[\partial_y\left(1/\gamma^2\right)\partial_z\widetilde{H}_x^d-\partial_z\left(1/\gamma^2\right)\partial_y\widetilde{H}_x^d\right] \quad (11a)$$

$$\partial_y\left[\partial_y\widetilde{H}_x^d/\gamma^2\right]+\partial_z\left[\partial_z\widetilde{H}_x^d/\gamma^2\right]-\widetilde{H}_x^d=\left[-\partial_y\left(\Delta\sigma\widetilde{E}_z^0/\gamma^2\right)+\partial_z\left(\Delta\sigma\widetilde{E}_y^0/\gamma^2\right)\right]$$
$$+k_x\left[\partial_y\left(1/\gamma^2\right)\partial_z\widetilde{E}_x^d-\partial_z\left(1/\gamma^2\right)\partial_y\widetilde{E}_x^d\right]/\left(\omega\mu_0\right) \quad (11b)$$

The ($\sim$)-symbol has been introduced to remind that the field components are spatially Fourier transformed. To compute the response of a finite source, a range of non-zero k_x-values must be considered to represent the fields in the FT-domain. In the actual implementation we use a total number of 12 equally spaced (in the logarithmic domain) wavenumbers ranging from 0 to 0.013 m^{-1}. Now assume the along-strike total field components $\widetilde{E}_x=\widetilde{E}_x^0+\widetilde{E}_x^d$ and $\widetilde{H}_x=\widetilde{H}_x^0+\widetilde{H}_x^d$ are given. The other components can be calculated by numerical differentiation of these principal components employing the following equations [derived from Eq. (10) after spatial FT]:

$$\begin{aligned}\widetilde{E}_y&=\left[-i\omega\mu_0\partial_z\widetilde{H}_x^d-ik_x\partial_y\widetilde{E}_x^d+i\omega\mu_0\Delta\sigma\widetilde{E}_y^0\right]/\gamma^2+\widetilde{E}_y^0,\\ \widetilde{E}_z&=\left[i\omega\mu_0\partial_y\widetilde{H}_x^d-ik_x\partial_z\widetilde{E}_x^d+i\omega\mu_0\Delta\sigma\widetilde{E}_z^0\right]/\gamma^2+\widetilde{E}_z^0\end{aligned} \tag{12a}$$

$$\begin{aligned}\widetilde{H}_y&=-\left[\sigma\partial_z\widetilde{E}_x^d+ik_x\partial_y\widetilde{H}_x^d+ik_x\Delta\sigma\widetilde{E}_z^0\right]/\gamma^2+\widetilde{H}_y^0,\\ \widetilde{H}_z&=\left[-ik_x\partial_z\widetilde{H}_x^d+\sigma\partial_y\widetilde{E}_x^d+ik_x\Delta\sigma\widetilde{E}_y^0\right]/\gamma^2+\widetilde{H}_z^0\end{aligned} \tag{12b}$$

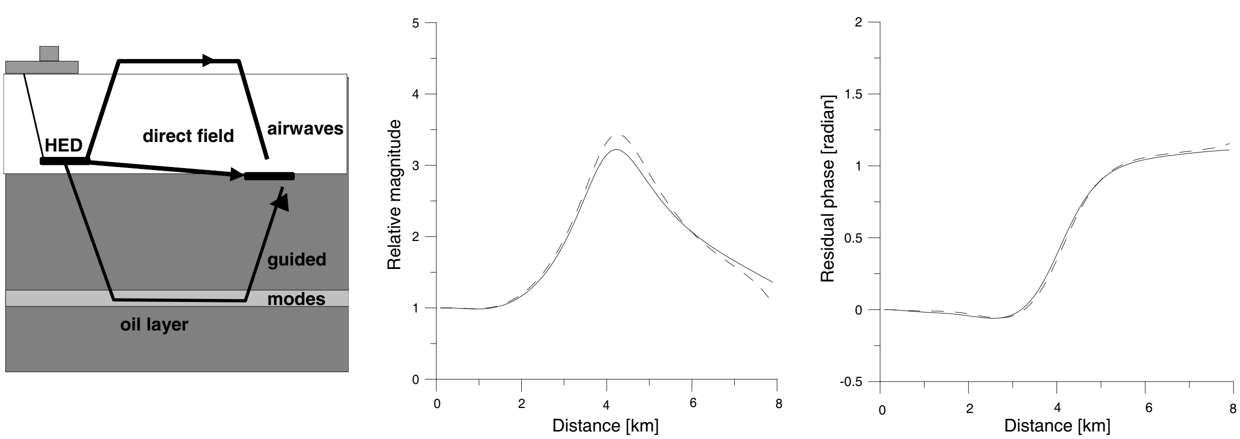
Equations. (11a) and (11b) are 2-D elliptical partial differential equations. They were solved numerically employing a finite-difference scheme with all first derivatives being represented by centered two-point operators as shown in Eq. (13):

$$\begin{aligned}&a(i+1/2,\,j)\cdot\{U(i+1,\,j)-U(i,\,j)\}\,\frac{1}{\Delta y^2}+a(i-1/2,\,j)\cdot\{U(i-1,\,j)-U(i,\,j)\}\,\frac{1}{\Delta y^2}\\ &+a(i,\,j+1/2)\cdot\{U(i,\,j+1)-U(i,\,j)\}\,\frac{1}{\Delta z^2}+a(i,\,j-1/2)\cdot\{U(i,\,j-1)-U(i,\,j)\}\,\frac{1}{\Delta z^2}\\ &+b(i,\,j)U(i,\,j)=s(i,\,j)\end{aligned} \tag{13}$$

where U represents the along-strike electric or magnetic field components, a and b are space-variant and complex coefficients, Δy and Δz represent grid sizes and s is the discrete representation of the generalized source term. The system in Eq. (13) can be written formally as $M\bar{u} = \bar{s}$, where M is a complex matrix containing the known space-variant FD-operators, $\bar{u}$ is the complex EM-wave vector (to be solved) and $\bar{s}$ is the known complex source vector. In the actual implementation, M is modified taking into account the boundary conditions. We used here a complex two-point absorbing boundary condition [4]. The matrix M will be a big and sparse matrix of dimension $N \times N$ where N is equal to the total number of grid points. Hence, one needs an efficient way of storing non-zero numbers. Here we chose to use the so called compressed sparse row (CSR) format. The system in Eq. (13) was solved iteratively employing the conjugate gradient method. Prior to this, the system was preconditioned using incomplete LU factorization (ILU), and transformed to its normal-equation form:

$$\hat{M}^{T^*}\hat{M}\bar{u} = \hat{M}^{T^*}\hat{\bar{s}} \quad \text{where } \hat{M} = K^{-1}M, \quad \hat{\bar{s}} = K^{-1}\bar{s} \tag{14}$$

where T means taking the matrix transpose the asterix denotes complex conjugating and K is the preconditioning matrix. The system matrix $\hat{M}^{T^*}\hat{M}$ in Eq. (14) is Hermitian. In such a case it can be shown that an iterative formulation of Eq. (14) based on the conjugate gradient method will be stable [8]. The coupled Eqs. (11a) and (11b) are solved in an alternating fashion by including the coupling terms on the right-hand sides of the equations in the source term.



(a) (b) (c)

Figure 2: (a) Schematics of test model, (b) Relative magnitude and (c) residual phase of inline polarized HED-source (e. g., with and without hydrocarbon layer). Maximum source-receiver offset was 8 km. Comparison with CSEM1D-program (solid curve).

3. VALIDITY OF THE METHOD

There is a general lack of analytical expressions for 2-D models, so our method was tested against the program CSEM1D, which calculates the response of a Hertz-dipole source in the frequency-domain for a 1-D model. The CSEM1D-program is based on the theory described by Chave and Cox (1982) [2]. Fig. 2(a) shows a sketch of the test model. The HED-source was placed 50 m above the seafloor and had an operating frequency of 0.25 Hz. A water depth of 200 m guaranteed a significant airwave contribution. The conductivities of the overburden and the 200 m thick oil layer were 1 Ω-m and 50 Ω-m, respectively.

Field components were computed for an inline polarized source and receivers placed on the seafloor. Our code showed good agreement with the CSEM1D-program (cf. Figs. 2(b) and (c)), where we have computed the relative (horizontal) electric field, e. g., the ratio between the E_y response associated with the hydrocarbon layer and a homogeneous subsurface, respectively.

ACKNOWLEDGMENT

This research work has been funded by The Norwegian Geotechnical Institute (NGI) and the Norwegian oil company Statoil.

REFERENCES

1. Alumbaugh, D. L., G. A. Newman, G. A. L. Prevost, and J. N. Shahid, "Three-dimensional wideband electromagnetic modeling on massively parallel computers," *Radio Science*, Vol. 31, 1–23, 1996.
2. Chave, A. D. and C. S. Cox, "Controlled electromagnetic sources for measuring electrical conductivity beneath the oceans, part 1: forward problem and model study," *Journal of Geophysical Research*, Vol. 87, 5327–5338, 1982.
3. 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, No. 3, 144–152, 2002.
4. Enquist, B. and A. Majda, "Absorbing boundary condition for the numerical simulation of waves," *Math. Comput.*, Vol. 31, 629–651, 1977.
5. Flosadottir, A. and L. M. MacGregor, "Applications of 3-Dimensional electromagnetic induction by sources in the ocean: development and testing of 3-D modeling code," *Second Annual Scientific Report, Mast III Project Contract No. MAS3-CT97-0120*, 1999.
6. Hohmann, G. W., "Three-dimensional induced polarization and electromagnetic modeling," *Geophysics*, Vol. 40, 309–324, 1975.
7. Unsworth, M. J., B. J. Travis, A. D. Chave, and A. D, "Electromagnetic induction by a finite electric dipole source over a two-dimensional earth," *Geophysics*, Vol. 58, 198–214, 1993.
8. Varga, R. S., *Matrix Iterative Analysis*, Springer Verlag, 1991.

Comparison of Different Finite-difference Schemes of Modeling Marine Controlled-Source Electromagnetic Fields for Hydrocarbon Exploration

Dagang Wu, **Ji Chen**, **and Ce Liu**

The University of Houston, USA

Abstract— In this paper, two different finite difference schemes, based on vector-scalar potential formulation and direct electric field formation, are developed for modeling marine controlled-source electromagnetic survey. These methods are applied to analyze a typical hydrocarbon exploration environment. Based on the results of our simulation, we compare these two methods in terms of their numerical accuracy and efficiency.

1. INTRODUCTION

In the past two decades, the marine controlled-source electromagnetic (MCSEM) method has been used to determine sub-seafloor electric resistivity for hydrocarbon explorations [1, 2]. A typical MCSEM survey configuration is illustrated in Figure 1. This method places receivers on the seafloor to collect signals generated from one or several transmitters towed and powered from a ship. The electric and magnetic fields, collected by receivers, are the results of interactions between the electromagnetic signals with the seawater and sub-seafloor. Various numerical approaches have been used to model such electromagnetic field response in MCSEM method [3, 4] and the focus of this paper is on finite difference (FD) methods [5–9].

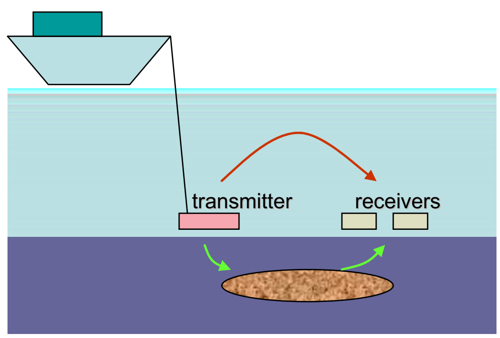


Figure 1: MCSEM survey configuration.

The conventional finite difference method, based on second-order Maxwell's equation [3, 5], has been widely applied to analyze geophysical and electromagnetic problems ranging from direct current (DC) to higher frequencies. Recently, a vector-scalar potential FD formulation has been proposed and applied for low-frequency magnetotelluric (MT) and electromagnetic induction problems [6–9]. In this paper, these two different approaches are implemented and used to analyze typical MCSEM applications.

2. METHODOLOGY

The conventional finite-difference method is based on the solution of Maxwell's equations [3, 5]. Since the operating frequencies of MCSEM survey are often below 1MHz, the effect of displacement current can be ignored from Maxwell's equations. Therefore, we can easily arrive at the following vector Helmholtz equation for electric fields as:

$$\nabla \times \nabla \times E' + j\omega\mu_0\sigma' E' = -j\omega\left(\sigma - \sigma^0\right)E^0 \tag{1}$$

where ω is the angular frequency, μ_0 is free space permeability, σ' is electrical conductivity, σ^0 is the conductivity of the background medium, and E^0 is the primary electric fields produced by the transmitter in the background medium. The Dirichlet boundary condition of electric fields

$(n \times E|_{\partial\Omega} = 0)$ is applied to terminate simulation domains. The total electric fields collected by receivers are computed by summing the primary electric fields and scattered electric fields E' calculated by Equation (1). The FD method used to solve Equation (1) is referred to as the direct FD method.

The second finite-difference method developed here is based on vector-scalar potential formulation as described in [6–10]. We refer this method as the mixed-potential FD method. Following the derivation described in [6–10], we can obtain the second-order partial differential equations for the mixed vector-scalar potentials as:

$$\left.\begin{array}{r} B = \mu_0 \nabla \times A \\ E' = -j\omega\mu_0 A - \nabla A \\ \nabla \cdot A = 0 \end{array}\right\} \Longrightarrow \left\{ \begin{array}{c} \nabla^2 A - j\omega\mu_0\sigma' A - \sigma' \cdot \nabla V = -\left(\sigma - \sigma^0\right) E^0 \\ j\omega\mu_0 \nabla \cdot (\sigma' A) + \nabla \cdot (\sigma' \cdot \nabla V) = \nabla \cdot \left[\left(\sigma' - \sigma^0\right) E^0\right] \end{array}\right. \tag{2}$$

To terminate simulation domains, we employ Dirichelet boundary condition for scalar potential V and mixed boundary conditions for vector potential A [10], shown as the following equations

$$\left.\begin{array}{r} V|_{\partial\Omega} = 0 \\ n \times A|_{\partial\Omega} = 0 \\ \dfrac{\partial(n \cdot A)}{\partial n}|_{\partial\Omega} = 0 \end{array}\right\} \tag{3}$$

Similar as the direct FD method, the total electric field is composed of primary electric field and scatter electric field obtained from (2) with boundary Equation (3).

Both these two finite-difference methods are employed on Yee's staggered grid to solve Maxwell's equation or vector-scalar potential formulation numerically [5–10]. After solving the system matrix generated from the direct FD method and the mixed-potential FD method, total electric fields and corresponding magnetic fields collected by the receivers can be obtained.

3. NUMERICAL EXAMPLE

A three-dimensional MCSEM survey example, shown in Figure 2, is analyzed using the direct FD method and the mixed-potential FD method described briefly above. A cylindrical disk model representing hydrocarbon reservoir is embedded in the seafloor. The radius of this disk model is 500 m. Other geometrical parameters along with the electrical conductivities of seawater, seafloor and disk model are illustrated in Figure 2. The transmitter is modeled as one electric dipole along the x direction and the receivers are offset from the transmitter from 80–1000 m in the x-direction. The operating frequency is chosen to be 1 Hz. The received electric fields along the x-direction at different receivers calculated by the direct FD method and the mixed-potential FD method are shown in Figure 3 and excellent agreement between these two methods is observed. This can be easily proven that these two methods have same accuracy theoretically since they are both derived directly from Maxwell's equations.

As mentioned in [7–10], in terms of solution speed, the mixed-potential FD method shows some advantages over the direct-FD method because it generates a system matrix that is more suitable

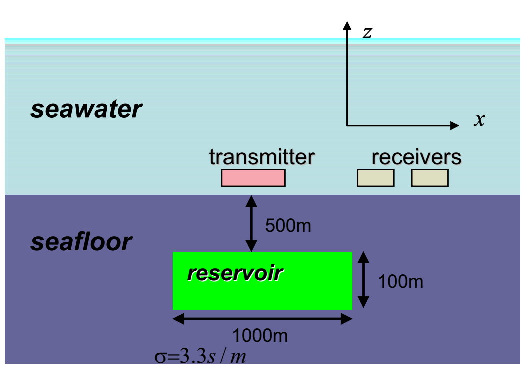


Figure 2: Simulation model.

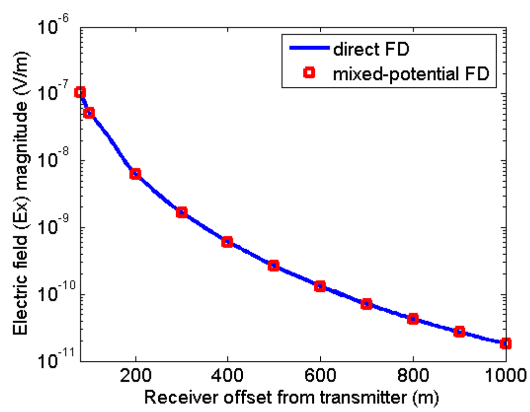


Figure 3: Simulation results.

for iterative matrix solvers. However, this has only been validated by electromagnetic induction problems and MT problems in land environment. In Table 1, we show the convergence rate and the CPU time of these two methods applied to the problem described in Figure 2. In the simulations, the same uniform grid is applied for both methods. BiCGStab method with Jacobi preconditioner is employed to solve matrices produced by these two FD methods and the tolerance of iterative solver is set to be 10^{-6}. The results at different frequencies are tabulated in Table 1.

Table 1: Convergence comparison of direct FD and mixed-potential FD methods.

		Iteration steps		CPU time(s)	
Grid size(m)	Frequency(Hz)	Direct FD	Mixed-potential FD	Direct FD	Mixed-potential FD
50	0.01	> 500	200	> 510	194
	0.10	> 500	208	> 510	202
	1.00	190	> 500	194	> 400
	5.00	55	> 500	58	> 400

From Table 1, it is observed that the mixed-potential FD has faster convergence rate than that of the direct FD at low frequencies, while this is no longer true at higher frequencies. This phenomenon might be explained as follows. In MCSEM survey modelings, large grid size (20∼100 m) is often used in mixed-potential FD formulations because simulation models have very large dimensions (1000∼4000 m). Unfortunately, at higher frequencies, this could make the mixed-potential FD system matrix no longer block-diagonal dominant [7, 8]. Consequently, this will increase the condition number of the system matrix and lead to slow convergence for iterative matrix solvers. On the other hand, the mixed-potential FD has shown an advantage over direct FD when small grid sizes were applied [7–10], where block-diagonal dominance is satisfied.

4. CONCLUSIONS

In this paper, two numerical finite-difference approaches, the direct FD method and the mixed-potential FD method, are formulated and implemented to simulate controlled-source electromagnetic survey in marine environment. Their accuracies and efficiencies are compared by numerical examples. From our simulation results, we observe that these two different methods have a similar numerical accuracy while the efficiencies of these two methods may depend on specific applications. Based on our preliminary numerical experiments, we found that the mixed-potential FD method may not always be advantageous than over the direct FD method in MCSEM survey modelings.

REFERENCES

1. Edwards, N., "Marine controlled source electromagnetics: principles, methodologies, future commercial applications," *Surveys in Geophysics*, Vol. 26, 675–700, 2005.
2. MacGregor, L. M. and M. C. Sinha, "Use of marine controlled source electromagnetic sounding for sub-basalt exploration," *Geophysical Prospecting*, Vol. 48, 1091–1106, 2000.
3. Evan, S. Um., "On the physics of galvanic source electromagnetic geophysical methods for terrestrial and marine exploration," M. S. Thesis, University of Wisconsin-Madison, 2005.
4. King, J. D., "Using a 3D finite element forward modeling code to analyze resistive structures with controlled-source electromagnetics in a marine environment," M. S. Thesis, Texas A&M University-College station, 2004.
5. Newman, G. A. and D. L. Alumbaugh, "Frequency-domain modeling of airborne electromagnetic responses using staggered finite differences," *Geophysical Prospecting*, Vol. 43, 1021–1042, 1995.
6. LaBrecque, D. J., "Finite difference modeling of 3-D EM fields with scalar and vector Potentials," Oritaglio, M., and Spies, B., Eds., *Three-dimensional Electromagnetics: Soc. Expl. Geophys*, 146–160, 1999.
7. Haber, E., U. M. Ascher, and D. W. Oldenburg, "Fast simulation of 3D electromagnetic problems using potentials," *J. Comput. Phys.*, Vol. 163, 150–171, 2000.

8. Haber, E. and U. M. Ascher, "Fast finite volume simulation of 3D electromagnetic problems with high discontinuous coefficients," *SIAM J. Scientific Compututing.*, Vol. 22, 1943–1961, 2001.
9. Hou, J., R. K. Mallan, and C. Torres-Verdin, "Finite-difference simulation of borehole EM measurements in 3D anisotropic media using coupled scalar-vector potentials," *Geophysics*, Vol. 71, G225–G233, 2006.
10. Fang, S. and T. Wang, "Wideband electromagnetic finite-difference modeling for 3D cylindrical geometries," *75th Ann. Internat. Mtg., Soc. Expl. Geophys.*, Expanded Abstracts, 348–351, 2005.

The New Progress in Near-surface Seafloor Exploration by Transient Electromagnetic Method

Jun Lin, **Changsheng Liu**, **Yan Wang**, **Fengdao Zhou**, **and Hui Li**
Key Lab of Geo-Exploration and Instrumentation, Jilin University, Ministry of Education
Changchun, China

Abstract— Transient electromagnetic method has been used to investigate the conductivity of seafloor since late $20th$ century. When the exploration was done, two antennas of large separation distance were placed over seafloor, and the step response is measured. In this paper, we study the exploration of seafloor conductivity by measuring the inductive voltage of receiver loop, which is the time derivative of step response. A configuration of central loop is used. Its transmitter loop is small and a smaller receiver loop is placed in the center. An instrument system which can be put into water and controlled from the computer on ship is developed for the exploration. Because the configuration is small and part function of the instrument system is specially designed for measurement during moving, high efficient, towed exploration can be easily achieved. Finally, an experiment of seafloor exploration was performed in Hulu Island, Liaoning province, and two oil pipelines are found.

1. INTRODUCTION

In recent years, oceanic resource exploitation and seafloor engineering are paid more attention to, and geophysical exploration in ocean is demanded heavily. Many geophysical methods were developed and used for ocean exploration [4]. As a basilic way of geophysical exploration, transient electromagnetic method was also attempted to be applied to seafloor exploration [6]. The research of transient electromagnetic exploration on seafloor mainly focused on various dipole-dipole systems, in which theoretical step responses for various dipole-dipole configurations located on seafloor surface were studied [1–3, 7]. And horizontal, in-line, electric dipole-dipole system and horizontal, coaxial, magnetic dipole-dipole system were applied to the exploration experiments of sea floor [5, 6]. In the exploration, configurations were located on seafloor, and instrument systems, commonly used for land exploration, were placed on ship. The survey data was finally inversed to get the conductivity of half-space seafloor. This paper studies the project of seafloor exploration by measuring the derivative of step response — inductive voltage of central loop which is commonly used on land, and putting both instrument system and antennas in water for continuously towed measurement.

2. THE INDUCTIVE VOLTAGE OF CENTRAL LOOP ON SEAFLOOR

Based on electromagnetic theory, taking both seawater and seafloor as half-space, the inductive voltage of central loop on seafloor can be computed. Assuming the seawater conductivity is 3 S/m which is close to the real value of ocean, when the seafloor conductivity is changed, the inductive voltage curves of receiver loop, which are function of time, are showed in Figure 1. The inductive voltage curve varies with seafloor conductivity. The seafloor of small conductivity can be distinguished in early time and the seafloor of high conductivity can be distinguished in late time. Thus by measuring the inductive voltage in receiver loop, the seafloor conductivity can be determined.

3. THE INSTRUMENT SYSTEM FOR SHALLOW SEAFLOOR EXPLORATION

During the exploration on seafloor, the antennas are located on seafloor. In order to avoid the negative effect of survey ship and long transmission cable between instrument and antennas, the instrument system should also be put into water close to the antennas. According to the requirements of seafloor exploration, a suit of instrument system, STEM-I, is developed by our lab. In STEM-I, transmitter and receiver are incorporate, so that the volume of the instrument system is small and can be placed in a small hermetical container. This system can output current of 70 A, and has maximum sampling rate of 5 μs, and is able to distinguish the minimum voltage of 1 μV. More than 10 G memory may be used to store measure data. When STEM-I is used for seafloor exploration, the instrument system in water can be controlled by upper computer on ship through cable, and the measure data can be transmitted to upper computer for display on time at a transmission rate more than 6 M/s. STEM-1 may start data measure by user's triggering. It can also do it by itself according the time that user set in advance.

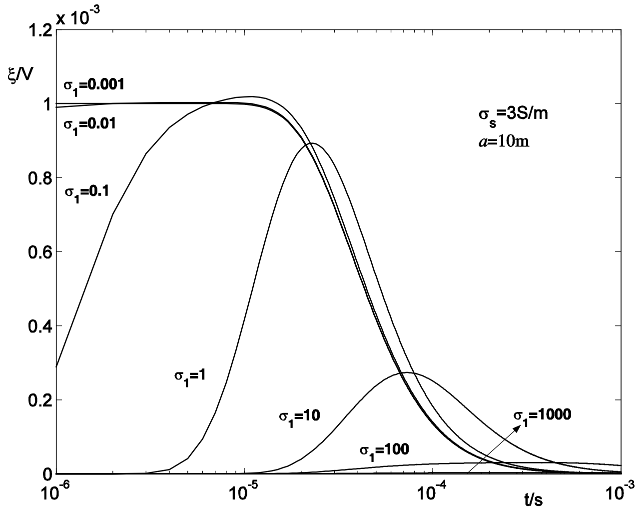


Figure 1: The inductive voltage of central loop. The radius of transmitter loop, a, is 10 meters. The seawater conductivity σ_s is 3 S/m, and the seafloor conductivity σ_1 is changed from 0.001 to 1000 S/m.

4. THE CONTINUOUS SCAN OF SEAFLOOR

When small body in seafloor or refined boundary of different seafloor conductivity is investigated, it is necessary to detect the seafloor at small station distance. If this is done by the traditional way, in which the antennas are placed on each station and stopped for measure, the work efficiency is very small, and the station distance can not be too small. So we measure the transient response of seafloor continuously during sailing. The survey ship tows the antennas and instrument system above seafloor along the survey line. Measure time interval which decides the station distance with the speed of ship is input through upper computer before exploration. When the antennas are on the start station of survey line, the measure begins and then repeats ceaselessly and automatically after each time interval until the whole line is finished. The time interval can also be zero, and then the measure will be uninterrupted.

5. A CASE OF SEAFLOOR PIPELINE EXPLORATION

An experiment was carried on in Hulu Island, Liaoning province. The depth of sea was about 20 meters and the surface of seafloor was covered by sand and silt and was flat. From the known information, we knew that there was oil pipeline in the seafloor, but didn't know its exact position. In order to find the pipeline, we used a small central loop with transmitter loop of 2×2 m, and continuously scanned the seafloor by STEM-1. The speed of survey ship is about 0.75 m/s. By the exploration, we found two pipelines in the seafloor and got it position. Figure 2 is the profile between 800 and 900 m in survey line. There is a peak around 827 m, which indicates the existence of a high conductive body, oil pipeline, in this position.

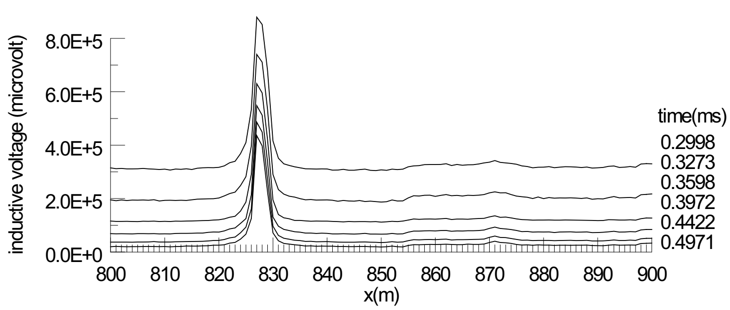


Figure 2: The profile of inductive voltage in Hulu island.

ACKNOWLEDGMENT

This research is funded by 863 Program of China under grant 2002AA615030. The authors are also grateful to all the members in EM group of the Key Lab of Geo-Exploration and Instrumentation, for their work in this research.

REFERENCES

1. Cheesman, S. J., R. N. Edwards, and A. D. Chave, "On the theory of sea-floor conductivity mapping using transient electromagnetic systems," *Geophysics*, Vol. 52, No. 2, 204–217, 1987.
2. Edwards, R. N. and A. D. Chave, "A transient electric dipole-dipole method for mapping the conductivity of the sea floor," *Geophysics*, Vol. 51, No. 4, 984–987, 1986.
3. Edwards, R. N., "Two-dimensional modeling of a towed in-line electric dipole-dipole sea-floor electromagnetic system: The optimum time delay or frequency for target resolution," *Geophysics*, Vol. 53, No. 6, 846–853, 1988.
4. He, J. S. and L. Z. Bao, "The situation and progress of marine electromagnetic method reseach (in Chinese with English abstract)," *Progress in Geophysics*, Vol. 14, No. 1, 7–39, 1999.
5. Wang, Y. X., J. L. Wang, J. Y. Wang, et al., "The research of conductivity on sea-floor by transient electromagnetic system (in Chinese with English abstract)," *Chinese Journal of Geophysics*, Vol. 41, No. 6, 841–847, 1998.
6. Yang, J. W. and R. N. Edwards, "Controlled source time-domain electromagnetic methods for sea-floor conductivity mapping (in Chinese with English abstract)," *The Chinese Journal of Nonferrous Metals*, Vol. 8, No. 4, 705–713, 1998.
7. Zheng, R. S., "Forward modeling for transient electromagnetic response of ocean bottom reservoir model (in Chinese with English abstract)," *Journal of Oil and Gas Technology*, Vol. 27, No. 2, 190-193, 2005.

Study of the High Frequency Magnetotelluric Sounding for Prospecting the Deep and Periphery Mine by RRI Inversion

Jingtian Tang, **Ye Wang**, **Xiao Xiao**, **Jifeng Zhang**
Huakun Du, and **Caikun Gao**

Info-physics and Geomatics Engineering School, Central South University, China

Abstract— With the development of our country's economy, it is necessary that more and more mine resource is produced. After exploiting the earth's surface mine many years, the target of mine prospecting is turn to deep, peripheral area. In this paper, the author expatiate the effect of the high frequency MT method by the stratagem$^{\mathrm{TM}}$ EH-4 electrical conductivity image-forming system, and use RRI inversion, the method can prospect the mine field in specifically depth, so this method can forecast advance the geologic struc-ture and mineralize abnormity of the work area .Whereafter, the paper bring forth the successful result by analyzing the field data.

1. INTRODUCTION

With the development of our country's economy, it is necessary that more and more mine resource is produced. So it is a hidden trouble for our economic development and our mine resource steady supply that little mine resource reserves. For the mine deep and periphery areas, it is difficult to seek mine resource by traditional tracing the earth's surface mineral way. The mine prospecting must close combine with geophysical method and drill information, etc. Obtaining more accurate and abundant geology information, synthesizing traditional geological method, geochemistry method and geophysical method, we can carry out important breakthrough in mine prospecting. For most reserves critical mine, it is best way for potential superseding mine resource that prospecting deep and periphery area on mine. Because the current mine has been a long time prospecting, the most potential mine resource in the deep and periphery area. In this case, it is a kind of fast, precision, economic method that estimating macroscopically the deep and periphery mineral body by making full use of various geophysics methods. This paper introduces a kind of fast geophysics method to prospecting the mine resource, it is High Frequency Magnetotelluric Method (HMT), the HMT method's frequency range from 10 Hz to 100 KHz, using the Stratage$^{\mathrm{TM}}$ EH-4 conductivity imaging system that made in USA. The paper analyses the field data in a mine in Inner Mongolia by RRI inversion. It indicate that the HMT method have the virtue of high accuracy, high resolution, fast and convenience.

2. BRIEF INTRODUCTION OF HMT AND EH — 4 CONDUCTIVITY IMAGING SYSTEM

2.1. The Principle of HMT Method

The MT method is a kind of geophysical method that studies the earth crust . The source is natural electromagnetism field, when the EM wave transmits in underground, because of the electromagnetic induction effect, the electromagnetism field of the ground will include the information

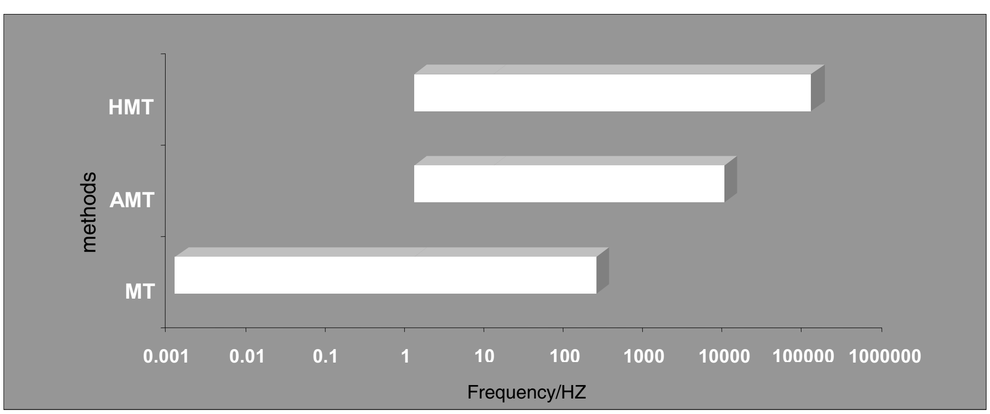


Figure 1: Frequency distribution of MT, AMT, HMT.

of geological body's resistivity. The difference of HMT, MT, AMT is the frequency range. MT method uses the lower frequency EM wave , range from $n \times 10^{-3}$ Hz to $n \times 10^{2}$ Hz, AMT method uses the EM wave frequency range from n Hz to 8192 Hz. Both MT and AMT use the lower frequency, so the exploratory depth is deeply; But the HMT method use the EM wave frequency range from 10 Hz to 100 kHz, higher frequency than the MT and AMT, The HMT method can prospect 1000 meters underground. At the same time, the higher frequency of HMT has more anti-interference ability. It's working frequency scope is in Figure 1.

2.2. The EH — 4 Conductivity Imaging System

The StratageTM EH - 4 conductivity imaging system is the geophysical instrument that made by the American company of EMI and Geometrics, it measures the electrical resistivity of the earth over depth ranges of a few meters to greater than one kilometers. This system uses natural and man-made electromagnetic signals. The instrument can be use for continuous electric conductivity section prospecting in steep area, That system measures the electric field ponderances EX, EY and magnetic field ponderances HX, HY, then obtains the resistivity sounding curve by FFT and power spectral density calculation, the higher frequency field data reflects the geology characteristic of the shallow earth, the lower frequency field data reflects the deep earth. Its work setting is in Figure 2.

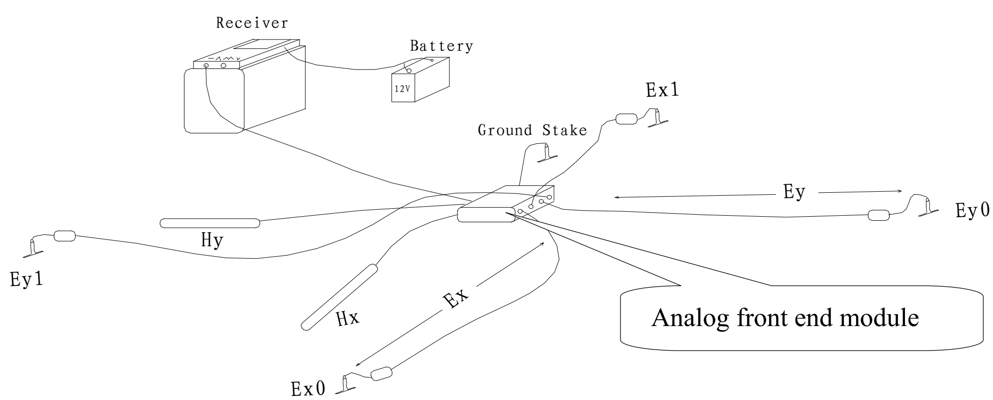


Figure 2: The layout map of EH-4 system.

3. THEORY OF RRI INVERSION

In the Cartesian system of coordinates, we suppose that the X axes parallel with the 2D structure, the Y axes perpendicularity the 2D structure, and the Z axes downhill. Neglecting the magnetism of iron ands displacement current, by the Maxwell square of low frequency EM wave, we can get the following formula through deduction for the TE and TM mode:

$$\begin{cases} \nabla^2 E_x = -i\omega\mu\sigma(y,\, z)E_x \\ \dfrac{\partial E_x}{\partial z} = i\omega\mu H_y \end{cases} \tag{1}$$

$$\begin{cases} \nabla^2 H_x + \nabla\rho \cdot \nabla H_x = -i\omega\mu_0 H_x \\ \rho\dfrac{\partial H_x}{\partial z} = E_y \end{cases} \tag{2}$$

Formula (1) and (2) are equations of TE and of TM mode, when we give the boundary condition, we can simulate forward model.

Define variable to the TE and mode of TMs respectively:

$$V = \frac{1}{E}\frac{\partial E_x}{\partial z} = i\omega\mu\frac{H_y}{E_x} \quad U = \frac{\rho}{H_x}\frac{\partial H_x}{\partial z} = \frac{E_y}{H_x} = Z_{yx}$$

Make the disturb analysis, and then establish the line integral calculus equations between data disturbing and model parameter disturbing:

$$\begin{cases} \delta d_{xy} = \dfrac{2}{V\left(y_i,\, 0\right)}\delta V = \displaystyle\int \dfrac{2\sigma_0(z)E_{x,0}^2\left(y_i,\, z\right)}{E_{x,0}\left(y_i,\, 0\right)H_{y,0}\left(y_i,\, 0\right)}\delta(\ln\sigma)dz \\ \delta d_{yx} = \dfrac{2}{U\left(y_i,\, 0\right)}\delta U = \displaystyle\int \dfrac{-2\sigma_0(Z)E_{y,0}^2\left(y_i,\, z\right)}{E_{y,0}\left(y_i,\, 0\right)H_{x,0}\left(y_i,\, 0\right)}\delta(\ln\sigma)dz \end{cases} \tag{3}$$

δd_{xy} and δd_{yx} are difference between the field data and the theories data for the TE and TM mode, $\sigma_0(z)$ is the electrical conductivity before change model, $H_{y,0}\left(y_i,\, 0\right)$, $E_{x,0}\left(y_i,\, 0\right)$, $H_{x,0}\left(y_i,\, 0\right)$ and $E_{y,0}\left(y_i,\, 0\right)$ are respectively the value of magnetic field and electric field in the i site that before change model, $E_{y,0}\left(y_i,\, z\right)$ and $E_{x,0}\left(y_i,\, z\right)$ is the theories electric field in the i site originally model or current iterative model.

In 2D inversion, we formulate the objective function considering horizontal and vertical of asymmetry

$$Q\left(y_i\right) = \int \left[\frac{\partial^2 m\left(y_i,\, z\right)}{\partial f^2(z)} + g(z)\frac{\partial^2 m(y,\, z)}{\partial y^2}\bigg|_{y=y_i}\frac{\partial^2 z}{\partial f^2(z)}\right]^2 df(z) \tag{4}$$

This is a Laplace norm number in all sites. The function $f(z)$ can control the length of a mark, and can be used to measure different depth of structure, take $f(z) = \ln\left(z + z_0\right)$, and constant z_0 is usually selected by surface resistivity and is the skin depth of the highest frequency. Take $m = \ln(\sigma) = -\ln(\rho)$. $g(z)$ is a punishment factor that control the horizontal direction structure.

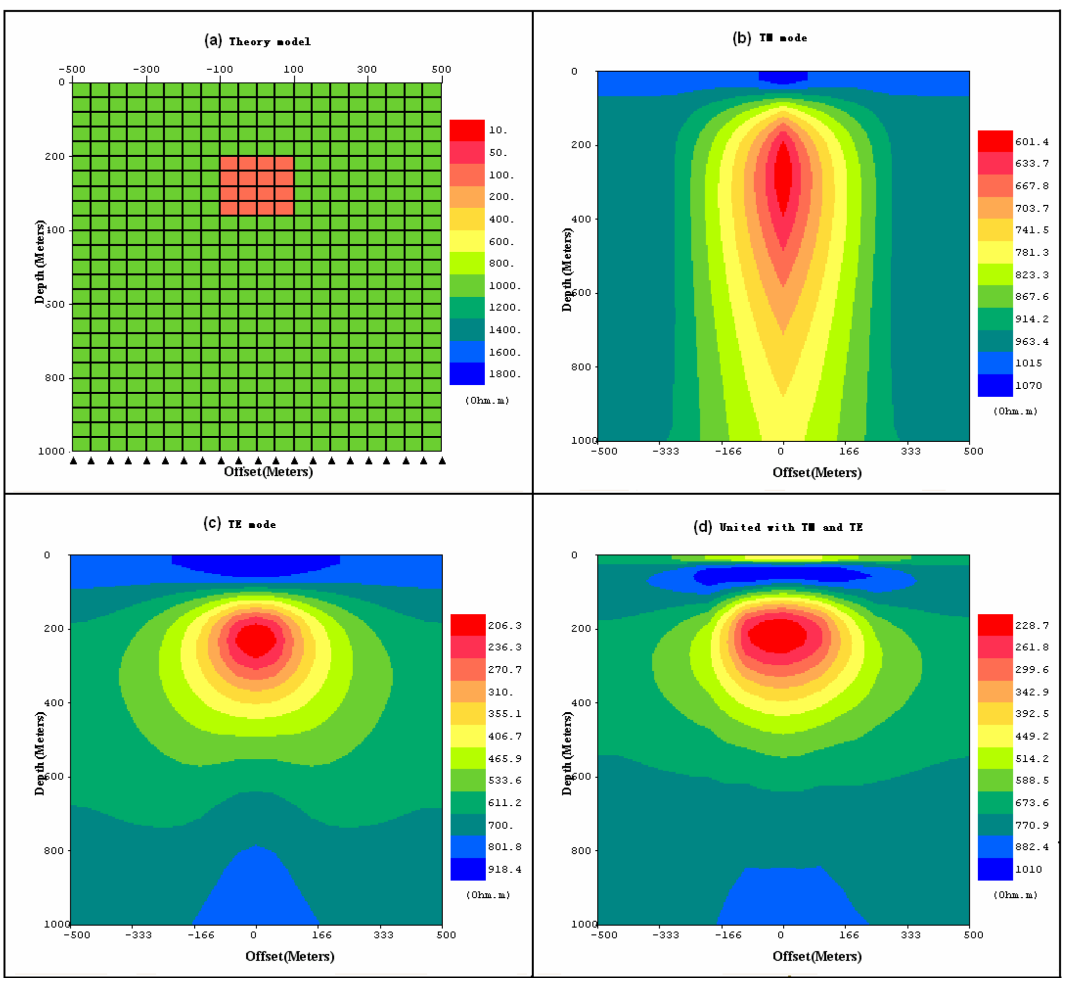


Figure 3: The inversion result of model with different type of data ((a)forward model, (b) TM mode, (c) TE mode, (d) united with TM and TE mode).

4. THE RESULT OF THEORETICAL MODEL INVERSION

For the sake of testing inversion result, validating inversion effect, this paper carries on 2D RRI inversion for two theories models.

In Figure 3(a), in homogeneousbe half space, the background resistivity is 1000 Ω·m, there has a low resistivity body cover 200 m×160 m underground 200 meters, the low resistivity body is 100 Ω·m. We adopt the mesh of 50 m×40 m in this model, taking the frequency scope as the 10 Hz∼100 kHz, we calculate the forward model. When we carry on the 2D RRI inversion use the result of forward model, we choose the original model's resistivity is 1000 Ω·m, establish its horizontal mesh size just equal to the forward model, the vertical mesh size is 10 m in the top three layers, the rest of the layer thickness will be 1.1 multiply step by step. We can get the inversion result of TM mode, TE mode, united with TM and TE mode, and will get such as the Figures 3(b) to 3(d) inversion result. Because of the mesh size of the model invariable, inversion time all about 8 minutes.

In Figure 3(b), the mode of TM ,there is a vertical low resistivity body , because of the influence of a low resistivity near earth surface, this is caused by the static effect obviously. In Figure 3(c), the mode of TE , can prospecting the site and size well .Contrast the mode of TM and mode of TE, we can know, for same section, the data influence of the mode of TM by the static effect bigger than in the data of the mode of TE. In Figure 3(d), the mode of united with TM and TE, can prospecting the site and size, too. Moreover, its background value even closes to the true circumstance.

According to the analysis of the three inversing models, we find that as long as we choose the reasonable original electric resistivity value, establishing the reasonable mesh size, we can get good inversion result. And that, we can carry on 2D RRI inversion in the common PC, the speed very quick (usually from several minutes to more than ten minutes), this makes 2D RRI inversion have greatly practical value.

5. FIELD DATA STUDY OF INNER MONGOLIA MINE

There has two HMT method lines in The Inner Mongolia mine, two lines (linn1, line2) are mutual parallelism, and two lines are apart from the 200 m mutually. There are four drills in line 1, there

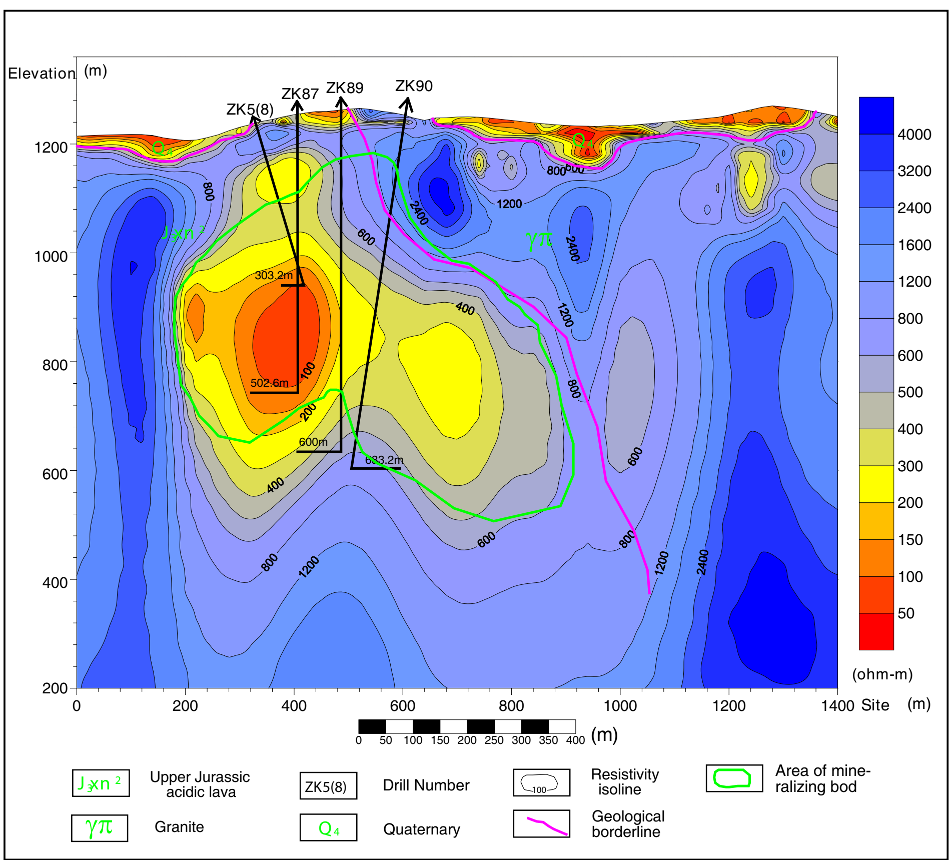


Figure 4: The synthetic cross section in line one.

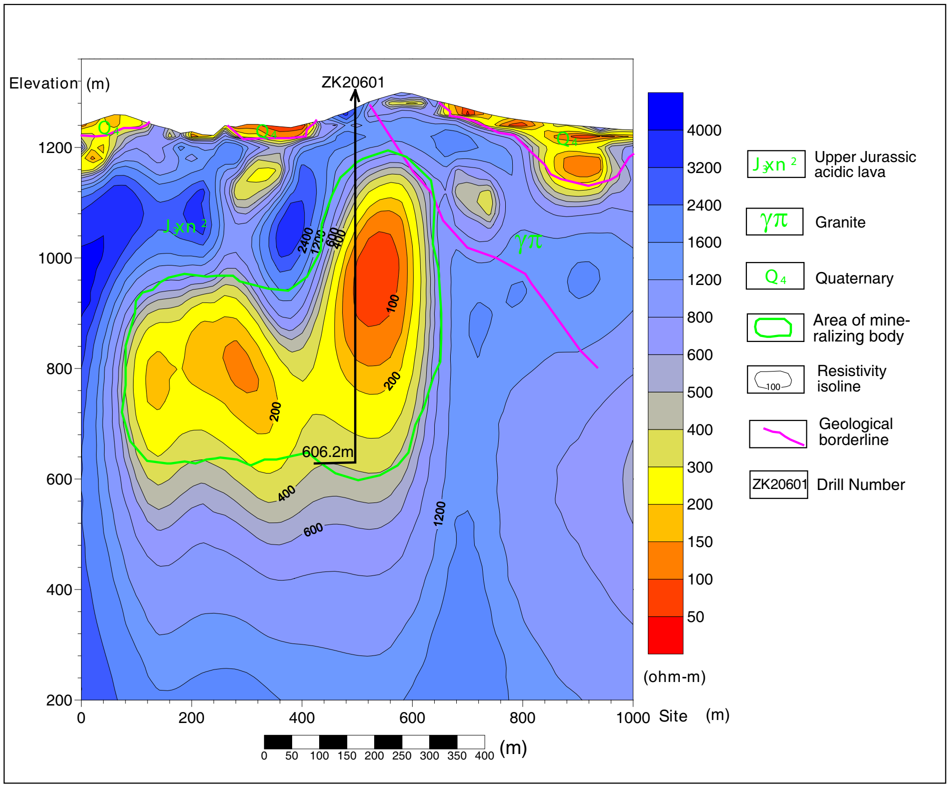


Figure 5: The synthetic cross section in line two.

are one drills in line 2. All drills information indicate that there are mineralize under ground 80 m, special below earth's surface range from 200 m to 500 m.

Figure 4 and Figure 5 are the scale united with the TE and TM RRI inversion maps respectively, and synthesizing the drill information and the geological data. In Figure 4, there is a 300 Ω·m low resistivity body in line 1 , it's in site 180∼880, the elevation is 560 m∼1150 m, we speculate that it is a mineral area. In Figure 5, there is a 300 Ω·m low resistivity body in line 2 , it's in site 80∼660, the elevation is 600 m∼1150 m, we speculate that it is a mineral area.

Contrasting the Figure 4 with Figure 5, we find the line 2 and line 1 very resemble of the electricity characteristic, both of them have a bigger low resistivity abnormality from 600 m to 1150 m at the elevation we speculate that it is a mineral area. Because the quartz is brittleness, there is much cranny, and well connection, so the mine of the low resistivity fill in the cranny, hence, the dense mineral body reflects the whole electricity characteristic of be like the Figure 4 and Figure 5.

According to the drill information, ZK87 hole, the mineralize body from underground 141.5 m to the 502.63 m, the thickness amounts to the 361.13 m; ZK89 hole, the mineralize body from underground 83.5 m to the 467.9 m, the thickness amounts to the 384.4 m; ZK90 hole, the mineralize body from underground 74.8 m to the 633.2 m, the thickness amounts to the 558.4 m; The ZK5(8) hole, the mineralize body from underground 110.4 m to the 302.4 m, the thickness amounts to the 192 m.ZK20601 hole of the line 2, the mineralize body from underground 97.3 m to 606.2 m, the thickness amounts to 472.5 m; All drills information indicate that the better mineralize body from underground 200 m to the 500 m. Figure 4, Figure 5 fit the drills information very well.

6. CONCLUSIONS

According to the mineral prospecting by using EH-4 electric conductivity imaging system in Inner Mongol, we draw a conclusion that the HMT method is effective in prospecting deep and peripheral area, the electromagnetism array profile (EMAP) can inhibit the static effect effectively, the space filter technical minishes the influence of the topography, Moreover, the HMT method has the advantages of prospecting large depth, high accuracy and high resolution. At the same time, contrasting the theoretical model with field data inversion, we can see that the RRI inversion is a steady algorithm, consuming little time. As long as we according to actual circumstance (the various foregone information), carry on the initialize model reasonably, and choose the better mesh density, can get good result.

The HMT method can not only be applied to the metallic mineral prospecting, but also can be applied to engineering geological survey, looking for the groundwater, prospecting the shallow geological structure. It is a kind of promising geophysical method.

REFERENCES

1. Xiao, X., "Study of MT 2D RRI inversion in Inner Mongolia mine," Master paper, Info-physics and Geomatics Engineering School, Central South University, 2006.
2. Keller, G. V. and A. A. Kaufman, *Frequency and Transient Sounding*, Geology Press, Beijing, 1987.
3. Smith, J. T. and J. R. Booker, "Rapid inversion of two- and three-dimensional magnetotelluric data," *J. of Geophys. Res.*, Vol. 96, 3905–3922, 1991.
4. Liu, G.-D. and T.-Y. Hao, "Searching of hidden mineral deposits by geophysical method," *Chinese Journal of Geophysics*, Vol. 36, No. 4, 850–854, 1995.
5. Liu, L.-M. and Z.-Q. Wang, etc., "Application of comprehensive information prediction for ore prospecting in deeper and outer of mines with critical reserves: an example from the Feng Huang Shan mine," Tong Ling, *Chinese Journal of Geology*, Vol. 37, No. 4, 444–452, 2002.
6. Shi, Y.-J.and G.-D. Liu, etc., *The Course of Magnetortelluric Sounding*, Earthquake Press, Beijing, 1985.
7. Tang, J.-T. and J.-S. He , "A new method to define the full-zone resistivity in horizontal electric dipole frequency soundings on a layered earth," *Chinese Journal of Geophysics*, Vol. 18, No. 2, 543–552, 1994.
8. Zhang, X. and R. Chen, "The successful utilization of IP relative phase method for distinguishing anomalies related to mineralization," *Geophysical* & *Geochemical Exploration*, Vol. 22, No. 4, 251 254, 1998.
9. Zhou, H., W. Q. Song, and B.-X. Yin, "The inversion of 2D MT data by the electromagnetic reverse scattering," *Progress in Geophysics*, Vol. 16, No. 2, 50–57, 2001.

Modeling of 3D MCSEM and Sensitivity Analysis

Zhanxiang He[1,2], Zhigang Wang[2], Gang Yu[3], Kurt Strack[3], and Haiying Liu[2]

[1]Department of Geosciences, University of Houston, USA
[2]BGP CNPC, China
[3]KMS Technologies — KJT Enterprises, Inc, USA

Abstract— After decades of research and development, Marine Controlled Source Electromagnetic (MCSEM) has come into the application phase. It has shown good practical effect, but the 3D modeling is far from practical stage. Based on the modeling of 3D frequency-domain MCSEM to a complicated target body and sensitivity analysis, a method which can quantificationally delimit boundary of anomaly body is described in the paper. For air-wave-dominated far field zone, we suggest to adopt the theory of plane wave to calculate apparent resistivity thereby improving the adaptability of the method.

1. INTRODUCTION

In the past decades of years, the theoretical research and practical application of MCSEM have undergone hard process. The initial research began early in 1970's [1], but only in the recent years when digital instrument [3, 6], data processing algorithm [7] and interpretation method are significantly developed, MCSEM came into application phase [2–5, 8]. Presently, the widely adopted inversion is 1D. Although 1D inversion result is approximate, it is very useful [9]. In 2004, Mittet and Ellingsrud (2004) released their 1D inversion algorithm and result, which indicates that 1D inversion can identify hydrocarbon layer.

In 2002, Eidesmo used 2D finite element algorithm to study boundary effect of MCSEM (Eidesmo, et al., 2001). Because the 2D EM modeling is not a suitable method to describe anomaly feature, it is seldom used in MCSEM, and the general used is 3D modeling. In recent years, 3D modeling of MCSEM has been developed rapidly. Numerous 3D modeling algorithms are brought forward, thereinto Zhdanov from Utah University is a representative. Ueda adopted quasi-linear approximate algorithm to do rapid CSEM modeling [11] thereby remarkably speeding calculation and ensuring precision. The algorithm gives good result to practical data. When there is more than one source, it consumes more time. Zhdanov and Wan (2005) use 3D integral equation algorithm to do CSEM modeling to complex model with inhomogeneous background. As a result, the hydrocarbon reservoir is accurately imaged with constrain from known salt dome or volcanic body. Zhdanov and Wan (2005) also used the algorithm to do rapid migration image to MCSEM data. The result indicates that the algorithm is an effective method for imaging submarine resistive body. In addition, Zhdanov and Yoshioka (2005) have done 3D iteration inversion to MCSEM data.

Modeling of time-domain MCSEM is late than that of frequency-domain, but according to 1D and 3D modeling results released by Um and Alumbaugh (2005), the time-domain EM is sensitive to small-sized reservoir. When pulse period and T-R spacing are selected suitably, noise can be significantly suppressed and data precision can be improved. It can also reduce the influence of air wave to time-domain EM field in shallow water.

The paper presents the remarkable improvement on the practicability of the algorithm based on the 3D integral equation algorithm presented by Zhdanov. The improved algorithm can realize the modeling of complex practical data. In addition, an approximate quantitative method delimiting resistive body is presented in the paper based on the modeling to different models. The paper also discussesed how to make frequency-domain CSEM method effective in the air-wave-controlled environment.

2. CREATING OF COMPLEX MODEL FOR INTEGRAL EQUATION ALGORITHM

Zhdanov recently presented an integral equation algorithm which can realize the modeling to a model with inhomogeneous background. The algorithm adopts Green function to calculate the response of an 1D horizontal layered model. It can also conduct 3D EM modeling to the model with inhomogeneous resistivity background. The modeling result to a geology model composing several anomaly bodies which have the same geo-electric feature also is presented in the paper. The algorithm shows considerable application foreground. But in the paper there is no detailed description about how to create practically complex model.

The program to rapidly create a 3D resistivity model form reservoir model or from practical 3D seismic-data is presented in the paper. At first, 1D background model and inhomogeneous resistivity background model are constructed individually based on known data. Secondly, construct a target model or practical model based on 3D seismic data and logging data on reservoir zone. Regional or known geology body is usually taken as inhomogeneous resistivity background thereby minimizing target model. For example, for the model composing only one reservoir zone, not only all meshes within reservoir but also those out of reservoir may be evaluated by variable resistivity according to the actual condition. If the model composes several reservoir zones, meshes within different reservoir zone may be evaluated individually.

3. HOW TO MAP THE BOUNDARY OF RESISTIVE RESERVOIR

For frequency-domain CSEM, the most significant figure to study resistive reservoir is MVO (Magnitude Versus Offset) curve which can display the variation of anomaly magnitude or phase as a function of offset. Whatever the geo-electric structure is, magnitude of electrical field will attenuate when source and receiver offset (Tx-Rx spacing) increases. But hydrocarbon model will exhibit significant different character from non-hydrocarbon model on MVO curve. Figure 1 displays two different MVO curves. The boundary of reservoir is at 5 km. The black curve corresponds to hydrocarbon model while the red one to the non-hydrocarbon model. The two curves exhibit significant difference. The generalization curve (Figure 2) shows apparent high magnitude indicating the resistive reservoir. Either MVO curve or generalized curve can only indicates the existence of hydrocarbon reservoir and can not map the reservoir boundary, but on the MVO second derivative curve (MVOSD) (Figure 3) it can be found that reservoir boundary exhibit apparent and approximately stable character, i.e., the first zero point is located within the resistive body while the second extremum appears abound 0.5 km away from hydrocarbon boundary. To get the common rule, we modeled numerous models. Figure 4 displays a model composing a $50 \times 50\,\text{km}^2$ square reservoir. The source is positioned at the center of one boundary. When source-receiver offset crosses the boundary, the MVO curves (Figure 5) for models with different lateral extension (from 3 km to 11 km) show different characters. When the offset is small, the dominated field is the reflected wave. At this small offset, the MVO curves for half-space model, 2D model and 3D model overlay each other. With the increasing of offset, the reflected field from underlying layer gradually begins to influence the MVO curve thereby making the curve going upwards. The wider the side length of the reservoir, the higher the curve goes upwards. But with the continual increasing of the offset, the influence from air wave increases accordingly. When air wave takes the dominative position, MVO curves overlay again. On the figure, it is indicated that the curves for models with different lateral extensions and 1D model joint at a point which corresponds to the boundary of resistive reservoir. Figure 6 is the MVOSD curve. The MVOSD curves for reservoirs with different scale show approximately similar character while neither half-space model nor 2D model exhibits this kind of character. Especially, the second extreme point always appears abound 0.5 km away from the boundary of resistive reservoir. The third extreme point indicates another boundary of reservoir. Based on these characters, reservoir boundary can be delimited. Much modeling calculation indicates that the second extreme point of MVOSD curve will not move when source frequency, target depth and thickness, seabed depth, even the resistivity of target vary if the resistive reservoir is detectable. Therefore, the extremeum is an indicator of the boundary of resistive reservoir.

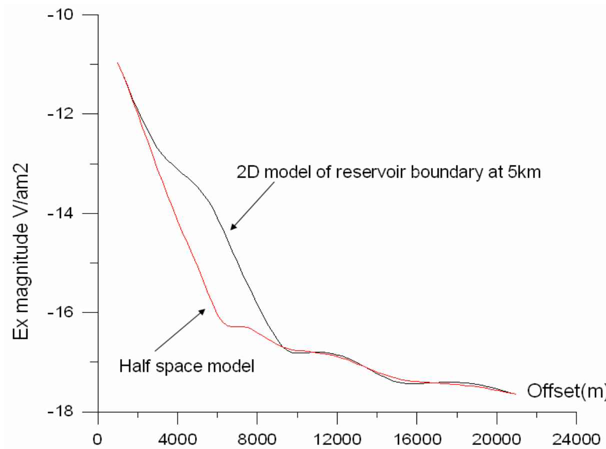


Figure 1: MVO curves.

4. MAPPING THE BOUNDARY OF SEVERAL RESERVOIR LAYERS ALONG VERTICAL DIRECTION

How to recognize anomaly bodies by MVO and MVOSD curves when there is more than one resistive body at different depths? Under the situation, MVO curve is powerless. It can only indicate the existence of resistive bodies but can't recognize these anomaly bodies individually. As

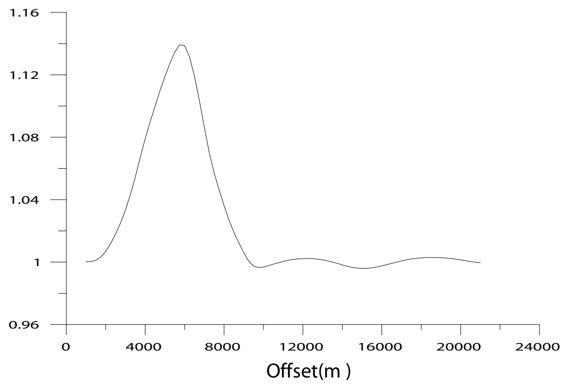


Figure 2: Generalized MVO curve.

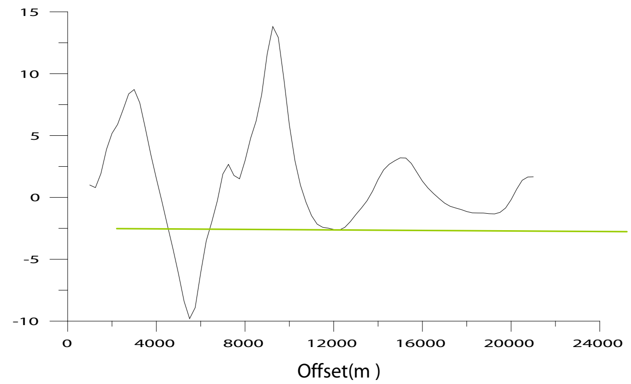


Figure 3: MVOSD curve for boundary at 5 km model.

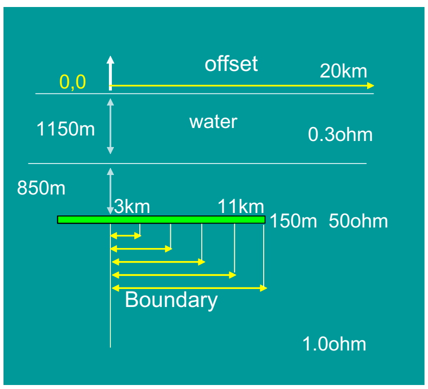


Figure 4: Sketch for models with different size.

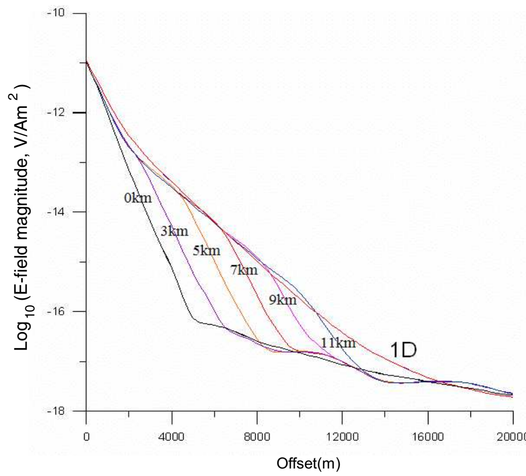


Figure 5: MVO curves for models with different size.

for MVOSD curve, it can not recognize the lower anomaly bodies if the upper resistive body is larger than the lower one(s). But if the upper body is smaller than the lower one and the boundary of the lower body is located out of the boundary of the upper body, the second derivative curve will appear two apparent extreme points which correspond to the boundaries of the two resistive bodies respectively. If the two resistive bodies are located near, the two extreme points will joint at one point which indicates the boundary of the larger anomaly body. Figure 7 displays the MVOSD curves for three models. It clearly displays the phenomenon. (Explain Figure 7 here).

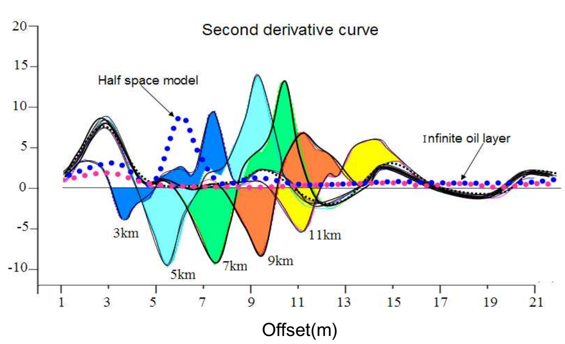


Figure 6: Second derivative curve of MVO for models with different size.

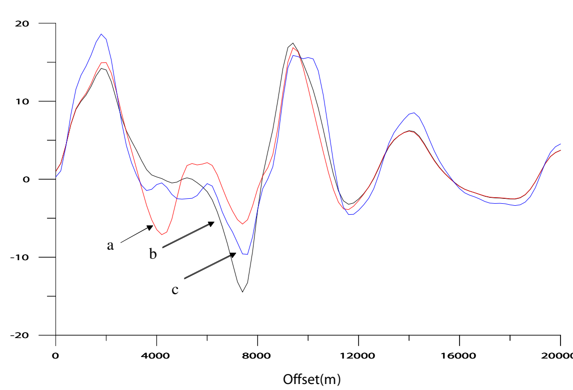


Figure 7: MVOSD curves for a model composing three resistive bodies Vertically. a. The upper resistive body is 4 km wide, the lower 7 km wide, b. The upper resistive body is 7 km wide, the lower 4 km wide, c. The upper resistive body is 3 km wide, the middle 5 km wide.

5. HOW TO DELIMIT RESISTIVE BODY WHEN T-R SET IS POSITIONED FAR FROM IT

If T-R set does not cross resistive body or the source is positioned far away from the resistive body and source-receiver spacing is large, the signal will be dominated by air wave after a long transmission. Therefore, MVO or MVOSD curve does not exhibit any anomaly. How to recognize the resistive body under this situation? Someone may think that the frequency-domain CSEM is powerless at this situation. But our modeling results still can recognize the resistive body under this situation. Because the excited field is mainly plane wave, Cagniard resistivity can be calculated just like ground CSAMT, and the resistive body can be delimited based on apparent resistivity profile. If we continue to do inversion on the profile, spatial distribution of the resistive body can be mapped. Figure 8 shows the apparent resistivity profile. On the profile a high resistivity zone appears at the center. The generalization profile (Figure 8(b)) indicates boundary of the resistive body in more detail.

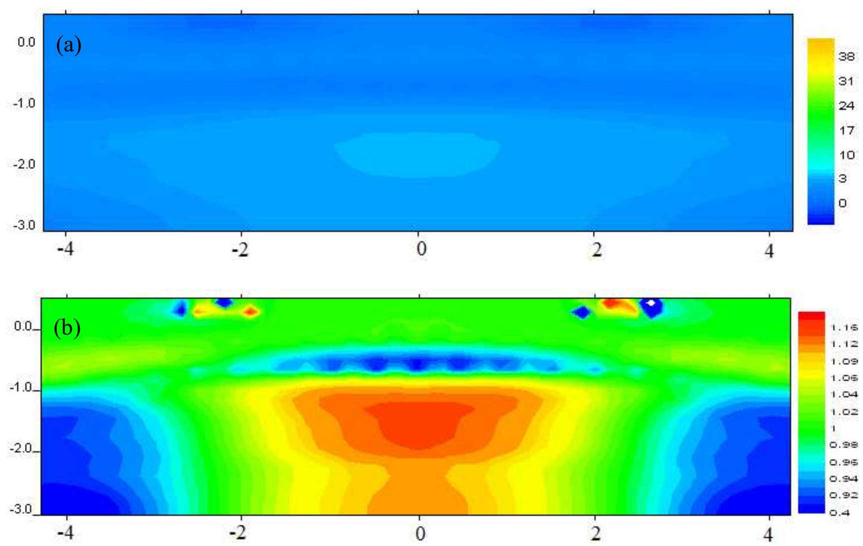


Figure 8: (a) Apparent resistivity profile (b) Generalized resistivity profile.

6. CONCLUSION

Frequency-domain MCSEM is one of the major EM methods applied on marine oil and gas exploration. It has been proved successfully by many oil companies and service companies. But the data processing still stays in qualitative stage. Theoretically, MVOSD curve can accurately delimit reservoir, but practically the precision is limited by the precision of observed data. Therefore, if observed data have high quality and precision, it is an effective means to delimit resistive body. When MVO curve is powerless due to air wave, theory of plane wave can help to calculate apparent resistivity. At all events, rapid 3D inversion is an important step of data processing and interpretation for successful MCSEM survey.

ACKNOWLEDGMENT

We greatly appreciate Dr. Shenghui Li for his helpful advice and technical discussion and Prof. Hua-Wei Zhou as advisor & his help.

REFERENCES

1. Cox, C. S., J. H. Filloux, and J. Larsen, "Electromagnetic studies of ocean currents and electrical conductivity below the ocean floor," *The Sea*, Vol. 4, Part I, 637–693, ed. Maxwell (ed.), Wiley, 1971.
2. Constable, S. and C. S. Cox, "Marine controlled source electromagnetic sounding 2, The PEGASUS experiment," *Journal of Geophysical Research*, Vol. 101, 5519–5530, 1996.
3. Constable, S., A. Orange., G. M. Hoversten, and H. F. Morrison, "Marine magnetotellurics for petroleum exploration 1. A seafloor instrument system," *Geophysics*, Vol. 63, 816–825, 1998.

4. Hoversten, G. H., H. F. Morrison, and S. Constable, "Marine magnetotellurics for petroleum exploration 2. Numerical analysis of subsalt resolution," *Geophysics*, Vol. 63, 826–840, 1998.
5. Hoversten, G. H., S. Constable, and H. F. Morrison, "Marine magnetotellurics for base salt mapping: Gulf of Mexico field-test at the Gemini structure," *Geophysics*, Vol. 65, 1476–1488, 1999.
6. Heinson, G., S. Constable, and A. White, "Episodic melt transport at a mid-ocean ridge inferred from magnetotelluric sounding," *Geophys. Res. Lett.*, Vol. 18, 1917–1920, 2000.
7. Tompkins, M. J., A. Greer, N. Barker., R. Weaver, and L. M. MacGregor, "Effects of vertical anisotropy on marine active source electromagnetic data and inversions," *66th EAGE Annual Meeting*, Paris, France, 7–10 June, 2004.
8. MacGregor, L., M. Sinha, and S. Constable, "Electrical resistivity structure of the Valu Fa Ridge, Lau Basin, from marine controlled-source electromagnetic sounding," *Geophys. J. Int.*, Vol. 146, 2001.
9. Eidesmo, T., S. Ellingsrud, etc., "Sea Bed Logging (SBL), a new methods for remote and direct identification of hydrocarbon filled layers in deepwater areas," *First Break*, Vol. 20, 144–152, 2002.
10. Mittet, R., L. Løseth, and S. Ellingsrud, "Inversion of SBL data acquired in shallow waters Paris," *66th EAGE Conference & Exhibition*, France, June 2004.
11. Ueda, T. and M. S. Zhdanov, "Fast numerical modeling of marine controlled-source electromagnetic data using quasi-liner," *75th SEG*, Houston, 2005.
12. Zhdanov, M. S. and S. K. Lee, "Integral equation method for 3D modeling of EM field in complex structures with inhomogeneous background conductivity in marine CSEM applications," *75th SEG*, Houston, 2005.
13. Zhdanov, M. S. and L. Wan, "Rapid seabed imaging by frequentcy domain EM migrateon," *75th SEG*, Houston, 2005.
14. Zhdanov, M. S. and K. Yoshioka, "3D iterative inverseon of the marine CSEM data," *75th SEG*, Houston, 2005.
15. Um, E. and D. Alumbaugh, "On the physics of the marine-time-domain controlled source electromagnetic method for detecting hydrocarbon reservoirs," *75th SEG Annual Conference*, 6–11 November, Houston, TX, 2005.

The Electromagnetic Responses of Under Seabed Layer and Inversion Method Study

Z.-F. Zeng[1,2], L.-H. Wang[1], S.-T. Zheng[1], F.-S. Liu[2], and Y.-M. Wang[3]

[1]College of Geoexploration Sci. & Tech., Jilin University, 130026, China
[2]Applied Mathematic Research Center (AMRC), Delaware State University, 19901, USA
[3]Geophysical exploration Inc., Daqing Petroleum Administrative Bureau, 163357, China

Abstract— Seabed hydrocarbon detection by electromagnetic method is widely used for its effectiveness and efficiency. In this paper, we analyze the seabed survey and geometry from the reflection coefficient, develop a new algorithm and complete the software to calculate the EM response of VMD and HED source which is widely used in the seabed detection. At the end of the paper, we inverse the 1D response by least square method based on Monte Carlo method for initial selection. The simulation result shows that our algorithm is stable and efficient.

1. INTRODUCTION

In the hydrocarbon exploration, the seismic method is very important and can provide most information of the stratum and structure of the reservoir. But it is difficult to predict the existence of hydrocarbon with the elastic properties. For the relationship between the hydrocarbon and resistivity, the low frequency electromagnetic sounding has the potential of directly indicating the oil and gas [1, 12].

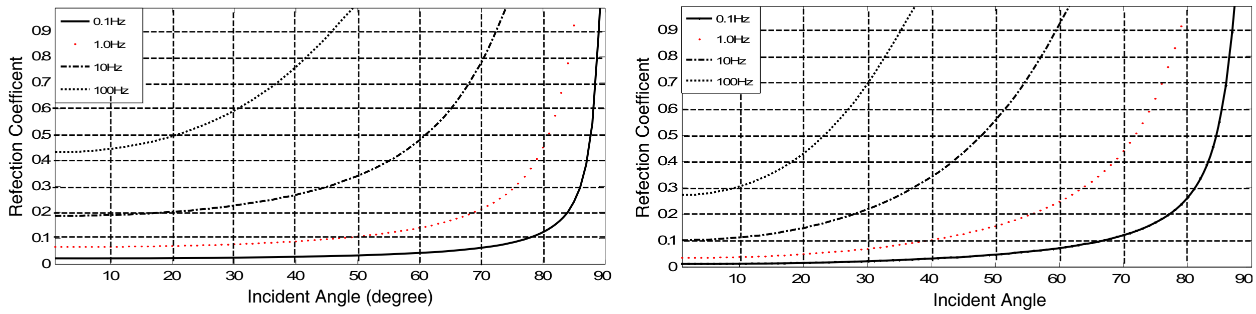
The frequency and time domain electromagnetic sounding such as Long-Offset Transient EM (LOTEM), Controlled-Source Audio-Frequency Magneto-Telluric (CSAMT), are widely used in the hydrocarbon detection on land and seabed. In the recent years, the seabed logging provides the theory and cases for hydrocarbon detection [6–9] using the refraction EM wave to detect high resistivity stratum at seabed. The seabed logging is widely accepted and it has the great potential in the expensive seabed hydrocarbon detection.

In this paper we analyze the reflection coefficient of high resistivity thin layer and discuss the potential to detect the layer by different EM modes which can be used for analyzing the measurement configuration. We calculate the responses of the TM and TE wave modes and analyze the difference. We develop a 1D inversion algorithm for multiple-layers by multiple frequencies using least square method based on the Monte Carlo method to avoid the inefficient and instability of the 2D and 3D inversion. The result shows that our algorithm is efficient and stable.

2. THE REFLECTION COEFFICIENT OF THIN RESISTIVITY LAYER

In the hydrocarbon detection, the layer bearing oil or gas always show high resistivity and thin in thickness. The low frequency EM with long wavelength has low resolution. In order to study the possibility of using EM wave to detection the layers and to analyze the survey configuration, we calculate the reflection coefficient of three layer model which consists of the upper layer and the lower half space (layer 1 and 3) separated by a thin resistivity layer (layer 2). According to [3], the three layer EM plane-wave reflection coefficient is $\tilde{R}_{12} = R_{12} + \dfrac{T_{12}R_{23}T_{21}\exp(2ik_{2z}l_2)}{1 - R_{12}R_{23}\exp(2ik_{2z}l_2)}$, where R_{ij} is the reflection coefficient of layer i to layer j, T_{ij} is the transmission coefficient of layer i to layer j, and $R_{ij} = 1 - T_{ij}$. For the TE and TM wave mode, $R_{ij}^{TE} = \dfrac{\mu_j k_{iz} - \mu_i k_{jz}}{\mu_j k_{iz} + \mu_i k_{jz}}$ and $R_{ij}^{TM} = \dfrac{\varepsilon_j k_{iz} - \varepsilon_i k_{jz}}{\varepsilon_j k_{iz} + \varepsilon_i k_{jz}}$. Due to the low frequency and low resistivity of the layer, the wave number can be written as $k_{iz} = \sqrt{i\omega\sigma_i\mu_i}\cos\theta_i = \sqrt{i\omega\sigma_i\mu_i}\left(1 - \dfrac{\sigma_1}{\sigma_2}\sin^2\theta_1\right)^{1/2}$, where σ, μ, ε are conductivity, susceptibility and permittivity respectively.

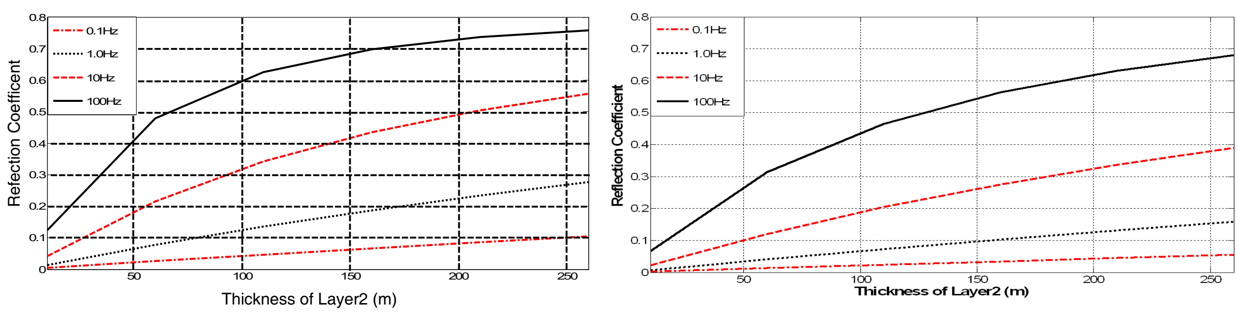
From (1), we can calculate the refection coefficient of the three-layer model. The model is analog to seabed logging [7] . The first and third layer have the resistivity of $1\,\Omega\cdot\text{m}$, the dielectric constant of 20, and second layer has the resistivity of $50\,\Omega\cdot\text{m}$, and the dielectric constant of 6. Figure 1 shows the reflection coefficient vs. the incident angle. We can see that the reflection coefficient increases



TE mode TM mode

Figure 1: The reflection coefficient vs. incident angle for different frequency.

very fast as the incident angle increases up to a value and also increases greatly as the frequency increases. The $\tilde{R}_{12}$ up to 1 by TE mode is faster than the one by TM mode. This means that when the offset increases to a certain level, the received EM energy decreases greatly which will be in favor to the refraction energy receiving. Figure 2 shows that the reflection response increases as the thickness of layer 2 increases. The two modes show that they are different in amplitude but same in shape.



(TE mode) (TM mode)

Figure 2: The reflection coefficient vs. thickness of layer 2 for different frequency.

From these two figures, we can conclude that the thin high resistivity layer 2 can have large response in the low frequency EM detection and the large offset can be in favor to the refraction measurement which can be calculated from Figure 1, and the frequency must be low enough to detect the deep reservoir. But for the complex resistivity structure of the reservoirs, it is difficult to apply the Low frequency EM wave to detect by refraction wave.

3. THE RESPONSE OF TE AND TM MODE FROM 1D MODEL

For the seabed hydrocarbon detection, the model can be simplified as layer model with the EM source embedded in the first layer [3, 4]. In our simulation, the model is four-layer model, and the source embeds in the bottom of the first layer [7]. We use the layer parameters as in [7]. The layer 1, layer 2 and layer 3 are 1000 m, 1000 m and 100 m in thickness and have resistivity of 0.1, 1, and 50 Ω·m respectively, and layer 4 has a resistivity of 1 Ω·m. In the calculation, the diameter of VMD and the length of HVD is 50 m. The current is 50 A.

Figure 3 is the amplitude of magnetic field of VMD source vs. frequency in different offsets. We can see that the amplitude decreases as the frequency increases, and when the frequency increases to a value, the amplitude changes very small, which means that when the reflection coefficient increases to 1, the amplitude changes very little as the frequency increases.

Comparing Figure 3 with Figure 4, we can see that the amplitude of Electric filed has the complex shape and change more in the higher frequency range. Figure 5 shows the responses as seawater depth changed of HED. We can see when the depth of seawater increases to 1000 m, the response does not change much comparing at 1500 m. We can conclude that, when the depth of seawater increases to 1000 m, it does not affect the response of seabed EM detection under this condition.

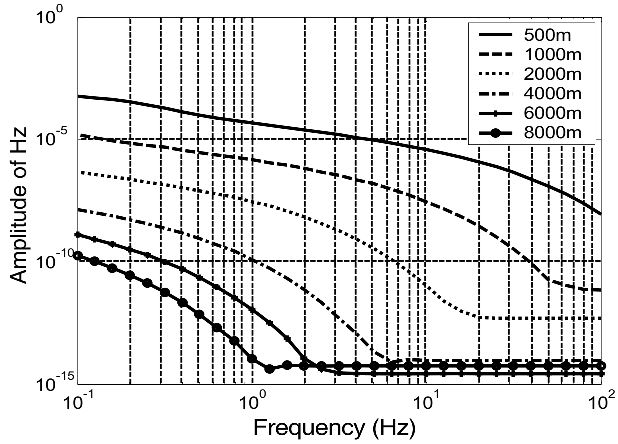


Figure 3: The responses Hz of VMD.

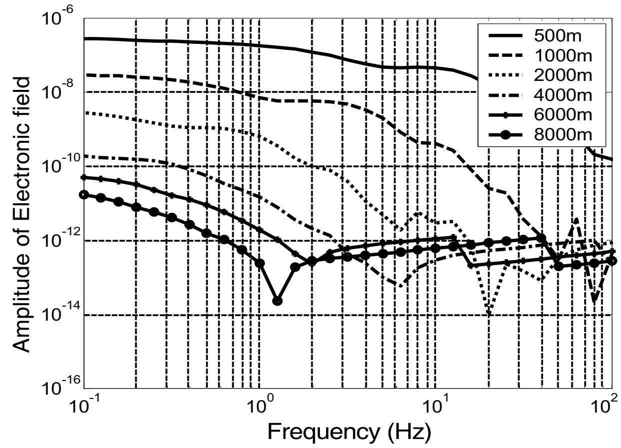


Figure 4: The responses Ez of HED.

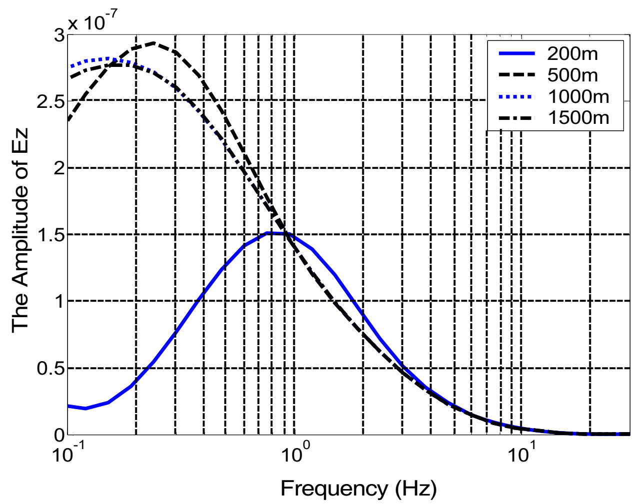


Figure 5: The responses of electrical field as sea water depth changed.

From the simulation, we can calculate the responses of the magnetic field and the electrical field faster and accurately, which prodi the basis of interpretation and inversion.

4. THE REAL-TIME INVERSION METHOD AND SIMULATION RESULTS

The interpretation of low frequency electromagnetic detection is needed to find the geometry and properties of the underground structure and layers. The inversion is a very important for above purpose. But for the multi-frequency detection, 2D and 3D inversion methods usually cost too much computation time or unstable. So 1D inversion is also very important in detection. But for multilayer seabed hydrocarbon detection, there are many parameters to be inverted such as layer thickness and properties. In this paper we use the least square method based on Monte Carlo method to reduce the instability and to increase the convergence rate.

Let $f(X)$ be the objective function, D be the measure data, $G(X)$ be the response of the model, and $R(X)$ be the fit residual vector, and X be the vector. Then the inverse problem can be induced as the optimal problem:

$$\min f(X) = (D - G(X))^T (D - G(X)) = R(X)^T R(X). \tag{1}$$

From (1), we use the Marqardt damping least square method [5, 10] to get ΔP, and get the $k+1$th iterative model vector $X_{k+1} = X_k + \Delta P_k$. If $J(X) = R'(X)$, We have

$$(J(X)^T J(X) + \alpha I)\Delta P = -J(X)^T R(X), \tag{2}$$

where I is the unit vector, and α is the damping factor.

In our problem, there are 4 or more layers, the problem is nonlinear and there are more parameters of the responses. Equation (1) has more solutions and therefore the problem is ill-posed. We apply the Monte Carlo method to find an optimal [11]. The solution from the Monte Carlo method is used as the initial value for the Marqardt least square method.

According to the method mentioned above, we invert the layer 3 using the HED source length of 50 m. The transmitter and receiver is located on the seabed at same the height of $z = h = 1000\,\text{m}$. The model parameters are $\rho_1 = 0.1\,\Omega{\cdot}\text{m}$, $\rho_2 = \rho_4 = 1\,\Omega{\cdot}\,\text{m}$, $\rho_3 = 50\,\Omega{\cdot}\,\text{m}$, $h_1 = 1000\,\text{m}$, $h_2 = 1000\,\text{m}$, and $h_3 = 100\,\text{m}$.

We can see that the algorithm is accurate and also very fast. In the calculation, we obtain most of the parameters. From the result, we can see the 1D inversion method can be used in real-time detection.

Table 1: The inverse result of layer 3.

Parameters	Resistivity of layer 3 ($\Omega{\cdot}$m)	Depth (m)	Thickness (m)
Model value	50	1000	100
Initial value	80	800	50
Inversion result	50.74	1002.54	98.60
error	1.48%	0.31%	1.4%
Initial logarithm fit covariance 11.23The covariance after 15th iteration: 0.039			

5. CONCLUSIONS

We study low frequency EM detection in the seabed in the following three aspects: the reflection coefficient calculation, field responses simulation and inversion. From our study, we conclude that: (1) by the study of reflection coefficient, we can determine the survey geometry and the offset to detect the seabed layers; (2) we also complete the software in MatLab to calculate the response of different dipoles and different fields; and (3) the 1D inversion of the layer model shows that our method is stable and efficient which can be used in the real-time interpretation of survey result in the field.

ACKNOWLEDGMENT

The research supported by the DoD grant DAAD 19-03-1-0375.

REFERENCES

1. Andrieux, H., N. Rüter, and Vozoff, "A first attempt at monitoring underground gas storage by means of time-lapse multichannel transient electromagnetics," *Geophysical Prospecting*, Vol. 48, No. 3, 2000.
2. Chave, A. D., "Numerical integration of related Hankel transforms by quardrature and continued fraction expansion," *Geophysics*, Vol. 48, 1671–1686, 1983.
3. Chew, W.-C., *Wave and Fields in Inhomogeneous Media*, IEEE Press, New York, 1995.
4. Chew, W. C. and S. Y. Chen, "Response of a point source embedded in a layered medium," *IEEE Antennas Wireless Propagat. Lett.*, accepted for publication.
5. Dennis, J. E., D. M. Gay, and R. E. Welsch, "An adaptive nonlinear least-squares algorithm," *ACM Trans. on Math. Software*, Vol. 7, 348–368, 1981.
6. Eidesmo, T., S. Ellingsrud, L. M. MacGregor, et al., "Sea Bed Logging (SBL), a new method for remote and direct identification of hydrocarbon filled layers in deepwater areas," *First Break*, Vol. 20, No. 3, 144–150, 2002.
7. Ellingsrud, S., M. C. Sinha, S. Constable, L. M. MacGregor, T. Eidesmo, and S. Johansen, "Remote sensing of hydrocarbon layers by Sea Bed Logging (SBL): results from a cruise offshore Angola," *The Leading Edge*, Vol. 21, 972–982, 2002.
8. Kong, F. N., H. Westerdahl, S. Ellingsrud, et al., "Seabed logging: A possible direct hydrocarbon indicator for deepsea prospects using EM energy," *Oil & Gas Journal*, May 13, 2002.

9. MacGregor, L. M. and M. C. Sinha, "Use of marine controlled source electromagnetic sounding for sub-basalt exploration," *Geophysical Prospecting*, Vol. 48, 1091–1106, 2000.
10. Marquardt, D. W., "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, Vol. 11, 431–441, 1963.
11. Tarantola, A., *Inverse Problem Theory*, Society for Industrial and Applied Mathematics, 2005.
12. Wright, D., A. Ziolkowski, and B. Hobbs, "Hydrocarbon detection and monitoring with a multicomponent transient electromagnetic (MTEM) survey," *The Leading Edge*, Vol. 21, No. 9, 2002.

Remote-control in Transient Electromagnetic System for Shallow Seafloor

Y. Wang, J. Lin, G. H. Zhou, F. D. Zhou, C. S. Liu, and H. Li
Key Lab of Geo-Exploration and Instrumentation, Jilin University
Ministry of Education, China

Abstract— Sea environment is complex, and the work on seafloor is difficult and dangerous. In order to improve the efficiency and reduce the unnecessary waste and investment, remote-control has been used in Transient Electromagnetic system design. Network and RS485 communication are adopted in receiver and transmitter, and finally real-time data observation, large data storage, power management and easy transmission of data are realized.

1. INTRODUCTION

The development and exploration on seafloor resources have been high concerned by all over the world. The system used on the seafloor has been manufactured and reported by many researchers [1, 2, 4, 5]. However, up to now, most the electromagnetic measurement systems operate as follows: there is a microcomputer which can measure automatically in receiver and transmitter respectively and the observed data will be stored in a device with large storage. Then the instruments placed in a sealed body are sunk into the seafloor under a heavy anchor, and its position can be determined by Global Positioning System or some specialized sonar equipment. The instruments will reject the anchor automatically and float up to the sea when the observation is finished. At last, the operators will salvage the system, transfer the data and interpret and so on. In this way, the availability of data can not been determined, and the system sinks and floats with power on will increase the dangerous degree. Although in China TEMS-3S system has been used on research of conductivity on sea-floor, and they can observe the transient signal real-time, the system did not put on the seafloor [3], so the remote-control is not realized actually. In order to realize the remote-control and then improve the productivity, remote-control system has been studied.

2. DESIGN REALIZATION

The whole transient electromagnetic system includes three main parts: the upper computer, the towed body (the receiver and the transmitter) and the configuration. Each of them are connected by communication cable and towed cable. The schematic sketch is shown in Figure 1. The receiver and transmitter coils are located in a bracket, and then floated at a fixed height in the sea by a buoy. The synchronization is wiring manner because the short separation of receiver and transmitter, and the storage battery is adopted for power supply. For the upper computer, the parameters set, control and display are all finished here. In order to realize the real-time observation of received signal, network transmission is chosen. Network transmission has high speed and efficiency, and its transmission distance is long suitable for shallow seafloor prospecting. In the receiver, embedded PC104 (HXL/P300) — Industrial Personal Computer with network interface is taken as the main control unit. For lower software development cost and period, Windows 98 operating system is installed on a notebook PC hard disk which storage is 20 G. Such large storage also has another function that it can store more received data. Because PC104 uses Windows 98 operating system, we hope that when the measured work is over, this system can be powered off normally not the supplied power be cut off suddenly. If this function can be realized, PC104 can be shut off any time we want to reduce power cost. Several methods have been tried, such as ActiveX control

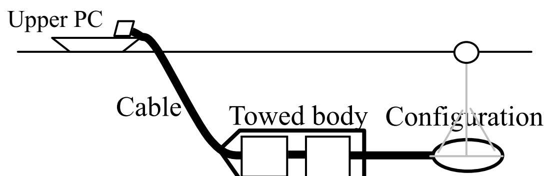


Figure 1: The schematic sketch of whole transient electromagnetic system.

tool in Visual Basic language, but all those methods can not realize those functions completely. Finally, a commercial software (PcAnywhere) has been found, this software not only can display the interface of another remote computer but also can do any operations in this interface to control this computer including power off and reset computer. In addition, if you want to copy any data from this computer to your computer, this software also can help you.

Figure 2: Transmitter control software.

Figure 3: PcAnywhere remote-control interface.

For the transmitter, we also want to know the work state, such as the transmitter current, the ramp time, and the current curve. Single Chip Microprocessor (SCM) AT90S8515L is the main control unit of transmitter, so full duplex RS485 mode is adopted for easy program on SCM. The electrical level conversion between 485 and TTL uses the chip SN75176. Although PcAnywhere software can power off and reset the PC104, the function — powering on can not be realized. So this function will be realized by transmitter. All those functions are realized by C++ Builder language. In order to suit the notebook PC without serial port, we use a transfer interface between RS485 and USB interface.

3. RESULTS

The control software of transmitter is shown in Figure 2. This software has the following functions that RS485 transmission parameters set, transmitting current parameters set, the state of

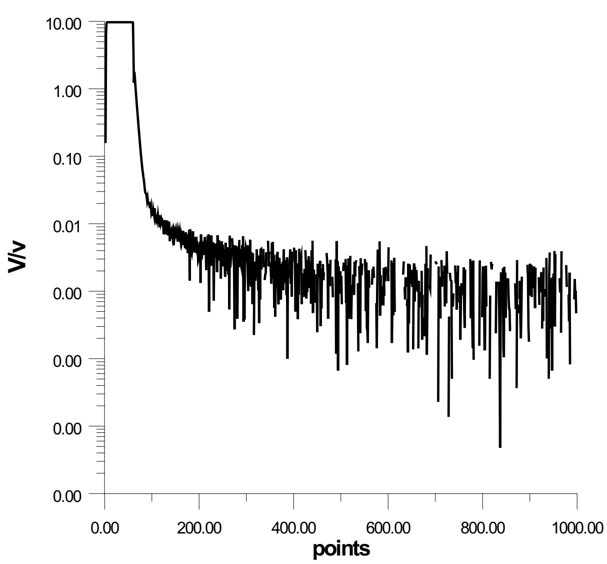


Figure 4: Transient signal remotely observed.

transmitter display, transmitting synchronization and receiver startup power control, transmitting power control and current display and so on.

Figure 3 is an example of PcAnywhere software to control the power of PC104. From this figure you can see the lower computer interface can be displayed on the upper computer. In this condition, when the sampling program displays the acquisition is over, and then we can display the transient signal on the lower computer or on the upper computer after data transmission operation. Figure 4 is a transient signal received on a lake in Changchun city.

4. CONCLUSIONS

Through network and RS485 modes, the transmitter and receiver are controlled by upper computer. Real-time data observation is realized and this has made the experiments on shallow seafloor successfully. At the same time, large storage ability, power management, and easily data transmission also have been realized.

ACKNOWLEDGMENT

This work is funded by hi-tech research and development program of China. Teacher Liu Hanjiang is gratefully acknowledged for his work on sealed cabin manufacture. The experiment has been done on the aid of all members in TEM lab. The author would like to thank all of them.

REFERENCES

1. Edwards, R. N., "Two-dimensional modeling of a towed in-line electric dipole-dipole sea-floor electromagnetic system: The optimum time delay or frequency for target resolution," *Geophysics*, Vol. 53, No. 6, 846–853, 1988.
2. Yang, J. W. and R. N. Edwards, "Controlled source time-domain electromagnetic methods for sea-floor conductivity mapping," *The Chinese Journal of Nonferrous Metals*, Vol. 8, No. 4, 705–713, 1998.
3. Wang, Y. X., J. L. Wang, and J. Y. Wang, "The research of conductivity on sea-floor by transient electromagnetic system," *Chinese Journal of Geophysics*, Vol. 41, No. 6, 841–847, 1998.
4. Eidesmo, T., L. M. MacGregor, and M. C. Sinha, "Sea Bed Logging (SBL), a new method for remote and direct identification of hydrocarbon filled layers in deepwater areas," *News Feature*, Vol. 20, 144–152, 2002.
5. Constable, S. C., "Marine electromagnetic methods — A new tool for offshore exploration," *The Leading Edge*, 438–444, 2006.
6. Wang, Y., J. Lin, and B. Liu, "Communication design of towed seafloor TEM system by network card PC/104," *Journal of Test and Measurement Technology*, Vol. 3, No. 19, 351–354, 2005.

Modeling of Seabed Logging Data for a Sand-shale Reservoir

Zhong Wang and Leiv-J. Gelius
Department of Geosciences, University of Oslo, Norway

Abstract— a new rock-physics modeling tool was developed to describe the electric properties of a sand-shale reservoir. Four types of clay distribution models were implemented to describe possible sand-shale reservoirs including anisotropy. The corresponding algorithms gave as output an estimate of the efficient-medium conductivity of the reservoir. The tool was interfaced to a 1.5-D EM-modeling program used to simulate Seabed logging data. The potential of this integrated modeling approach was demonstrated by calculating the variation in the EM response associated with a petroleum reservoir, due to different clay distributions as well as different brine salinities.

1. INTRODUCTION

Seabed logging (SBL) is a new method employing EM energy to detect and characterize hydrocarbon bearing reservoirs in marine environments [8]. However, compared with other more mature EM methods, the representation of the target (e.g., the petroleum reservoir) is still rather crude in standard SBL-modeling programs. The reservoir zone is often assigned a fixed conductivity value without any link to a rock physics description. However, the actual petroleum reservoir is a complex mixture of fluid, sand, clay and gas. Different formations and structures will give rise to different EM properties and at the end to different measurements. An accurate and efficient EM description of the reservoir zone is therefore necessary to further understand and develop the SBL technique.

2. ROCK-PHYSICS MODELING

This study is limited to the sand-shale petroleum reservoir. In the sand-shale formation, clay minerals have a substantial effect to the overall equivalent conductivity of the rock. We have implemented four types of mixing-models to describe different clay distributions. In most mixing models one will calculate the effective conductivity of the water and hydrocarbons before calculating the whole rock conductivity. This is the hydrocarbons first method which we also employ here. Assuming zero conductivity of the oil particles, it follows from an Archie type of formula [1]:

$$\sigma_f = \sigma_w S_w^n \tag{1}$$

where σ_f is the effective fluid conductivity, σ_w is the water conductivity and S_w is the water saturation. The factor n is the saturation exponent which is typically between 1.7 and 2. Note that the salinity of the formation water strongly determines its conductivity.

2.1. Mixing-model 1: Structural Clays

In this model the clay grains act as framework grains without altering the reservoir properties. Hence, none of the pore space is occupied by clay. For structural clays we can employ the Bussian/Hanai-Bruggeman equation [2] with low-frequency limit.

$$\sigma^* = \sigma_f \phi^m \left(\frac{1-\sigma_s/\sigma_f}{1-\sigma_s/\sigma^*}\right)^m \tag{2}$$

where σ_s is the effective mean volume conductivity of the grains:

$$\sigma_s = p\sigma_c + (1-p)\,\sigma_{sa} \tag{3}$$

and p is the volume fraction of clay in the solid portion, σ_c is the conductivity of the clay and σ_{sa} is the conductivity of the sand (quartz) grains. In Eq. (2) ϕ is the porosity and m is the so called cementation factor assumed to be in the range from 1.7 to 2.3 for a consolidated material. For a clay-rich sand where the water/fluid is such that conduction is dominated by the grains ($\sigma_s \gg \sigma_f$) a simplified version of Eq. (2) can be derived [5].

2.2. Mixing-model 2: Coated Clays

In the *coated-clay* model the clay grains actually coat the sand grains. When clay coat the sand grains, the irreducible water saturation of the formation increases, dramatically lowering the resistivity values. We assume that the grain consist of a non-conducting silicate core coated with a conductive clay. Lima and Sharma [5] proposed to employ the Hashin-Shtrikman upper bound for a coated sphere to obtain the effective electrical conductivity for the grains:

$$\sigma_s = \frac{\sigma_c \left[2p\sigma_c + (3-2p)\,\sigma_{sa}\right]}{(3-p)\,\sigma_c + p\sigma_{sa}} \tag{4}$$

where p is the volume fraction of the coating clay. For a non-conducting core ($\sigma_{sa} = 0$) as would be the case for most sandstone minerals, we have

$$\sigma_s = \frac{2p\sigma_c}{3-p} \tag{5}$$

This last equation can also be generalized to take into account different grain shapes [4]. As for Mixing-model 1 the effective conductivity of the whole rock is given by Eq. (2).

2.3. Mixing-model 3: Dispersed Clays

In this mixing-model the clay grains fill the pore space between sand grains. We will assume that the composite medium is built from an initial fixed volume of fluid by adding to it, in steps, infinitesimal amounts of insulating sand grains and clay aggregates. This procedure was proposed by Lima and Sharma [5] who adapted the incremental method first introduced by Feng and Sen [3]. The incremental method solution can be written as (assuming non-conducting core $\sigma_{sa} = 0$):

$$\sigma^* = \sigma_f \phi^{3/2} \left[\frac{1 + (1-3p)\,\sigma_c/2\sigma^*}{1 + (1-3p)\,\sigma_c/2\sigma_f}\right]^{3p/(1-3p)} \tag{6}$$

where p again is the volume fraction of clay in the solid portion. For the special case of $\sigma_f \gg \sigma_c$ Eq. (6) can be simplified to a Waxman-Smits type of equation [9], which also can be modified to handle the case of non-spherical grains.

2.4. Mixing-model 4: Laminar Shales

The last type is the *laminar-shale* model, which consist of sequences of shale layers between sand layers. The effect of thinly bedded sand-shale sequences on a macroscopic scale is electrical anisotropy. Hence, the effective conductivity parallel with the bedding will be different from the effective conductivity normal to the layers. The ratio between vertical and horizontal resistivity can be 2–10 for hydrocarbon reservoirs (up to 100 reported). For this mixing-model we first compute the effective conductivity of the sand and fluid system σ^*_{sand} employing Bussian/Hanai-Bruggeman theory. The effective conductivities (vertical and horizontal) of the laminar shale is now given by the Wiener bounds [10]:

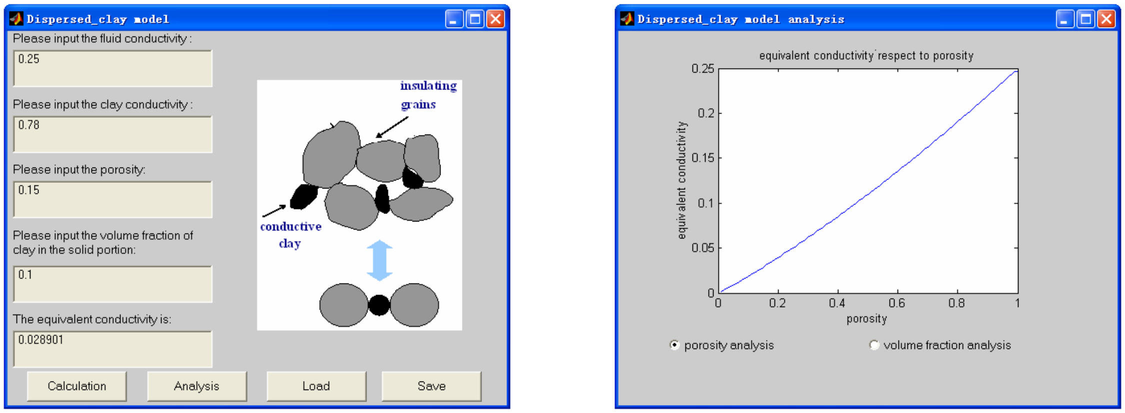
$$\sigma^*_h = v_{\text{sand}}\sigma^*_{\text{sand}} + v_{\text{shale}}\sigma_{\text{shale}}\,, \quad \frac{1}{\sigma^*_v} = \frac{v_{\text{sand}}}{\sigma^*_{\text{sand}}} + \frac{v_{\text{shale}}}{\sigma_{\text{shale}}}\,, \quad v_{\text{sand}} + v_{\text{shale}} = 1 \tag{7}$$

where v_{sand} and v_{shale} are the volume fractions of sand and shale, respectively.

We now illustrate the basic functions of the modeling tool through an example of dispersed-clay type. Fig. 1(a) shows the calculation interface of this model option. The right part illustrates the basic physical principle of the modeling. The corresponding parameters can be input or loaded (by pressing the ‘load’ button) from an input file. The equivalent conductivity of the mixture can be calculated and displayed in a separate window. The calculation result together with the input parameters can be saved as an output file. Pressing the ‘Analysis’ button will open a new window (shown in Fig. 1(b)) where the sensitivity of the effective conductivity with respect to some key input parameters can be analyzed. The example in Fig. 1(b) shows the effect of varying the porosity.

3. EM-MODELING PROGRAM

We have developed a 1.5-D EM-modeling tool that calculates the EM response of four types of dipole sources in a layered medium: horizontal electrical dipole (HED), vertical electrical dipole (VED),



(a) (b)

Figure 1: Example of rock-physics modeling. (a) Calculation interface of dispersed clay model and (b) analyzing window showing the actual computations.

horizontal magnetic dipole (HMD) and vertical magnetic dipole (VMD). In SBL, the transmitter antenna is normally towed by a vessel at a depth just above the seabed. The receiver antenna (or eventually an array) is placed on the seafloor (in-line or cross-line polarized). The 1.5-D simulation software includes a graphical user interface and is based on J. A. Kong's formulations [6, 7]. In order to be as realistic as possible, the influence of the air layer in the limit of shallow water depths as well as the possibility of anisotropy caused by a finely layered reservoir (e.g., mixing-model 4) are also taken into account.

4. EXAMPLES OF THE INTEGRATED SYSTEM

In the following we limit our discussion to an isotropic case, hence mixing-models 1, 2 and 3. First, we employed our integrated modeling tool to study the variation in the EM response caused by the different clay distributions. The following parameters were used in the simulations: water conductivity: $\sigma_w = 15.3846$ S/m (with salinity of 100 kppm); effective porosity: $\phi = 0.15$; water saturation $S_w = 0.15$; conductivity of clay: $\sigma_c = 1.0$ S/m; saturation exponent: $n = 2$; volume fraction of clay in the solid portion: $p = 0.1$. The effective conductivity of the reservoir rock for the three different mixing-models were found to be: *structural clay* $\sigma^*_{\text{str.}} = 0.1219$ S/m; *coated clay* $\sigma^*_{coat..} = 0.0903$ S/m; *dispersed clay* $\sigma_{disp..} = 0.0397$ S/m.

Next, we established a 1-D layered-media model for the actual SBL simulations as showed in Table 1 below.

Table 1: The stratified media model for SBL

Layer	**Thickness(m)**	**Conductivity(S/m)**
Sea water	500	3.2
Layer1 overburden	1000	1
Layer2 reservoir	100	σ^* (coated, structural or dispersed clay)
Layer 3 half space	∞	1

In the simulations we assumed a fixed HED-source placed 50 meter above the seabed with an operating frequency of 0.25 Hz and 100 receivers deployed on the seabed (in-line direction and with a receiver interval of 100 m). Fig. 2 shows the magnitude of the E_ρ-*field* (i.e., the horizontal electric component along the in-line direction) normalized by the response from a homogenous subsurface for all three clay distributions. According to Fig. 2, all mixing-models are sensitive to the oil layer, but the dispersed clay model shows a much larger anomaly than the other two mixing model. The anomaly decreases as the offset increases (e.g., right half of the curves) due to strong airwaves. In the second experiment we considered the dispersed-clay model only and studied the effect of varying the brine salinity. Fig. 3 shows the normalized E_ρ-field of the HED source for three different salinity values. The water conductivities corresponding to these salinity values are summarized in Table 2. From Fig. 3, it is evident that a high salinity corresponds to a large effective conductivity of the reservoir rock and consequently a low EM detectability.

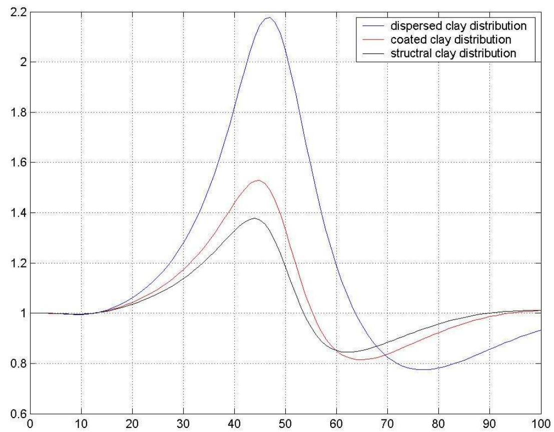


Figure 2: Normalized E_ρ-field for various clay models. Dispersed-clay (blue), coated-clay (red), structural-clay (black).

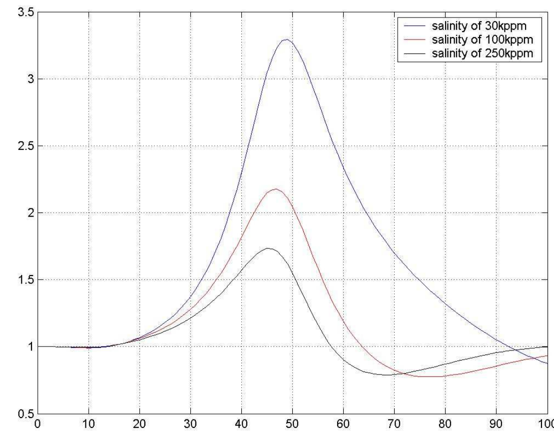


Figure 3: Normalized E_ρ-field for different salinities, (dispersed-clay model). 30 kppm (blue), 100 kppm (red) and 250 kppm (black).

Table 2: Variations in the effective conductivity of the reservoir rock and water conductivity due to different brine salinities.

Salinity(kppm) / Cond.(S/m)	30	100	250
σ_w	5.2632	15.3846	27.7778
σ^*	0.014979	0.039676	0.066

5. CONCLUDING REMARKS

From the examples above, it is evident that different clay distributions, keeping the volume fraction of clay constant, give rise to very different effective conductivities of the reservoir rock, and hence EM response. We have also seen that the salinity of brine in a sand-shale reservoir rock also affects the EM response considerably. In reality a hydrocarbon reservoir is more complex than the models discussed in this paper. However, we feel that the reservoir rock physics description introduced here will make EM-modeling more realistic offering useful information for further development of the SBL technique as well as for interpretation of SBL data.

ACKNOWLEDGMENT

We wish to thank the Norwegian Research Council, Hydro and Statoil for funding through a Petromaks project.

REFERENCES

1. Archie, G. E., "The electrical resistivity log as an aid in determining some reservoir characteristics," *Tran. AIME*, Vol. 146, 54–62, 1942.
2. Bussian, A. E., "Electrical conductance in a porous medium," *Geophysics*, Vol. 48, 1258–1268, 1983.
3. Feng, S. and P. N. Sen, "Geometrical model of conductive and dielectric properties of partially saturated rocks," *J. Appl. Phys.*, Vol. 58, 3236–3243, 1985.
4. Fricke, H., "A mathematical treatment of the electric conductivity and capacity of disperse system," *Phys. Rev.*, Vol. 24, 575–587, 1924.
5. Lima Olivar, A. L. and M. M. Sharma, "A grain conductivity approach to shaly sandstones," *Geophysics*, Vol. 55, 1347–356, 1990.

6. Kong, J. A., *Electromagnetic Wave Theory*, (ISBN 0-471-52214-7), 312–321, 1990.
7. Kong, J. A., "Electromagnetic fields due to dipole antennas over stratified anisotropic media," *Geophysics*, Vol. 37, 985–996, 1972.
8. Kong, F. N., H. Westerdahl, S. Ellingsrud, T. Eidesmo, and S. Johansen, "'Seabed logging': A possible direct hydrocarbon indicator for deepsea prospects using EM energy," *OIL & GAS Journal*, 2002.
9. Waxman, M. H. and L. J. M. Smits, "Electrical conductivities in oil-bearing shaly sands," *Soc. Petr. Eng. J.*, Vol. 8, 107–122, 1968.
10. Wiener, O., *Abh. Math. Phys. Kl. Sachs. Akad. Wiss.*, Lpz., Vol. 32, 509, 1912.

Multi-component Processing of Sea Bed Logging Data

L.-J. Gelius
Department of Geosciences, University of Oslo, P. O. Box 1047 Blindern, N-0316, Norway

Abstract— Marine remote sensing of hydrocarbons based on EM-soundings does not work very well in case of shallow waters due to strong airwave contributions. In this paper we investigate a multi-component signal processing approach which has the potential to attenuate such airwaves. Tests of the technique employing synthetic North Sea data gave significant improvements.

1. INTRODUCTION

In oil and gas exploration seismic is the far most important type of data used to image the subsurface. While the seismic waves have the ability to detect gas, they often fail to discriminate between water and oil as a pore fluid. On the contrary, electromagnetic waves are able to separate between the two fluids, based on their large differences in resistivity value. The reason for this is based on the fact that water supports free ions and easily transports electric current whereas oil acts as an insulator. This observation has motivated the development of a marine EM-based remote sensing method for hydrocarbons denoted Sea Bed Logging (SBL) [3]. This technique employs a mobile horizontal electric dipole (HED) source towed by a vessel and an array of seafloor electric field receivers. However, in the limit of shallow water (e.g., 200 m or less), the subsurface responses from potential high-resistivity zones like hydrocarbon reservoirs are masked by airwaves. The airwaves are energy which diffuses from the source to the air-sea interface, propagates as a lateral wave along this interface, and then diffuses downward to the receiver. In this paper we discuss how to minimize the influence of these waves employing multi-component processing of SBL-data.

2. MULTI-COMPONENT PROCESSING

As shown in Fig. 1, the generated electromagnetic soundings employing the SBL-method can in general be divided into three main contributions: direct EM field, guided modes (associated with high-resistivity zones like hydrocarbons) and airwaves. If the distance (offset) between the transmitting and receiving antenna is large enough (approx. 3 times the target depth) the direct field can be neglected. In case of deep water the guided modes will then dominate the large-offset measurements. However, when moving to shallow water depths this is no longer the case and the airwaves will mask the subsurface responses.

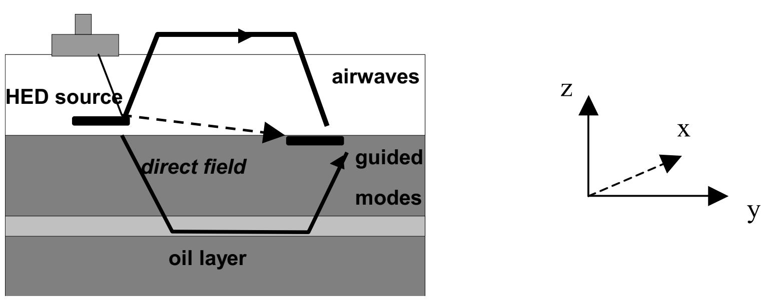


Figure 1: Main EM contributions recorded by the SBL-method (inline antenna).

These airwaves will diffuse in the downward direction and be almost normal incident to the seafloor, whereas the guided modes associated with a hydrocarbon reservoir will leak out and diffuse in the upward direction. Hence, these two contributions have the potential to be separated at the seafloor sensor if the measured EM field can be decomposed in upward and downward traveling modes. In seabed logging, both electric and magnetic field components are recorded at the seafloor. Amundsen et al. [1] have introduced a rigorous approach to EM field decomposition making use of both E and H field components. We investigate this idea further and introduce here a more straightforward approach to the same problem. In the following we assume that all EM modes exist along the vertical direction only. We introduce a Cartesian coordinate system as shown in Fig. 1, with the z-axis defining the vertical direction, the y-axis defining the inline horizontal direction and where the x-axis falls along the cross-line direction. Moreover, we assume that the EM fields can be approximated by plane-waves. In the case of an inline polarized HED-source, it follows from Maxwell's equations after Fourier transformation of the fields in space (x and y) as well as in time:

$$-k_y^2\hat{E}_z + \partial_z\hat{E}_y = -i\omega\mu_0\hat{H}_x \tag{1}$$

However, since only normal incident EM fields are assumed, this equation can be further simplified by setting $k_y = 0$

$$\partial_z\hat{E}_y = -i\omega\mu_0\hat{H}_x \tag{2}$$

Let the received signal at the electric sensor be decomposed into upward (U) and downward (D) travelling modes, i.e.,

$$\hat{E}_y = \hat{U} + \hat{D} \tag{3}$$

Since $\hat{U}$ and $\hat{D}$ represent plane-wave modes propagating along the vertical, they are given by the following expressions (simplest form):

$$\hat{U} = \exp\left[ik_z z\right], \quad \hat{D} = \exp\left[-ik_z z\right] \tag{4}$$

where the sign convention of the vertical wavenumber is chosen as follows

$$k_z = k = \omega\sqrt{\mu_0\varepsilon^*}, \quad \varepsilon^* = \varepsilon + i\sigma/\omega \tag{5}$$

By combining Eqs.(3)–(5) we obtain for an attenuating medium

$$\hat{H}_x = \frac{i}{\omega\mu_0}\frac{\mathrm{d}\hat{E}_y}{\mathrm{d}z} = \frac{i}{\omega\mu_0}\frac{\mathrm{d}\left(\hat{U}+\hat{D}\right)}{\mathrm{d}z} = \frac{k_z}{\omega\mu_0}\left(\hat{D}-\hat{U}\right) = \sqrt{\frac{\varepsilon^*}{\mu_0}}\left(\hat{D}-\hat{U}\right) \cong e^{i\pi/4}\sqrt{\frac{\sigma}{\omega\mu_0}}\left(\hat{D}-\hat{U}\right) \tag{6}$$

Finally, Eqs.(3) and (6) can be combined to give (after double inverse spatial Fourier transform)

$$U = \frac{1}{2}\left[E_y - e^{-i\pi/4}\sqrt{\frac{\mu_0\omega}{\sigma}}H_x\right] \tag{7}$$

where U represents the 'airwave corrected' version of the electric field E_y. The factor that is multiplying the magnetic field we recognize as the intrinsic plane-wave impedance in case of a well conducting medium. A similar result has also been derived by [1].

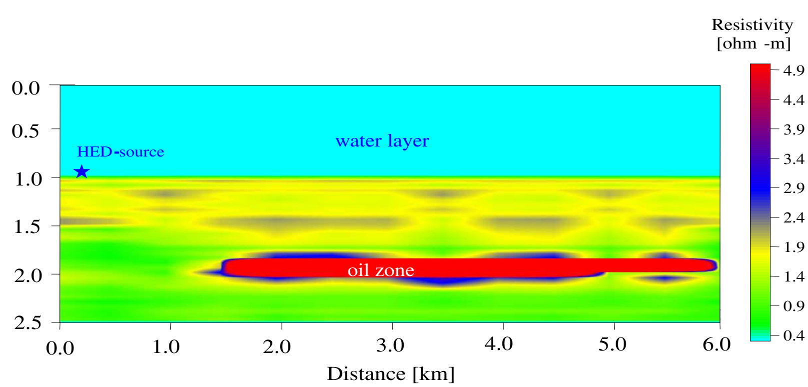


Figure 2: 2-D North-Sea resistivity model.

3. NORTH SEA TEST MODEL

A 2-D resistivity model was constructed based on typical resistivity-log values from the North Sea. An oil-reservoir of thickness 100 metres and thinning out to 50 metres was assumed at a depth of 0.9 km below the seafloor. The resistivity of the thin oil zone was varying between 40 and 50 Ω-m. We considered two different water depths: 200 m and 1000 m. Fig. 2 shows a plot of the 2-D resistivity model (deep-water case).

The HED-source, with an operating frequency of 1Hz, was placed 20 meters above the seafloor and at a lateral distance of 200 meters from the left boundary of the model (see Fig. 2). We considered the case of an inline polarized (i.e., along the y-direction) antenna and computed synthetic data employing a 2.5D hybrid EM-modeling program [2] tailored for the SBL-case. Fig. 3 shows a (logarithmic) plot of the magnitude of the electric field component E_y (which is polarized in the same direction as the source). The figure represents a vertical slice through the 3-D computational volume (with the plane including the source).

Figure 3 clearly demonstrates that strong guided modes leaking from the thin oil zone can be detected at the sea floor, especially for source-receiver offsets larger than approximately 3.5–4.0 km (e.g., about 3–4 times the target depth).

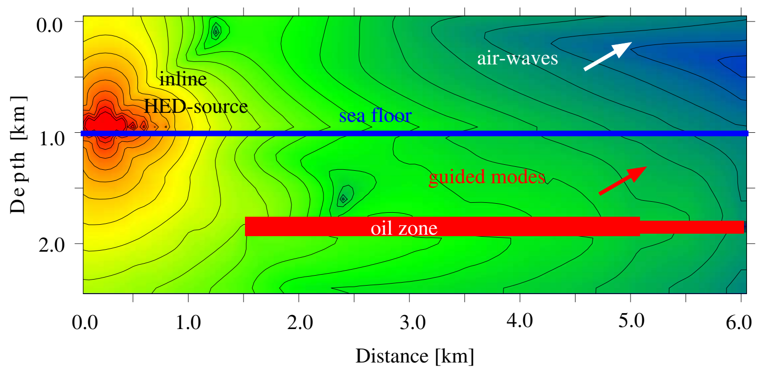


Figure 3: Plot of the magnitude (after taking the logarithm) of the electric field component E_y. Inline polarized HED-source and deep water.

Up till now we have considered the deep-water resistivity model as given by Fig. 1. The result obtained in Fig. 3 shows that for such water depths the airwaves are so attenuated that when they reach the sea floor only negligible energy remains at larger offsets (compared to the guided EM energy leaking from the resistive layer). However, if we assume a shallow water depth this will no longer be the case. Fig. 4 is identical to Fig. 3 except that the water depth now has been set to 200 meters. Note how the airwaves dominate along the sea floor, masking the leaking guided modes almost completely. In order to reveal the signature of the oil zone, these airwaves should be removed or at least strongly attenuated.

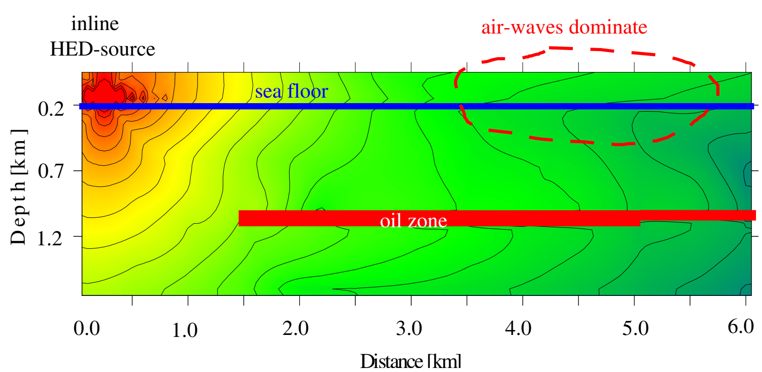
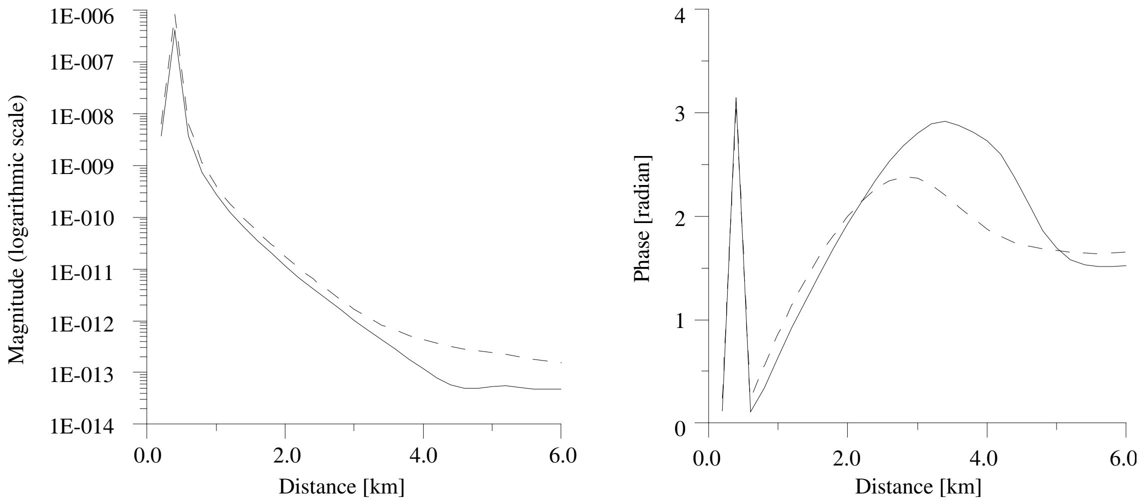


Figure 4: Same as Fig. 3 except that the water depth now is 200 m.

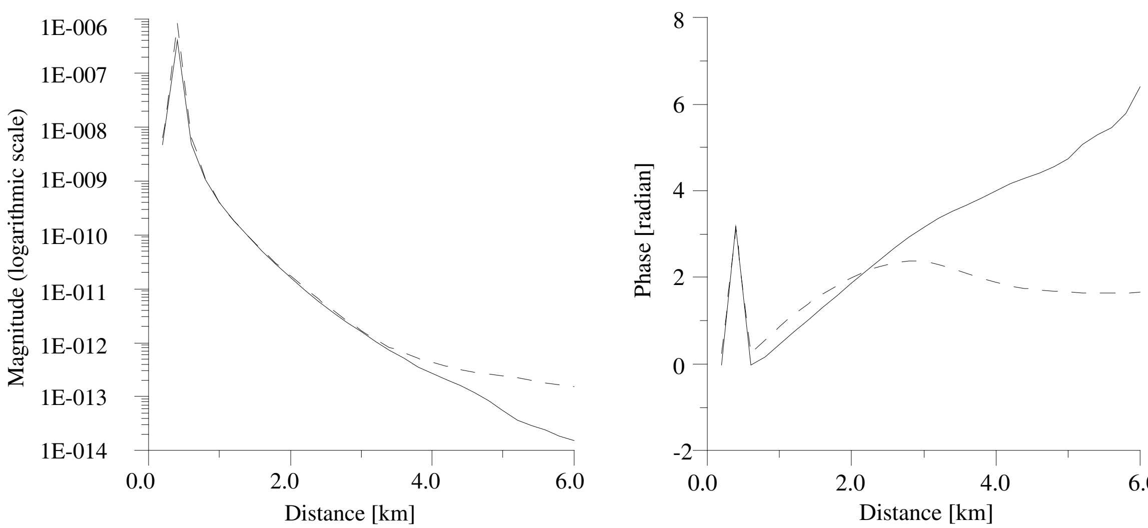
Let us now investigate the potential of the correction scheme given by Eq. (7) employing the same North Sea test model as above and a shallow water depth of 200 meters. Hence, in addition to computing the inline horizontal electric field (cf. Fig. 4), the cross-line horizontal magnetic field was also computed. First, we carried out tests with the decomposition carried out just above the seabed. Figs. 5(a) and (b) show that the method is not working very well for this case.

A better idea is to carry out the decomposition just below the seafloor, since it is known from EM geophysical techniques that the plane-wave impedance senses the material below the sensor and has no sensitivity to the material above it [4]. Using this idea we obtained the results shown in Figs. 6(a) and (b). We now easily see that the EM field decomposition method has worked satisfactorily and that the contribution from the airwaves has been efficiently attenuated (note the fairly straight phase curve after correction in Fig. 6(b) as expected).



(a) (b)

Figure 5: (a) Magnitude (log scale) and (b) phase of the inline field component E_y for a shallow-water case of 200 m. Before (broken curve) and after (solid curve) application of the air-removal technique. Wavefield decomposition **above** seafloor.



(a) (b)

Figure 6: Magnitude (log scale) and (b) phase of the inline field component E_y for a shallow-water case of 200 m. Before (broken curve) and after (solid curve) application of the air-removal technique. Wavefield decomposition **below** seafloor.

4. CONCLUSIONS

Marine remote sensing of hydrocarbons based on EM-soundings gives strongly distorted subsurface measurements in case of shallow waters. However, by combining the measured electric and magnetic field components the contribution from the airwaves can be strongly attenuated.

ACKNOWLEDGMENT

This research has been partially funded by The Norwegian Geotechnical Institute (NGI).

REFERENCES

1. Amundsen, L., L. Løseth, R. Mittet, S. Ellingsrud, and B. Ursin, "Decomposition of electromagnetic fields into upgoing and downgoing components," accepted for publication in *Geophysics*, 2006.
2. Gelius, L., "Modeling the response of a seafloor antenna in the limits of low frequency and shallow water," paper accepted for *PIERS 2007*.
3. Kong, F., H. Westerdahl, S. Ellingsrud, T. Eidesmo, and S. Johansen, "'Seabed logging': A possible direct hydrocarbon indicator for deepsea prospects using EM energy," *Oil & Gas Journal*, May 13 edition, 2002.
4. Wait, J. R., "Theory of magnetotelluric fields," *Radio Science*, Vol. 66D, 509–541, 1962.

Comparison of Antenna Types and Orientations for Detecting Hydrocarbon Layers in Seabed Logging

S. E. Johnstad[1], **H. Westerdahl**[2], **F. N. Kong**[2], **and Z. Wang**[2]

[1]Hydro Oil & Energy, Norway
[2]Norwegian Geotechnical Institute, Norway

Abstract— In seabed logging applications, only three types of sources: VED, HED-R (in-line) and HMD-P (cross line) are sensitive in detecting the target — a thin horizontal hydrocarbon layer. The other three types of the sources: VMD, HED-P (cross line), HMD-R are not sensitive in detecting the target.

1. INTRODUCTION

For seabed logging measurements, there can be following basic types of sources: VED, VMD, HED and HMD. For horizontal antennas, there are two directions: in-line and cross-line. We use '-R' to represent in-line, and '-P' to represent cross-line. Hence there are six types of sources, i.e., VED, VMD, HED-R, HED-P, HMD-R, HMD-P. Obviously there are also six types of receiver antennas. Hence we have 6*6 = 36 source/receiver combinations. In this paper we are going to investigate which source/receiver combinations are sensitive in detecting the thin hydrocarbon layers.

2. SENSITIVITY TO DETECT HYDROCARBON LAYER BASED ON GUIDE WAVE ASSUMPTION

We use the following table to show the sensitivity to detect hydrocarbon layer for different source/receiver combinations, where we use the words 'strong' and 'weak' to characterize the sensitivity in detecting the existence of a thin oil layer, and use 'zero' to mean no field is received for that combination.

Table 1: Sensitivity to detect hydrocarbon layer for different source/receiver combinations.

	Receiver VED	Receiver HED-R	Receiver HED-P	Receiver VMD	Receiver HMD-R	Receiver HMD-P
Source: VED	strong	strong	0	0	0	strong
Source: HED-R	strong	strong	0	0	0	strong
Source: HED-P	0	0	weak	weak	weak	0
Source: VMD	0	0	weak	weak	weak	0
Source:HMD-R	0	0	weak	weak	weak	0
Source:HMD-P	strong	strong	0	0	0	strong

The above table is based on our assumption that the main field components, which propagate inside a thin horizontal hydrocarbon layer, are the vertical E field and cross H field, referred to as the guide wave field components in [1]. Based on this assumption, the in-line HED and VED sources are considered being sensitive in detecting oil layers, since they can generate the guide wave fields inside the oil layer. The cross-line HED is not sensitive in detecting oil layers, since it can not generate the guide wave field components inside the oil layer.

The duality principle states that the fields generated by a magnetic source can be obtained from the fields generated by an electric source as long as following replacements are made: $E \to H$, $H \to -E$, $\mu \to \varepsilon$, $\varepsilon \to \mu$. Hence HMD-R generates the fields as HED-P, and HMD-P generates the fields as HED-R. And VMD generates the fields that VED doesn't generate. We have then derived the results for the magnetic dipole sources shown in the table. In the next section we will use analytical results to verify the results shown in the table.

3. VERIFICATION BY ANALYTICAL RESULTS

The analytical solution for dipole antenna embedded in anisotropic layers (1D target) was developed by Wait [5] and J. A. Kong [2]. Wait's method uses the potential vector and J. A. Kong's method does not use the potential vector. Instead he uses the vertical E and H fields to represent the E and H fields in other directions. In our software implementation we have used J. A. Kong's method [2–4], which calculates all six field components for both electric and magnetic sources along any direction. Hence the software is suitable for verifying our comments on the comparison of antenna types and orientations for detecting hydrocarbon layers.

The model is shown in Figure 1, where the sea (0.3 ohm-m) depth is 2000 m, the target (an oil layer, 50 ohm-m) is at 1000 m below the seabed and 50 m thick, and the overburden has a resistivity of 1 ohm-m. The source is at 50 m above the seabed and the receiver array is placed on the seabed. Frequency 1 Hz is used for modeling.

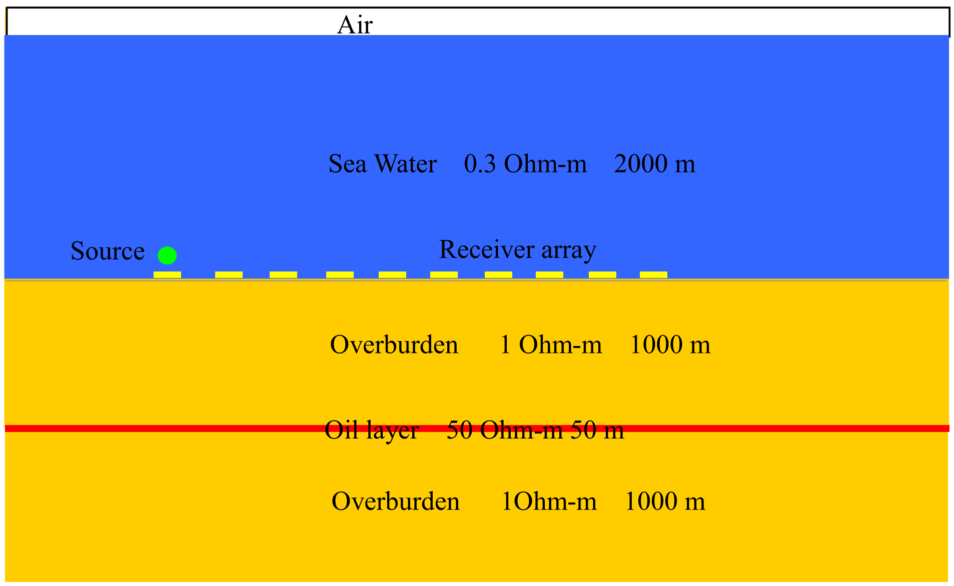


Figure 1: A seabed logging model.

Table 2: Corresponding figure numbers and curve colours for different source/receiver combinations.

	Receiver VED	Receiver HED-R	Receiver HED-P	Receiver VMD	Receiver HMD-R	Receiver HMD-P
Figure 6: Source, VED	(blue) strong	(red) strong	0	0	0	(green) strong
Figure 3: Source, HED-R	(blue) strong	(red) strong	0	0	0	(green) strong
Figure 2: Source: HED-P	0	0	(blue) weak	(red) weak	(green) weak	0
Figure 7: Source: VMD	0	0	(blue) weak	(red) weak	(green) weak	0
Figure 5: Source: HMD-R	0	0	(blue) weak	(red) weak	(green) weak	0
Figure 4: Source: HMD-P	(blue) strong	(red) strong	0	0	0	(green) strong

Figures 2–7 show the modeling results for different source types and orientations. To describe those figures, we have added the corresponding figure numbers and curve colors to Table 1, which leads to Table 2, for the modeling results for different source/receiver combinations.

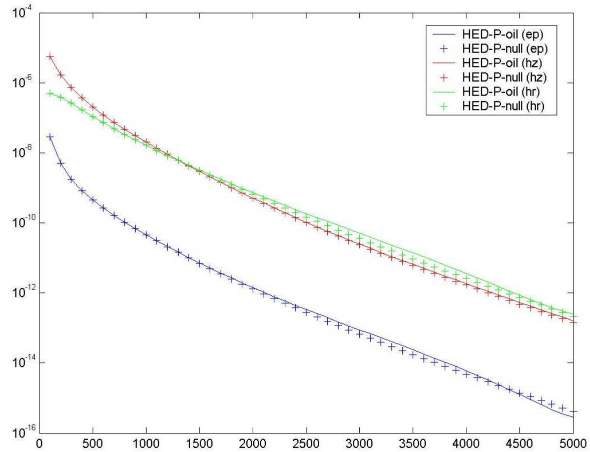


Figure 2: Received fields for HED source (cross). x-axis: receiver offset in meters.

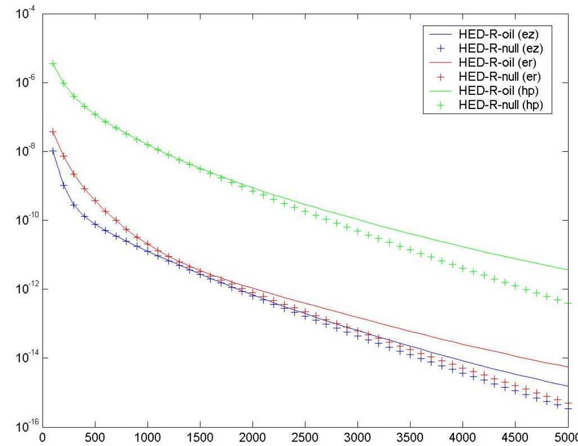


Figure 3: Received fields for HED source (in-line).

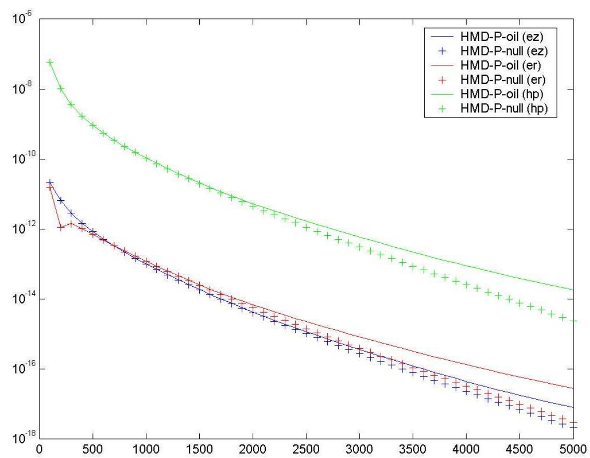


Figure 4: Received fields for HMD source (cross).

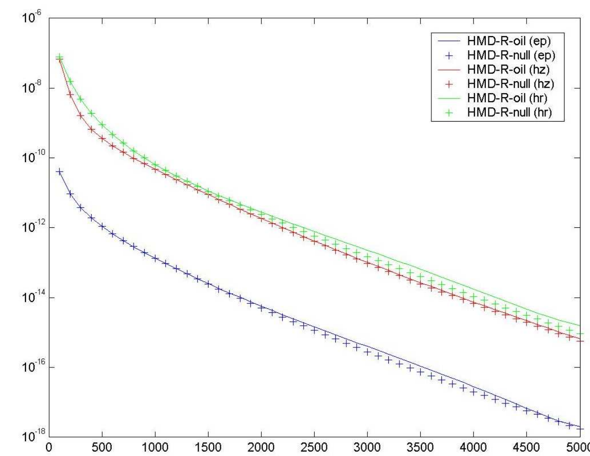


Figure 5: Received fields for HMD source (in-line).

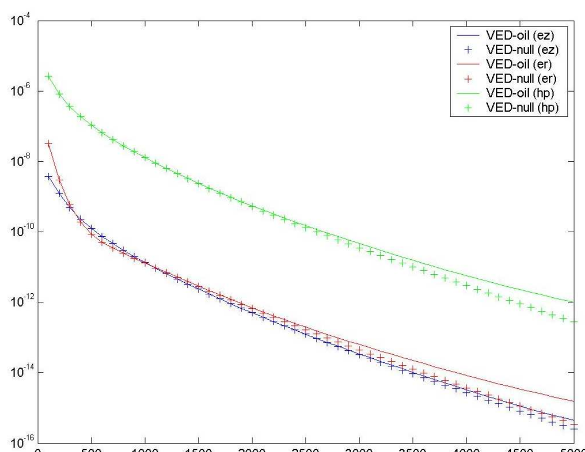


Figure 6: Received fields for VED source (cross).

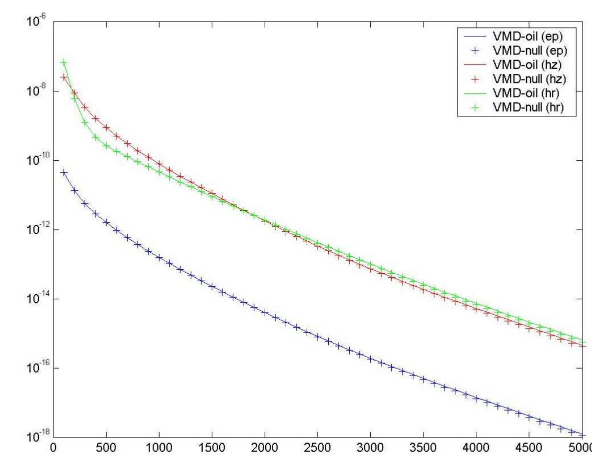


Figure 7: Received fields for VMD source.

In those figures, we use the dotted lines to show the fields received for the case without the oil layer, and use the solid lines to show the fields received for the case with the oil layer. When a field received with target has a larger difference to the field received without the target, we say this source/receiver combination has a strong response to the existence of the target. Otherwise we say the response is weak. From those figures we can see that our conclusion on the sensitivity of detecting a hydrocarbon layer shown in Table 1 is correct.

4. CONCLUSION

In seabed logging applications where a target is a thin horizontal hydrocarbon layer, The VED, HED-R (in-line) and HMD-P (cross line) are sensitive in detecting the target. The VMD, HED-P (cross line), HMD-R are not sensitive in detecting the target.

ACKNOWLEDGMENT

The authors wish to thank the Norwegian Research Council, Hydro Oil & Energy and Statoil for the finance support through a Peromaks project.

REFERENCES

1. Kong, F. N., H. Westerdahl, S. Ellingsrud, T. Eidesmo, and S. E. Johansen, "Seabed logging: A possible direct hydrocarbon indicator for deepsea prospects using EM energy," *Oil & Gas Journal*, May 13, 2002.
2. Kong, J. A., "Electromagnetic fields due to dipole antennas over stratified anisotropic media," *Geophysics*, Vol. 37, No. 6, 985–996, 1972.
3. Kong, J. A., *Electromagnetic Wave Theory*, Second Edition, John Wiley & Sons, Inc., 1990.
4. Tang, C. H., "Electromagnetic fields due to dipole antennas embedded in stratified anisotrpic media," *IEEE Trans. Antenna and Prop.*, Vol. AP-27, No. 5, Sep. 1979.
5. Wait, J. R., "Fields of a dipole over an homogeneous anisotropic half-space," *Can. Jour. Phys.*, Vol. 44, 2387–2401, 1966.

Nonlinear Dynamics of Temporal Optical Soliton Molecules in Lasers

Ph. Grelu[1], **M. Grapinet**[1], **J. M. Soto-Crespo**[2], **and N. Akhmediev**[3]

[1]Laboratoire de Physique de l'Université de Bourgogne
UMR 5027 CNRS, B. P. 47870, 21078 Dijon, France
[2]Instituto de Optica, CSIC, Serrano 121, 28006 Madrid, Spain
[3]Optical Sciences Group, Research School of Physical Sciences and Engineering
The Australian National University, Canberra ACT 0200, Australia

Abstract— Recent experiments demonstrate that fiber laser cavities are able to support various multisoliton complexes, analogous to soliton molecules. These advances, which could have impact on optical information transmission or storage, are guided by the concept of dissipative soliton and supported by numerical simulations.

As passively mode-locked lasers rely strongly on nonlinear dissipation, there is a growing interest in understanding various pulse dynamics in terms of the dynamics of dissipative solitons [1]. In particular, the interaction between dissipative temporal solitons can lead to the formation of stable multi-soliton complexes. The stability of multi-soliton complexes arises from the existence of attractors that are able to bind stably several pulses in a way somehow analogous to the formation of molecules.

The simplest of these complexes is the soliton pair, or stable bound state of two solitons. It was predicted in the frame of the Ginzburg-Landau propagation equation model and characterized by a fixed distance, as well as a fixed phase relationship of $\pm\pi/2$ between the two solitons [2], then found experimentally in a passively mode-locked erbium-doped fiber ring laser [3]. A soliton pair uses the dissipative nonlinear dynamics of the active cavity in order to remain stable for hours without the need of external feed back stabilization. The formation of soliton molecules is not limited to diatomic molecules: triatomic and other polyatomic molecules were also reported [4]. Moreover, the ability to form soliton molecules is not limited to a very specific cavity design, it is rather a general feature that make them potentially observable in most passively mode-locked laser cavity schemes, since it relies mainly on the strength of nonlinear dissipation, which in turn provides attractor sets for soliton molecules formation. For instance, soliton molecules were found with negative and with positive path-averaged chromatic dispersions, as well as with low and high pumping powers [1, 4].

In the following, after having briefly presented the experimental setup used in the experiments, we shall discuss the latest issues discovered with soliton molecules in fiber lasers. They concern the possible control of formation and dissociation of soliton molecules, as well as vibrational motions.

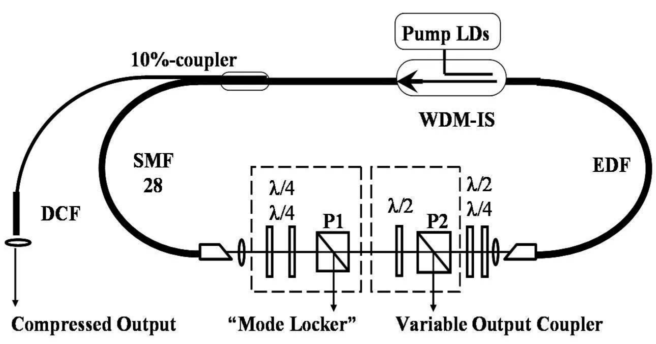


Figure 1: Sketch of the fiber laser experimental setup.

1. FIBER LASER EXPERIMENTAL SETUP

As sketched in Fig. 1, we use a dispersion-managed fiber ring laser which series a 1.6-meter-long 1400-ppm erbium-doped fiber (EDF) in the normal dispersion regime ($D = -35\,\text{ps.nm}^{-1}.\text{km}^{-1}$) which provides laser emission around $1.5\,\mu\text{m}$, a polarization-insensitive optical isolator, a given length of standard telecom fiber (SMF) in the anomalous dispersion regime ($D = \pm 17\,\text{ps. nm}^{-1}.\text{km}^{-1}$), and an open-air section that includes retarding wave plates and a polarizing beam splitter cube. Nonlinear polarization evolution that takes place in both fibers, along with polarization discrimination by the polarizing cube, provides the equivalent of an ultrafast saturable absorber. Under a 980-nm pumping power in the 100-300 mW range, stable mode locking is achieved with appropriate adjustment of the retarding wave plates that precede the polarizing cube. According to the length of SMF chosen for a given experiment, the averaged cavity dispersion can be varied between anomalous or normal. Under adjustments of pump power and retarding wave plates, pulse durations are in the 100-600 fs range, with intracavity energies in the 50-500 pJ range. Using these adjustments of pump power and retarding wave plates, we are able to switch from single soliton to multiple soliton operation. Part of the multiple soliton operations lead to the formation of stable multisoliton complexes, such as the stable soliton pair [3].

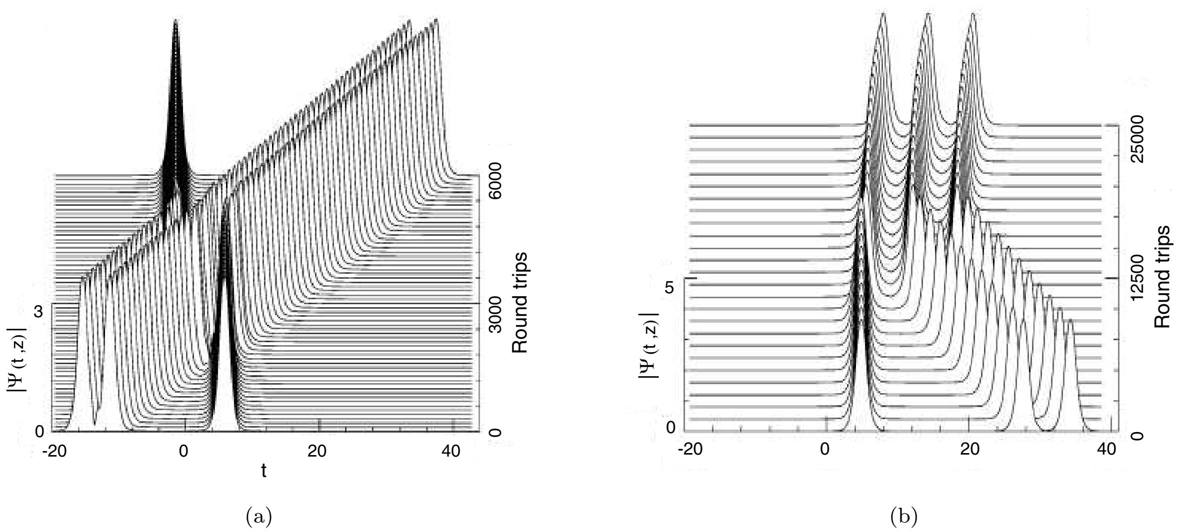


Figure 2: (a) Collision between soliton pair and soliton singlet of elastic type, (b) inelastic type with triplet formation.

2. CONTROL OF FORMATION AND DISSOCIATION

Besides stable multisoliton complexes, large varieties of pulse behavior are accessible with mode-locked lasers. Recently, study of collisions between a soliton pair and a single soliton was reported [5], showing that some amount of control is possible in the formation of these "soliton molecules".

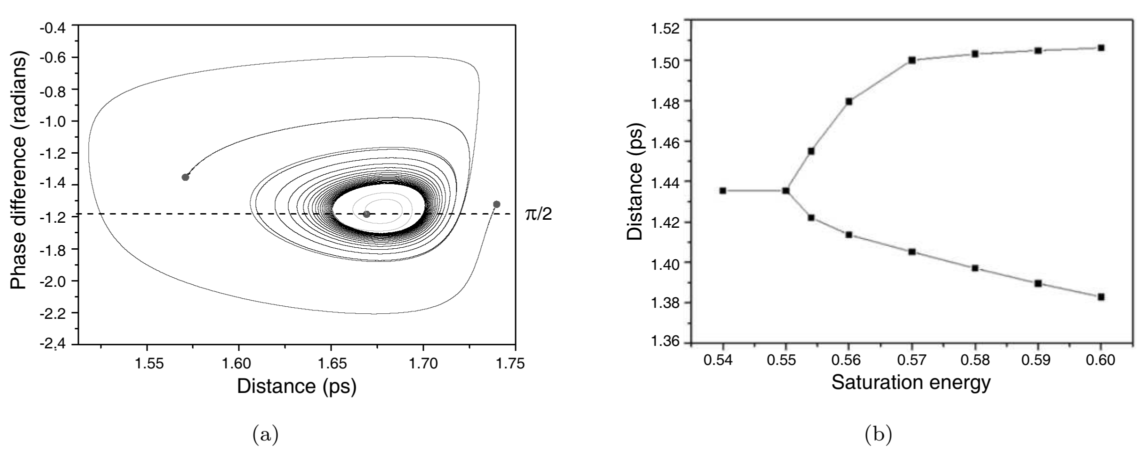


Figure 3: (a) limit cycle attractor which attracts soliton pairs into a vibrating state, in which the relative distance and phase relationship oscillate. For each trajectory, the initial condition is marked by a large dot, and continued with one small dot per cavity roundtrip, (b) starting from non vibrating soliton pairs, limit cycles appear after a Hopf bifurcation takes place along with the tuning of a cavity parameter, such as the saturation energy.

What can be the outcome of the interaction between a dissipative soliton pair and a soliton singlet? Tuning a cavity parameter it was possible experimentally to switch from collisions of "elastic" type, to "inelastic" collisions ending with formation of stable triplet soliton states. These observations were supported with numerical simulations based on the dispersion (parameter) managed cubic-quintic Ginzburg-Landau equation model that are illustrated in Fig. 2.

3. VIBRATING SOLITON MOLECULES

In analogy with molecules, soliton pairs and other bound multisoliton complexes can have vibrational motion. This depends on the set of cavity parameters, which in turn defines the type of attractor which provides stabilization of the "molecular structure". We studied the simplest case of a soliton pair, and followed the evolution of the relative distance and phase relationship. In such representation, the fully stable soliton pair i. e., with no relative motion in stationary regime corresponds to a point attractor [2, 3], whereas the vibrating soliton pair corresponds to a limit cycle attractor, as appears in Fig. 3(a). The switching between point attractor and limit cycle attractor (see Fig. 3(b)) can follow a supercritical Hopf bifurcation along with cavity parameter change [7]. Numerical simulations are consistent with experimental observations based on the analysis of optical spectra and optical autocorrelation recordings.

4. CONCLUSIONS

Nonlinear dissipation offers efficient ways to control or stabilize limited sets of light pulses, using the well defined properties of dissipative solitons. We have shown that the concept of a "soliton molecule", although it should be used carefully, is a useful guide to discover and understand new properties of dissipative multi-soliton complexes, such as formation or dissociation, and vibrations. The present state of the art is still at the fundamental level, but future applications could be found in optical regeneration or in optical buffer memories. In the context of propagation in dispersion-managed fiber links, optical soliton molecules have also been proposed as a way to increase channel transmission capacity in providing upper bits, and stability of a soliton molecule was demonstrated experimentally recently [8].

ACKNOWLEDGMENT

The work of Ph. G. was supported by Agence Nationale de la Recherche, and the work of J. M. S. C. was supported by the M. C. y T. under contract BFM2003-00427 and FIS2006-03376. N. A. acknowledges support from the Australian Research Council.

REFERENCES

1. Akhmediev, N. and A. Ankiewicz, editors, *Dissipative Solitons*, Springer-Verlag, Berlin, 2005.
2. Akhmediev, N. N., A. Ankiewicz, and J. M. Soto-Crespo, "Stable soliton pairs in optical transmission lines and fiber lasers," *J. Opt. Soc. Am. B*, Vol. 15, 515–523, 1998.
3. Grelu, P., F. Belhache, F. Gutty, and J. M. Soto-Crespo, "Relative phase-locking of pulses in a passively mode-locked fiber laser," *J. Opt. Soc. B*, Vol. 20, 863–870, 2003.
4. Grelu, P. and J. M. Soto-Crespo, "Multisoliton states and pulse fragmentation in a passively mode-locked fibre laser," *J. Opt. B: Quantum Semiclass. Opt.*, Vol. 6, S271–S278, 2004.
5. Grelu, P. and N. Akhmediev, "Group interactions of dissipative solitons in a laser cavity: the case of 2+1," *Opt. Express*, Vol. 12, 3184–3189, 2004. Akhmediev, N., J. M. Soto-Crespo, M. Grapinet, and P. Grelu, "Dissipative soliton interactions inside a fiber laser cavity," *Opt. Fiber Technol.*, Vol. 11, 209–228, 2005.
6. Soto-Crespo, J. M., M. Grapinet, P. Grelu, and N. Akhmediev, "Bifurcations and multiple-period soliton pulsations in a passively mode-locked fiber laser," *Phys. Rev. E*, Vol. 70, 066612, 1–11, 2004.
7. Grapinet, M. and P. Grelu, "Vibrating soliton pairs in a mode-locked laser cavity," *Opt. Lett.*, Vol. 31, 2115–2117, 2006.
8. Stratmann, M., T. Pagel, and F. Mitschke, "Experimental observation of temporal soliton molecules," *Phys. Rev. Lett.*, Vol. 95, 143902, 1–4, 2005.

Realization of a Cavity-soliton Laser Based on Broad-area Vertical-cavity Devices with Frequency-selective Feedback

T. Ackemann, Y. Tanguy, A. Scroggie, A. Yao, W. J. Firth, and G.-L. Oppo
SUPA, Department of Physics, University of Strathclyde, 107 Rottenrow, Glasgow G4 0NG, Scotland, UK

P. Paulau, A. V. Naumenko, and N. A. Loiko
Institute of Physics, NASB, 220072, Nezalezhnasti ave., 68, Minsk, Belarus

R. Jäger
ULM Photonics, Lise-Meitner-Str. 13, 89081 Ulm, Germany

Abstract— The spontaneous formation of small-area bistable lasing spots is observed at different positions within the aperture of a broad-area VCSEL with frequency-selective feedback. These spots can be switched on and off with an incoherent injected field. They are interpreted as spatial dissipative solitons. Approaches to model these devices are presented and first results on the occurrence of localization are reported.

1. INTRODUCTION

Recent years showed remarkable progress in achieving, controlling and understanding of bistable soliton-like emission states, which are usually referred to as cavity solitons (CS), in semiconductor microcavities [1–4]. They are discussed as 'bits' for future all-optically and potentially massively parallel information processing schemes. Currently realized schemes rely on a passive or active semiconductor microcavity driven by a broad-area holding beam (HB) of high spatial and temporal coherence [1–4]. This HB and its frequency detuning with the vertical-cavity surface-emitting lasers (VCSEL) cavity resonance induce bistability, a necessary condition for the appearance of CS. We are following a different approach for a self-sustained cavity soliton laser pumped only by incoherent means and investigate a VCSEL with frequency-selective feedback, therefore removing the need of a HB and simplifying the scheme. Such a device would support bistable self-localized emission states purely by electrical pumping. Previous theoretical analysis showed that this system exhibits bistability between the non-lasing off-state and lasing states [5], a potentially suitable situation for CS.

2. EXPERIMENT

The experiment is based on broad-area bottom-emitting VCSELs. It emits at 980 nm, and is electrically pumped through a 200 μm circular oxide aperture. The experimental apparatus is shown in Figure 1. The external cavity includes two lenses and a holographic grating in a Littrow configuration and is adjusted to be self-imaging in order to keep the high Fresnel number of the VCSEL.

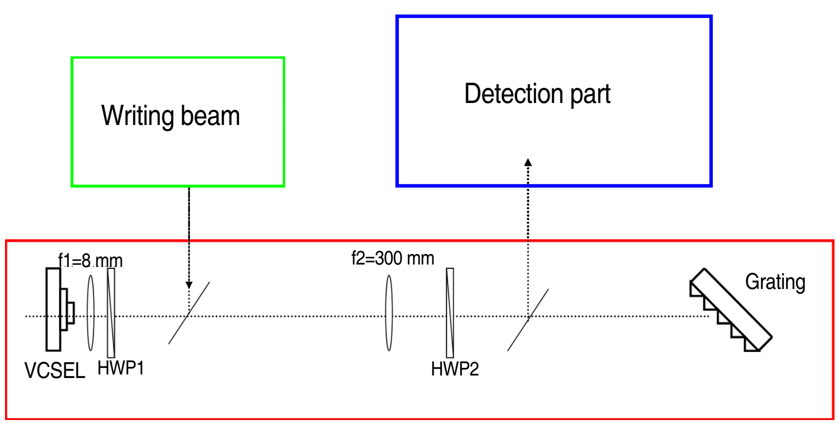


Figure 1: Experimental set-up. The external cavity is self-imaging, and beam samplers are positioned to couple out a fraction of the beam to the detection part, and also to allow the injected field to enter the cavity.

The grating frequency was chosen to coincide approximately with the longitudinal resonance of the VCSEL cavity below the solitary laser threshold. With a careful alignment of the feedback from the grating, several bistable spots appeared spontaneously in the near-field. Their switch-on and switch-off are abrupt, and their bistability range on the order of several mA. The spot size is about $10\,\mu$m. Frequency spectra of one spot demonstrated that it can operate on a single longitudinal mode, with a 10 MHz linewidth and 10 dB side lobe attenuation. These spots can exist at many though not all positions within the aperture and seem to be pinned by local defects in the near-fields, such as defects lines or other local inhomogeneities in the active layer.

Thereafter we examined the possibility of switch-on and switch-off of these localized emissions with an injected field, or writing beam (WB). The beam of a tunable laser was shaped to obtain a $20\,\mu$m beam waist onto the VCSEL, and its wavelength adjusted to be near the grating wavelength. Two of the spontaneously appearing spots were chosen, at locations away from the near-field boundaries and defect lines. The VCSEL was biased to be within the bistability range of these two spots. Figure 2 displays near-field intensity distributions during the experiment. The spots switch-on could be obtained by applying the WB directly onto the spot locations, while for the switch-off the WB is slightly misaligned to be located 10–15 μm away from the spots. We note that both switch-on and switch-off are incoherent processes, as the WB and spot frequencies do not need to be identical. This makes the scheme robust which is appealing for applications.

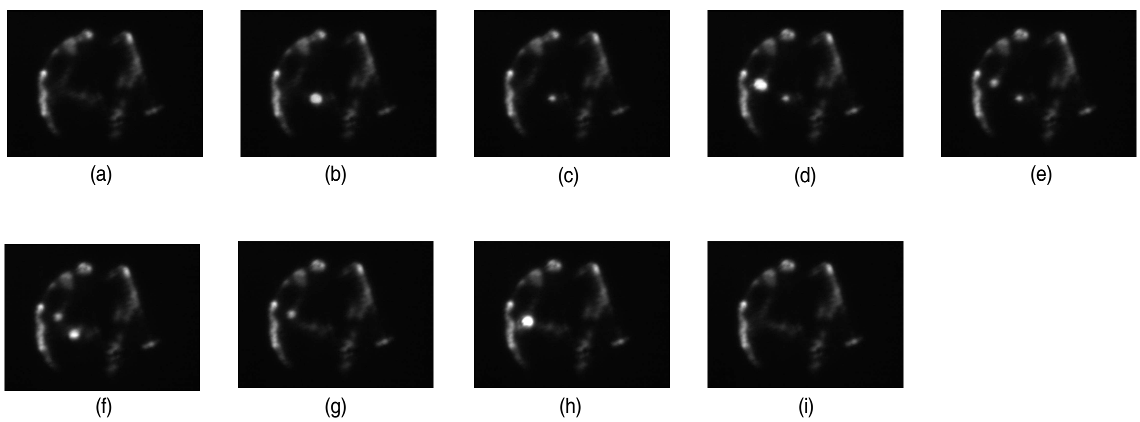


Figure 2: Near-fields showing the successive switch-on of two spots with an injected incoherent field (brightest spot), and their switch-off. The later is accomplished by injecting the WB beside the spot locations. a) Both spots are off, b) injection of WB, c) one spot was switched-on and stays, d) injection of WB at another location, e) second spot switched-on, f) WB injected beside the first spot, g) first spot was switched-off and does not reappear, h) injection of writing beam to switch-off the second spot, i) second spot remains off.

3. MODELING

Based on previous models for freely-running broad-area lasers [6] and small-area lasers with frequency-selective feedback [5] a model was developed which allows for the treatment of feedback effects in broad-area lasers. The model is based on a complex equation for the dynamics of the intra-cavity light field taking into account the finite width of the gain curve and a real equation for the carrier dynamics. Delayed feedback is taken into account on the level of a single round-trip in the external cavity (so-called Lang-Kobayashi approximation).

The stability of the non-lasing off-state versus spatially homogeneous and inhomogeneous perturbations was analyzed. As expected there is bistability (actually multistability) between the off-state and different homogeneous lasing states (corresponding to the one found earlier in [5]), if only homogeneous states are considered. However, it is found also that the off-state is unstable versus inhomogeneous perturbations, i.e., pattern formation. Hence no stable localized states are expected. In order to re-establish a bistable situation, a spatial Fourier filter was introduced in the feedback loop, which suppresses the instability of the off-state to all spatial perturbations except for a certain wavenumber. In this situation a traveling wave extending over all the aperture is found if the simulations are started from noisy initial conditions.

If this solution is used as a localized seed on an otherwise zero background for parameters where the background is linearly stable, localized states are obtained in numerical simulations.

The localized states have the form of a patch of a traveling wave pattern embedded in the zero background solution. Both rotational symmetric and asymmetric states are obtained depending on the exact initial conditions.

Current investigations are directed at clarifying their properties in dependence on parameters and the connection between the experimental and the numerical observations.

Further research is directed on a more detailed modeling of the cavity structure including the intracavity optical elements. This is possible by master equations approach [7] taking into account the ABCD matrix of the optical system. Consideration of the effects of misalignment (including deliberately-introduced asymmetric elements such as gratings) necessitates an extension of the usual ABCD matrix methods, typically based on 3×3 matrices [8]. Application of this approach to the experimental system is ongoing.

4. CONCLUSIONS

We demonstrated the existence of small-area bistable lasing spots in a frequency selective feed-back scheme. These could be switched on/off with an incoherent injected field. These spots seem to possess all properties expected from cavity solitons, though we caution that the possible role of material defects in localization needs to be addressed by future studies.

ACKNOWLEDGMENT

This work was supported by the EU-STREP FunFACS, INTAS and a starter grant by the Faculty of Science of the University of Strathclyde.

REFERENCES

1. Barland, S., J. R. Tredicce, M. Brambilla, L. A. Lugiato, S. Balle, M. Giudici, T. Maggipinto, L. Spinelli, G. Tissoni, T. Knödel, M. Miller, and R. Jäger, "Cavity solitons as pixels in semiconductors," *Nature*, Vol. 419, 699–702, 2002.
2. Taranenko, V. B. and C. O. Weiss, "Incoherent optical switching of semiconductor resonator solitons," *Appl. Phys. B*, Vol. 72, No. 7, 893–895, 2001.
3. Akhmediev, N. and A. Ankiewicz, editors, "Dissipative solitons," *Lecture Notes in Physics*, Vol. 661, Springer, New York, 2005.
4. Barbay, S., Y. Ménesguen, X. Hachair, L. Leroy, I. Sagnes, and R. Kuszelewicz, "Incoherent and coherent writing and erasure of cavity solitons in an optically pumped semiconductor amplifier," *Opt. Lett.*, Vol. 31, 1504–1507, 2006.
5. Naumenko, A., N. A. Loiko, M. Sondermann, K. F. Jentsch, and T. Ackemann, "Abrupt turn-on and hysteresis in a VCSEL with frequency-selective optical feedback," *Opt. Commun.*, Vol. 259, 823–833, 2006.
6. Loiko, N. A. and I. V. Babushkin, "Competition of orthogonally polarized transverse Fourier modes in a VCSEL," *J. Opt. B: Quantum Semiclass. Opt.*, Vol. 3, S234–S243, 2001.
7. Dunlop, A. M., W. J. Firth, and E. M. Wright, "Master equation for spatio-temporal beam propagation and Kerr lens mode-locking," *Opt. Commun.*, Vol. 138, 211–226, 1997.
8. Firth, W. J. and A. Yao, "Master equation for misaligned cavities," *J. Mod. Opt.*, Vol. 53, 2005–2010, 2006.

Dynamics of a Semiconductor Laser with Two External Cavities

A. N. Pisarchik and F. R. Ruiz-Oliveras
Centro de Investigaciones en Optica, Mexico

Abstract— We study numerically complex dynamics of a semiconductor with one and two external cavities. We show that dynamical regimes in the laser with one external cavity can be controlled by adding another external cavity and properly adjusting its length and feedback strength. We demonstrate the existence of different stable periodic, quasiperiodic, chaotic, and steady-state regimes which form Arnold's tongues in bi-dimensional parameter spaces of the length and feedback strength of the external cavities and the pump parameter.

The dynamics of a semiconductor laser subjected to external optical feedback has been studied by many researchers, since this laser has many applications in optical communications, interferometric sensors, frequency stability, etc. (see, for example, Ref. 1 and references therein). For moderate and strong feedback strengths this laser can display a very rich dynamical behavior, from periodic and quasiperiodic oscillations to chaos and coherence collapse [2, 3]. When the laser injection current is close to the solitary laser threshold, the laser operates in a stable steady-state regime. As the pump current is increased, intermittent drops of the laser intensity appear, that gives rise to low-frequency fluctuations (LFF). At higher currents, the laser optical bandwidth broadens that is known as coherence collapse [4, 5].

The dynamical behavior of a semiconductor lasers with a single feedback was studied extensively and the basic mechanism for LFF is well understood [6, 7]. The possibility for controlling LFF in a semiconductor laser with an external cavity by using a second delayed optical feedback was initially deduced by Liu and Ohtsubo [8]. Later their idea was developed by Rogister, et al., [9, 10], who realized the suppression of antimodes, responsible for LFF crises by properly adjusting the second feedback strength. Recently Mendez, et al., [11] have demonstrated experimentally that the frequency of LFF can be locked by external periodic modulation applied to the feedback strength of the external cavity. They have found different periodic and quasiperiodic regimes which formed Arnold's tongues in the space of the amplitude and frequency of the external modulation.

In this paper we demonstrate that the locking effect can be achieved without any external forcing. We show how LFF in a semiconductor laser with external cavity can be adequately controlled by adjusting properly both the length and the feedback strength of the second external cavity. The feedback time and the feedback strength of the second external cavity act in a similar manner as the period and amplitude of external modulation. We demonstrate that the variation of the parameters of the second external cavity allows one to obtain different dynamical regimes of the laser operation without any modifications in the solitary laser with a single external cavity.

A single-mode semiconductor laser with delayed feedback is usually modeled by the Lang-Kobayashi rate equations [12]. Due to the infinite dimension of the system, an analytical study is very difficult. Therefore, to study the dynamics of a semiconductor laser with two external cavities shown in Figure 1, we make numerical calculations of the modified equations similar to those explored previously by Sivapakrasam, et al., [13] and Carr [14]. For weak or moderate feedback the modified equations can be written as follows

$$\begin{aligned} d\dot{E}_0(t) &= (1/2)\,G\left[N(t)-N_{th}\right]E_0(t)+\kappa_1 E_0(t-\tau_1)\cos\left[\psi_1+\phi(t)-\phi(t-\tau_1)\right] \\ &\quad +\kappa_2 E_0(t-\tau_2)\cos\left[\psi_2+\phi(t)-\phi(t-\tau_2)\right], \end{aligned} \tag{1}$$

$$\begin{aligned} d\dot{\phi}(t) &= (\alpha/2)\,G\left[N(t)-N_{th}\right]E_0(t)-\kappa_1\left[E_0(t-\tau_1)/E_0(t)\right]\sin\left[\psi_1+\phi(t)-\phi(t-\tau_1)\right] \\ &\quad -\kappa_2\left[E_0(t-\tau_1)/E_0(t)\right]\sin\left[\psi_2+\phi(t)-\phi(t-\tau_2)\right], \end{aligned} \tag{2}$$

$$d\dot{N}(t) = P-N(t)/\tau_s-G\left[N(t)-N_0\right]E_0^2(t). \tag{3}$$

The Equations (1)–(3) are expressed with the variables of the amplitude $E_0(t)$ of the complex electric field $E(t)=E_0(t)\exp\left[i(\omega_0)t+\phi(t)\right]$ (ω_0 being the angular frequency of the solitary laser), the average carrier density in the active region $N(t)$ and the phase $\phi(t)$. The carrier densities at threshold $N_{th}=N_0+(\tau_p G)^{-1}$, where N_0 is the carrier density at transparency, τ_p is the photon

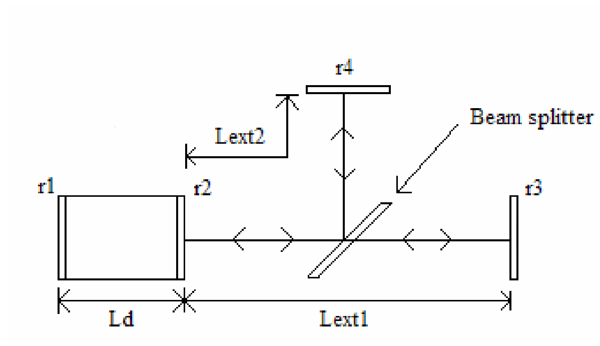


Figure 1: Scheme of laser diode with two external cavities. Ld, Lext1, and Lext2 are the lengths of the internal and two external cavities, r1 and r2 are the laser facets, and r3 and r4 are the external mirrors.

lifetime and G is the modal gain coefficient. The initial phases for the first and second external cavities $\psi_1 = \omega_0\tau_1$ and $\psi_2 = \omega_0\tau_2$, where τ_1 and τ_2 are the round trip times for the first and second external cavities. The other parameters are τ_s is the carrier lifetime, P is the pumping term, α is the linewidth enhancement factor, and κ_1 and κ_2 are the feedback strengths of the first and second external cavities. The last parameters are calculated with the following formula: $\kappa_{1,2} = (1 - r_{3,4}) / \left(\tau_{1,2}\sqrt{r_2/r_{3,4}}\right)$. In the simulation we use the following parameter values [15]: $\alpha = 3.5$, $G = 5 \times 10^{25}\,\mathrm{m}^3\,\mathrm{ns}^{-1}$, $\tau_s = 1\,\mathrm{ns}$, $\tau_s = 1\,\mathrm{ps}$, $P = 8 \times 10^{24}\,\mathrm{m}^{-3}$, $N_{th} = 5 \times 10^{24}\,\mathrm{m}^{-3}$ and $N_0 = 3 \times 10^{24}\,\mathrm{m}^{-3}$. In all simulations with two external cavities, the parameters of the first external cavity, ψ_1, τ_1 and κ_1, are fixed.

First, we consider a semiconductor laser with one external cavity ($\kappa_2 = 0$) and then we will show how a second external cavity affects the laser dynamics. The analysis is performed with codimensional-two bifurcation diagrams in the parameter spaces of the length and feedback strength of the external cavity and the pump parameter. Starting from a certain initial optical phase $\psi_1 = 31.49$, the dynamical behavior of the 837-nm wavelength semiconductor laser with one external cavity of 3.3-cm length exhibits a 2π-cyclic behavior [15]. Thus, we obtain a reduced angular frequency of the solitary laser to be $\omega_0 = 143\,\mathrm{GHz}$. The system Equations (1)–(3) with above parameters yield a chaotic solution in the form of LFF shown in Figure 2.

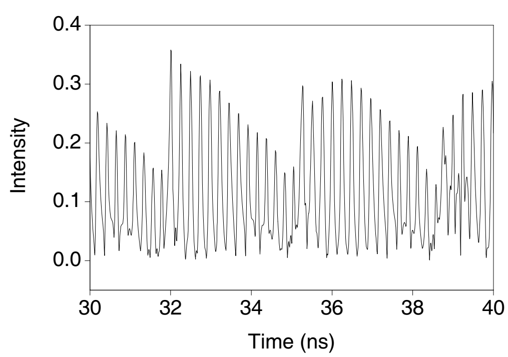


Figure 2: Chaotic time series in semiconductor laser with one external cavity.

In Figures 3 we plot the state diagram of the laser with one external cavity in the bi-dimensional parameter space of the external cavity length $l_1 = c\tau_1/2$ (c being the speed of light) and the feedback strength. The diagram represents the structure of Arnold tongues. The analysis of the state diagrams in Figure 3 shows that with increasing either l_1 or κ_1, the laser undergoes a periodic cascade of bifurcations: a Hopf bifurcation, a period-doubling bifurcation(s), a torus bifurcation after which a quasi-periodic regime is developed. Then quasi-periodicity is converted to chaotic LFF terminated in crisis where a new cascade is initiated. The black and yellow regions are bounded by Hopf bifurcation lines and the yellow and blue regions by period-doubling bifurcation lines. One can see that for very low feedback strengths ($\kappa_1 < 1\,\mathrm{ns}^{-1}$), the laser stays in a cw (steady-state) regime for any length of the external cavity. Another important result is that for relatively strong feedback strengths, the width of the chaotic tongues increases with increasing the cavity length.

Chaotic oscillations are not observed in the laser with a very short external cavity, i.e., the laser is always stable. Instead, the laser with a very long external cavity is always unstable.

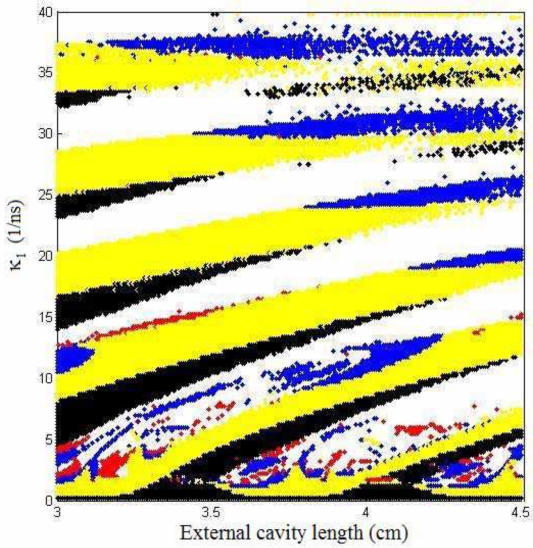


Figure 3: State diagram of semiconductor laser with one external cavity in parameter space of its length and feedback strength. The black dots indicate fix points or cw laser operation, the yellow, blue, and red regions are respectively period-1, period-2, and period-3 solutions, and the white regions indicate quasiperiodic and chaotic solutions.

The influence of the pumping on the dynamical states of the laser is illustrated in Figures 4 and 5. The laser with a relatively short external cavity ($l_1 < 1.7\,\mathrm{cm}$) and a weak feedback strength ($\kappa_1 < 1\ \mathrm{ns}^{-1}$) works in a stationary regime independent on how much is the pumping. However, for a longer external cavity the laser dynamics is very rich. The state diagrams in both graphics form the structure of Arnold tongues which are distributed almost equidistantly along the abscissa axes. One can also see that while P is increasing, the chaotic tongues become wider and finally at a very strong pump, the chaotic regime dominates.

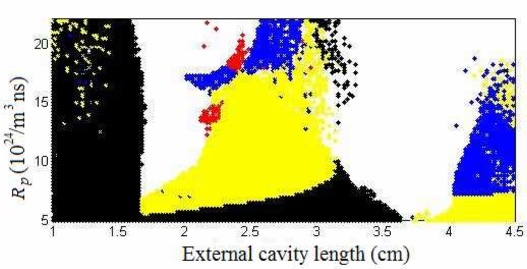


Figure 4: State diagrams of semiconductor laser with one external cavity in space of cavity length and pump parameter. $\kappa_1 = 25\,\mathrm{ns}^{-1}$.

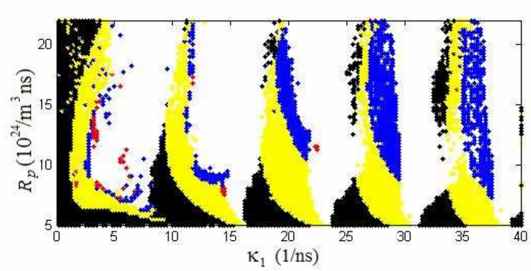


Figure 5: State diagrams of semiconductor laser with one external cavity in space of feedback strength and pump parameter. $l_1 = 3.3\,\mathrm{cm}$.

Figures 6 and 7 display the state diagrams of the laser with two external cavities in the parameter spaces of ratios of the round trip times and feedback strengths of the external cavities, τ_1/τ_2 and κ_1/κ_2. Although the parameters of the first external cavity are hold constant ($\psi_1 = 31.49$, $\tau_1 = 0.22\,\mathrm{ns}$ and $\kappa_1 = 25\,\mathrm{ns}^{-1}$), we plot the graphics in the coordinates of the ratios of the parameters of the two cavities for better illustration of the distribution of the locking regions in the parameter space. In Figure 6 the Arnold tongues represented different dynamical regimes are distributed almost periodically along both axes. So, we have cascades of bifurcations in both directions. This means that the phase of the second external cavity locks the phase of the first external cavity at certain ratios of their parameters. As seen from Figures 6 and 7, the minima of the tongues are located at rational numbers of the phases, $\tau_1/\tau_2 = n/m$ (both n and m being integers), i.e., at $\tau_1/\tau_2 = 1/1,\ 1/2,\ 1/4,\ 2/1,\ 3/2, \ldots$. It is particularly remarkable that the diagrams around $\tau_1/\tau_2 = 1$ are almost symmetrical. Thus, the second external cavity can be used to control dynamical regimes and chaos in a semiconductor laser with external cavity. For example, LFF shown in Figure 2 can be suppressed by adding a second external cavity and properly adjusting its length and feedback strength.

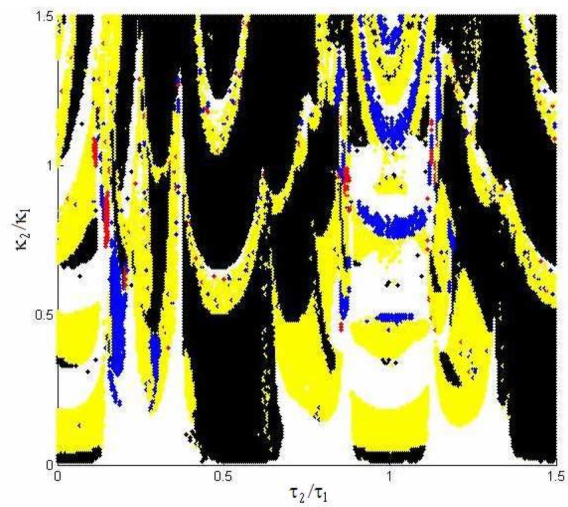


Figure 6: Space diagram of semiconductor laser with two external cavities in space of ratios of their round trip times and feedback strengths. $P = 8 \times 10^{24}\,\mathrm{m}^{-3}\mathrm{ns}^{-1}$.

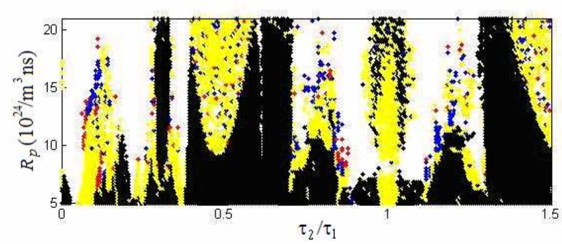


Figure 7: Space diagram of semiconductor laser with two external cavities in space of ratios of their round trip times and pump parameter. $\kappa_1/\kappa_2 = 1$.

In conclusion, in this work we have studied complex dynamics in a semiconductor laser with one and two external cavities. We have shown that dynamical states of the laser with one external cavity in the parameter spaces of the external cavity length, feedback strength and pump parameter have a form of Arnold tongues. A second external cavity provides an additional possibility to control laser dynamics by varying its parameters. The analysis of bi-dimensional state diagrams of the laser with two external cavities have demonstrated that steady-state and different periodic regimes can be locked at certain ratios of the lengths and feedback strengths of the two external cavities. We believe that these results can be of interest for laser engineering and some technological applications, as well as for optical communications.

ACKNOWLEDGMENT

The work was supported by the Mexican Council of Science and Technology (CONACYT), Project No. 46973. A. N. P. acknowledges support from the Spanish Ministry of Education and Science, Project No. SAB2004-0038.

REFERENCES

1. Ohtsuko, J., *Semiconductor Lasers: Stability, Instability and Chaos, Springer Series in Optical Sciences*, Vol. 111, Springer-Verlag, Berlin, 2005.

2. Besnard, P., B. Meziane, and G. M. Stephan, "Feedback phenomena in a semiconductor laser induced by distant reflectors," *IEEE J. Quantum Electron.*, Vol. 29, No. 5, 1271–1284, 1993.
3. Giudici, M., C. Green, G. Giacomelli, U. Nespolo, and J. R. Tredicce, "Andronov bifurcation and excitability in semiconductor laser with optical feedback," *Phys. Rev. E*, Vol. 55, No. 6, 6414–6418, 1997.
4. Salathe, R. P., "Diode lasers coupled to external resonators," *Appl. Phys.*, Vol. 20, 1–18, 1979.
5. Lenstra, D., B. H. Verbeek, and A. J. den Boef, "Coherence collapse in single-mode semiconductor lasers due to optical feedback," *IEEE J. Quantum Electron.*, Vol. 21, 674–679, 1985.
6. Mørk, J., B. Tromborg, and J. Mark, "Chaos in semiconductor laser with optical feedback: theory and experiment," *IEEE J. Quantum Electron.*, Vol. 28, No. 1, 93–108, 1992.
7. Sano, T., "Antimode dynamics and chaotic itinerancy in the coherence collapse of semiconductor lasers with optical feedback," *Phys. Rev. A*, Vol. 50, No. 3, 2719–2726, 1994.
8. Liu, Y. and J. Ohtsubo, "Dynamics and chaos stabilization of semiconductor lasers with optical feedback from an interferometer," *IEEE J. Quantum Electron.*, Vol. 33, No. 7, 1163–1169, 1997.
9. Rogister, F., P. Mégret, O. Deparis, M. Blondel, and T. Erneux, "Suppression of low frequency fluctuations and stabilization of a semiconductor laser subjected to optical feedback from a double cavity: theoretical results," *Optics Letters*, Vol. 24, No. 17, 1218–1220, 1999.
10. Rogister, F., D. W. Sukow, A. Gavrielides, P. Mégret, O. Deparis, and M. Blondel, "Experimental demonstration of suppression of low-frequency fluctuations and stabilization of an external cavity laser diode," *Optics Letters*, Vol. 25, No. 11, 808–810, 2000.
11. Mendez, J. M., R. Laje, M. Giudici, J. Aliaga, and G. B. Mindlin, "Dynamics of periodically forced semiconductor laser with optical feedback," *Phys. Rev. E*, Vol. 63, No. 6, 066218, 2001.
12. Lang, R. and K. Kobayashi, "External optical feedback effects on semiconductor injection laser properties," *IEEE J. Quantum Electron.*, Vol. 16, No. 3, 347–355, 1980.
13. S. Sivapakrasam, E. M. Shahverdiev, and K. A. Shore, "Experimental verification of the synchronization condition for chaotic external cavity diode lasers," *Phys. Rev. E*, Vol. 62, No. 5, 7505–7507, 2000.
14. Carr, T. W., "Onset of instabilities in self-pulsing semiconductor lasers with delayed feedback," *Eur. Phys. J. D*, Vol. 19, No. 2, 245–255, 2002.
15. Heil, T., I. Fischer, W. Elsäβer, B. Krauskopf, K. Green, and A. Gavrielides, "Delay dynamics of semiconductor lasers with short external cavities: Bifurcation scenarios and mechanisms," *Phys. Rev. E*, Vol. 67, No. 6, 066214, 2003.

Incoherent Writing and Erasure of Cavity Solitons in an Optically Pumped Vertical-cavity Semiconductor Optical Amplifier

S. Barbay, Y. Ménesguen, X. Hachair, L. Leroy, I. Sagnes, and R. Kuszelewicz
Laboratoire de Photonique et de Nanostructures, LPN-CNRS UPR20
Route de Nozay, 91460 Marcoussis, France

Abstract— Cavity solitons appear as self-localized bright spots sitting on a dark background in a nonlinear cavity that can be written or erased by a local excitation. In this paper, we show that beyond the classical mechanism using coherent field injection, incoherent writing/erasure of cavity solitons is possible using carrier injection mechanisms in an optically pumped vertical-cavity semiconductor amplifier. Independence of two solitons with and without the presence of other solitons is demonstrated.

Cavity solitons (CSs) are self-localized spots formed in the transverse plane of a nonlinear cavity. They have been observed in various macroscopic systems, and predicted [1] and observed in microscopic, semiconductor-based systems too [2]. They are formed in a spatially extended, bistable, and modulationally unstable system driven by a coherent field, the holding beam (HB) and appear generally as bright spots sitting on a dark background. Their excitation is accomplished either spontaneously, by the noise present in the system, or by a local excitation, the writing beam (WB). CSs can be independently addressed with a control beam and can be manipulated with the aid of phase or amplitude gradients of some control parameters. These properties make them interesting objects for all-optical information processing applications, for which they can be thought of as logical bit units for parallel information processing with reconfigurable capabilities. In that respect, semiconductor materials are particularly suited thanks to the time scales (1 ns or less) and spatial scales (~ 10 mm) that are involved.

Cavity solitons result from the intricated interaction of a coherent field with a nonlinear material whose excitation is mediated by the carrier population. Classically and historically, the inscription and erasure of such inhomogeneous states was realized with the help of coherent optical pulses addressed to the nonlinear cavity. Then, when favourable parameter conditions are met, the field/carrier interplay leads to a local switching of a homogeneous state towards an inhomogeneous localized state or to its erasure. In practical conditions, this can be achieved only if the control of the relative phase between the holding beam and the writing/erasure beam is tight, with an in-phase situation for writing and a π-phase shift for erasure.

Switching can be also achieved using the other internal variable of the system i.e., the carrier excitation. The control of the excitation can be achieved via an optical injection of higher energy photons that produces hot carriers and results in an incoherent mechanism for writing and erasing a CS. This possibility reveals very interesting because it eliminates the need for a controlled phase

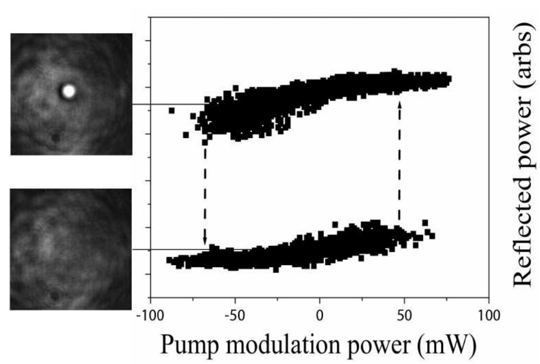


Figure 1: Hysteresis cycle recorded by ramping the pump power and monitoring the local intensity at a CS. The corresponding near field images of the sample in the upper and lower branches are shown on the left.

relation between the excitation and the driving field. Moreover, it allows a wavelength conversion process associated to the CS excitation, a property which could be potentially exploited in applications.

We have shown the incoherent and coherent writing and erasure of CSs in a broad-area, optically pumped, vertical-cavity semiconductor amplifier [3]. This result was obtained in an optically pumped microresonator because of an original cavity design [4] aimed at reducing the heat produced by the pumping. In some parameter ranges, When the holding beam wavelength is blue detuned with respect to the cavity resonance, we observed the formation of broad patterns with hexagonal or circular symmetry [5]. In other ranges, the patterns reduce to one or several self-localized spots. These spots are bistable as shown on Fig. 1 and can be repeatedly excited by a short optical pulse whose wavelength is different from the pump and holding beam wavelength [3]. We show the independant manipulation of two close spots establishing their CS nature. Surprisingly, we show that it is also possible to incoherently erase a CS (Fig. 2(b)), although this process is not fully understood yet.

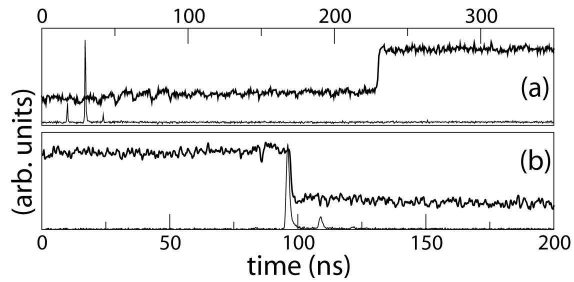


Figure 2: Local intensity at a CS (thick line) during the switch on and off of a CS (a and b resp.) induced by a 60 ps optical pulse (thin line) incoherent with the pump and holding beams.

One of the main properties of CS resides in the existence of a regime of mutual independence of two CS. The possibility of controlling two CS in an independent way using a coherent writing beam, has been shown in [2, 6]. Incoherent mechanisms should lead to the same CSs as those obtained with a coherent mechanism. This demonstration was also possible in our system [5]. Doing so, we also showed the influence of impurities in the localization an trapping of CSs. In order to realize the CS control sequence with an incoherent WB, we fixed all parameters to values such that two solitons display hysteresis cycles with overlapping PB power range. Starting with no spot, we injected the incoherent WB at a first location where a CS was generated. Then, we repeated the same operation at a different location to generate a second CS. As explained in the previous section, switch-on and

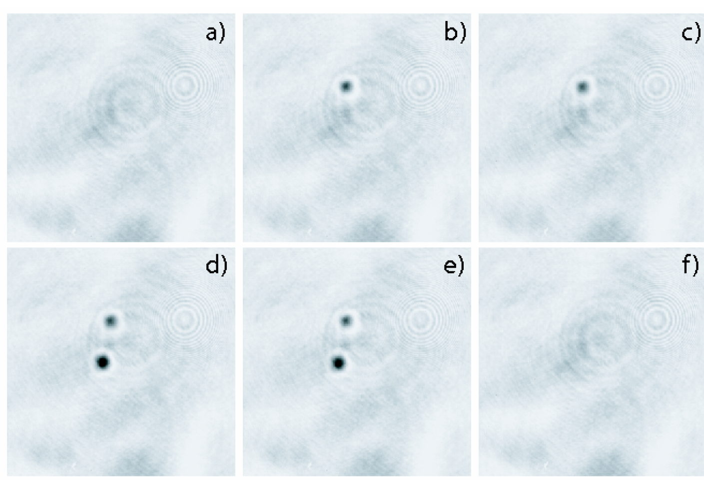


Figure 3: Inverted contrast, intensity distribution of the output field ($\lambda = 877.42$ nm). (a) The holding beam is on and the WB is blocked; (b) a 10 μm incoherent focused WB ($P_{WB} = 53$ mW) targets a point into the homogenous region; it induces the creation of one CS; (c) the WB is blocked again and the CS remains; (d) The WB is displaced and switched-on again to generate a second CS; (e) we blocked the WB and the two CS coexist; (f) after the switch off of the HB we come back to the initial situation (a) by restoring the HB.

-off of a CS with an incoherent WB cannot occur in the same parameter range. Thus, switch-off for these two CS was obtained by removing the HB for a few milliseconds, which finally cycled the system back to the homogeneous situation. The full series, showing the successive incoherent writing of two independent CS, is displayed in Fig. 3. Theory explains the trapping role of sample roughness and impurities [6, 7]. In our case, the thermal gradient forces the presence of CS in the center of the pumping region. When we chose a region on the sample where the number and the strength of impurities were high enough to compensate for the thermal gradient, we were able to obtain up to five cavity solitons (Fig. 4).

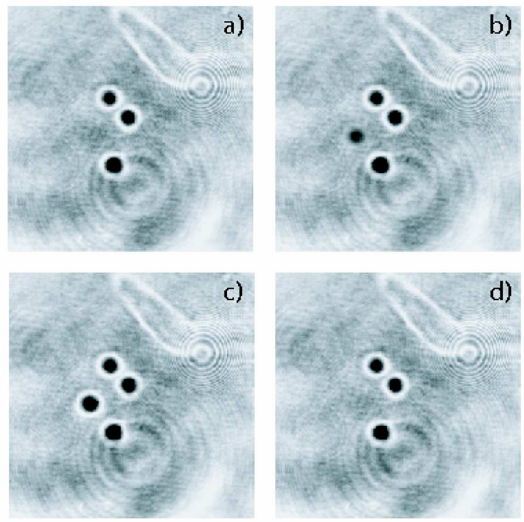


Figure 4: Inverted contrast, intensity distribution of the output field. (a) The HB is ON and the WB is blocked (presence of three CS; (b) a 10 μm incoherent focused writing beam ($P_{WB} = 53\,\mathrm{mW}$) targets a point into the homogenous region; (c) it induces the creation of a fourth CS; the WB is blocked again and the fourth CS remains; (d) after having switched off the HB, we come back to the initial situation (a) with the three CS by restoring the HB.

The positions of these CS are determined by those of the strongest impurities. In this case, CS were spontaneously created by increasing the pump beam (PB) power. Indeed, we knew that CS may be switched on by a WB as well as they could be induced by the noise present in the system. Moreover, it was not possible to control all five CS within the same range of parameters. This was due to the fact that they did not appear spontaneously for the same PB power. When we increased the PB power, those that appeared previously were expelled from the bistable range within which they are controllable.

In a subsequent experiment, we showed the ability to control a CS in the presence of three others (Fig. 4). Up to four CS were obtained as we increased the PB power. After the first one appeared, the second one did for a slightly higher PB power. This difference was sufficiently small to allow the simultaneous control of these two CS as in Fig. 3. The third CS, though, appeared for a PB power even higher, for which value the two first CS were not controllable anymore. Finally, a fourth CS appeared for a PB power which again made the third CS lose its controllability. But when we drove them back within their bistabilty range, they showed all the characteristics of a CS and in particular the possiblity to be controlled using a WB.

From a prospective viewpoint, new strategies can be considered for optical information processing with cavity solitons by the additional possibility to use writing beams free of phase control. This mechanism also introduces the convenience to realize wavelength conversion. Finally, the demonstration of the coexistence of up to five cavity solitons was performed, even though their simultaneous control was not yet achieved. This is a first step in the use of sets of independently controlled cavity solitons. An actual achievement requires the reduction of inhomogeneities and thermal gradients in the system. This can be expected by the increase of the transverse size, rejecting sources of non uniformities at the periphery of the system.

ACKNOWLEDGMENT

This work was partially supported by the European Commission through the FP6-IST-STREP 4868 Project FunFACS and by the région Ile de France.

REFERENCES

1. Brambilla, M., L. A. Lugiato, F. Prati, L. Spinelli, and W. Firth, *Phys. Rev. Lett.*, Vol. 79, 2042, 1997.
2. Barland, S., J. Tredicce, M. Brambilla, L. Lugiato, S. Balle, M. Giudici, T. Maggipinto, L. Spinelli, G. Tissoni, T. Knödel, M. Miller, and R. Jäger, *Nature*, Vol. 419, 699, 2002.
3. Barbay, S., Y. Ménesguen, X. Hachair, L. Leroy, I. Sagnes, and R. Kuszelewicz, *Opt. Lett.*, Vol. 31, 1504, 2006.
4. Barbay, S., Y. Ménesguen, I. Sagnes, and R. Kuszelewicz, *Appl. Phys. Lett.*, Vol. 86, 151119, 2005.
5. Ménesguen, Y., S. Barbay, X. Hachair, L. Leroy, I. Sagnes, and R. Kuszelewicz, to appear in *Phys. Rev. A*, 2006.
6. Hachair, X., S. Barland, L. Furfaro, M. Giudici, S. Balle, and J. Tredicce, *Phys. Rev. A*, Vol. 69, 043817, 2004.
7. Kuszelewicz, R., I. Ganne, I. Sagnes, G. Slekys, and M. Brambilla, *Phys. Rev. Lett.*, Vol. 84, 6006, 2000.

Excitability Mediated by Localized Structures in Kerr Cavities

Damià Gomila, Adrian Jacobo, Manuel A. Matías, and Pere Colet

Instituto Mediterraneo de Estudios Avanzados, IMEDEA (CSIC-UIB)
Campus Universitat Illes Balears, 07122 Palma de Mallorca, Spain

Abstract— We characterize a scenario where localized structures in nonlinear optical cavities display an oscillatory behavior which becomes unstable leading to an excitable regime. Excitability emerges from spatial dependence since the system locally is not excitable. We show the existence of different mechanisms leading to excitability depending on the profile of the pump field.

Localized structures (LS) in dissipative optical cavities arise as a consequence of the interplay between diffraction, nonlinearity, driving, and dissipation [1]. These structures, also known as cavity solitons, are unique once the parameters of the system have been fixed. This fact makes this structures potentially useful in optical storage and processing of information [2, 3]. LS may develop a number of instabilities, for instance their amplitude can oscillate in time while remaining static in space. Here we report on a novel regime of excitability associated to the existence of localized structures in a nonlinear optical system [4, 5]. Excitability has been found in a variety of systems [6], including optical systems [7], and is characterized by a nonlinear response under applied external perturbation. Perturbations exceeding a certain threshold are able to elicit in the system a full and well defined response. Furthermore after one perturbation the system cannot be excited again within a refractory period of time. Excitability is behind excitation waves in heart tissue and the existence of action potentials in neurons, and, so, may confer new computational capabilities to optical systems beyond information storage.

In this paper we show the existence of different mechanisms leading to excitability depending on the profile of the pump field. For a homogeneous pump the mechanism leading to excitable behavior is a saddle-loop bifurcation through which an stable oscillating LS collides with an unstable LS [4]. For a system pumped by a localized Gaussian beam on top of homogeneous background the scenario is richer and one finds two different mechanisms leading to excitability. One is based on a saddle-loop bifurcation as above while the other takes place through a saddle-node in the invariant circle (SNIC) bifurcation. This second mechanism has excitability threshold which can be much lower.

We consider a ring cavity filled with a nonlinear self-focusing Kerr medium pumped by an external field. In the mean field approximation, the dynamics of the electric field inside the cavity can be described by a single partial differential equation for the scaled slowly varying amplitude $E(\vec{x}, t)$ [8]

$$\partial_t E = -(1 + i\theta)E + i\nabla^2 E + E_0 + i\,|E|^2\,E, \tag{1}$$

where E_0 is the homogeneous driving field and θ is the cavity detuning. The homogeneous steady state solution of Eq. (1) is implicitly given by $E_s = E_0/(1 + i(\theta - I_s))$, where $I_s = |E_s|^2$. We use the intra-cavity background intensity I_s together with θ as convenient control parameters. The homogeneous solution has a modulational instability (MI) at $I_s = 1.0$ leading to the formation of subcritical hexagonal patterns.

The existence and dynamical properties of localized structures in this system, the so called Kerr cavity solitons (KCS), have been studied in [9] and references therein. Fig. 1 shows the region of existence of KCS in the $I_s - \theta$ parameter space. The KCS branch starts subcritically at the MI point. The middle-branch KCS is unstable and it has only a single unstable mode. There is a turning point (saddle-node) where an stable upper-branch KCS starts. The bifurcation diagram as function of I_s for fixed θ is shown in Fig. 1. A similar diagram can be obtained for fixed I_s as function of θ. Increasing I_s or θ the upper-branch KCS undergoes a Hopf bifurcation resulting in a periodically oscillating localized structure. The oscillation is such that it approaches the stable manifold of the middle-branch KCS and then escapes along the unstable manifold. As the detuning is increased the limit cycle goes closer and closer to the middle-branch KCS as illustrated in Fig. 2. At a critical value θ_c a global bifurcation takes place: the cycle touches the middle-branch KCS and becomes a homoclinic orbit [Fig. 2(c)]. This is an infinite-period bifurcation called saddle-loop or homoclinic bifurcation. The saddle-loop has a characteristic *scaling law* that govern the period

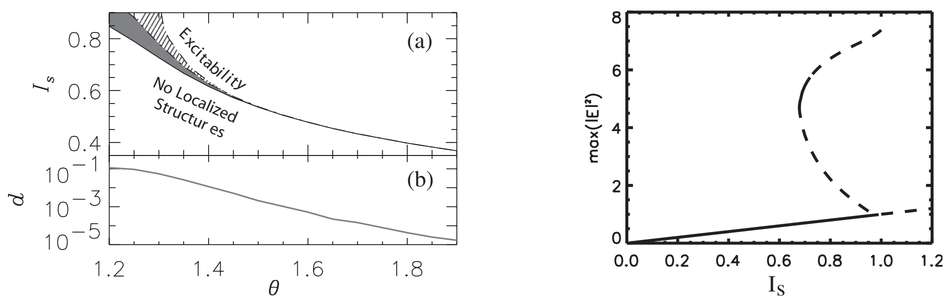


Figure 1: Left: (a) Phase diagram: I_s vs. θ showing the different regimes. LS are stable in the shaded region and oscillate in the dashed one. The solid line indicates the saddle-node bifurcation where the LS are created, the dot-dashed the Hopf, and the dashed the saddle-loop where the oscillation is destroyed. (b) Distance between the saddle-node and Hopf lines. Right: Bifurcation diagram of KCS for $\theta = 1.34$. The upper (middle) branch corresponds to stable (unstable) KCS. These branches are originated at a saddle-node bifurcation. The upper branch becomes Hopf unstable for larger values of I_s.

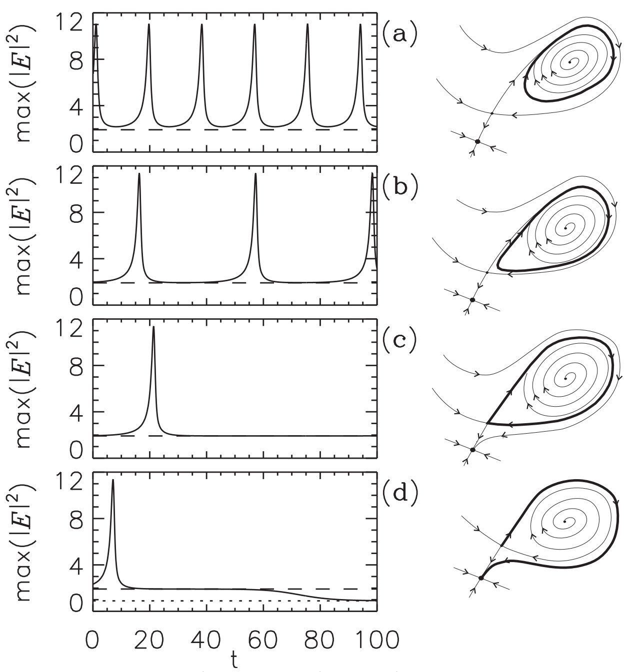


Figure 2: Left: LS maximum intensity as a function of time for increasing values of the detuning parameter θ. From top to bottom $\theta = 1.3$, 1.3047, 1.30478592, 1.304788. $I_s = 0.9$. Right: Sketch of the phase space for each parameter value. The thick line shows the trajectory of the LS in phase space.

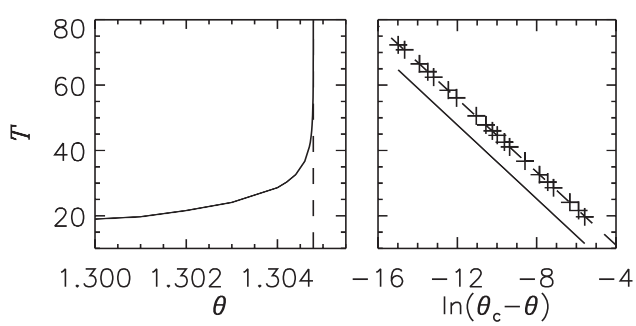


Figure 3: (a) Period of the limit cycle T as a function of the detuning θ for $I_s = 0.9$. The vertical dashed line indicate the threshold of the saddle-loop bifurcation $\theta_c = 1.30478592$. (b) Period T as a function of $\ln(\theta_c - \theta)$. Crosses correspond to numerical simulations while the solid line has a slope $1/\lambda_1$ with $\lambda_1 = 0.177$ obtained from the linear stability analysis of the middle-branch LS.

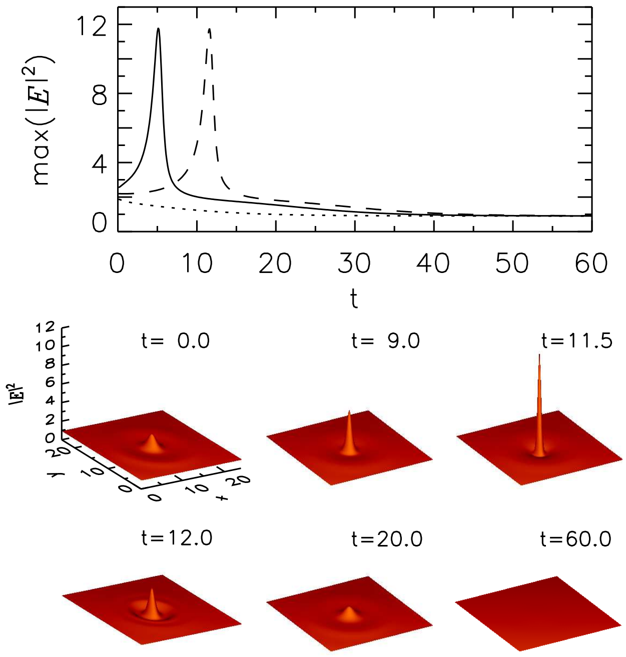


Figure 4: Evolution of a sub-threshold (dotted line) and two above-threshold perturbations, one close to the threshold (solid line) and one well above (dashed line). Top panel shows the time evolution of the maximum intensity, while the 3D plots show the transverse profile at different times for the dashed trajectory.

T of the limit cycle as the bifurcation is approached: $T \sim -\ln(\theta_c - \theta)/\lambda$, where λ is the unstable eigenvalue of the middle-branch KCS. We show that this scaling is verified in our system (Fig. 3).

Beyond the saddle-loop bifurcation the phase space shows a typical configuration presenting excitability [Fig. 2(d)]: it has a globally attracting fixed point (homogeneous solution), but localized disturbances (above the middle-branch KCS) can send the system on a long excursion in the phase space before returning to the fixed point (Fig. 4). The peak grows to a large value until the losses dominate the dynamics and stop the growth. Then decays exponentially until it disappears. A remnant wave is emitted out of the center dissipating the remaining energy. The long excursion in phase space is reminiscent of the coherence collapse phenomenon that arises in the 2D nonlinear Schrödinger equation. The Hopf instability of KCS in this limit has been studied in [10]

All this scenario is organized by a co-dimension two Takens-Bogdanov bifurcation point. In the limit of large detuning, the saddle-node, Hopf and saddle-loop bifurcation lines meet asymptotically, at $I_s = 0$ as shown in Fig. 1. It is known that the intersection of a saddle-node line with a Hopf line is a Takens-Bogdanov (TB) codimension-2 bifurcation point [11]. The unfolding around a TB point leads to a saddle-loop bifurcation line [11]. So, this unfolding fully explains the observed scenario, where our formally infinite-dimensional system appears to be perfectly described by a dynamical system in the plane.

We should emphasize that, in absence of spatial degrees of freedom, the system described by Eq. (1) is not an excitable system. Excitability arises here as an emergent property linked to the spatial degrees of freedom in the system. The mechanism for excitability is based on the dynamics of the localized structure by itself.

We consider in the following the case in which the system is pumped by a field E_I consisting of a narrow Gaussian beam on top of homogeneous background E_0: $E_I = E_0 + H\exp(-r^2/r_0^2)$ where $H = \sqrt{(I_s + I_{sh})[1 + (\theta - I_s - I_{sh})^2]} - E_0$. The control parameters are I_s associated to the background intensity, the detuning θ, and the new I_{sh} associated to Gaussian beam. Now the translational symmetry of the system (and also of its solutions) is broken, and the fundamental solution, no longer homogeneous, exhibits a small bump (small when compared to the truly localized structures), which is the system response to the Gaussian perturbation. The bifurcation diagram for fixed I_{sh} and θ is shown in Fig. 5.

The maximum of the fundamental solution increases as we increase I_s until a saddle-node bifurcation is encountered. This bifurcation, which is in fact a saddle-node on the invariant circle (SNIC) bifurcation is not present in the case of a homogeneous pump (cf. Fig. 1) for which the lower

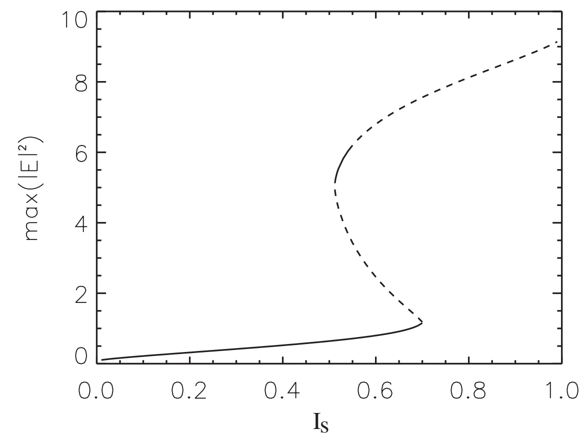


Figure 5: Bifurcation diagram $\max(|E|^2)$ vs I_s for pump consisting of a Gaussian beam on top of a homogeneous background ($I_{sh} = 0.7$, $\theta = 1.34$). Solid lines represent stable solutions and dashed lines unstable ones.

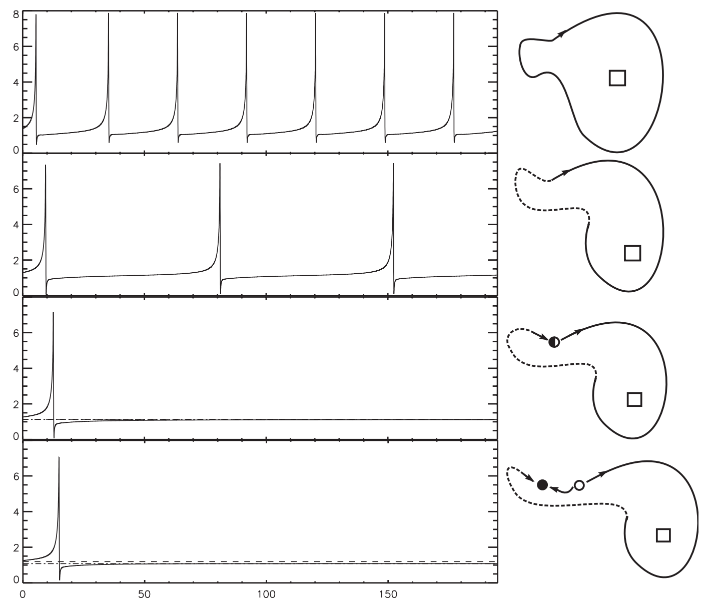


Figure 6: Left: Ls maximum intensity as a function of time for decreasing values of I_{sh}. From top to bottom, $I_{sh} = 0.545$, 0.475, 0.4471, 0.435. $I_s = 0.8050$, $\theta = 1.34$. Right: sketch of the phase space for each parameter value.

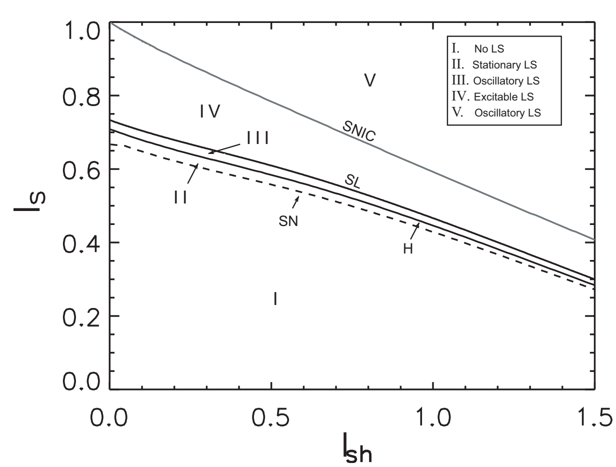


Figure 7: Phase diagram I_s vs. I_{sh} for $\theta = 1.34$.

branch is stable until $I_s = 1$. At the SNIC a middle branch unstable LS starts. As for homogeneous pump, there is a turning point (associated to the left saddle-node) where an stable upper-branch LS starts. Increasing I_s the upper-branch LS undergoes a Hopf bifurcation resulting in a periodically oscillating localized structure. As for homogeneous pump this oscillatory structure is destroyed in a saddle-loop bifurcation leading to excitable behavior. The saddle-loop takes place at a value of I_s below the SNIC. The excitable regime is possible only while the fundamental solution exists (I_s between the saddle-loop and the SNIC). After the SNIC both the fundamental and the upper branch LS are unstable, and a new, oscillatory, regime appears. For $I_s > 1$ the background becomes modulationally unstable leading to pattern formation. Fig. 6 shows the temporal evolution in the new oscillatory regime. As I_s is decreased the period of the oscillations becomes longer and it finally diverges when reaching the SNIC. The scaling of the period of oscillation confirms a saddle-node on the invariant circle bifurcation. A SNIC bifurcation induces excitable behavior for I_s below the critical value. This scenario is different from the previous one although the excitable regimes are similar. Here the excitable threshold can be controlled by the intensity of the addressing Gaussian beam that effectively approaches the fixed point and the saddle in the phase space.

Figure 7 shows the phase diagram as a function of I_s and I_{sh} with θ constant. $I_{sh} = 0$ corresponds to homogeneous pump. For small I_s the system has only one fixed point which is the fundamental solution (reg. I). Increasing I_s the occurrence of a saddle-node bifurcation leads to a stable and an unstable (saddle) branches of stationary LS solutions (reg. II). Further on, the stable branch becomes unstable in a Andronov-Hopf bifurcation and a cycle (oscillating LS) is created (reg. III). At this point a stable fixed point, a cycle and an unstable fixed point coexist in the system. If we further increase I_s the limit cycle approaches the saddle and collides in a saddle-loop bifurcation (SL line). Beyond this saddle-loop the fundamental solution becomes excitable in two possible ways (reg. IV). If the line indicated as SNIC is crossed, the fundamental solution (stable) and the lower LS stationary solution (saddle) annihilate inside an invariant circle, leading to oscillatory LS behavior (reg. V). Region IV is excitable in the sense that suitable perturbations to the fundamental solution lead to long excursions in phase space, in two possible ways, depending on whether the system is close to the SL or to the SNIC bifurcation lines.

ACKNOWLEDGMENT

We thank G. Orriols and D. Pazó for useful discussions. We acknowledge financial support from MEC (Spain) and FEDER: Grants BFM2001-0341-C02-02, FIS2004-00953 (CONOCE2) and FIS2004-05073-C04-03. AJ acknowledges financial support from the MEC.

REFERENCES

1. Feature section on cavity solitons, L. A. Lugiato ed., *IEEE J. Quant. Elect.*, Vol. 39, #2, 2003; Tlidi, M., P. Mandel, and R. Lefever, *Phys. Rev. Lett.*, Vol. 73, 640, 1994; Schapers, B., et al., ibid., Vol. 85, 748, 2000.
2. Barland, S., et al., *Nature*, Vol. 419, 699, 2002.
3. Firth, W. J. and A. J. Scroggie, *Phys. Rev. Lett.*, Vol. 76, 1623, 1996; Firth, W. J. and C. O. Weiss, *Opt. Photon. News*, Vol. 13, 55, 2002.
4. Gomila, D., M. A. Matias, and P. Colet, *Phys. Rev. Lett.*, Vol. 94, 063905, 2005.
5. Gomila, D., M. A. Matias, A. Jacobo, and P. Colet, "Phase space structure of 2D excitable localized structures," submitted for publication.
6. Meron, E., *Phys. Rep.*, Vol. 218, 1, 1992; Murray, J. D., *Mathematical Biology*, 3rd Ed., Springer, 2002.
7. Plaza, F., et al., *Europhys. Lett.*, Vol. 38, 85, 1997; Dubbeldam, J. L. A., B. Krauskopf, and D. Lenstra, *Phys. Rev. E*, Vol. 60, 6580, 1999; Wunsche, H. J., et al., *Phys. Rev. Lett.*, Vol. 88, 023901, 2002; Barland, S., et al., *Phys. Rev. E*, Vol. 68, 036209, 2003.
8. Lugiato, L. A. and R. Lefever, *Phys. Rev. Lett.*, Vol. 58, 2209, 1987.
9. Firth, W. J., et al., *J. Opt. Soc. Am. B*, Vol. 19, 747, 2002.
10. Skryabin, D., *J. Opt. Soc. Am. B*, Vol. 19, 529, 2002.
11. Guckenheimer, J. and P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer, New York, 1983.

Dissipative Structures in Metamaterial Optical Resonators

P. Tassin[1], **P. Kockaert**[2], **N. Veretenov**[3], **G. Van der Sande**[1,3],
I. Veretennicoff[1], **and M. Tlidi**[3]

[1]Department of Applied Physics and Photonics, Vrije Universiteit Brussel
Pleinlaan 2, B-1050 Brussel, Belgium
[2]Optique et acoustique, Université Libre de Bruxelles
CP 194/5, 50, Av. F. D. Roosevelt, B-1050 Bruxelles, Belgium
[3]Optique non linéaire théorique, Université Libre de Bruxelles
CP 231, Campus Plaine, B-1050 Bruxelles, Belgium

Abstract— We consider a Kerr optical resonator containing a left-handed and a traditional right-handed material. This configuration allows for negative diffraction. We analyze the dynamical behavior of the system in this negative diffraction regime when the system has one, two and three transverse dimensions. The 3D case corresponds to spatiotemporal dynamics. We identify a modulational instability affecting the lower branch of the hysteresis cycle and observe dissipative patterns and an up-switching process.

1. INTRODUCTION

Today, left-handed materials attract a lot of attention not only from the electromagnetics, but also from the optics community. This is due to the fact that more and more left-handed materials are fabricated at optical frequencies [1]. These materials have the interesting property that their refractive index is negative [2]. Only a few works have considered nonlinearities in these materials. Recent works addressed the problem of light propagation in Kerr type nonlinear left-handed materials [3, 4]. It has been shown that nonlinear propagation in left-handed materials can be described by the nonlinear Schrödinger equation, with the difference that some of its coefficients have their sign reversed. Due to the similarity between the propagation equations in left- and right-handed materials, the evolution of an optical field in both materials is qualitatively identical. We propose a system in which both right-handed and left-handed materials are combined. In this way, it is possible to observe new physics, as will be reported below.

In order to enhance the nonlinear effects in our system, we use an optical resonator (see Fig. 1) whose behavior can be described by a modified Lugiato-Lefever (LL) model [5, 6]. The resonator is made of two layers containing a left-handed and a right-handed material. It is driven by a continuous wave homogeneous input field at optical frequency ω. The LL-model reveals the existence of homogeneous steady state solutions. Below, we address the stability of these solutions and their relation with the number of transverse dimensions. The 1D and 2D cases correspond respectively to dispersionless cavities with one and two spatial dimensions orthogonal to the propagation direction, while the 3D case includes dispersion effects, leading to spatiotemporal dynamics (self-organized light bullets).

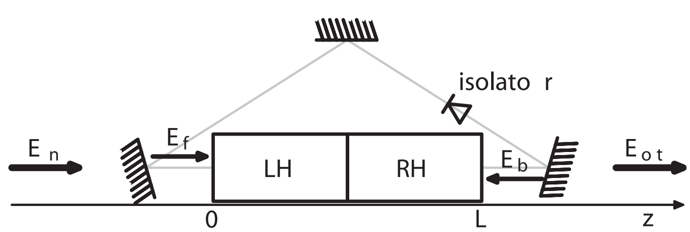


Figure 1: The setup: a driven ring resonator filled with left- (LHM) and right-handed (RHM) materials.

2. MODEL

We derived a mean-field model for this resonator based on the following three assumptions: (1) reflections at the interfaces of the layers can be neglected; (2) the dissipative Fresnel number is large; (3) the cavity is shorter than the diffraction, dispersion and nonlinearity lengths. The first of these approximations needs some more explanation as such reflections and the associated counterpropagating beam could alter the dynamics considerably. Having two independent material parameters (ϵ and μ), the structure of the LHM can be tuned to have an impedance $\eta = \sqrt{\epsilon/\mu}$ equal to that of the RHM, i.e., the RHM and the LHM are impedance matched. Note that this, in general, does not mean that the LHM will have equal but opposite index. In another possible realization, an optical isolator can be used to inhibit the counterpropagating signal. A slow dimensionless time scale T is introduced, where $T = \pi t/(F t_r)$, with t_r the cavity roundtrip time and F its finesse. The slow time evolution of the envelope $A(x, y, \tau, T)$ is then governed by

$$\frac{\partial A}{\partial T} = -(1 + \mathrm{i}\Delta) A + E + \mathrm{i}\Gamma |A|^2 A + \mathrm{i}\left[D\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + \beta \frac{\partial^2}{\partial \tau^2} \right] A. \tag{1}$$

The detuning Δ is related to the linear phase accumulated by the light during one roundtrip, $\Phi = \pi\Delta/F$. The input field amplitude has been scaled by $E_{\mathrm{in}} = \pi E/F$ and the coefficient of the nonlinear term describes the combined effect of the Kerr nonlinearities in both layers: $\Gamma = 3F(\gamma_R\, l_R + \gamma_L l_L)/\pi$. A self-focusing nonlinearity ($\Gamma > 0$) is assumed in the remainder of this text. Finally, the diffraction and dispersion coefficients are given by

$$D = \frac{F}{\pi}\left(\frac{l_R}{2k_R} + \frac{l_L}{2k_L} \right), \quad \beta = \frac{F}{\pi}\left(\beta_R l_R + \beta_L l_L \right). \tag{2}$$

The regime that we are interested in has a negative diffraction coefficient D. It is indeed possible to engineer the strength of the diffraction coefficient due to the balance of positive diffraction in the RHM and negative diffraction in the LHM. When $l_L > -l_R k_L/k_R$, it is even possible to change the sign of D.

The homogeneous steady state (HSS) solutions A_s of Eq. (1) are $E = [1 + \mathrm{i}(\Delta - \Gamma |A_s|^2)] A_s$ (Fig. 2). $|A_s|^2$ as a function of $|E|^2$ is single-valued for $\Delta < \sqrt{3}$ and multiple-valued for $\Delta > \sqrt{3}$. We have performed the stability analysis of these HSS. With periodic boundaries, the deviation from the steady state is taken proportional to $\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r} + \lambda T)$ with $\mathbf{k} = (k_x, k_y, k_\tau)$ and $\mathbf{r} = (x, y, \tau)$. The threshold associated with modulational instability is $|A_c| = 1/\sqrt{\Gamma}$ and $E_c = \sqrt{\Delta^2 - 2\Delta + 2}$. At that bifurcation point, the HSS becomes unstable with respect to modes satisfying the relation, $D(k_x^2 + k_y^2) + \beta k_\tau^2 = 2 - \Delta$.

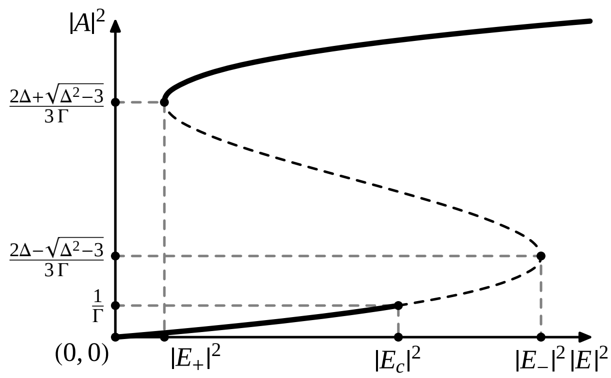


Figure 2: Bistability of the optical resonators. Stable homogeneous steady state solution are plotted with a full blue line, unstable solutions are plotted with a dashed red line.

3. RESULTS

When the driving field amplitude is in the unstable portion of the lower branch, an infinite number of linearly unstable Fourier modes exist, which trigger the spontaneous evolution of the intracavity field amplitude towards a spatially periodic distribution that occupies the whole cavity space available in the transverse directions. We study the formation of such patterns by the numerical simulation of Eq. (1). The initial condition consists of some small random noise added to the HSS and periodic boundary conditions are assumed.

An example of a dissipative structure in the one-dimensional system is shown in Fig. 3. We compare the wavelength given by the linear stability analysis ($\Lambda_c = 3.14$) with the wavelength

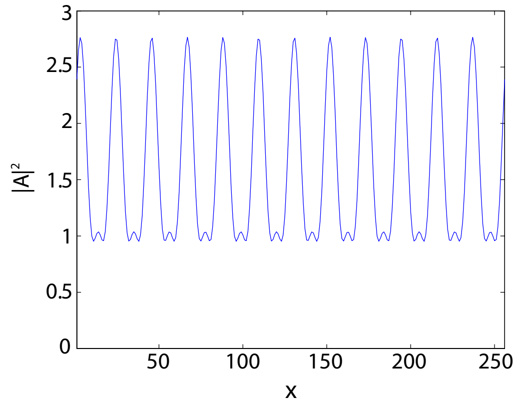


Figure 3: Stable dissipative structure emerging in the one-dimensional case. The intensity is plotted. Parameters are $\Delta = 10$, $\Delta = -2$, $\Gamma = 1$ and $E = 9.25$.

of the stable dissipative structure obtained numerically ($\Lambda_{\text{num}} = 3.23$). A very good agreement between the two wavelengths is obtained.

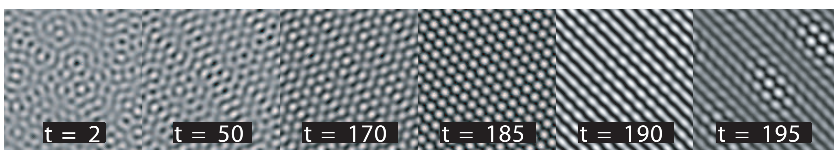


Figure 4: Dynamics of the pattern formation in the two-dimensional case. The patterns change from the initial state to stripes, hexagons and self-organized spots. The parameters are the same as those in Fig. 3. Maxima are plain white and the integration mesh is 256×256.

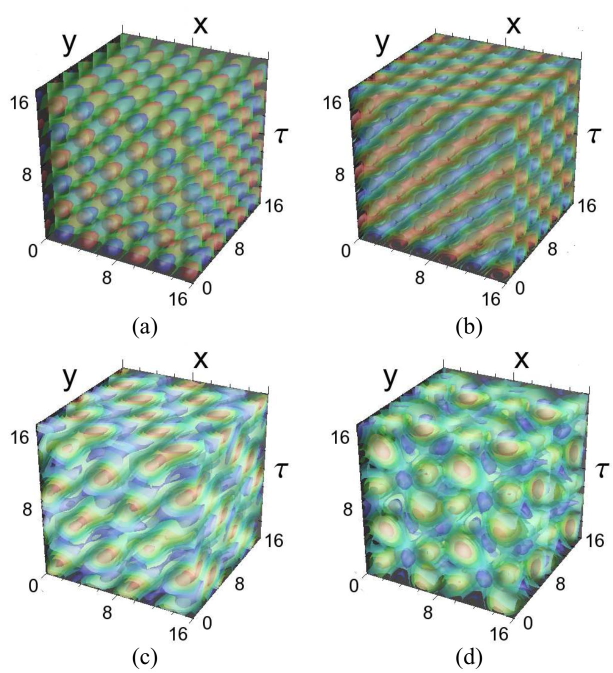


Figure 5: 3D structures observed in the resonator. Parameters are the same as in Fig. 3.

We perform similar simulations in a two-dimensional system using periodic boundary conditions. The following typical sequence of patterns is obtained: hexagons, stripes, and self-organized spots. This behavior is shown in Fig. 4, where we have plotted the real part of the intracavity field during the spatiotemporal evolution. We have observed that the average intensity grows exponentially during this process, which is clearly due to the amplification of unstable wavelengths, as can be expected from the linear stability analysis. Finally, the field amplitude reaches the high HSS. In the range $E_c < E < E_-$, the up-switching process dominates over the pattern forming process and therefore leads to the truncation of the homogeneous hysteresis cycle.

Similar behavior occurs in three dimensions [7], as shown in Fig. 5.

4. CONCLUSION AND PERSPECTIVES

We analyzed the dynamical behavior of the intracavity field of an optical resonator designed to have negative diffraction. We identified modulational instability appearing on the lower branch of the bistability curve of the homogeneous steady-state solutions. The pattern forming process was investigated numerically in one, two, and three transverse dimensions. In the 1D case, we found stable dissipative patterns, whereas an up-switching process inhibits the formation of stable light patterns in two and three dimensions. By tuning the diffraction coefficient to lower values, it is in principle possible to make localized structures as small as we want.

The understanding of modulational instability is of high importance for the study of localized structures (dissipative solitons), as it is well known that MI is a prerequisite for their formation. Another perspective of this work is the extension to other kinds of nonlinearities, like quadratic ones. Finally, we want to devise how the LL-model should be adapted when diffraction is very close to zero. This will introduce a limit on the size of the patterns that can effectively be formed in our setup.

ACKNOWLEDGMENT

This work was financially supported by the Belgian Science Policy Office under grant IAP-V18 (The Belgian Photon Network) and by the Research Council (OZR) of the Vrije Universiteit Brussel. M. T. is a Research Associate of the Fonds National de la Recherche Scientifique (FNRS). P. T. is a Research Assistant of the Research Foundation - Flanders (FWO - Vlaanderen).

REFERENCES

1. Grigorenko, A. N., A. K. Geim, H. F. Gleeson, Y. Zhang, A. A. Firsov, I. Y. Khrushchev, and J. Petrovic, "Nanofabricated media with negative permeability at visible frequencies," *Nature*, Vol. 438, 335–338, 2006.
2. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of ε and μ," *Sov. Phys. Usp.*, Vol. 10, 509–514, 1968.
3. Scalora, M., M. S. Syrchin, N. Akozbek, E. Y. Poliakov, G. D'Aguanno, N. Mattiucci, M. J. Bloemer, and A. M. Zheltikov, "Generalized nonlinear Schrödinger equation for dispersive susceptibility and permeability," *Phys. Rev. Lett.*, Vol. 95, 013902, 2005.
4. Kockaert, P., P. Tassin, G. van der Sande, I. Veretennicoff, and M. Tlidi, "Negative diffraction pattern dynamics in nonlinear cavities with left-handed materials," *Phys. Rev. A*, Vol. 74, 033822, 2006.
5. Lugiatio, L. A. and R. Lefever, "Spatial dissipative structures in passive optical systems," *Phys. Rev. Lett.*, Vol. 58, 2209–2212, 1987.
6. Tlidi, M., M. Haelterman, and P. Mandel, "Three-dimensional structures in diffractive and dispersive nonlinear ring cavities," *Europhys. Lett.*, Vol. 42, 505–509, 1998.
7. Tassin, P., G. van der Sande, N. Veretenov, P. Kockaert, I. Veretennicoff, and M. Tlidi, "Three-dimensional structures in nonlinear cavities containing left-handed materials," *Opt. Express*, Vol. 14, 9338–9343, 2006.

Stability of Dissipative Solitons as Solutions of Asymmetrical Complex Cubic-quintic Ginzburg-Landau Equation

V. Skarka[1], **N. B. Aleksić**[2], **D. Gauthier**[1], **and D. V. Timotijević**[2]

[1]Laboratoire POMA, UMR 6136 CNRS, Université d'Angers
2, boulevard Lavoisier, 49045 Angers, Cedex 1, France
[2]Institute of Physics, Pregrevica 118, 11000 Belgrade, Serbia

Abstract— The existence and nonlinear dynamics of optical dissipative spatial solitons asymmetrical with respect to transverse coordinates are studied. Steady state solutions of asymmetrical complex cubic-quintic Ginzburg-Landau equation are computed using the variational approach extended in order to treat asymmetrical input. Using established stability criterion a domain of dissipative parameters for stable solitonic solutions is fixed. Analytical predictions are confirmed by numerically simulated evolution of an asymmetrical input pulse towards a stable and robust dissipative soliton.

There is a growing interest for optical solitons as form preserving self-trapped structures. Spatial and spatiotemporal solitons are good candidates in all-optical signal processing since they are self-guided in bulk media [1, 2]. Stable operation of laser systems, closely related to the issue of dissipative soliton stability, is crucial for generating ultra-short pulses [3, 4].

In order to generate few-parameters family of two transverse dimensions solitons the diffraction has to be compensated by self-focusing [5]. Taken into account that real systems are generally dissipative, in order to conserve solitons an appropriate gain has to match linear and nonlinear losses. Another prerequisite for practical applications is to consider evolution starting from input asymmetrical with respect to transverse coordinates x and y [5]. Therefore, a corresponding equation, $(2 + 1)$-dimensional complex cubic-quintic Ginzburg-Landau equation (CQGLE) [6] has to contain a Laplacian with respect to x and y describing beam diffraction

$$i\frac{\partial E}{\partial z} + \frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + |E|^2 E + \nu|E|^4 E = \mathbb{Q}. \tag{1}$$

E is the normalized complex envelope of the optical field. In order to prevent the wave collapse the saturating nonlinearity is required. As a consequence, cubic and quintic nonlinearities have to have opposite signs, i.e., parameter ν is negative. Dissipative terms are denoted by $\mathbb{Q}$

$$\mathbb{Q} = i\left\{\delta_o E + \varepsilon_o|E|^2 E + \mu_o|E|^4 E + \beta_o\left(\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2}\right)\right\}. \tag{2}$$

The first term with parameter δ_o is linear loss. The cubic and quintic gain-loss terms contain respectively parameters ε_o and μ_o. The last term accounts for the parabolic gain if $\beta_o > 0$. The generation of dissipative solitons demands as prerequisite a simultaneous balance of diffraction with self-focusing as well as gain with loss. Therefore, for a given set of parameters the family of solutions reduces to a fixed solution. There are no exact analytical solutions of CQGLE except for particular sets of parameters [7]; one has to resort to computer simulations. In a recent paper [8] we extended the variational approach to complex dissipative systems described by symmetrical CQGLE. Based on this variational approach a general stability criterion for dissipative one-, two-, and three-dimensional solitons is established.

In this paper we extend the variational approach to dissipative systems described by asymmetrical $(2+1)$-dimensional CQGLE (ACQGLE) (see Eq. (1)). The set of Euler-Lagrange equations is solved in order to obtain steady state solutions. In order to study the stability of such solutions we extended our stability criterion to asymmetrical conditions. Asymmetrical input pulses generated in the proposed domain of dissipative parameters, evolve towards stable dissipative spatial solitons, as numerical simulations of ACQGLE confirm.

In order to establish the variational approach for ACQGLE we construct the total Lagrangian $\mathbb{L} = \mathbb{L}_c + \mathbb{L}_Q$ of the system described by Eq. (1) containing both, a conservative part

$$\mathbb{L}_c = \frac{i}{2}(E\frac{\partial E^*}{\partial z} - E^*\frac{\partial E}{\partial z}) + \left|\frac{\partial E}{\partial x}\right|^2 + \left|\frac{\partial E}{\partial y}\right|^2 - \frac{|E|^4}{2} - \frac{\nu|E|^6}{3} \tag{3}$$

and a dissipative one [8]

$$\mathbb{L}_Q = i\left(\delta_o\,|E|^2 + \frac{\varepsilon_o|E|^4}{2} + \frac{\mu_o|E|^6}{3} - \beta_o\left(\left|\frac{\partial E}{\partial x}\right|^2 + \left|\frac{\partial E}{\partial y}\right|^2\right)\right). \tag{4}$$

Then, an asymmetrical trial function is proposed

$$E = A\exp\left[-\frac{x^2}{2X^2} - \frac{y^2}{2Y^2} + i\left(Cx^2 + Dy^2 + \Phi\right)\right], \tag{5}$$

given as a functional of amplitude A, spatial widths X and Y, wave front curvatures C and D, and phase Φ. Following Kantorovitch, constant parameters of the Rayleigh-Ritz method are substituted here by functions η=A, X, Y, C, D, Φ of an independent variable z [8]. Optimization of each of these functions gives one of six Euler-Lagrange equations averaged, together with conservative Lagrangian $L_c = \int\int \mathbb{L}_c dxdy$, over transverse coordinates x and y

$$\frac{d}{dz}\left(\frac{\partial L_c}{\partial \eta'}\right) - \frac{\partial L_c}{\partial \eta} = 2\Re e\int\int dxdy\mathbb{Q}\frac{\partial E^*}{\partial \eta} \tag{6}$$

where $\Re e$ denotes the real part [8].

Within variational approximation, to the partial differential ACQGLE corresponds a set of six coupled first order differential equations (ODEs) resulting from the variation with respect to the amplitude

$$\frac{dA}{dz} = -\delta A + \frac{3\varepsilon A^3}{4} + \frac{5\mu A^5}{9} - \beta A\left(\frac{1}{X^2} + \frac{1}{Y^2}\right) - 2A(C+D) \tag{7}$$

widths

$$\frac{dX}{dz} = 4CX - \frac{\varepsilon A^2 X}{4} - \frac{2\mu A^4 X}{9} + \frac{\beta}{X} - 4\beta C^2 X^3, \tag{8}$$

$$\frac{dY}{dz} = 4DY - \frac{\varepsilon A^2 Y}{4} - \frac{2\mu A^4 Y}{9} + \frac{\beta}{Y} - 4\beta D^2 Y^3, \tag{9}$$

wave front curvatures

$$\frac{dC}{dz} = \frac{1}{X^4} - \frac{A^2}{X^2}\left(\frac{1}{4} + \frac{2\nu A^2}{9}\right) - 4C\left(C + \frac{\beta}{X^2}\right), \tag{10}$$

$$\frac{dD}{dz} = \frac{1}{Y^4} - \frac{A^2}{Y^2}\left(\frac{1}{4} + \frac{2\nu A^2}{9}\right) - 4D\left(D + \frac{\beta}{Y^2}\right), \tag{11}$$

and phase

$$\frac{d\Phi}{dz} = 2\beta(C+D) - \frac{1}{X^2} - \frac{1}{Y^2} + A^2\left(\frac{3}{4} + \frac{5\nu}{9}A^2\right). \tag{12}$$

In order to have a stable pulse background, the linear dissipation term has to correspond to loss, i.e., the parameter δ_o must be always negative [8]. All remaining dissipative parameters in Eqs. (7-12) are divided by $|\delta_o| = \delta$: $\varepsilon = \varepsilon_o/\delta$, $\mu = \mu_o/\delta$, and $\beta = \beta_o/\delta$.

The exact steady state solutions are obtained from Eqs. (7–11) for derivatives of amplitude, widths, and curvatures equal zero. The only possible steady state solutions are symmetrical with equal widths $X = Y$ and curvatures $C = D$. In the dissipative case the power $P = A^2XY$ is no longer a constant [5, 8]. However, in steady state the power $P = 9(2\nu A^2 + 2.25)^{-1}$, the width $X = 3A^{-1}(2\nu A^2 + 2.25)^{-\frac{1}{2}}$, and the propagation constant $\Omega = \frac{d\Phi}{dz} = A^2(0.11A^2 + 0.25)$ depend only on the amplitude as in the conservative case [5]. Variationally obtained family of conservative steady state solutions reduces to a fixed double solution for a given set of dissipative parameters. Indeed, the steady state amplitude has two discrete values A^+ and A^-

$$A^{\pm} = 1.125\sqrt{\frac{(\beta - 2\varepsilon) \pm \sqrt{(\beta - 2\varepsilon)^2 + 7.11(3\mu - 2\beta\nu)}}{(3\mu - 2\beta\nu)}}. \tag{13}$$

The existence of a double solution ($A^- > A^+$) in the ε,μ-domain in Fig. 1, implies a cubic gain $\varepsilon > 0$ and a quintic loss $\mu < 0$. The striking difference with conservative systems is the nonzero wave front curvature $C = 1.125A^2[1.125(\varepsilon-\beta)+A^2(\mu-\nu\beta)]$ [5]. The gain-loss balance together with the compensation of diffraction by saturating nonlinearity can be realized only for nonzero curvature fixed steady state solutions. This curvature is negative in the considered domain. Therefore, the self-focusing is dominating the diffraction which would provoke raising of amplitude without presence of loss and gain terms. As a consequence, a cross compensation appears involving together self-focusing, loss and gain.

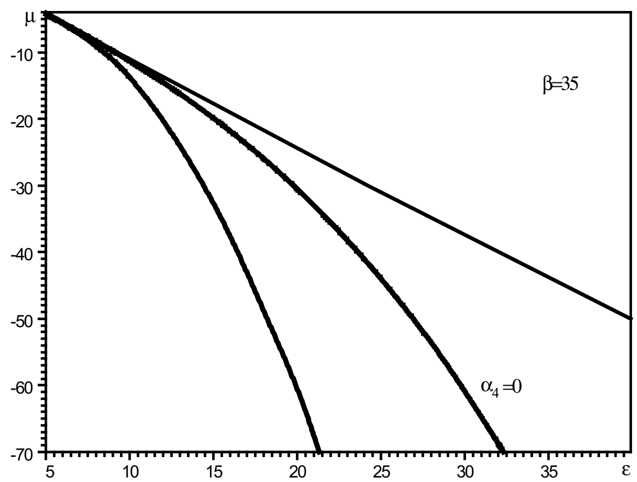


Figure 1: Domain of stable steady state solutions A^-.

Only stable steady state solutions can be solitons. In order to check the stability of solutions our stability criterion based on the method of Lyapunov's exponents is extended [8]. Steady state solutions are stable if and only if Hurwitz's conditions on coefficients α_1, α_2, α_3, α_4, and α_5 of equation having solutions with negative real part, are fulfilled. The stability criterion for variationally obtained steady state solutions of (2+1)-dimensional ACQGLE is expressed as follows,

$$\alpha_1 = \left(\varepsilon A^2 - 0.44\mu A^4 - 4\right)\delta > \frac{\alpha_3}{\alpha_2}, \tag{14}$$

$$\alpha_2 = -2\nu A^6(0.025\nu A^2 + 0.028) > 0, \tag{15}$$

$$\alpha_3 = A^4(0.125 + \nu A^2)^2((2\varepsilon - 0.198\beta)A^2 - 1.58)\delta > 0, \tag{16}$$

$$\alpha_4 = \left(0.5\,(\varepsilon+\beta) + 0.44\,(\mu+\nu\beta)\,A^2\right)\delta > 0, \tag{17}$$

and $\alpha_5 = 0$. As a consequence, only solution A^- is stable in the domain between the parabola and the line $\alpha_4 = 0$ in Fig. 1. Input pulse chosen in the stable domain is not yet a stable soliton since the variationally obtained v-curve in Fig. 2, corresponding to the power P as a function of the parameter ε, is only a good approximation of numerically obtained exact n-curve. Following Prigogin's theory of dissipative structures and self-organization the curve in Fig. 2 can be interpreted as a bifurcation curve with the control parameter ε [9]. For illustration, for parameters $\delta = 0.001$, $\mu = -70$, $\beta = 35$, and $\varepsilon = 44.1$, amplitudes A^+ and A^- are denoted respectively by a triangle in Fig. 2. The triangle corresponding to A^+ is on the lower unstable branch of the v-curve. The triangle associated to A^- is on the upper stable branch. If this solution is taken as the input in numerical simulations it will evolve towards the stable dissipative soliton represented by the triangle on the exact n-curve. However, the same final issue is obtained starting from an asymmetrical input pulse with the set of dissipative parameters belonging to the stability domain. Detailed analysis of numerical contributions of terms in ACQGLE confirmed the cross-compensation as the mechanism of stabilization. Indeed, the self-focusing excess resulting from the sum of all real terms on the left hand side of ACQGLE (Eq. (1)) is compensated by the real part of parabolic terms on the right hand side (Eq. (2)). The cross-compensation is achieved by matching of all imaginary terms from the right hand side with the imaginary part of diffraction term.

In conclusion, the asymmetrical $(2+1)$-dimensional CQGLE is considered using jointly numerical and analytical approach based on the extension of the variational method. Stability of obtained

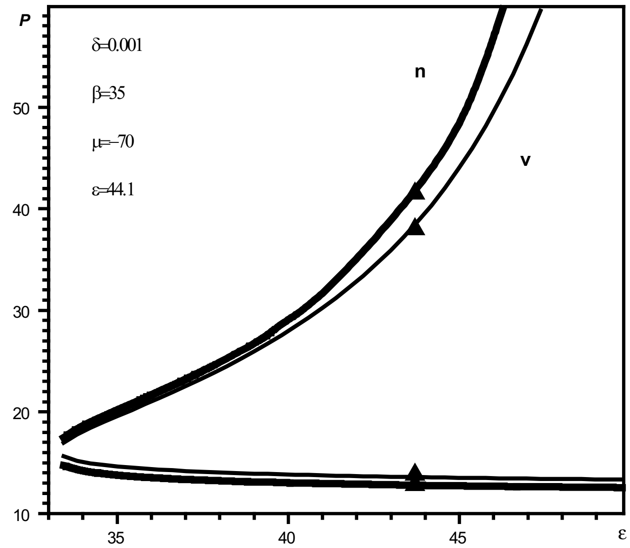


Figure 2: Upper stable and lower unstable branch of variational v curve and numerical n curve.

steady state solutions is established following extended stability criterion. The choice of asymmetrical input pulses with dissipative parameters belonging to the stability domain fixed by this criterion, insures generation of stable dissipative solitons. The possibility to control asymmetrical dissipative pulses, which during the propagation become either spatial or spatiotemporal stable dissipative solitons, opens the way for diverse practical applications concerning signal processing and mode-locked laser systems.

REFERENCES

1. Mihalache, D., et al., *Phys. Rev. Lett.*, Vol. 97, 073904, 2006.
2. McLeod, R., K. Wagner, and S. Blair, *Phys. Rev. A*, Vol. 52, 3254, 1995.
3. Chen, C.-J., P. K. A. Wai, and C. R. Menyuk, *Opt. Lett.*, Vol. 19, 198, 1994.
4. De Angelis, C., M. Santagiustina, and S. Wabnitz, *Opt. Commun.*, Vol. 122, 23, 1995.
5. Skarka, V., V. I. Berezhiani, and R. Miklaszewski, *Phys. Rev. E*, Vol. 56, 1080, 1997. V. Skarka, V. I. Berezhiani, and R. Miklaszewski, *Phys. Rev. E*, Vol. 59, 1270, 1999.
6. Akhmediev, N. N. and A. A. Ankiewicz, *Solitons, Nonlinear Pulses and Beams*, Chapman and Hall, London, 1997.
7. Skarka, V. and N. B. Aleksic, *Phys. Rev. Lett.*, Vol. 96, 013903, 2006.
8. Nicolis, G. and I. Prigogine, *Self-Organization in Nonequilibrium Systems*, John Wiley and Sons, New York, 1977.

Improved Interpolation of Evanescent Plane Waves for Fast Multipole Methods

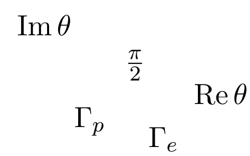
Henrik Wallén
Electromagnetics Laboratory, Helsinki University of Technology, Finland

Abstract— One way to implement a low-frequency or broadband fast multipole method is to use the spectral representation, or inhomogeneous plane-wave expansion, of the Green's function. To significantly improve the error-controllability of the method, we propose a new interpolation and anterpolation scheme for the evanescent part.

The fast multipole method (FMM) can be used to accelerate the iterative solution of integral equations in electromagnetics and acoustics. In particular, the multilevel fast multipole algorithm (MLFMA) by Chew et al., is very efficient for solving large scale electromagnetic scattering problems [1]. However, the dynamic FMM and MLFMA are based on a plane-wave expansion of the Green's function that is not error-controllable in the sub-wavelength scale. To implement a stable FMM in the sub-wavelength scale, there is basically two different approaches: either based multipole-series [2–4]; or based on a low-frequency stable inhomogeneous plane-wave expansion of the Green's function [5–8].

The spectral representation of the Green's function,

$$G(\mathbf{r}) = \frac{e^{ikr}}{4\pi r} = \frac{ik}{8\pi^2} \int_{-\pi}^{\pi} \int_{\Gamma} e^{i\mathbf{k}(\theta,\varphi)\cdot\mathbf{r}} \sin\theta \, d\theta \, d\varphi, \qquad z > 0, \tag{1}$$



gives an inhomogeneous plane-wave expansion that is accurate at any frequency, but unfortunately, the representation is direction dependent. The integration path Γ is split in two parts: a propagating part Γ_p, corresponding to ordinary propagating plane waves; and an evanescent part Γ_e, corresponding to evanescent plane waves (complex $\mathbf{k}$). For the propagating part, we can embed the direction dependency into the translation function and get excellent error controllability using an FFT-based approach based on [9]. The evanescent part is, however, more problematic.

Using the substitution $\sigma = -ik\cos\theta$, we can express the evanescent part of (1) as

$$G_e(\mathbf{D}+\mathbf{d}) = \int_{-\pi}^{\pi} \int_{0}^{\infty} T_e(\sigma,\varphi)\, e^{i\mathbf{k}(\sigma,\varphi)\cdot\mathbf{d}} \, d\sigma \, d\varphi, \qquad T_e(\sigma,\varphi) = \frac{1}{8\pi^2} e^{i\mathbf{k}(\sigma,\varphi)\cdot\mathbf{D}}, \tag{2}$$

where the wave vector is

$$\mathbf{k}(\sigma,\varphi) = \sqrt{\sigma^2+k^2}\,(\hat{\mathbf{x}}\cos\varphi + \hat{\mathbf{y}}\sin\varphi) + i\sigma\hat{\mathbf{z}}. \tag{3}$$

Using a generalized Gaussian quadrature rule [7, 10] for the σ-integral and a trapezoidal rule for the φ-integral, we can get up to 10 digits accuracy in (2). Unfortunately, a straightforward interpolation in σ, as proposed in [7], is not as accurate.

For a scalar source q in a cube Q, the evanescent radiation-pattern is

$$F^e_\infty(\sigma,\varphi) = \int_Q e^{-i\mathbf{k}(\sigma,\varphi)\cdot\mathbf{r}'}\, q(\mathbf{r}')\, dV'. \tag{4}$$

The sample points in σ are approximately scaled by a factor 2 between adjacent levels, and the square root $\sqrt{\sigma^2+k^2}$ in (3) makes the radiation patterns hard to interpolate accurately near $\sigma = 0$. To overcome this problem, we exploit the properties of the spherical harmonics

$$Y_{nm}(\theta,\varphi) = c_{nm}(\sin\theta)^{|m|}\, U_n^{|m|}(\cos\theta)\, e^{im\varphi}, \tag{5}$$

where c_{nm} is a constant and $U_n^{|m|}$ is a polynomial of degree $n-|m|$. The spherical harmonic $Y_{nm}(\theta,\varphi)$ is a polynomial of degree n in $\cos\theta$ if m is even, and $Y_{nm}(\theta,\varphi)/\sin\theta$ is a polynomial of degree $n-1$ in $\cos\theta$ if m is odd.

A propagating radiation pattern can be expanded using a series of spherical harmonics. Furthermore, the spherical harmonics extend to complex θ, and so we can also approximate an evanescent radiation pattern using a truncated series of spherical harmonics as

$$F_\infty^e(\sigma,\varphi) \approx \sum_{m=-N}^{N} b_m(\sigma)\, e^{im\varphi} = \sum_{m=-N}^{N} \sum_{n=|m|}^{N} a_{mn}(1+\sigma^2/k^2)^{|m|/2}\, U_n^{|m|}(i\sigma/k)\, e^{im\varphi}. \tag{6}$$

Here, $b_m(\sigma)$ is a polynomial of degree N in σ if m is even, while $b_m(\sigma)/\sqrt{1+\sigma^2/k^2}$ is a polynomial of degree $N-1$ in σ if m is odd.

As we use the samples of the radiation pattern to represent the outgoing field, the above suggests the following interpolation procedure for the evanescent part:

- Use a discrete Fourier transform (DFT) in φ to get the samples $b_m(\sigma_j)$.
- Scale $b_m(\sigma_j) \leftarrow b_m(\sigma_j)/\sqrt{1+\sigma_j^2/k^2}$ for all odd m.
- Interpolate in σ using ordinary Lagrangian interpolation to get the samples $b_m(\tilde{\sigma}_j)$ for the parent level.
- Scale back $b_m(\tilde{\sigma}_j) \leftarrow b_m(\tilde{\sigma}_j)\sqrt{1+\tilde{\sigma}_j^2/k^2}$ for all odd m.
- Zero-pad if we need to interpolate in φ, and use an inverse DFT to get the samples of $F_\infty^e(\sigma,\varphi)$ for the parent level.

As usual, the anterpolation procedure is a straightforward adjoint of the interpolation procedure.

In theory, this new interpolation procedure should be very accurate, provided that we use enough sample points. In practise, the resulting interpolations are clearly more accurate than the previously used ones [7], but not significantly slower.

As a benchmark, to verify the accuracy of the new interpolation scheme, we use the N-body problem

$$F_m = F(\mathbf{r}_m) = \sum_{\substack{n=1\\ n\neq m}}^{N} G(\mathbf{r}_m - \mathbf{r}_n)\, q_n, \qquad \text{for } m = 1,\ldots,N, \tag{7}$$

with $N = 10\,000$ and a 6-level MLFMA using the spectral representation of the Green's function, as in [7], both using the interpolation scheme of [7] and the new interpolation scheme presented here. The points $\mathbf{r}_n$ are randomly distributed on a sphere with a radius of one wavelength (λ) and the scalar sources q_n has random amplitude. The used levels have cube side-lengths between $\lambda/2$ and $\lambda/16$, which is the most problematic region for the interpolation in σ. The results are summarized in Table 1. The new interpolation scheme significantly improves the error controllability of the translation procedures proposed in [7], so that we can obtain up to 8 digits accuracy in the scalar Green's function.

Table 1: Relative L_∞-errors in F_m for the benchmark problem (7) for different target accuracies d_0 and both the old and new interpolation scheme.

d_0	Old interp.	New interp.
2	4.33×10^{-3}	4.11×10^{-3}
4	6.51×10^{-5}	1.57×10^{-5}
6	3.24×10^{-5}	5.17×10^{-7}
8	1.36×10^{-5}	7.72×10^{-9}

REFERENCES

1. Chew, W. C., J. M. Jin, E. Michielssen, and J. M. Song, Eds., *Fast and Efficient Algorithms in Computational Electromagnetics*, Ch. 3–4, Artech House, 2001.
2. Greengard, L., J. Huang, V. Rokhlin, and S. Wandzura, "Accelerating fast multipole methods for the Helmholtz equation at low frequencies," *IEEE Comput. Sci. Eng. Mag.*, Vol. 5, No. 3, 32–38, Jul.–Sep. 1998.
3. Zhao, J.-S. and W. C. Chew, "Three-dimensional multilevel fast multipole algorithm from static to electrodynamic," *Microwave Opt. Tech. Lett.*, Vol. 26, No. 1, 43–48, July 2000.
4. Jiang, L. J. and W. C. Chew, "A mixed-form fast multipole algorithm," *IEEE Trans. Antennas Propagat.*, Vol. 53, No. 12, 4145–4156, Dec. 2005.
5. Jiang, L. J. and W. C. Chew, "Low-frequency fast inhomogeneous plane-wave algorithm (LF-FIPWA)," *Microwave Opt. Tech. Lett.*, Vol. 40, No. 2, 117–122, Jan. 2004.
6. Darve, E. and P. Havé, "A fast multipole method for Maxwell equations stable at all frequencies," *Phil. Trans. R. Soc. Lond. A*, Vol. 362, No. 1816, 603–628, Mar. 2004.
7. Wallén, H. and J. Sarvas, "Translation procedures for broadband MLFMA," *Progress in Electromagnetics Research*, Vol. PIER 55, 47–78, 2005.
8. Wallén, H., S. Järvenpää, and P. Ylä-Oijala, "Broadband multilevel fast multipole algorithm for acoustic scattering problems," *Journal of Computational Acoustics*, accepted for publication, 2006.
9. Sarvas, J., "Performing interpolation and anterpolation entirely by fast fourier transform in the 3-D multilevel fast multipole algorithm," *SIAM J. Numer. Anal.*, Vol. 41, No. 6, 2180–2196, 2003.
10. Yarvin, N. and V. Rokhlin, "Generalized Gaussian quadratures and singular value decompositions of integral operators," *SIAM J. Sci. Comput.*, Vol. 20, No. 2, 699–718, 1998.

Searching for Electrostatic Resonances in Metamaterials Using Surface Integral Equation Approach

A. Sihvola, P. Ylä-Oijala, S. Järvenpää, and M. Taskinen

Helsinki University of Technology, Electromagnetics Laboratory
P.O. Box 3000, FI–02015 TKK, Espoo, Finland

Abstract— In this paper, the possibilities of finding out surface plasmons in negative-permittivity metamaterial scatterers, using surface integral equations, are discussed. Electrostatic resonances manifest themselves as very high values of the polarizability for a certain (negative) value of the permittivity. Another way to locate the position of the resonances is to calculate the eigenvalues of the source-free integral equation operator for the static potential on the surface of the particle.

1. INTRODUCTION

A dielectric sphere with relative permittivity ε_r behaves like an electric dipole when it is excited by a long-wavelength electric field. The amplitude of this dipole moment is proportional to the electric field and the coefficient of proportionality is the polarizability. In a form normalized by the volume of the sphere and free-space permittivity, the polarizability in normalized form reads [1].

$$\alpha_n = 3\frac{\varepsilon_r - 1}{\varepsilon_r + 2} \tag{1}$$

This expression reveals fascinating properties when we allow the relative permittivity ε_r to take negative values (this is of interest within the study of metamaterials). Obviously, the case $\varepsilon_r = -2$ leads us into a singularity. This infinity, known for metals, for example, in the high optical or ultraviolet frequencies, carries several names: surface plasmon, Fröhlich mode, electrostatic resonance.

Despite this singularity, the case of sphere (and also, by the way, ellipsoid, for which the polarizability can also be calculated in closed form) can be said to behave "reasonably" over the whole range of real values for the relative permittivity. Namely, only one infinity appears and its position and amplitude are well defined and connected to the axis ratios of the scatterer. However, when the shape deviates from these canonical forms, very strange things seem to happen. In our experience, all numerical and semianalytical efforts fail to reconstruct the whole range of the polarizability for negative values, at least in all nontrivial cases that we have been studying. Several singularities appear, some of which seem to be physical, but many perhaps numerical artifacts.

In this presentation, we will present our findings on these electrostatic resonances for negative-permittivity materials. We use surface-integral equation with method of moments to solve the polarizability for different shapes: in addition to sphere and ellipsoid, we study the shapes of cube and semisphere for which the wedges and corners may be one source of difficulties in the solution of the electrostatic potential. Another approach to locate the electrostatic resonances is to find eigenvalues to the source-free integral equation for the surface charge distribution [2]. This method gives the positions of the singularities directly in contrast to the straightforward evaluation of the polarizability over the whole domain of negative values of permittivity. However, this calculation returns a large set of eigenvalues, and it needs certain efforts to identify which of the eigenvalues corresponds to a given multipole resonance. In the following, we will present results from our calculations of the behavior of electrostatic resonances when the shape starts to deviate from the perfect spherical form towards a cube. Also, the effect of a finite frequency on the resonances is studied.

2. SURFACE INTEGRAL EQUATION FORMULATION

Let D be a homogeneous dielectric object in a homogeneous medium with a surface S and exterior unit normal vector $\boldsymbol{n}$. The exterior and interior of D are characterized with constant electromagnetic parameters ε_l and $\mu_l = \mu_0$, $l = 1, 2$, respectively. In addition, let $\boldsymbol{E}^p$ and $\boldsymbol{H}^p$ denote the primary fields with sources in the exterior.

In the electrostatic case ($\omega = 0$) an integral equation for the unknown potential function ϕ on the surface S reads [3].

$$\left((\varepsilon_1 - \varepsilon_2)\,\mathcal{K} - \frac{\varepsilon_1 + \varepsilon_2}{2}\,\mathcal{I} \right) \phi = -\varepsilon_1\,\phi^p, \tag{2}$$

where

$$(\mathcal{K}\phi)(\boldsymbol{r}) = \int\limits_S \frac{\partial G_0(\boldsymbol{r}, \boldsymbol{r}')}{\partial n'}\,\phi(\boldsymbol{r}')\,dS', \tag{3}$$

$\mathcal{I}$ is the identity operator, $G_0 = 1/(4\pi|\boldsymbol{r} - \boldsymbol{r}'|)$ is the static Green's function and ϕ^p is the scalar potential of the primary electric field.

In the dynamic case ($\omega > 0$) the integral equation system of the N-Müller formulation for the unknown equivalent magnetic and electric surface current densities, $\boldsymbol{M} = -\boldsymbol{n} \times \boldsymbol{E}$ and $\boldsymbol{J} = \boldsymbol{n} \times \boldsymbol{H}$, reads [4].

$$\left(\begin{bmatrix} \boldsymbol{n} \times (\frac{\varepsilon_1}{\eta_1}\boldsymbol{\mathcal{D}}_1 - \frac{\varepsilon_2}{\eta_2}\boldsymbol{\mathcal{D}}_2) & \boldsymbol{n} \times (\varepsilon_1\boldsymbol{\mathcal{K}}_1 - \varepsilon_2\boldsymbol{\mathcal{K}}_2) \\ \boldsymbol{n} \times (\mu_1\boldsymbol{\mathcal{K}}_1 - \mu_2\boldsymbol{\mathcal{K}}_2) & \boldsymbol{n} \times (\eta_2\mu_2\boldsymbol{\mathcal{D}}_2 - \eta_1\mu_1\boldsymbol{\mathcal{D}}_1) \end{bmatrix} \right.$$
$$\left. -\frac{1}{2} \begin{bmatrix} (\varepsilon_1 + \varepsilon_2)\boldsymbol{\mathcal{I}} & 0 \\ 0 & (\mu_1 + \mu_2)\boldsymbol{\mathcal{I}} \end{bmatrix} \right) \begin{bmatrix} \boldsymbol{M} \\ \boldsymbol{J} \end{bmatrix} = \begin{bmatrix} \varepsilon_1\,\boldsymbol{n} \times \boldsymbol{E}^p \\ -\mu_1\,\boldsymbol{n} \times \boldsymbol{H}^p \end{bmatrix} \tag{4}$$

where

$$(\boldsymbol{\mathcal{K}}_l\boldsymbol{F})(\boldsymbol{r}) = \int\limits_S \nabla\,G_l(\boldsymbol{r}, \boldsymbol{r}') \times \boldsymbol{F}(\boldsymbol{r}')\,dS', \tag{5}$$

$$(\boldsymbol{\mathcal{D}}_l\boldsymbol{F})(\boldsymbol{r}) = \frac{-1}{ik_l}\,\nabla \int\limits_S G_l(\boldsymbol{r}, \boldsymbol{r}')\nabla'_s \cdot \boldsymbol{F}(\boldsymbol{r}')\,dS' + ik_l \int\limits_S G_l(\boldsymbol{r}, \boldsymbol{r}')\boldsymbol{F}(\boldsymbol{r}')\,dS', \tag{6}$$

$G_l = e^{ik_l|\boldsymbol{r}-\boldsymbol{r}'|}/(4\pi|\boldsymbol{r} - \boldsymbol{r}'|)$ is the dynamic (free space) Green's function of region D_l, $\eta_l = \sqrt{\mu_l/\varepsilon_l}$ and $k_l = \omega\sqrt{\varepsilon_l\mu_l}$, $l = 1, 2$.

As has been shown in [4] the N-Müller formulation remains stable even at very low frequencies and thus can be used to study the behavior of electrostatic resonances as the frequency is increased. In fact, at $\omega = 0$, the N-Müller formulation decouples into two separate equations for the electrostatics and magnetostatics cases

$$\left((\varepsilon_1 - \varepsilon_2)\boldsymbol{\mathcal{K}}_0 - \frac{1}{2}(\varepsilon_1 + \varepsilon_2)\boldsymbol{\mathcal{I}} \right) \boldsymbol{M} = \varepsilon_1\boldsymbol{n}_1 \times \boldsymbol{E}^p, \tag{7}$$

$$\left((\mu_1 - \mu_2)\boldsymbol{\mathcal{K}}_0 - \frac{1}{2}(\mu_1 + \mu_2)\boldsymbol{\mathcal{I}} \right) \boldsymbol{J} = -\mu_1\boldsymbol{n}_1 \times \boldsymbol{H}^p, \tag{8}$$

where $\boldsymbol{\mathcal{K}}_0$ is the following integral operator

$$(\boldsymbol{\mathcal{K}}_0\boldsymbol{F})(\boldsymbol{r}) = \boldsymbol{n} \times \int\limits_S \nabla G_0(\boldsymbol{r}, \boldsymbol{r}') \times \boldsymbol{F}(\boldsymbol{r}')\,dS'. \tag{9}$$

The above equations, (2) and (4), are discretized with the method of moments using Galerkin method and linear continuous basis functions for the static potential and Rao–Wilton–Glisson basis functions for the equivalent surface currents densities in the dynamic case. More details are presented in [3] and [4].

3. EIGENVALUE PROBLEM

In the static case the resonances at negative values of ε_2 can be found by studying the eigenvalues of the integral operator of (2) [2]. By rewriting the equation as

$$(\mathcal{I}\phi) + 2\,\frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2}(\mathcal{K}\phi) = 0 \tag{10}$$

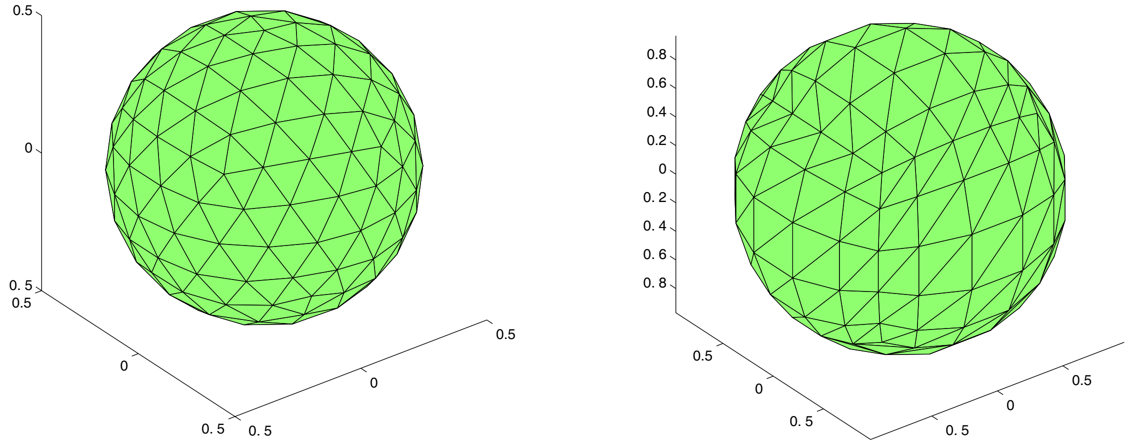


Figure 1: Triangular mesh on a sphere (left) and on a smooth cube with $n = 2.1$ (right).

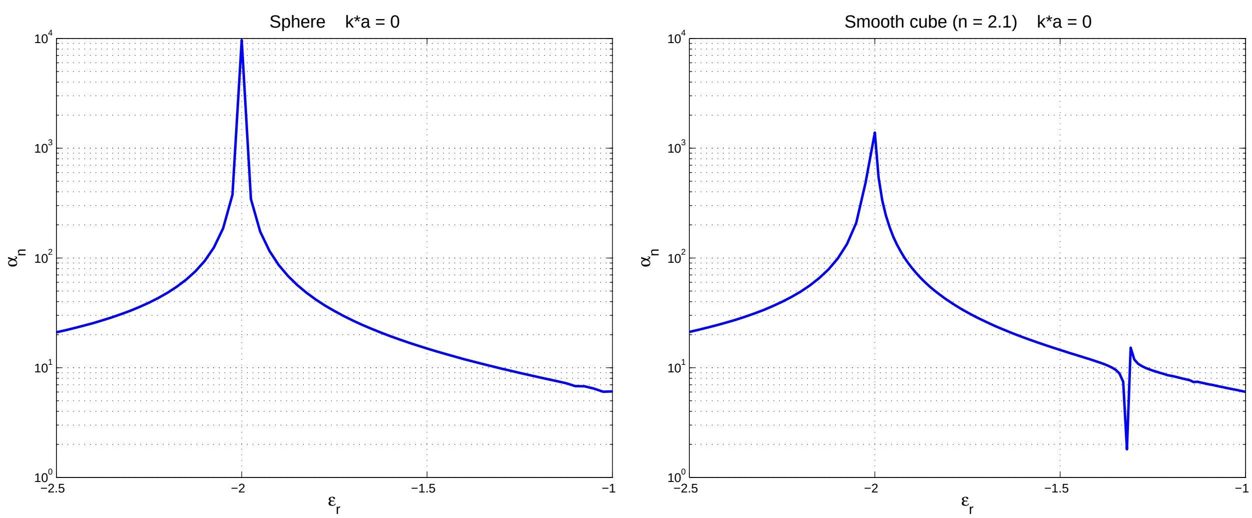


Figure 2: Absolute value of the electrostatic polarizability of a sphere (left) and a smooth cube with $n = 2.1$ (right) at zero frequency as a function of ε_r.

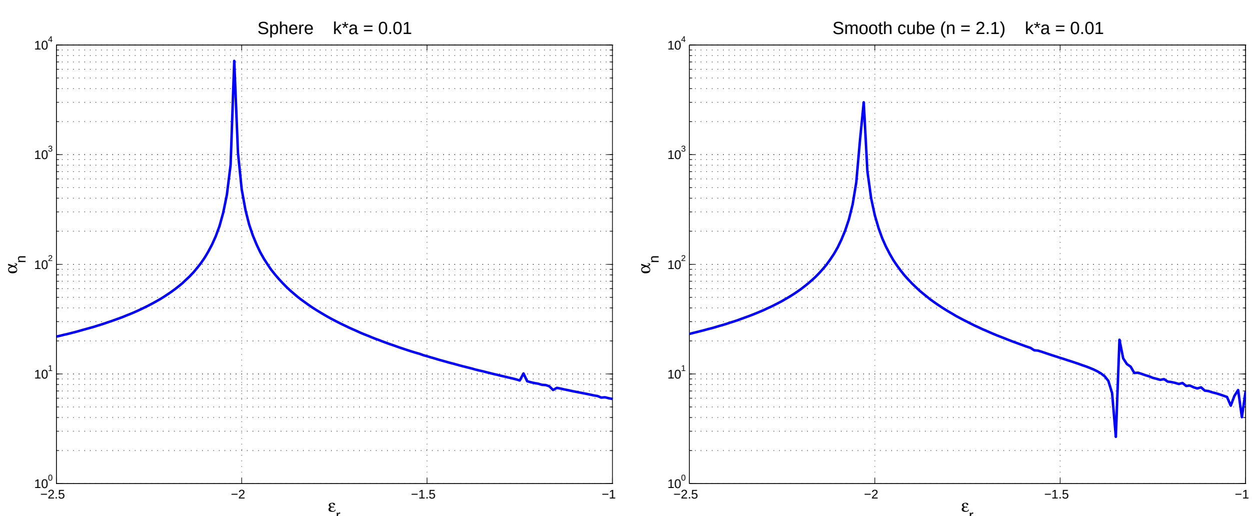


Figure 3: Absolute value of the electrostatic polarizability of a sphere (left) and a smooth cube with $n = 2.1$ (right) as a function of ε_r at $k_0 a = 0.01$.

this can be formulated as a generalized eigenvalue problem

$$A\,x + \lambda B\,x = 0 \tag{11}$$

where

$$\lambda(\varepsilon_2) = -2\,\frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2}. \tag{12}$$

4. NUMERICAL RESULTS

We present numerical results for a sphere and a smoothened cube defined as

$$|x|^n + |y|^n + |z|^n = a^n, \quad n \geq 2, \qquad (a > 0). \tag{13}$$

The case $n = 2$ corresponds to a sphere and $n = \infty$ corresponds to a cube. Figure 1 displays the triangular surface meshes of a sphere and a smoothened cube with $n = 2.1$.

Figures 2–3 show the (absolute value of the) normalized polarizability of a sphere and a smoothened cube at zero frequency and at a frequency corresponding to $k_0 a = 0.01$ as a function of the relative permittivity of the object. Here a is the radius of the sphere and k_0 is the wave number in vacuum. In the dynamic case the electrostatic normalized polarizability is calculated from the monostatic radar cross section σ as [5].

$$|\alpha_n| = \frac{\sqrt{4\pi\sigma}}{k_0^2 V}, \tag{14}$$

where V is the volume of the object.

5. CONCLUSIONS

The results for the polarizability curves as function of the negative permittivity show clearly the changes of the electrostatic resonance peak as the shape of the scatterer starts to deviate from sphere, and also in the case when the frequency becomes finite and the situation is no longer electrostatic. The main plasmon ($\varepsilon_r = -2$) begins to move into smaller values ($\varepsilon_r < -2$). But at the same time new resonances appear at permittivity points $-1-1/m$, $(m = 2, 3, 4, \ldots)$, and these values also start to float towards more negative values. An interesting observation is that although the "smoothed cube" with $n = 2.1$ is a shape visually nearly identical to a sphere (Figure 1), its resonance behavior is already much different. Furthermore, our preliminary studies about the behavior of the static surface integral operator show that these resonances correspond exactly to the positions of the eigenvalues of the source-free integral operator eigenproblem.

REFERENCES

1. Bohren, C. F. and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, Wiley, New York, 1983.
2. Fredkin, D. R. and I. D. Mayergoyz, "Resonant behavior of dielectric objects (electrostatic resonances)," *Physical Review Letters*, Vol. 91, No. 25, 253902, 2003.
3. Sihvola, A., P. Ylä-Oijala, S. Järvenpää, and J. Avelin, "Polarizabilities of platonic solids," *IEEE Transactions on Antennas and Propagation*, Vol. 52, No. 9, 2226–2233, Sep. 2004.
4. Ylä-Oijala, P. and M. Taskinen, "Well-conditioned Müller formulation for electromagnetic scattering by dielectric objects," *IEEE Transactions on Antennas and Propagation*, Vol. 53, No. 10, 3316–3323, Oct. 2005.
5. Sihvola, A., T. K. Sarkar, and B. Kolundzija, "From radar cross section to electrostatics," *IEEE Antennas and Propagation Letters*, Vol. 3, Issue 16, 324–327, Oct. 2004.

Efficient Preconditioning Strategies for the Multilevel Fast Multipole Algorithm

Levent Gürel[1,2], Tahir Malas[1], and Özgür Ergül[1]

[1]Department of Electrical and Electronics Engineering, Bilkent University
TR-06800, Bilkent, Ankara, Turkey
[2]Computational Electromagnetics Research Center (BiLCEM), Bilkent University
TR-06800, Bilkent, Ankara, Turkey

Abstract— For the iterative solutions of the integral equation methods employing the multilevel fast multipole algorithm (MLFMA), effective preconditioning techniques should be developed for robustness and efficiency. Preconditioning techniques for such problems can be broadly classified as fixed preconditioners that are generated from the sparse near-field matrix and variable ones that can make use of MLFMA with the help of the flexible solvers. Among fixed preconditioners, we show that an incomplete LU preconditioner depending on threshold (ILUT) is very successful in sequential implementations, provided that pivoting is applied whenever the incomplete factors become unstable. For parallel preconditioners, sparse approximate inverses (SAI) can be used; however, they are not as successful as ILUT for the electric-field integral equation. For a remedy, we employ variable preconditioning, and we iteratively solve the near-field system in each major iteration. However, for very large systems, neither of these methods succeeds to reduce the iteration counts as desired because of the thinning of the near-field matrices for increasing problem sizes. Considering this fact, we develop a preconditioner using MLFMA, with which we solve an approximate system. Respective advantages of these different preconditioners are demonstrated on a variety of problems ranging in both geometry and size.

1. INTRODUCTION

In this paper, we consider fast iterative solutions of the integral-equation methods, which yield dense $n \times n$ linear systems in the form of $\bar{\mathbf{Z}} \cdot \mathbf{x} = \mathbf{b}$. These methods are widely used to solve scattering and radiation problems in computational electromagnetics (CEM) because of their applicability to complex geometries and high accuracy. With the help of the multilevel fast multipole algorithm (MLFMA) [1], iterative methods can be used with $O(n \log n)$ computational complexity provided that iteration counts are limited with proper preconditioners. This important problem is addressed successfully by only a few CEM researchers [2–4].

Two formulations of the integral equations are widely used for the radiation and scattering problems. The electric-field integral equation (EFIE) is used for geometries involving open surfaces. On the other hand, the combined-field integral equation (CFIE) is the appropriate formulation for targets with closed surfaces since it is free from the internal resonance problem that EFIE may suffer. Moreover, CFIE yields linear systems that are easier to solve iteratively compared to EFIE.

In each step, iterative methods require matrix-vector products with the coefficient matrix $\bar{\mathbf{Z}}$. MLFMA provides this operation to the solver by computing the near-field entries exactly and the far-field entries approximately but with controllable error. Hence, only the near-field part of the coefficient matrix $\bar{\mathbf{Z}}^{\mathrm{NF}}$ is stored and the interactions of the far-field entries are computed "on the fly". For this purpose, the computational domain is located in a cube, and then the cube is divided recursively into smaller ones. This partitioning of the computational domain defines the levels of MLFMA. The near-field part corresponds to the interactions of the smallest cubes with their neighbors and their self-interactions.

In this paper, we present our work for preconditioning of the integral-equation methods. First, we consider sequential implementations of MLFMA and concentrate on the incomplete LU (ILU) preconditioners. We show that ILU preconditioners can be very successful, provided that pivoting is applied in case of instability. Because of the difficulty of the parallelization of the ILU preconditioners, we construct an efficient sparse approximate inverse (SAI) preconditioner for the parallel implementations. For EFIE, SAI is not as successful as ILU in reducing the iteration counts; hence we use it as a preconditioner for the solution of the near-field system. Then, the solution of the near-field system is used as the preconditioner of the original system. For very large problems, on the other hand, due to their increasing sparseness, the near-field matrices do not serve as good approximations to the dense system matrices. Therefore, we use an incomplete version of MLFMA to obtain a more powerful preconditioner. With these efforts, we are able to solve large-scale problems using a 16-processor computer in moderate solution times.

2. INCOMPLETE LU PRECONDITIONING

During factorization of sparse matrices, sparsity is lost in general. However, by sacrificing some of the nonzero elements of the exact LU factorization, incomplete LU (ILU) preconditioners can be constructed. This idea is the most established preconditioning method with many freely available implementations [9] and it has proven to be successful in the iterative solution of sparse linear systems [6].

Consider an incomplete factorization of the near-field matrix, $\bar{\mathbf{Z}}^{\mathrm{NF}} = \bar{\mathbf{L}} \cdot \bar{\mathbf{U}}$. If we retain the nonzero values of $\bar{\mathbf{L}}$ and $\bar{\mathbf{U}}$ only at the nonzero positions of $\bar{\mathbf{Z}}^{\mathrm{NF}}$, we end up with the no-fill ILU method, or ILU(0). For problems that are not far from being diagonally dominant, such as the ones resulting from CFIE, this simple idea usually works well. On the other hand, since ILU(0) does not consider the numerical values of the entries, it becomes ineffective in predicting the locations of the largest entries for matrices that are far from being diagonally dominant and highly indefinite.

For such problems, a second class of ILU preconditioners is developed based on the principle of dropping matrix elements depending on their magnitudes. Among such methods, ILUT (τ, p) [6] drops matrix elements that are smaller than τ times the 2-norm of the current row and of all the remaining entries at most p largest ones are kept. ILUT is known to yield more accurate and stable factorizations compared to ILU(0) with the same amount of fill-in [7]. However, for some cases, although the factorization terminates normally, the incomplete factors sometimes turn out to be unstable. When the problem is related to the small pivots, this problem can be avoided using partial pivoting as in the complete factorization case.

Table 1: ILU results for CFIE.

Problem	Number of unknowns	LU	No PC		BJ		ILU(0)		
		Iter	Iter	Time	Iter	Time	Iter	Setup	Time
Sphere	132,003	29	49	1,103	32	684	29	23	665
Thin box	147,180	37	158	1,965	106	1,290	45	271	1,025
Wing	117,945	31	86	1,110	52	779	32	46	542
Flamme	78,030	63	229	2,138	115	1,096	66	43	694
Helicopter	183,546	42	253	7,338	106	3,081	44	145	1,739

Table 2: ILU results for EFIE.

Problem	Number of unknowns	LU	Jacobi		ILUTP		
		Iter	Iter	Time	Iter	Setup	Time
Patch	137,792	53	833	16,209	81	661	2,167
Open cube	171,655	332	-	-	376	2,243	9,833
Open prism	163,871	195	-	-	253	996	6,883
Half sphere	116,596	93	1,052	25,947	110	1,353	3,579

We realize that ILU(0) and ILUT works very well for the CFIE and EFIE systems, respectively. In Table 1 and Table 2, we show the number of iteration counts and solution times with GMRES no-restart on a solver containing Opteron 244 processors. Zero initial guess is used and 10^{-6} residual error is aimed in maximum 1500 iterations. We compare the ILU preconditioners with the following ones.

- **LU:** Denotes the exact solution of the near-field matrix, which is used as a benchmark preconditioner. Due to its excessive computational requirements, LU is not a practical preconditioner; it is presented merely for comparison of iteration numbers.
- **Block-Jacobi (BJ):** The blocks in this widely used preconditioner correspond to self-interactions of the smallest clusters in MLFMA. Exact inverse of each small block is obtained by LU decomposition. There are n such blocks with fixed sizes; hence, the cost of this preconditioner is $O(n)$.

- **Jacobi:** Contains only the diagonal of the near-field matrix. In Table 2, this preconditioner is included instead of BJ, since the latter performs poorer for EFIE.

From Table 1, we recognize that particularly for real-life problems, such as Flamme [8] and the helicopter, ILU(0) decreases the iteration counts and the solution times significantly. Moreover, the iteration counts turn out to be very close to that of LU, signaling the optimality of ILU(0). For open geometries presented in Table 2, first we note that for EFIE systems, even with a robust solver, either the numbers of iterations are very high or no convergence is seen at all. ILUTP refers to ILUT preconditioner with pivoting, which is required for the robustness of ILUT. The parameters of ILUTP are chosen so that the memory requirement does not exceed that of the near-field matrix. From the table one can see that ILUTP reduces the iteration numbers by an order of magnitude compared to Jacobi when the latter succeeds to converge.

3. SPARSE APPROXIMATE INVERSES FOR PARALLEL COMPUTING

The main disadvantage of the ILU preconditioners is the lack of parallelizability of the factorization and the application phases. To overcome this limitation, SAI preconditioners are developed, which are based on approximating the inverse of the matrix directly [3, 4]. For this purpose, an approximate inverse $\bar{\mathbf{M}} \approx \left(\bar{\mathbf{Z}}^{\mathrm{NF}}\right)^{-1}$ is explicitly constructed and stored. Then, application of the preconditioner is carried out by a sparse matrix-vector multiplication, whose parallelization can be performed efficiently.

The results in Table 3, obtained on a 16-processor parallel computer, indicate that SAI is much more successful in reducing the iteration counts and solution times compared to BJ in closed geometries. Even though total solution times are higher, only a small number of right-hand-side (RHS) vectors suffice to compensate setup time of SAI. Hence, for multiple-illumination problems involving real-life closed targets, SAI should be preferred over BJ.

Table 3: Comparison of parallel SAI and BJ.

CFIE		BJ				SAI				
Problem	Number of unknowns	Iter	Setup	Solution Time	Total Time	Iter	Setup	Solution Time	Total Time	RHS
Sphere	829,881	40	0.18	882	882	37	532	838	1,370	12
Thin box	147,180	104	0.14	352	352	63	304	218	522	3
Wing	117,945	53	0.05	98	98	38	192	73	264	8
Flamme	895,407	185	0.68	8,096	8,097	151	3,239	6,588	9,827	3
Helicopter	739,404	107	0.30	3,698	3,699	77	1,840	2,739	4,578	2

However, SAI preconditioners are not as successful as ILU preconditioners for EFIE. For this purpose, we consider using the SAI preconditioner for the iterative solution of the near-field system, and then the near-field solution is used as a preconditioner to the original system. For the iterative near-field preconditioner, the preconditioning operation is not fixed. Hence, a flexible solver should be used for the system matrix, allowing the preconditining operation change from iteration to iteration. We use FGMRES [6] in our experiments. For the near-field system solution, we use GMRES because of its ability to reduce the residual error quickly on the first iterations.

This scheme yields a forward-type preconditioner such as the ILU preconditioner. The difference is that, in ILU preconditioning, the preconditioner is already in factorized form, i.e., $\bar{\mathbf{M}} = \bar{\mathbf{L}} \cdot \bar{\mathbf{U}} \approx \bar{\mathbf{Z}}^{\mathrm{NF}}$, and for a given vector $\mathbf{y}$, the system $\bar{\mathbf{M}} \cdot \mathbf{x} = \mathbf{y}$ is solved by using backward and forward solves. On the other hand, for the preconditioning scheme described, which we call the iterative near-field preconditioner, the preconditioner is the exact near-field matrix, i.e., $\bar{\mathbf{M}} = \bar{\mathbf{Z}}^{\mathrm{NF}}$, but we approximately solve the system $\bar{\mathbf{M}} \cdot \mathbf{x} = \mathbf{y}$ by an iterative method.

In Table 4, we compare LU, SAI, and iterative near-field (NF/SAI) preconditioners. With NF/SAI, we achieve very close iteration counts to that of LU. We note that, SAI-preconditioned GMRES accelerates the convergence of the near-field system dramatically. We set the maximum number of iterations to 5, and this suffices for an effective preconditioner.

Table 4: Comparison of LU, SAI, and NF/SAI for EFIE. 'MLE' stands for 'memory limitation is exceeded'.

EFIE		LU	SAI Setup	SAI		NF/SAI	
Problem	Number of unknowns	Iter		Iter	Solution Time	Iter	Solution Time
Patch	137,792	53	52	91	336	59	253
	719,000	MLE	214	190	4,091	141	3,369
Open prism	127,925	112	156	172	628	120	520
	409,514	MLE	336	389	4,601	209	3,476
Half sphere	9,911	38	7	60	24	40	17
	116,596	93	77	156	510	103	383
Reflector antenna	356,439	MLE	952	125	878	71	646

4. USING MLFMA FOR EFFECTIVE PRECONDITIONING

For problems involving large numbers of unknowns, near-field matrices become increasingly sparser and, beyond some level, it becomes nearly impossible to achieve low iteration counts with the preconditioners obtained from the near-field matrices. Therefore, for effective preconditioning of very large problems, we need more information than that is provided by the near-field matrix.

Table 5: Comparison of incomplete MLFMA preconditioner with NF/SAI and SAI.

EFIE		SAI setup	LU	SAI		NF/SAI			IMLFMA/SAI		
Problem	Number of unknwns		Iter	iter	Solution Time	Iter outer	Iter Inner	Solution Time	Iter outer	Iter Inner	Solution Time
Patch	137,792	52	53	91	336	59	176	253	14	134	172
	719,000	214	-	190	4,091	141	421	3,369	32	313	2,152
Open prism	127,925	156	-	172	628	120	360	520	30	300	442
	409,514	336	-	389	4,601	209	627	3,476	54	540	2,561
Half sphere	9,911	7	38	60	24	40	120	17	8	76	14
	116,596	77	93	156	510	103	309	383	21	208	307
Reflector Antenna	356,439	952	-	125	878	71	213	646	17	165	478

Table 6: Comparison of incomplete MLFMA preconditioner with others for CFIE.

CFIE		Block Jacobi			SAI			IMLFMA/SAI		
Problem	Number of unknwns	Setup	Iter	Solution Time	Setup	Iter	Solution Time	Iter Outer	Iter Inner	Solution Time
Thin box	147,180	0.14	104	352	304	63	218	11	103	120
Wing	117,945	0.05	53	98	191.536	38	73	8	49	62
Helicopter	739,404	0.30	107	3,698	1,840	77	2,739	13	118	1,161
Flamme	895,407	0.68	185	8,096	3,239	151	6,588	26	255	2,491

This information can only be provided by MLFMA in the form of matrix-vector products. Following this idea, in each step of the iterative solver, we solve a nearby system using MLFMA. This preconditioning scheme yields a nesting of the solvers. The inner solver is used for preconditioning purpose, hence we need to solve only an approximate system. In order to perform the matrix-vector multiplication faster, we lower the sampling densities of MLFMA compared to the high-accuracy MLFMA. In this way, it is possible to reduce the cost of the preconditioning operation, and thereby the total solution time significantly.

In Table 5 and Table 6, we present the total solution times of aforementioned parallel preconditioners and the preconditioner obtained from approximate MLFMA, for some of the large problems shown in Table 3 and Table 4. IMLFMA/SAI is the preconditioner obtained with an incomplete MLFMA. As can be seen from the tables, using MLFMA for preconditioning purposes reduces the total solution time so significantly that even the solutions of large-scale problems can be obtained in less than an hour.

5. CONCLUSION

In this work we show that, for the success of large-scale iterative solutions of integral equation methods, the key is effective preconditioning. Even with systems resulting from CFIE, when the size of the problem gets larger or the target geometry is complex as in real-life problems, preconditioning is required to achieve solutions with less time. On the other hand, severely ill-conditioned EFIE systems sometimes do not converge even with robust solvers such as the no-restart GMRES.

Our experiments with the sequential programs reveal that ILU preconditioners can be safely applied to integral equation methods, provided that pivoting is applied to ILUT for EFIE systems. A comparison with the exact solution of the near-field matrix shows the near optimality of them. Furthermore, ILU is the most established preconditioning technique, whose implementations are available in solver packages, such as PETSc [9]. Hence, we strongly recommend their use for sequential problems.

For larger problems, on the other hand, one should resort to SAI preconditioners. Though SAI works well up to certain problem sizes, more effective preconditioning strategies, such as the use of iterative solution of the near-field system or the use of incomplete MLFMA for preconditioning help to solve even larger systems with low-memory and solution-time requirements.

ACKNOWLEDGMENT

This work was supported by the Scientific and Technical Research Council of Turkey (TUBITAK) under Research Grant 105E172, by the Turkish Academy of Sciences in the framework of the Young Scientist Award Program (LG/TUBA-GEBIP/2002-1-12), and by contracts from ASELSAN and SSM.

REFERENCES

1. Lu, C.-C. and W. C. Chew, "Multilevel fast multipole algorithm for electromagnetic scattering by large complex objects," *IEEE Trans. Antennas Propagat.*, Vol. 45, No. 10, 1488–1493, 1997.
2. Lee, J., J. Zhang, and C.-C. Lu, "Incomplete LU preconditioning for large scale dense complex linear systems from electromagnetic wave scattering problems," *J. Comput. Phys.*, Vol. 185, 158–175, 2003.
3. Lee, J., J. Zhang, and C.-C. Lu, "Sparse inverse preconditioning of multilevel fast multipole algorithm for hybrid integral equations in electromagnetics," *IEEE Trans. Antennas and Propagation*, Vol. 52, No. 9, 158–175, 2004.
4. Carpentieri, B., I. S. Duff, L. Giraud, and G. Sylvand, "Combining fast multipole techniques and an approximate inverse preconditioner for large electromagnetism calculations," *SIAM J. Sci. Comput.*, Vol. 27, No. 3, 774–792, 2005.
5. Benzi, M., "Preconditioning techniques for large linear systems: a survey," *J. Comput. Phys.*, Vol. 182, No. 2, 418–477, 2002.
6. Saad, Y., *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, ABJ, 2003.
7. Chow, E. and Y. Saad, "Experimental study of ILU preconditioners for indefinite matrices," *J. Comput. Appl. Math.*, Vol. 86, No. 2, 387–414, 1997.
8. Gürel, L., H. Bağcı, J.-C. Castelli, A. Cheraly, and F. Tardivel, "Validation through comparison: measurement and calculation of the bistatic radar cross section of a stealth target," *Radio Science*, Vol. 38, No. 3, 1046–1058, 2003.
9. Balay, S., W. D. Gropp, L. C. McInnes, and B. F. Smith, "PETSc users manual," Tech. Report ANL-95/11 Revision 2.1.5, Argonne National Laboratory, 2004.

The Edge Finite Element Method for Magnetic Fields Excited by Artificial Source in Frequency Domain

Zhifeng Xu and Baiyao Ruan
Guilin University of Technology, China

Abstract— In this paper, the total magnetic fields are decomposed into the magnetic fields excited by the artificial source in vacuum space and the abnormal magnetic fields excited by the inhomogeneous median or topography. The functional equation of abnormal magnetic fields has been constructed. Because the electromagnetic problem is open, it is localized by boundary integral method. Since the abnormal magnetic fields on the artificial boundary are weak, it can be neglected and stiffness matrix becomes symmetric. The interpolation functions of the abnormal magnetic fields have been constructed by edge finite element method. For simplifying the calculation of functional integration, the local edge finite element space in the reference coordinate that is equivalent to the edge finite element space in global coordinate has been constructed.

1. INTRODUCTION

In the process of modeling the electromagnetic fields excited by the artificial source, the common finite element method has many faults, for examples: introduction of the pseudo-solution and the magnetic fields are not free-divergent [8]. The edge finite element method can overcome those difficulties. In Ref. [7], Hassler Whitley introduced an approximate method to integrate the differential equation on the manifold. This method is low order approximation. In 1980, J. C. Nedelec introduced a new family of finite element in his paper [2]; this finite element is equivalent to the Whitney element in low order. In 1986, he continued to develop the mixed finite elements on conforming in $\mathbf{H}\,(\mathrm{div})$ and $\mathbf{H}\,(\mathrm{curl})$ respectively in paper [3].

In this paper, the magnetic fields is taken as working variable [4], the total magnetic fields are decomposed into the fields excited by the artificial source in vacuum space and the abnormal fields [1]. Then the functional equation of abnormal magnetic fields has been constructed. By using the artificial boundary, the open electromagnetic problem is localized, then the functional integration is carried out in finite domain. The stiffness becomes symmetric when the integration of the abnormal magnetic fields in boundary integration is neglected. In order to keep the boundary conditions and $\nabla \cdot \mathbf{H} = 0$ at every point of the given domain, the approximation of abnormal magnetic fields has been constructed by the edge finite element method.

2. THE FUNCTIONAL EQUATION OF MAGNETIC FIELD

Let

$$\frac{1}{Z(\mathbf{r})} = i\omega\mu(\sigma(\mathbf{r}) + i\omega\varepsilon(\mathbf{r})) \tag{1}$$

Then the magnetic fields satisfied the equation as below [1]:

$$\nabla \times (Z(\nabla \times \mathbf{H})) + \mathbf{H} = Z\nabla \times \mathbf{j}_f = \mathbf{J}_s \tag{2}$$

By general variational principle, the functional equivalent with Equation (1) can be got as below:

$$\Phi(\mathbf{H}) = \iiint_V \left\{ \frac{1}{2} \left[Z\left(\nabla \times \mathbf{H}\right) \cdot \left(\nabla \times \mathbf{H}\right) - \mathbf{H} \cdot \mathbf{H} \right] + \mathbf{J}_s \cdot \mathbf{H} \right\} dv \tag{3}$$

The whole space is divided into two sub-domains Ω_1 and Ω_2 by the artificial boundary Γ. Because the source is closed in finite domain Ω_1 and in the domain Ω_2 the magnetic fields satisfy the equation as below:

$$\nabla \times (Z(\nabla \times \mathbf{H})) + \mathbf{H} = \mathbf{0} \tag{4}$$

Then Equation (3) can be written as follows:

$$\begin{aligned}\Phi(\mathbf{H}) \;=\; & \iiint_{\Omega_1} \left\{ \frac{1}{2} [Z(\nabla \times \mathbf{H}) \cdot (\nabla \times \mathbf{H}) - \mathbf{H} \cdot \mathbf{H}] + \mathbf{J}_s \cdot \mathbf{H} \right\} dv \\ & + \iint_{\Gamma} \frac{1}{2} \left[Z(\nabla \times \mathbf{H}) \times \mathbf{H} \right] \cdot \mathbf{n} ds \end{aligned} \tag{5}$$

The total magnetic fields are decomposed into magnetic fields excited by the artificial source in vacuum and abnormal fields, namely:

$$\mathbf{H} = \mathbf{H}_f + \mathbf{H}_a \tag{6}$$

The Equation (6) is substituted into Equation (5). Because $\mathbf{H}_f$ doesn't change for the same source, the variation is only for the abnormal magnetic fields $\mathbf{H}_a$, then items including combination of $\mathbf{H}_f$ and $\mathbf{J}_s$ and the quadratic of $\mathbf{H}_f$ are neglected. Due to the boundary integration $\iint_\Gamma \frac{1}{2}\left[Z(\nabla\times\mathbf{H}_a)\times\mathbf{H}_a\right]\cdot\mathbf{n}ds$, the stiffness is not symmetric. But $\mathbf{H}_a$ is small, the integration can be neglected, then equation can be given as below:

$$\begin{aligned}\Phi(\mathbf{H}_a) =& \iiint_{\Omega_1}\frac{1}{2}\left\{Z(\nabla\times\mathbf{H}_a)\cdot(\nabla\times\mathbf{H}_a)+2Z(\nabla\times\mathbf{H}_a)\cdot(\nabla\times\mathbf{H}_f)-\mathbf{H}_a\cdot\mathbf{H}_a-2\mathbf{H}_a\cdot\mathbf{H}_f+2\mathbf{J}_s\cdot\mathbf{H}_a\right\}dv\\ &+\iint_\Gamma\frac{1}{2}\left\{Z\left[\mathbf{n}\times(\nabla\times\mathbf{H}_f)\right]\cdot\mathbf{H}_a+Z[\mathbf{n}\times(\nabla\times\mathbf{H}_a)]\cdot\mathbf{H}_f\right\}ds\end{aligned} \tag{7}$$

In the global coordinate, by using tetrahedral element, the variational (7) can be wrote into discrete formulas as below:

$$\begin{aligned}\Phi(\mathbf{H}_a) =& \sum_i\Bigg(\iiint_{\Omega_i}\frac{1}{2}\Big\{Z^i\big(\nabla\times\mathbf{H}_a^i\big)\cdot\big(\nabla\times\mathbf{H}_a^i\big)+2Z^i\big(\nabla\times\mathbf{H}_a^i\big)\cdot\big(\nabla\times\mathbf{H}_f^i\big)-\mathbf{H}_a^i\cdot\mathbf{H}_a^i-2\mathbf{H}_a^i\cdot\mathbf{H}_f^i+2\mathbf{J}_s^i\cdot\mathbf{H}_a^i\Big\}dv\\ &+T_c\iint_{\Gamma_j}\frac{1}{2}\Big\{Z^i\big[\mathbf{n}\times(\nabla\times\mathbf{H}_f^i)\big]\cdot\mathbf{H}_a^i+Z^i\big[\mathbf{n}\times(\nabla\times\mathbf{H}_a^i)\big]\cdot\mathbf{H}_f^i\Big\}ds\Bigg)\end{aligned} \tag{8}$$

where T_c is 1 if the element includes artificial boundary, and otherwise zero.

3. THE PRINCIPLES OF COORDINATE TRANSFORMING

It is difficult to calculate the variation (8) in the global coordinate. For every element, the standard reference element can be constructed. The edge finite elements in the two coordinates are equivalent. It is easy to analysis the element in reference. The old and new coordinates satisfy the transforming relationship as below:

$$\mathbf{X} = \mathbf{T}\mathbf{x} + A \tag{9}$$

Let

$$\mathbf{T}_1 = (T_{11},\, T_{12}\, T_{13})^T, \quad \mathbf{T}_2 = (T_{21},\, T_{22},\, T_{23})^T, \quad \mathbf{T}_3 = (T_{31},\, T_{32},\, T_{33}),$$

the differential surface element satisfies the transforming relationship as below:

$$dS = Kds \tag{10}$$

where

$$K = \sqrt{(\mathbf{T}_1\times\mathbf{T}_3)^2\cos^2\theta_2 + (\mathbf{T}_1\times\mathbf{T}_2)^2\cos^2\theta_3 + (\mathbf{T}_2\times\mathbf{T}_3)^2\cos^2\theta_1}$$

For tetrahedral element, its surfaces is plane, and then the normal direction of the given surface is stable. $\mathbf{n}$ is the unit vector of direction, $\cos\theta_1 = \frac{(\mathbf{T_2}\times\mathbf{T_3})\cdot\mathbf{n}}{|\mathbf{T_2}\times\mathbf{T_3}|}$, and others can be obtained similarly. For a given boundary element, K is constant.

The differential volume element satisfies the transforming relationship as follows:

$$dV = \det(T)dv \tag{11}$$

For equivalence of the two edge finite space, the fields must satisfy the transforming relationship as below [5]:

$$\mathbf{U} = \mathbf{T}^{-t}\mathbf{u} \tag{12}$$

It can be proven that the fields also satisfy the relationship as below [6]:

$$\nabla_x\times\mathbf{U} = \frac{1}{\det(\frac{DX_k}{D\xi})}\nabla_\xi\times\mathbf{u} \tag{13}$$

4. ELEMENT ANALYSIS

In the reference coordinate ,$\Phi^{(j)}$can be written as:

$$\Phi^{(j)} = \frac{1}{2}\sum_{i=1}^{5} A_i \tag{14}$$

where

$$\begin{aligned}
A_1 &= \iiint_{\Omega_{it}} \frac{Z^i(\mathbf{x})}{\det(T)}(\nabla_x \times h_a^i)\cdot(\nabla_x \times h_a^i)dv\\
A_2 &= 2\iiint_{\Omega_t^i} \frac{Z^i(\mathbf{x})}{\det(T)}(\nabla_x \times h_a^i)\cdot(\nabla_x \times h_f^i)dv\\
A_3 &= -\iiint_{\Omega_t^i} \det(T)(T^{-t}h_a^i)\cdot(T^{-t}h_a^i)dv\\
A_4 &= -2\iiint_{\Xi_t^i} \det(T)(T^{-t}h_a^i)\cdot(T^{-t}h_f^i)dv\\
A_5 &= 2\iiint_{\Omega_t^i} \det(T)(T^{-t}j_s^i)\cdot(T^{-t}h_a^i)dv\\
A_6 &= \iint_{\Gamma_i'} Z^i(x)L_{mk}a_{fk}^i t_{lm}h_{al}^i K ds\\
A_7 &= \iint_{\Gamma_i'} Z^i(x)L_{mk}a_{ak}^i t_{lm}h_{fl}^i K ds\\
T^- &= \{t_{ij}\}, \quad \mathbf{a} = \nabla_x \times \mathbf{u}, \quad L_{mk} = \frac{1}{\det(T)}\varepsilon_{ijk}t_{li}T_{jm}T_{js}n_s'
\end{aligned}$$

The dumb index means sum for subscript.

The standard tetrahedral element is displayed in Figure 1, the number and the orientation of its edge also are illustrated in Figure 1. The unit vector of the edge direction are shown as following [5]:

$$\mathbf{t}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \mathbf{t}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \mathbf{t}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}, \mathbf{t}_4 = \begin{pmatrix}\frac{-\sqrt{2}}{2}\\0\\\frac{\sqrt{2}}{2}\end{pmatrix}, \mathbf{t}_5 = \begin{pmatrix}\frac{\sqrt{2}}{2}\\\frac{-\sqrt{2}}{2}\\0\end{pmatrix}, \mathbf{t}_6 = \begin{pmatrix}0\\\frac{-\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\end{pmatrix}$$

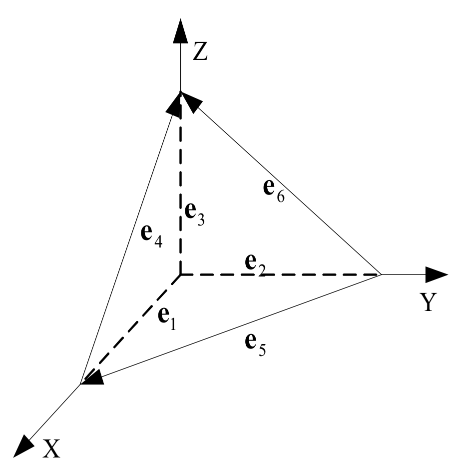


Figure 1: The standard tetrahedral element and orientation of edges.

By using the normalized basic vector as below:

$$\theta_1 = \begin{pmatrix}1-z-y\\x\\x\end{pmatrix}, \theta_2 = \begin{pmatrix}y\\1-z-x\\y\end{pmatrix}, \theta_3 = \begin{pmatrix}z\\z\\1-y-x\end{pmatrix}, \theta_4 = \begin{pmatrix}-z\\0\\x\end{pmatrix}, \theta_5 = \begin{pmatrix}y\\-x\\0\end{pmatrix}, \theta_6 = \begin{pmatrix}0\\-z\\y\end{pmatrix},$$

the vector in the standard tetrahedral element can be approximated with one order accuracy by the following formula:

$$\mathbf{u} = L_i \theta_i \tag{15}$$

Substitute Equation (15) into Equation (14), the following formula can be obtained:

$$\begin{aligned} A_1 &= L_j L_k K_{jk}^{(1)} \\ &\quad K_{jk}^{(1)} = \iiint_{K_t} \frac{Z^i(x)}{\det(T)} (\nabla \times \theta_j) \cdot (\nabla \times \theta_k) dv \end{aligned} \tag{16}$$

$$\begin{aligned} A_2 &= 2 L_k Z_k^{(1)} \\ &\quad Z_k^{(1)} = \iiint_{K_t} \frac{Z^i(x)}{\det(T)} (\nabla \times \theta_k) \cdot (\nabla \times \mathbf{h}_f^i) dv \end{aligned} \tag{17}$$

$$\begin{aligned} A_3 &= L_j L_k K_{jk}^{(2)} \\ &\quad K_{jk}^{(2)} = -\iiint_{K_t} T^{-t}\theta_j \cdot T^{-t}\theta_k dv \end{aligned} \tag{18}$$

$$\begin{aligned} A_4 &= 2 L_k Z_k^{(2)} \\ &\quad Z_k^{(2)} = -\iiint_{K_t} T^{-t}\theta_k \cdot T^{-t}\mathbf{h}_f^i dv \end{aligned} \tag{19}$$

$$\begin{aligned} A_5 &= 2 L_k Z_k^{(3)} \\ &\quad Z_k^{(3)} = \iiint_{K_t} T^{-t}\mathbf{j}_s^i \cdot T\theta_k dv \end{aligned} \tag{20}$$

$$\begin{aligned} A_6 &= L_n Z_n^{(4)} \\ &\quad Z_n^{(4)} = \iint_{\Gamma_t} Z^i(x) L_{mk} a_{fk}^i t_{lm} \theta_{nl} K ds \end{aligned} \tag{21}$$

$$\begin{aligned} A_7 &= L_n Z_n^{(5)} \\ &\quad Z_n^{(5)} = \iint_{\Gamma_t} Z^i(x) L_{mk} a_{nk} t_{lm} h_{fl}^i K ds \\ &\quad a_{nk} = (\nabla \times \boldsymbol{\theta}_n)_k \end{aligned} \tag{22}$$

Then

$$\Phi^s = \left(K_{ij} L_i L_j + 2 Z_k L_k\right)/2 \tag{23}$$

where

$$\begin{aligned} K_{ij} &= K_{ij}^{(1)} + K_{ij}^{(2)} \\ Z_k &= Z_k^{(1)} + Z_k^{(2)} + Z_k^{(3)} + T_c \left(Z_k^{(4)} + Z_k^{(5)}\right)/2 \end{aligned}$$

By using variational method and $K_{ij} = K_{ji}$, the following equation can be got from Equation (23):

$$K_{ij} L_j = Z_i \tag{24}$$

Let

$$\mathbf{K} = (K_{ij}), \quad \mathbf{L} = (L_1, L_2, L_3, L_4, L_5, L_6)^t, \quad \mathbf{Z} = (Z_1, Z_2, Z_3, Z_4, Z_5, Z_6)^t,$$

the Equation (24) can be written into matrix form as below:

$$\mathbf{KL} = \mathbf{Z} \tag{25}$$

5. CONCLUSION

The total magnetic fields are decomposed into the fields excited by the artificial source and the abnormal magnetic fields excited by the inhomogeneous median or topography, then the functional of the abnormal fields has been constructed. By using artificial boundary integral method, the open electromagnetic problem is localized. When item related to the boundary integration of abnormal fields on the artificial surface is neglected, the stiffness matrix becomes symmetric.

REFERENCES

1. Ruan, B. Y. and Y. X. Wang, "A boundary element modeling method for the electromagnetic field by artificial source in frequency domain with 3-Dtopography," *Chinese J. Geophys.*, (in Chinese), Vol. 48, No. 5, 1197–1204, 2005.
2. Nedelec, J. C., "Mixed finite element in $\mathbf{IR}^3$," *Numer. Math*, Vol. 35, 315–341, 1980.
3. Nedelec, J. C., "A new family of mixed finite element in $\mathbf{IR}^3$," *Numer. Math*, Vol. 50, 57–81, 1986.
4. Xie, G., J. Li, and etc., "3-D electromagnetic modeling and nonlinear inversion," *Geophysics*, Vol. 65, No. 3, 804–822, 1999.
5. Solin, P., *Partial Differential Equations and the Finite Element Method*, Wiley-Iinterscience, New York, 2006.
6. Solin, P., K. Segeth, and I. Dolezel, *Higher-Order Finite Element Methods*, Chapman & Hall/CRC Press, 2003.
7. Whitney, H., *Geometric Integration Theory*, Princeton University Press, Princeton, 1957.
8. Jin, J. M., *The Finite Element Method in Electromagnetics*, John Wiley & Sons, New York, 1993.

The Distributions of Electromagnetic Fields Excited by Magnetic Dipole on Ragged Surface

Baiyao Ruan and Zhifeng Xu
Guilin University of Technology, China

Abstract— In the electromagnetic surveying, the ragged topography usually affected those results. When designing work plans and analyzing field data, the topography effect on electromagnetic signs must be carefully considered. In this paper, the pure terrain effect on electromagnetic signs excited by artificial magnetic dipole was studied by boundary element method. By modeling results, the topography effects on electric and magnetic factors are different, but compensated. When the artificial source located at the appropriate position over the surface, the terrain effect on magnetic factors along the profile is almost the same.

1. INTRODUCTION

When carrying out electromagnetic surveying on ragged surface, for example, in electromagnetic surveying on surface or airborne electromagnetic prospecting, the electromagnetic signs are distorted by ragged terrain. In [1] and [2], Ruan put forward boundary element method to calculate the electromagnetic fields on complex surface excited by artificial source.

In this paper, the regular method is introduced to calculate the singular integral in the surface integral. By this method, the electromagnetic field distribution excited by artificial source along given profile is apparently discovered.

2. ELECTROMAGNETIC FIELD BOUNDARY INTEGRAL EQUATION

Supposed that the region Ω is closed by the boundary Γ and occupied by the homogeneous and isotropic media with conductivity σ and permittivity ε, permeability (μ) of medium is thought as the vacuum permeability, the time factor of electromagnetic fields is $e^{i\omega t}$, so the electromagnetic field on the surface satisfied the boundary integral equation as follows [1, 2]:

$$\begin{cases} \mathbf{H}(p) + \int_\Gamma \{H_n \nabla(\varphi_1 - \varphi_2) + (\mathbf{n}\times\mathbf{H}) \times \nabla(\varphi_1-\varphi_2) \\ \qquad +(\mathbf{n}\times\mathbf{E})[i\omega\varepsilon(\varphi_1-\varphi_2) - \sigma\varphi_2]\} d\Gamma = \mathbf{H}_p(p) \\ \mathbf{E}(p) + \alpha\mathbf{E}_n(p)\left(1-\dfrac{\omega_p}{4\pi}\right) + \int_\Gamma \{E_n[\nabla(\varphi_1-\varphi_2) - \alpha\nabla\varphi_2] \\ \qquad +(\mathbf{n}\times\mathbf{E})\times\nabla(\varphi_1-\varphi_2) - i\omega\mu(\mathbf{n}\times\mathbf{H})(\varphi_1-\varphi_2)\} d\Gamma = \mathbf{E}_p(p) \end{cases} \tag{1}$$

where $\mathbf{H}_n$ is normal component of magnetic fields and $\mathbf{E}_n$ is normal component of electric fields. By solving this boundary integral equation, the response of ragged surface can be got.

It is difficult to directly solve Equation (1), to solve the Equation (1) by numerical method, the surface is divided into M triangular elements, and the center in each triangular are taken as nodes, then Equation (1) is transformed into linear equation set as below:

$$\mathbf{K}\cdot\mathbf{H} = \mathbf{S} \tag{2}$$

where $\mathbf{H} = (H_{x1}, H_{y1}, H_{z1}, E_{x1}, E_{y1}, E_{z1}, \ldots, H_{xm}, H_{ym}, H_{zm}, E_{xm}, E_{ym}, E_{zm})^T$, $\mathbf{S}$ is vector relative to artificial source. $\mathbf{K}$ is a factor matrix got by fundamental solution. It is proved that $\mathbf{K}$ is a diagonally dominant matrix, so the Equation (2) can be solved by SSOR method.

3. SINGULAR INTEGRATION IN ELEMENT METHOD

When dealing the diagonal element in the factor matrix $\mathbf{K}$, the singular integration are usually met in the boundary method as below

$$\Gamma_1 = \iint_\Delta \frac{\exp(-ikr)}{r} ds \qquad\qquad \mathbf{\Gamma} = \iint_\Delta \nabla\left(\frac{\exp(-ikr)}{r}\right) ds \tag{3}$$

where the integral region has different shapes according to the elements. When taking the triangular element, the center of the triangular element is the integral function singular point. According to the center of the element, the triangular element is divided into three singular elements, then those integration are thought in those sub-triangular-elements.

Using coordinate transform, the local polar coordinate is related to every sub-triangular-element, so the integration Γ_1 can be written as below in the local polar coordinate:

$$\Gamma_1 = \int_0^{\theta_{\max}} \frac{1}{-ik} \left(\exp(-ikr(\theta)) - 1\right) d\theta \tag{4}$$

where the integration has no singularity.

In the global Descartes coordinate, supposed:

$$\mathbf{A} = \nabla \left(\frac{\exp(-ikr)}{r} \right) \tag{5}$$

So in the local polar coordinate, the components of $\mathbf{A}$ can be written as below:

$$A_i = f(r)\varphi_i(\theta) \tag{6}$$

where $f(r) = \dfrac{\exp(-ikr)}{r^2}(-ikr - 1)$, $\varphi_i(\theta) = a_{1'i}\cos\theta + a_{2'i}\sin\theta$

a_{ij} is factor related to coordinate transforming. Thus,

$$\Gamma_i = \int_0^{\theta_{\max}} F(r(\theta))\varphi_i(\theta)d\theta \tag{7}$$

where

$$F(r(\theta)) = \Phi_1(\theta) - \Phi_2(\theta) \tag{8}$$

and $\Phi_1(\theta) = \exp(-ikr(\theta)) - 1$, $\Phi_2(\theta) = \int_0^{r(\theta)} \dfrac{\exp(-ikr)}{r} dr$.

The integral function in $\Phi_2(\theta)$ at the point $r = 0$ is not apparently definite. In order to get the definite value, $\Phi_2(\theta)$ is regularized as below [3]:

$$\overline{\Phi}_2(\theta) = \int_0^{r(\theta)} \frac{\exp(-ikr) - H(1-r)}{r} dr \tag{9}$$

where $H(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases}$. Then the regularized $\overline{\Phi}_2(\theta)$ can be apparently written as below:

$$\overline{\Phi}_2(\theta) = \ln r(\theta) + \sum_{j=1}^{\infty} \frac{(-ikr(\theta))^j}{j!j} \tag{10}$$

Then

$$\Gamma_k = \int_0^{\theta_{\max}} (\exp(-ikr(\theta)) - 1)\varphi_k(\theta)d\theta - \int_0^{\theta_{\max}} \varphi_k(\theta)\ln r(\theta)d\theta - \sum_{j=1}^{\infty} \frac{(-ik)^j}{j!j} \int_0^{\theta_{\max}} \varphi_k(\theta)r^j(\theta)d\theta \tag{11}$$

where every item of the above equation can be integrated by Gauss numerical method, the infinite sum can be truncated according to the limitation of error. When $|a_j| < \varepsilon$, the sum can be stopped, where $a_j = \dfrac{(-ik)^j}{j!j} \displaystyle\int_0^{\theta_{\max}} \varphi_k(\theta)r^j(\theta)d\theta$, ε is special limitation of error.

4. RESULTS

In order to understand the redistribution of the electromagnetic fields due to the effect of the topography, the boundary model is taken as Figure 1, it is V-shape inclined valley, the source located at the point $(0, -80)$, the resistivity of the earth median is $100\,\Omega$m. The two factors related to electromagnetic fields are defined as below:

$$\eta_1 = \left| \frac{\mathbf{H}_t}{\mathbf{H}_f} \right|, \quad \eta_2 = \left| \frac{\mathbf{E}_t}{\mathbf{E}_f} \right| \tag{12}$$

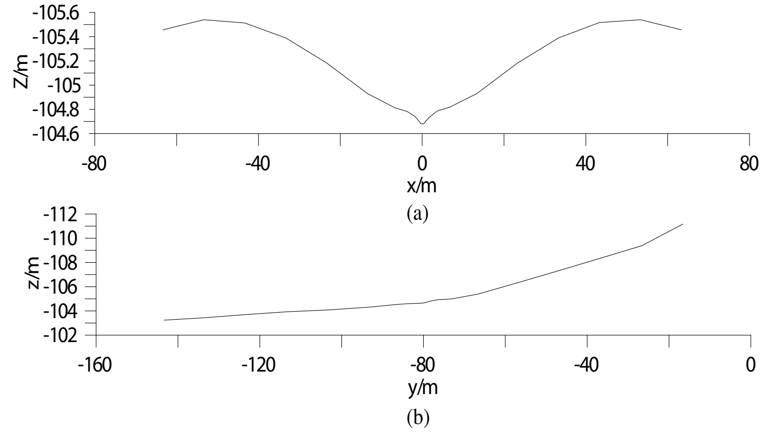


Figure 1: The profiles of the terrain along x and y direction.

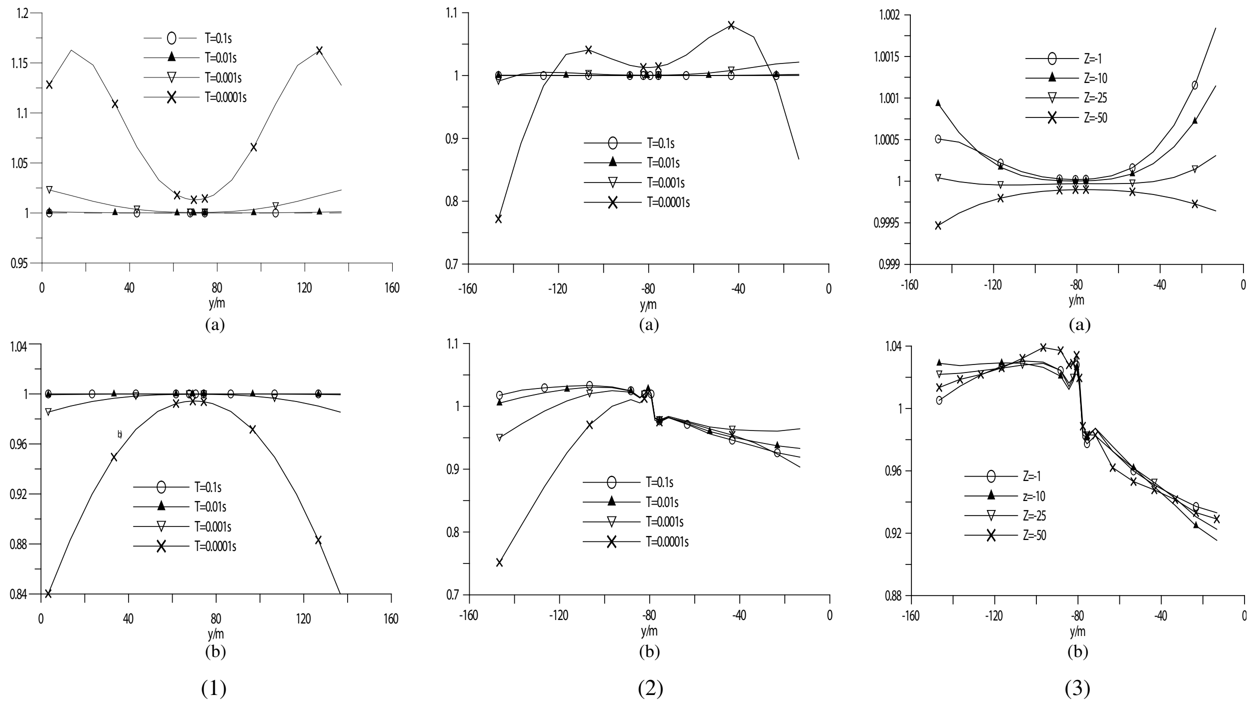


Figure 2: The distributions of electromagnetic factors excited by artificial source. (1) (a) The distribution of magnetic factor along the profile on horizontal surface; (b) The distribution of electric factor along the profile on horizontal surface. (2) (a) The distribution of magnetic factor along the profile on ragged surface; (b) The distribution of electric factor along the profile on ragged surface. (3) (a) The distribution along the profile on horizontal surface of magnetic factor excited by the source located at different position; (b) The distribution along the profile on horizontal surface of electric factor excited by the source located at different position.

where $\mathbf{E}_t$, $\mathbf{H}_t$ is the electromagnetic fields excited by the source above surface, $\mathbf{E}_f$, $\mathbf{H}_f$ is the electromagnetic fields excited by the same source in vacuum. Apparently, the two parameters are related to energy density of electric and magnetic field respectively.

Figure 2 displays the electromagnetic factor distribution of the different topography along the profile $x = -16.7\,\mathrm{m}$, the electromagnetic factors are excited by artificial magnetic dipole. The topographic profile $x = -16.7\,\mathrm{m}$ of ragged surface is same with Figure 1(b).

The electromagnetic factor distribution excited by different periodical sources along the profile in the horizontal surface shown in Figure 2(1). Because the magnetic dipole is symmetric, the electromagnetic factor along the profile is symmetric. In case of the long period, the topographic effect on electromagnetic factor is not big. When the period is short, the effect is strong. But the effects on electric and magnetic factors is compensated. According to Figure 2(1), magnetic fields are stronger than the exciting magnetic fields, but electric fields are weaker than exciting electric fields. By the view of energy transforming, the electric fields energy translates into magnetic energy. It is related to that electric fields excite current in the conductive media and the current excites the magnetic fields. In case of short period, the influence of dissipative media on the electromagnetic fields is apparent. When the distance between the source and the recording point increases, the electromagnetic factors always decrease.

Figure 2(2) illustrates the results of the relief surface. The compensation between electric and magnetic factor is very apparent. Due to the non-symmetry of the surface, the electromagnetic factors are already not symmetric. The electric factors in $y < -80\,\mathrm{m}$ are bigger than in $y > -80\,\mathrm{m}$, which is very apparent. For magnetic factors, the factors in $y < -80\,\mathrm{m}$ are smaller than in $y > -80\,\mathrm{m}$,but that is not apparent.

Figure 2(3) shows the distribution of the electromagnetic factors excited by the source located at the different position over the ragged surface. The period of exciting source is $T = 0.01\,\mathrm{s}$. According to the figure, the electric factor variation related to source altitude is not large, but the magnetic factors is. When the position of source is low, the magnetic factor variation along the profile is large. When position is high, the variation is small.

5. CONCLUSION

According to those above results, the topographic influence on the electromagnetic fields excited by the artificial magnetic dipole source is very big, and the effect on electric and magnetic is different, but compensated. When the source locates at appropriate position over the surface, the variation of the influence on magnetic factors along the given profile is almost the same.

REFERENCES

1. Ruan, B. and Y. Wang, "New topography inversion using EM field," *PIERS 2005*, 79–83, Hangzhou, China, August 2005.
2. Ruan, B. and Y. Wang, "A boundary element modeling method for the electromagnetic field by artificial source in frequency domain with 3-D topography," *Chinese J. Geophys.*, (in Chinese), Vol. 48, No. 5, 1197–1204, 2005.
3. Yu, D. H., *Mathematic Theory in Natural Boundary Method*, (in Chinese), Science Press, Beijing, 1993.

A 3D-2D AGILD EM Modeling and Inversion Imaging

J. Li[1], **G. Xie**[1], **M. Oristaglio**[2], **L. Xie**[1], **and F. Xie**[1]

[1]GL Geophysical Laboratory, USA
[2]Schlumberger-Doll Research, USA

Abstract— In this paper, we propose a 3D-2D Advanced Global Integral and Local Differential (AGILD) Electromagnetic (EM) modeling and inversion imaging algorithm in the cylindrical coordinate system. Transmitting sources are excited in the center hole and the EM field datas are received in many surrounded receiver holes. We use the 3D AGILD EM inversion and the above data configuration to make the multiple holes' image. Our algorithm is called as the 3D-2D AGILD EM modeling and inversion imaging. This algorithm effectively reduces the computational cost of the full 3D nonlinear inversion and increases the resolution of the 2D inversion. The multiple 2D cross holes' inversion iteration are developed. 3D AGILD EM modeling method reduces matrix solving cost and eliminates the error reflection on the artificial boundary. The 3D-2D AGILD EM inversion reduces the 3D inversion ill posed and the 3D full matrix solving cost. Many synthetic data inversion tests show that the 3D-2D multiple cross holes inversion can obtain high resolution image. Our 3D-2D AGILD EM modeling and inversion imaging method will be useful for oil exploration, earthquake exploration, geophysical engineering, environment characteristic monitoring, nondestructive testing, medical imaging, and material and nanometer materials sciences and engineering.

1. INTRODUCTION

There have been several research papers on the electromagnetic modeling and inversion and its application in the geophysics and other scientific and engineering fields. Habashy et al. proposed 'Beyond the Born and Rytov Approximation' in 1993 [1]. Habashy and Oristaglio proposed 'Simultaneous Nonlinear Reconstruction of Two-dimension Permittivity and Conductivity' in 1994 [2]. 'Some Uses (and abuses) of Reciprocity in Wave Field Inversion' was also proposed by Oristaglio and Habashy in 1996 [3]. Hohmann proposed 'Three-dimensional Induced Polarization and Electromagnetic Modeling' in 1975 [4]. We have proposed the Global Integral and Local Differential (GILD) EM modeling and inversion method since 1997. These papers have been published in the Geophysics [5], Physical D [6], SEG book, and PIERS conference proceedings [7–10]. In [5], we detailedly described the GILD EM modeling and inversion method. We proposed the advanced GILD modeling and inversion in PIERS 2005 in Hangzhou [7] and PIERS 2006 in Cambridge [8]. These methods are sorted as AGILD methods. The Stochastic AGILD algorithm is called as SAGILD EM modeling and inversion [9, 10]. According to the AGILD EM modeling and inversion, the 3D-2D AGILD EM modeling and inversion method for the imaging of multiple cross holes is proposed in this paper.

We consider the data site configuration in which the sources are in the center hole and the receivers are in the surrounded holes, as shown in Figure 1. This data site configuration is useful for oil geophysical exploration, environment engineering, and mine exploration. Similar data site is available for the medical detection, material study and nondestructive testing. When a laser source is excited in the center and micro and nanometer probes are in the surrounded sensors, we can construct a nano-device to study nanometer materials. We call this data configuration as Center Source and Multiple Receivers in the Surrounded site (CSMRS).

The general 3D EM inversion has high computational cost for the CSMRS configuration and the 2D and 2.5D EM inversion approaches can not interpret the spatial abnormality of 3D parameters in the rotational direction. In this paper the 3D-2D AGILD EM modeling and inversion imaging method can greatly reduce the computational cost, eliminate the error reflection on the artificial boundary, and obtain reasonable resolution. Using cylindrical coordinate system, the strip magnetic field differential integral equation in the boundary of cylindrical strip and center hole sub domain, and the Galerkin magnetic field equation in the internal cylindrical domain are coupled to construct 3D AGILD EM modeling. By restricting the 3D modeling data and Green's function into 2D cylindrical sectors, the 3D into 2D restricting strip magnetic field differential integral equation in the strip boundary cylindrical sector and the 3D into 2D restricting Galerkin magnetic field equation in the internal cylindrical sector are coupled to construct the 3D-2D AGILD EM inversion. Three adjacent 2D section AGILD EM inversions are coupled to construct the 3D-2D AGILD EM inversion iteration.

We use the 3D-2D AGILD EM inversion imaging method to make multiple holes' images. The synthetic data and field data interpreting tests show that the 3D-2D AGILD EM inversion has reasonable resolution of 3D parameter in the radial, vertical and rotational direction and reduces the computational cost.

The arrangement of this paper is as follows. Introduction has been described in Section 1. The 3D AGILD EM modeling in cylindrical coordinate is described in Section 2. In Section 3, we describe the 3D AGILD EM inversion. The 3D-2D AGILD multiple cross holes' inversion and their multiple cross holes data configuration are described in Section 4. The 3D-2D AGILD multiple cross hole images are presented in Section 5. In Section 6, we describe the conclusion.

2. 3D AGILD EM FIELD MODELING

2.1. The 3D Strip Magnetic Field Differential Integral Equation

By substituting the field and coordinate transformation between the rectangle and cylindrical coordinate system, we derive the 3D strip magnetic field differential integral equation [7] in the cylindrical coordinate system as follows

$$\begin{aligned}\boldsymbol{H}(\rho,\theta,z) =& \boldsymbol{H}_b(\rho,\theta,z)\\ &+\int\limits_{\Omega}\frac{((\sigma+i\omega\varepsilon)-(\sigma_b+i\omega\varepsilon_b))}{(\sigma+i\omega\varepsilon)}\boldsymbol{E}_b^M\left(\rho',\theta',z',\rho,\theta,z\right)(\nabla\times\boldsymbol{H})\rho' d\rho' d\theta' dz'\\ &+\int\limits_{\partial\Omega^-}\boldsymbol{E}_b^M\times\boldsymbol{H}\cdot d\vec{S}-\int\limits_{\partial\Omega^-}\frac{1}{(\sigma+i\omega\varepsilon)}\boldsymbol{H}_b^M\times(\nabla\times\boldsymbol{H})\cdot d\vec{S},\end{aligned} \tag{1}$$

where

$$\nabla\times\boldsymbol{H} = \nabla\times\boldsymbol{H}\left(\rho',\theta',z'\right), \tag{2}$$

$$\nabla\times\boldsymbol{H}(\rho,\theta,z) = \frac{1}{\rho}\begin{vmatrix}\vec{\rho} & \rho\vec{\theta} & \vec{z}\\ \dfrac{\partial}{\partial\rho} & \dfrac{\partial}{\partial\theta} & \dfrac{\partial}{\partial z}\\ H_\rho & \rho H_\theta & H_z\end{vmatrix}, \tag{3}$$

$\boldsymbol{E}$ is the electric field, $\boldsymbol{H}$ is the magnetic field, $\boldsymbol{E}_b^M$ and $\boldsymbol{H}_b^M$ are 3×3 Green's tensor function excited by the magnetic dipole source. $\boldsymbol{E}_b^M(r',r)$ has integrative singular at $r=r'$, and $r=(\rho,\theta,z)$ locates in the outside boundary of the strip or in the center with $\rho=0$. The r' locates in $\partial\Omega$-, the internal boundary of the strip, therefore, the 3D strip magnetic field differential integral equation has no coordinate singular at the pole $\rho'=0$. It only has weak integrative singular kernel.

2.2. The 3D Magnetic Field Galerkin Equation

The 3D magnetic field Galerkin equation in the rectangular coordinate system is as follows

$$\int\limits_{\Omega_e}\left(\left(\frac{1}{\sigma+i\omega\varepsilon}\nabla\times\boldsymbol{H}(r)\right)(\nabla\times\phi\boldsymbol{I}(r))+i\omega\mu\boldsymbol{H}(r)\phi(r)\right)dr=\int\limits_{\Omega_e}\boldsymbol{S}(r)\phi(r)dr. \tag{4}$$

By substituting the field and coordinate cylindrical transformation formulas into the above Galerkin equation (4), we have the 3D magnetic field Galerkin equation in the cylindrical coordinate [7, 8],

$$\begin{aligned}
&\int_{\Omega} \frac{1}{\sigma + i\omega\varepsilon} \left(\left(\frac{\partial H_\rho}{\partial z} - \frac{\partial H_z}{\partial \rho} \right) \frac{\partial \phi}{\partial z} - \left(\frac{\partial}{\partial \rho} \rho H_\theta - \frac{\partial H_\rho}{\partial \theta} \right) \frac{1}{\rho^2} \frac{\partial \phi}{\partial \theta} \right) \rho d\rho d\theta dz \\
&+ i\omega \int_{\Omega} \mu H_\rho \phi \rho d\rho d\theta dz = -i\omega \int_{\Omega} \mu M_\rho \phi \rho d\rho d\theta dz, \\
&\int_{\Omega} \frac{1}{\sigma + i\omega\varepsilon} \left(-\left(\frac{1}{\rho} \frac{\partial H_z}{\partial \theta} - \frac{\partial H_\theta}{\partial z} \right) \frac{\partial \phi}{\partial z} + \left(\frac{\partial}{\partial \rho} \rho H_\theta - \frac{\partial H_\rho}{\partial \theta} \right) \frac{1}{\rho^2} \frac{\partial \rho \phi}{\partial \rho} \right) \rho d\rho d\theta dz \\
&+ i\omega \int_{\Omega} \mu H_\theta \phi \rho d\rho d\theta dz = -i\omega \int_{\Omega} \mu M_\theta \phi \rho d\rho d\theta dz, \qquad (5) \\
&\int_{\Omega} \frac{1}{\sigma + i\omega\varepsilon} \left(\left(\frac{1}{\rho} \frac{\partial H_z}{\partial \theta} - \frac{\partial H_\theta}{\partial z} \right) \frac{1}{\rho} \frac{\partial \phi}{\partial \theta} - \frac{\partial \phi}{\partial \rho} \left(\frac{\partial H_\rho}{\partial z} - \frac{\partial H_z}{\partial \rho} \right) \right) \rho d\rho d\theta dz \\
&+ i\omega \int_{\Omega} \mu H_z \phi \rho d\rho d\theta dz = -i\omega \int_{\Omega} \mu M_z \phi \rho d\rho d\theta dz,
\end{aligned}$$

where ϕ is a trilinear basic test function.

2.3. 3D AGILD EM Field Modeling

The 3D cylindrical domain is divided into a set of cylindrical cube elements. Each element has 8 nodes $(\rho_j,\ \theta_j,\ z_j)$, $j = 1, 2, \ldots, 8$. The magnetic field differential integral equation (1) on the strip boundary zone and the center cylindrical sub domains and the Galerkin equations in the internal sub domains are coupled to construct the 3D AGILD EM field modeling [7, 8].

3. 3D AGILD EM INVERSION

3.1. The 3D Strip EM Parameter Variance Differential Integral Equation

The 3D magnetic field strip differential integral equation (1) in the cylindrical coordinate is the second type of integral equation for the magnetic field. From equation (1), we can derive the 3D strip EM parameter variance differential integral equation

$$\begin{aligned}
\delta \boldsymbol{H}(\rho, \theta, z) \; = & \int_{\partial\Omega^-} \frac{\delta(\sigma + i\omega\varepsilon)}{(\sigma + i\omega\varepsilon)^2} \boldsymbol{H}_b^M \times (\nabla \times \boldsymbol{H}) \cdot d\vec{S} \\
& - \int_{\Omega} \frac{\delta(\sigma + i\omega\varepsilon)}{(\sigma + i\omega\varepsilon)^2} \boldsymbol{E}_b^M \left(\rho', \theta', z', \rho, \theta, z\right) (\nabla \times \boldsymbol{H}) \rho' d\rho' d\theta' dz', \qquad (6)
\end{aligned}$$

where $\delta \boldsymbol{H}$ is the data misfit. $\delta\sigma$ is the conductivity variance, and $\delta\varepsilon$ is the dielectric variance. Equation (6) is the first type of integral equation for the conductivity and dielectric variances and is of the ill posed condition.

3.2. The 3D EM Parameter Variance Galerkin Equation

From the 3D magnetic field Galerkin equation (4), the 3D EM parameter variance Galerkin equation can be derived as follows,

$$\begin{aligned}
&\int_{\Omega_e} \left(\frac{\delta(\sigma + i\omega\varepsilon)}{(\sigma + i\omega\varepsilon)^2} \nabla \times \boldsymbol{H}(r) \right) (\nabla \times \phi \boldsymbol{I}(r)) dr = \\
&\int_{\Omega_e} \left(\left(\frac{1}{(\sigma + i\omega\varepsilon)} \nabla \times \delta \boldsymbol{H}(r) \right) (\nabla \times \phi \boldsymbol{I}(r)) + i\omega\mu\delta \boldsymbol{H}(r) \phi(r) \right) dr, \qquad (7)
\end{aligned}$$

where $\delta\sigma$ is the conductivity variance, $\delta\varepsilon$ is the dielectric variance, $\delta \boldsymbol{H}$ is the magnetic field variance, and $r = (\rho, \theta, z)$. The weak ill posed variance Galerkin equation (7) is for the conductivity and dielectric variances.

3.3. 3D AGILD EM Inversion

The 3D strip EM parameter variance differential integral equation (6) on the boundary strip and central cylindrical hole and the 3D EM parameter variance Galerkin equation on the internal subdomains are coupled to construct the 3D AGILD EM inversion [7].

4. 3D-2D AGILD MULTIPLE CROSS HOLES INVERSION

4.1. The Multiple Cross Holes' Data Configuration

In the geophysical exploration, oil exploration, and laser nanometer material detection, we describe the multiple cross holes' data configuration in the Figure 1 and Figure 2. The sources are excited in the center hole and the receivers are placed in the surrounded holes to receive the EM data. The receiver holes can be numbered. The sources and receivers data configuration is shown in Figure 2. The data configuration generates multiple cross hole sections, as shown in Figure 1.

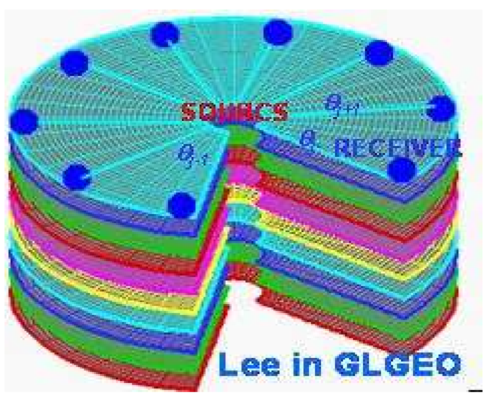


Figure 1: The multiple cross holes cylinder.

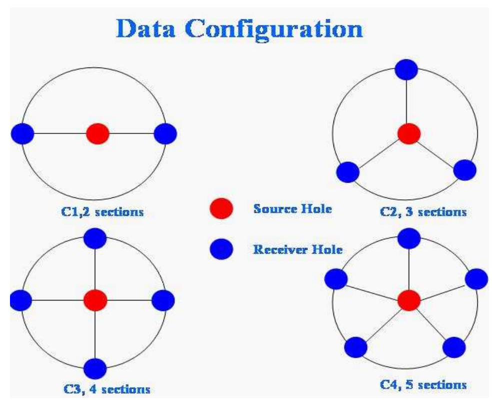


Figure 2: Multiple cross holes data configuration.

4.2. 3D-2D AGILD Multiple Cross Holes Inversion

To reduce the computational time cost, we propose the 3D-2D AGILD multiple cross holes inversion in this section. Because the transmitting sources are excited in the center cylindrical hole and data are received in the surrounded holes, we restrictively map the 3D AGILD EM inversion onto the 2D AGILD EM inversion on the three adjoining section sub domains with θ_{j-1}, θ_j, θ_{j+1}. Let $AGI(\delta\sigma, \delta\varepsilon; \boldsymbol{H}, \delta\boldsymbol{H}, \theta)$ be the AGILD inverse operation for the increment of conductivity $\delta\sigma$ and dielectric $\delta\varepsilon$, where the magnetic field $\boldsymbol{H}$ and its variance $\delta\boldsymbol{H}$ are given by previous iteration, and θ denotes the cross hole section with space variable ρ and z. By induction, suppose the $\delta\sigma_n$ and $\delta\varepsilon_n$ in the θ_{j-1} cross hole section are obtained in this iteration, the $\delta\sigma_{n-1}$ and $\delta\varepsilon_{n-1}$ in the θ_{j+1} cross hole section are obtained by the previous $(n-1)$th iteration. To find the $\delta\sigma_n$ and $\delta\varepsilon_n$ in the θ_j cross hole section in Figure 1, we solve the 3D-2D AGILD inversion

$$\begin{aligned}
&AGI\left(\delta\sigma_n, \sigma\varepsilon_n; \boldsymbol{H}_{n-1}, \delta\boldsymbol{H}_{n-1}, \theta_{j-1}\right) \\
&+AGI\left(\delta\sigma_n, \sigma\varepsilon_n; \boldsymbol{H}_{n-1}, \delta\boldsymbol{H}_{n-1}, \theta_{j}\right) \\
&+AGI\left(\delta\sigma_{n-1}, \delta\varepsilon_{n-1}; \boldsymbol{H}_{n-1}, \delta\boldsymbol{H}_{n-1}, \theta_{j-1}\right) = 0,
\end{aligned} \tag{8}$$

5. THE 3D-2D AGILD MULTIPLE CROSS HOLES' IMAGE

The cylindrical domain is divided into N cross holes sections. The 3D-2D AGILD EM inversion is used to recover the electric conductivity, dielectric, and magnetic permeability. In the strip boundary zone and center hole sub domain (Figure 1), we use collection Finite Element Method (CFEM) to discrete the magnetic field and parameter variance differential integral equation. The Galerkin FEM (GFEM) is used to discrete the Galerkin magnetic field and parameter variance equation in the internal sub domains. The strip boundary zone, center hole zone, and internal discrete FEM equation are coupled to construct the 3D-2D AGILD discrete inversion equation,

$$
\begin{aligned}
& AGI\left(\delta\sigma_n^h, \delta\varepsilon_n^h; \boldsymbol{H}_{n-1}^h, \delta\boldsymbol{H}_{n-1}^h, \theta_{j-1}\right) \\
& +AGI\left(\delta\sigma_n^h, \delta\varepsilon_n^h; \boldsymbol{H}_{n-1}^h, \delta\boldsymbol{H}_{n-1}^h, \theta_{j}\right) \\
& +AGI\left(\delta\sigma_{n-1}^h, \delta\varepsilon_{n-1}^h; \boldsymbol{H}_{n-1}^h, \delta\boldsymbol{H}_{n-1}^h, \theta_{j-1}\right) = 0,
\end{aligned} \tag{9}
$$

where $\delta\sigma_n^h$ is the piecewise FEM discretization of the $\delta\sigma_n$ in the nth iteration, the $\delta\varepsilon_n^h$, $\boldsymbol{H}_{n-1}^h$, and $\delta\boldsymbol{H}_{n-1}^h$ are the discretization of the $\delta\varepsilon_n$, $\boldsymbol{H}_{n-1}$, and $\delta\boldsymbol{H}_{n-1}$, respectively.

There are only unknown conductivity variance $\delta\sigma_n^h$ and $\delta\varepsilon_n^h$ in the cross hole section θ_j to be solved. There are 12×16, i. e., 172 elements in each cross hole section. The strip boundary zone full matrix and internal sparse matrix are fastly solved by the AGILD algorithm [7, 8]. After a few iteration steps of the 3D-2D AGILD inversion, we obtained reasonably high resolution cross hole images. The resistivity image in the cross sections 1, 3, 5, and 7 is plotted in Figures 4, 6, 8, and 10, respectively. The synthetic model is plotted in the left graph, the 3D-2D AGILD cross holes' image is plotted in the right graph. The magnetic field data in the cross sections 1, 3, 5, and 7 are plotted in the Figures 3, 5, 7, and 9, respectively, the amplitude of the field is plotted in the left graph, and the phase of the field is plotted in the right graph. The two frequencies are 9600 Hz and 12000 Hz. The regularizing parameter in this paper is 1.9835. For some field data, the regularizing parameter will be larger than 2.

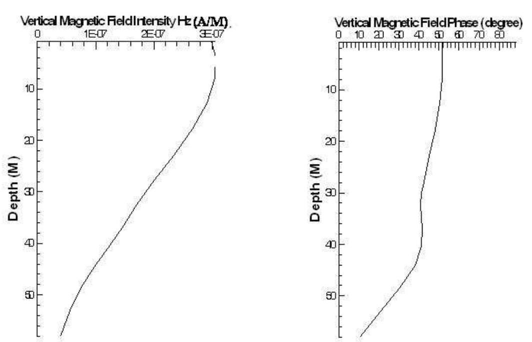


Figure 3: Magnetic field in the hole 1 which is excited by source 1, amplitude is in left, side phase is in right side in Figures 3, 5, 7, and 9.

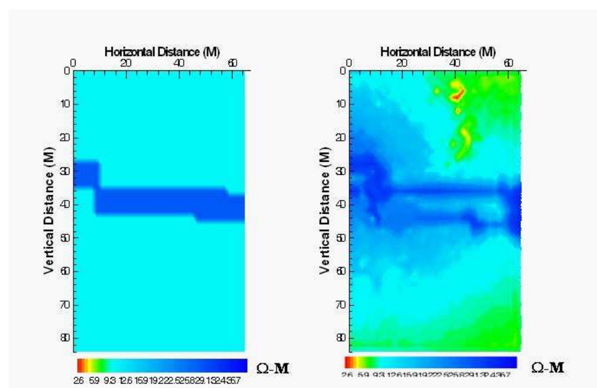


Figure 4: Resistivity imaging in cross section 1 by 3D-2D AGILD multiple holes' image software, model is in left, the cross hole image is in right.

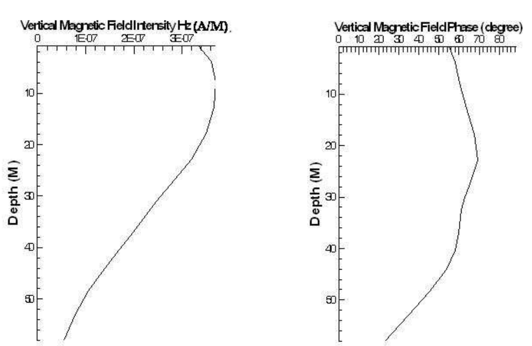


Figure 5: Magnetic field in the hole 3.

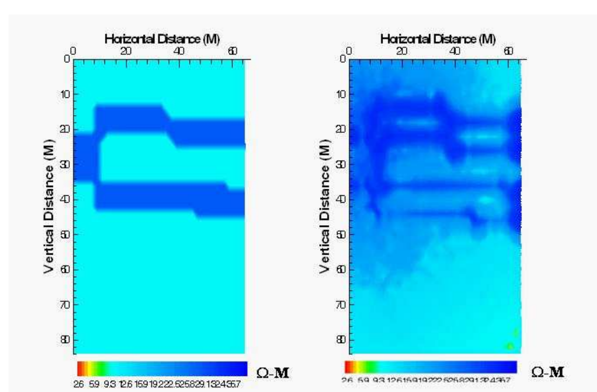


Figure 6: Resistivity imaging in cross section 3.

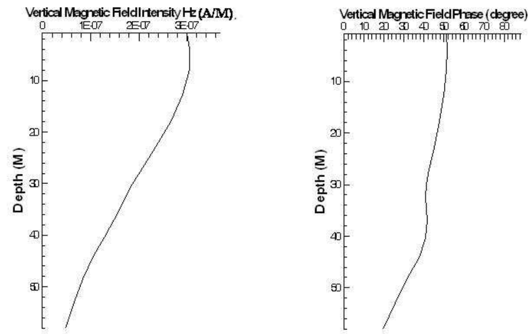


Figure 7: Magnetic field in the hole 5.

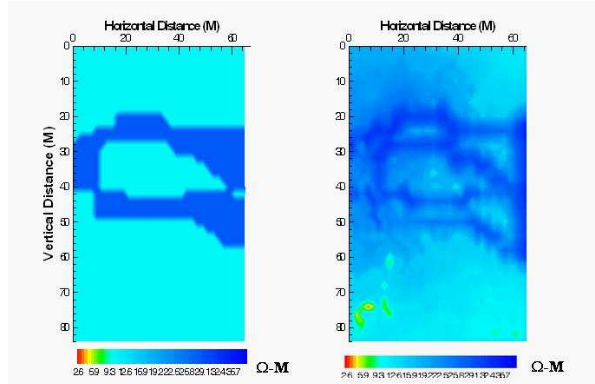


Figure 8: Resistivity imaging in cross section 5.

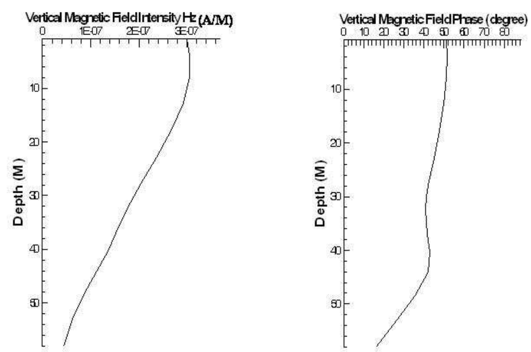


Figure 9: Magnetic field in the hole 7.

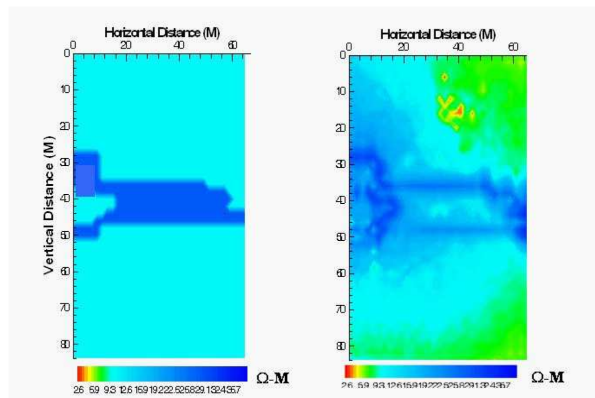


Figure 10: Resistivity imaging in cross section 7.

6. CONCLUSION

Many synthetic model datas and some field datas are interpreted by the 3D-2D AGILD multiple holes inversion. These images show that the 3D-3D AGILD multiple holes inversion method is fast, stable and has reasonably high resolution. The resistivity images in Figures 4, 6, 8, and 10 have obvious space variation in the θ direction, which shows the 3D-2D AGILD multiple inversion method has high resolution in the 3D space. The 3D-2D parallel AGILD inversion algorithm validation and the multiple cross holes imaging are performed by using 3D-2D AGILD software 32DAGILDINV which is made by Lee Xie in GLGEO. The cross checking validation between the 3D-2D AGILD inversion and GL modeling and inversion [11, 12] shows that 3D AGILD, 3D-2D AGILD, and GL EM modeling and inversion softwares are useful for oil exploration, earthquake exploration, geophysical engineering, environment characteristic monitoring, nondestructive testing, medical imaging, and material and nanometer materials sciences and engineering.

REFERENCES

1. Habashy, T. M., R. W. Groom, and R. R. Spies, "Beyond the Born and Rytov approximation," *J. Geophys. Res.*, Vol. 10, No. 98-B2, 1759–1775, 1993.
2. Habashy, T. M., M. L. Oristaglio, and A. T. de Hoop, "Simultaneous nonlinear reconstruction of two dimensional permittivity and conductivity," *Radio Sci.*, Vol. 29, 1101–1118, 1994.
3. Oristaglio, M. L. and T. M. Habashy, "Some uses (and abuses) of reciprocity in wave field inversion," *Wavefields and Reciprocity*, 1–22, P. M. van den Berg, H. Blok, and J. Fokkema (Eds.), Delft University Press, 1996.
4. Hohmann, G. W., "Three-dimensional induced polarization and electromagnetic modeling," *Geophysics*, Vol. 40, 309–324, 1975
5. Xie, G., J. Li, M. Majer, D. Zuo, and M. Oristaglio, "3D electromagnetic modeling and nonlinear inversion," *Geophysics*, Vol. 65, No. 3, 804–822, 2000.
6. Xie, G. and J. Li, "New parallel stochastic global integral and local differential equation modeling and inversion," *Physics D*, 133, 477–487.

7. Xie, G., J. Li, and F. Xie, "Advanced GILD EM modeling and inversion," *Proc. PIERS 2005*, 105, Hangzhou, 2005.
8. Xie, G., J. Li, and F. Xie, "2.5D AGLID EM modeling and inversion," *Proc. of PIERS 2006*, 390–394, Cambridge, 2006.
9. Li, J. and G. Xie, "A cubic-hole finite element for 3D resistivity modeling," *Three-dimensional Electromagnetics*, No. 7, 591–599, edited by Michael Oristaglio, Geophysical Developments.
10. Li, J., G. Xie, and F. Xie, "New stochastic AGLID EM modeling and inversion," *Proc. of PIERS 2006*, 490–494, Cambridge, 2006.
11. Xie, G., J. Li, and F. Xie, "New global and local electromagnetic field (GL) modeling and inversion," *Proc.of PIERS*, 68, Hangzhou, 2005.
12. Xie, G., F. Xie, L. Xie, and J. Li, "GL method and its advantages for resolving historical difficulties," *Progress In Electromagnetics Research, PIER*, 63, 141–152, and *Proc. of PIERS*, 306-311, Cambridge, 2006.

Application of CSAMT Method for Exploring Coal Mine in Fujian Province, Southeastern China

Zhiguo An and Qingyun Di

Institute of Geology and Geophysics, Chinese Academy of Sciences, China

Abstract— The coal resource reserve in Fujian province is the third of nine provinces in South China. In several decades, most areas with the distributing of uncovered and semi-covered coal stratum have already been explored. Therefore the exploration work in future should mainly be concealed coal stratum, i.e., under the red-bed, old stratum and volcanic. Because of the thick vegetation, grave hypsography and high deepness prospect target, CSAMT method is used for prospecting on eight prospecting lines in four representative coal stratums.

The coal stratum that can be excavated in this section is Lower Permian Series, Tungtzeyan Formation. The thickness of this Formation is more than 680 m. It is the medium coal-contain section in the coal-forming environment, and is the major coal-containing section in Lower Permian Series. The aim of this prospecting is to ascertain the upper and lower depth of the coal stratum, and the prospecting target depth is 1200 m. For this target depth, frequency range between 9600 and 1 Hz is used to get good interpretation, the vertical distance R from transmitter to the survey line is 8 km, the power supply current is 10 A, and voltage is 600 V. Transmitter and receiver time synchronization is controlled by GPS.

The difficulties of the CSAMT method in this project are terrain problems and serious static problems. Therefore, in the data processing, major works has been done for terrain correction and static correction, which can smooth inverted results for better interpretation.

After comparing the interpretation result of CSAMT data with the investigation result of engineer geology, it is found that there is distinct difference in dielectric properties between coal stratum, upper old stratum, red-bed and volcanic. The result shows that CSAMT method is helpful to identify coal stratum under this geological condition.

1. INTRODUCTION

This paper deals with the Controlled Source Audio-frequency Magnetotelluric method (CSAMT) applied to coal mine in the Fujian province, southeastern China. This is the very first time that CSAMT has been used for precious energy sources exploration in Fujian province. Coal is significant resource and material in China, so it plays a vital role in Chinese economic progress. With geological exploration on coal mine for several decades, the distributing areas of the uncovered and semi-covered coal stratum have already been detected and exploited. Nevertheless, according to the geological theory and borehole data, there is still concealed coal stratum, such as, the coal stratum under red-bed, old stratum and volcanic. Because of the thick vegetation, grave hypsography and high deepness prospect target, the conventional geophysical methods, including electric sounding and seismic methods, are not very useful for exploring coal in Fujian province.

Because of the considerable advantages of CSAMT sounding, consisting of great depth of detection, excellent lateral resolution, flexible survey design, and little topographic effect, it has been applied in Fujian area to understand more extended coal deposits. Based on the CSAMT sounding data in this survey, a total of eight survey lines and eight cross-section profiles have been made in four typical areas. According to the geological interpretation, the main factors governing the localization of the mineral bodies are resistivity and rock contacts. Fortunately, the coal stratums that can be excavated in this section belong to Lower Permian Series, Tungtzeyan Formation, whose electric property shows low resistivity, and the thickness of this Formation is more than 680 m, thus there is obvious difference between coal stratum and its upper stratums, which satisfies the electric foundation of carrying out CSAMT work in Fujian province. This field survey was carried out in four areas shown in Figure 1.

2. PRINCIPLE OF THE CSAMT METHOD

Theoretically, the CSAMT method is used to receive the E-field and the orthogonal H-field which are produced by the electromagnetic field. The electromagnetic field is induced by the alternating current source, which is transmitted into the subsurface by grounded dipole. For calculating the

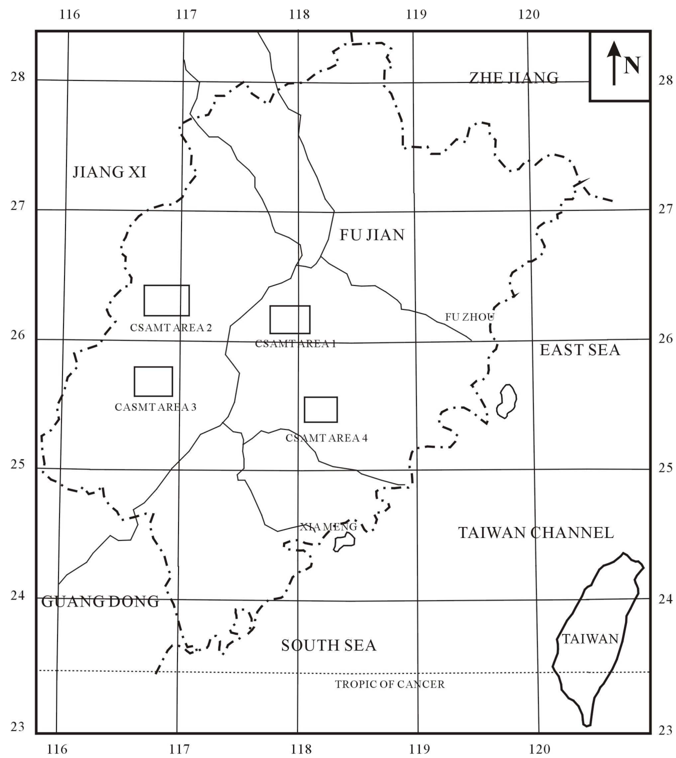


Figure 1: CSAMT survey areas in Fujian province. The four CSAMT areas locate in the center and west of Fujian province, and the blocks indicate four survey areas.

apparent resistivity, the Cagniard resistivity formula has been employed. Finally, the resistivity layers of each study area will be obtained by means of matching the field apparent resistivity curve with the model.

CSAMT measures the electromagnetic field that imparted to the earth by an electric dipole source shown in Figure 2. The source consists of a 1 to 2 km long insulated wire, grounded removable dipole at both ends that insert into the ground, a high-powered transmitter, and motor generator set also. The frequency of the signal is varied in order to vary the penetration depth of survey (the

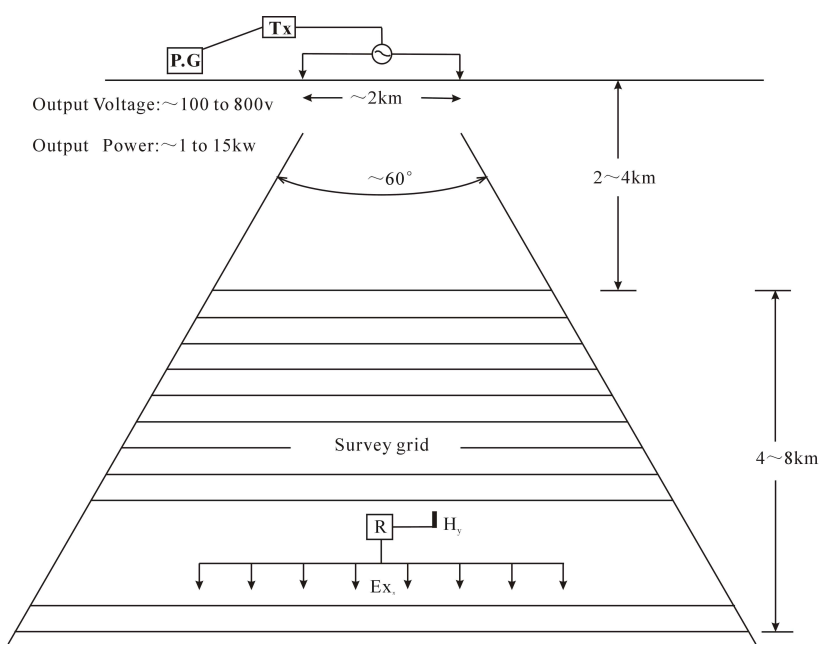


Figure 2: Geometry of CSAMT measurement.

lower the frequency, the deeper the penetration). The measurements are made in an area that is 2 to 4 km away from the source, where the source field in far zone approximates a plane wave.

The CSAMT method, introduced by Goldstein (1971) and Goldstein and Strangway (1975), is an electromagnetic sounding technique which measures resistivity variations in the ground. Its deep penetration and high resolution make it an excellent tool for exploring resource and engineer geological problems. It is also broadly applicable to exploration for base and strategic metals, geothermal and petroleum resources, and a variety of long deep tunnel studies.

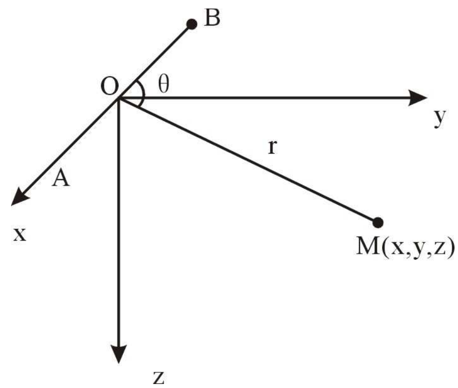


Figure 3: Horizontal dipole and coordinate geometry: θ: angle of AB and M point; o is the center of AB, r is the distance of AB and M; μ_0: permeability.

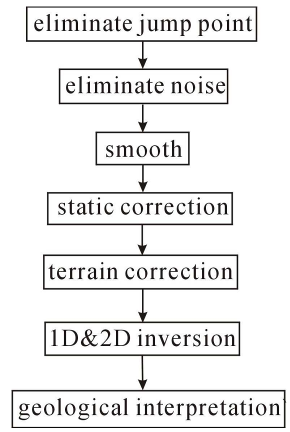


Figure 4: Data processing flow.

CSAMT measures the electric (E) and magnetic (H) components of the electromagnetic field. The E-field is sensed as a voltage across a grounded wire, while the H-field is sensed as a voltage in a high-gain magnetic antenna. Based on electromagnetic theory and Maxwell Equation, three components of E and H could be calculated, and then the ratio of horizontal perpendicular components, E_x and H_y, gives the Cagniard resistivity (Cagniard, 1953) of the ground with the coordination in Figure 3.

$$E_x = \frac{I \cdot AB \cdot \rho_1}{2\pi r^3} \cdot \left(3\cos^2\theta - 2\right) \tag{1}$$

$$E_y = \frac{3 \cdot I \cdot AB \cdot \rho_1}{4\pi r^3} \cdot \sin 2\theta \tag{2}$$

$$E_z = (i-1)\frac{I \cdot AB \cdot \rho_1}{2\pi r^2} \cdot \sqrt{\frac{\mu_0\omega}{2\rho_1}} \cdot \cos\theta \tag{3}$$

$$H_x = -(1+i)\frac{3I \cdot AB}{4\pi r^3} \cdot \sqrt{\frac{2\rho_1}{\mu_0\omega}} \cdot \cos\theta \cdot \sin\theta \tag{4}$$

$$H_y = (1+i)\frac{I \cdot AB}{4\pi r^3} \cdot \sqrt{\frac{2\rho_1}{\mu_0\omega}} \cdot \left(3\cos^2\theta - 2\right) \tag{5}$$

$$H_z = i\frac{3I \cdot AB \cdot \rho_1}{2\pi\mu_0\omega r^4} \cdot \sin\theta \tag{6}$$

the ratio of E_x and H_y:

$$\frac{E_x}{H_y} = \frac{\frac{IAB\rho_1}{2\pi r^3}\left(3\cos^2\theta - 2\right)}{(1+i)\frac{IAB}{4\pi r^3}\sqrt{\frac{2\rho_1}{\mu_0\omega}}\left(3\cos^2\theta - 2\right)} = \frac{2\rho_1}{(1+i)\sqrt{\frac{2\rho_1}{\mu_0\omega}}} \tag{7}$$

in brief:

$$\rho_s = \frac{1}{\mu_0\omega}\left|\frac{E_x}{H_y}\right|^2, \tag{8}$$

and then Cagniard resistivity

$$\rho_s = \frac{1}{5f}\frac{|E_x|^2}{|H_y|^2} \qquad (R \text{ in ohm-m})$$

Their phases yield another quantity known as phase difference or impedance phase:

$$P = E_{\text{phase}} - H_{\text{phase}} \quad (P \text{ in milliradian}),$$

The depth of CSAMT data is related to signal frequency f and to resistivity R:

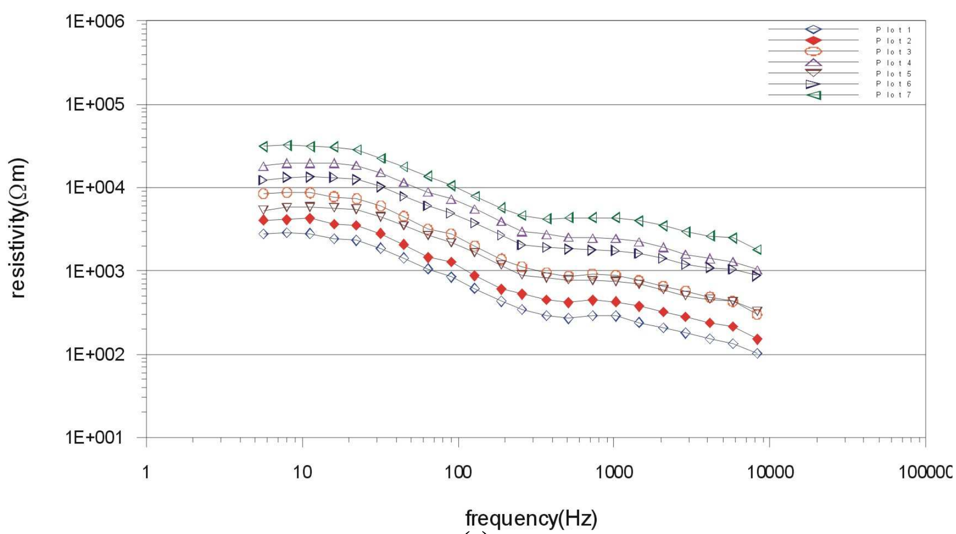
$$D = 502 \times \sqrt{R/f} \quad (D \text{ in meters})$$

CSAMT surveys generally take data in the range of 0.125 to 8192 Hz. Thus, in typical mining environments, penetration down to 1 or 2 km is easily achievable. Lateral resolution is controlled by the E-field dipole size, so CSAMT can use a very small electric field dipole to achieve high lateral resolution while still getting deep penetration by using low frequencies.

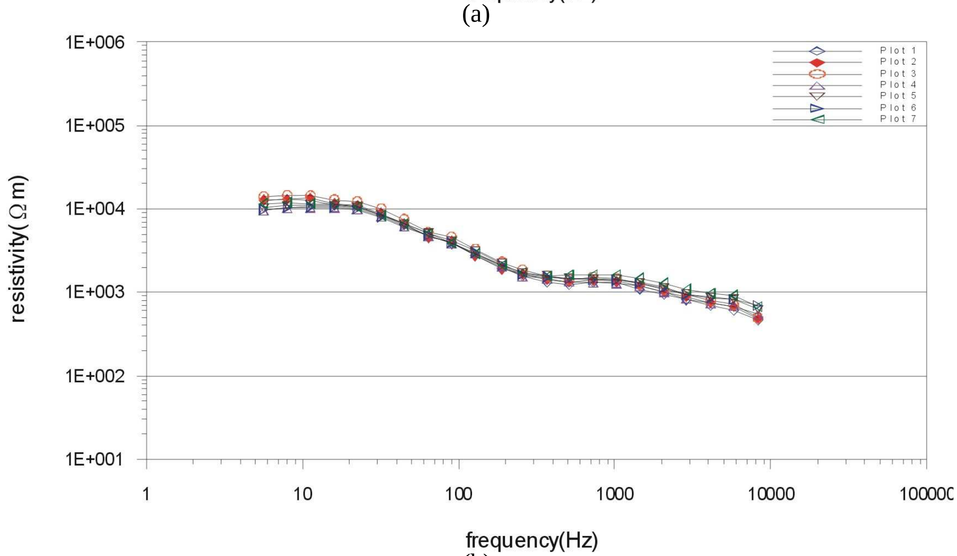
It is worth noting that CSAMT measurements are made far from the source, in the far-filed zone, so the sounding path is directed vertically into the ground, and penetration depends upon ground resistivity and signal frequency, not on receiver-transmitter separation. Therefore, CSAMT has inherently less topographic effect than traditional sounding methods.

3. DATA ACQUISITION AND PROCESSING

CSAMT is the abbreviation for controlled source audio-frequency magnetotellurics. The transmitter can transmit the electromagnetic signals with frequency from 8192 Hz to 0.125 Hz. The receiver at



(a)



(b)

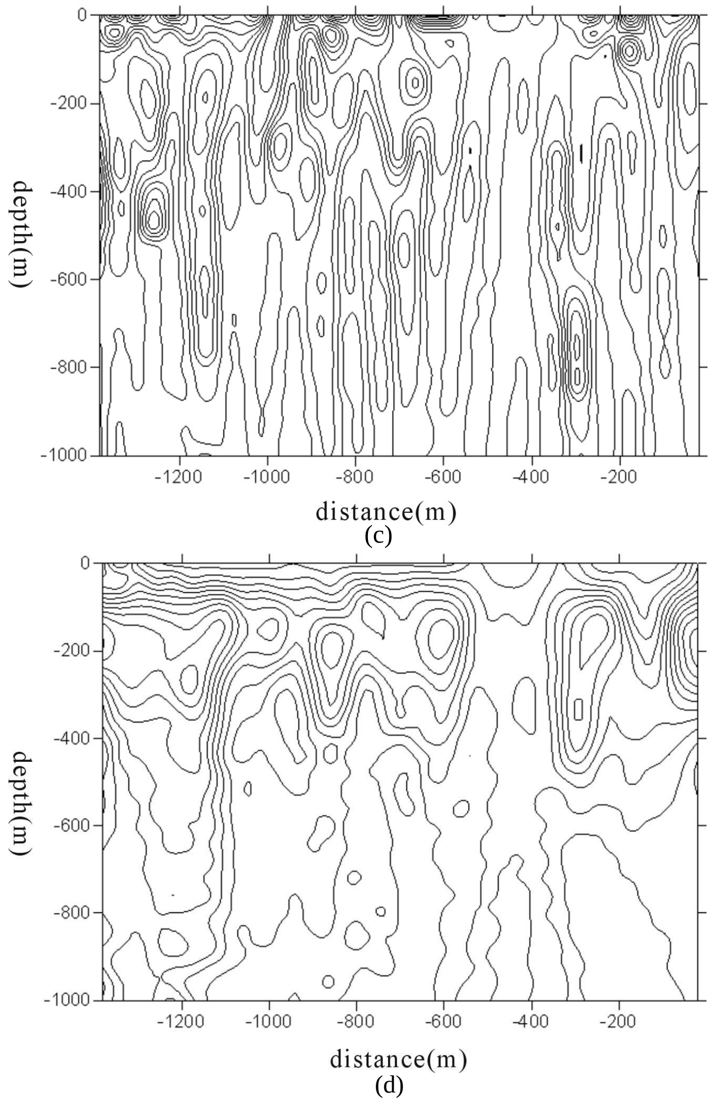


Figure 5: Comparison of before static correction and after static correction: (a) because of horizontal heterogeneity, static shift was produced, (b) after static correction, the data of adjacent seven points were coincident, (c) in the contour map, there was much strap anomaly for static correction, it could not be interpreted correctly, (d) with adopting Hanning filter, the static correction has obviously been reduced, which presented real underground structure.

the survey area will receive the electromagnetic signals, which are containing the subsurface geology information. After completing static correction and terrain correction, the data can be inverted, and then the result profiles for geology interpretation are plotted in Figure 4.

The instrument is V6A system made by Canada Phoenix Geophysics Ltd. Company. The distance between two adjacent survey stations is 40 m. The frequency used is from 9600 to 1 Hz. Because of exploration depth of 1500 m at this area, we adapt 4F30 automatic transmitting-receiving frequency series. Before we make the CSAMT measurements, the calibration of box and coil should be done for the coincidence and correction.

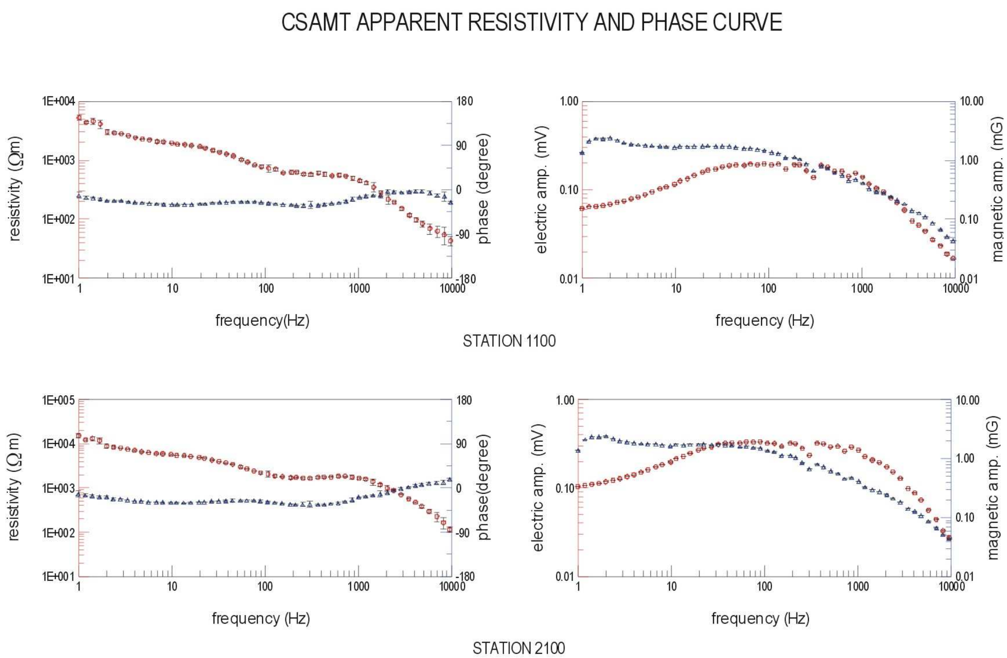
3.1. Static Correction

It is well known that one of the disadvantages of the CSAMT method is static shift, so under such circumstances the calculated apparent resistivity must be corrected for static effect. As for CSAMT method, just like other electromagnetic methods, static problem is usually common because the horizontal heterogeneity near the surface may bring the static effect. There are several methods to induce the static effect. To compensate for this effect, it is necessary to use some regularizing algorithms, in simple cases, which is accomplished with low-pass filtering of the field data.

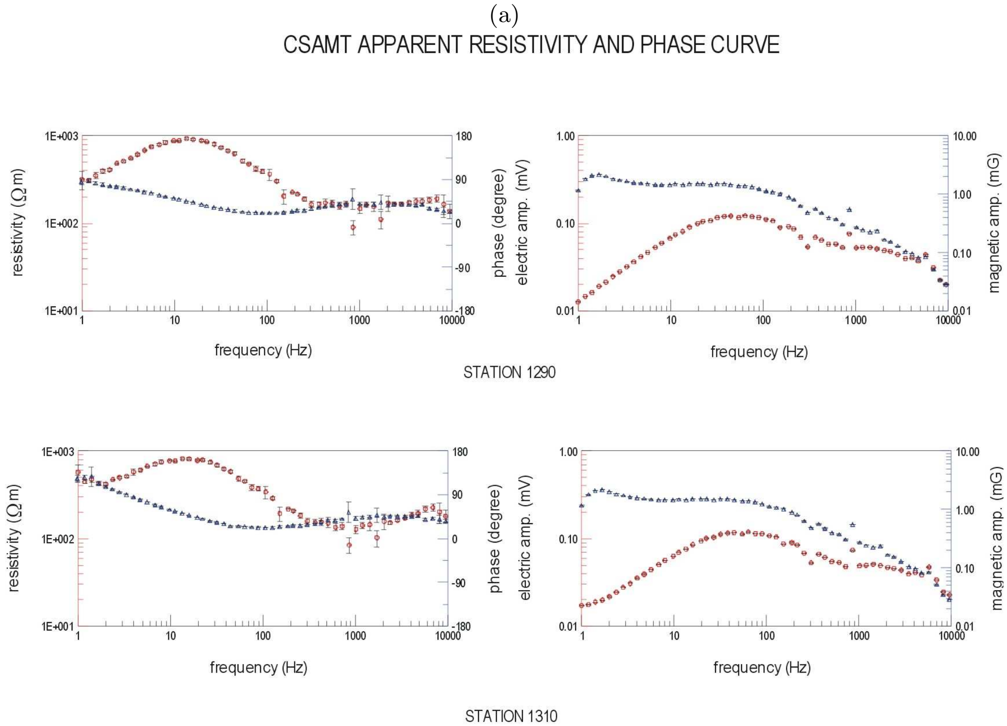
The filtering used here is Hanning window filtering:

$$h(x) = \begin{cases} \frac{1}{\omega}\left(1+\cos\frac{2\pi x}{\omega}\right) & |x| \le \omega/2 \\ 0 & |x| > \omega/2| \end{cases}$$

where ω is the filtering window width. In practice, $h(x)$ is dispersed to seven point filtering. Now take actual data for example, as follows in Figure 5.



(a)



(b)

Figure 6: Comparison of normal stratum and coal stratum about electric and magnetic field: (a) the two maps are CSAMT apparent resistivity and phase curves on normal stratum, respectively, (b) the two maps are CSAMT apparent resistivity and phase curves on coal stratum.

3.2. Terrain Correction

Two techniques are used to carry out topography correction. One is space correction about survey points, and the other is 2-D topography correction with using forward modeling.

The horizontal distance of two electrodes is fixed in the engineering exploration. The slope distance between adjacent survey points will be not equal because of different slope angles. Therefore in order to be processed and interpreted well, the original data of electric and magnetic field should be corrected.

3.3. Analysis of Typical Curve

The resistivity curve should be analyzed before geological interpretation, and the electric difference between normal stratum and coal stratum should be distinguished clearly as shown in Figure 6. Figure 6(a) are the two maps of CSAMT apparent resistivity and phase curves on normal stratum, respectively. The resistivity value is about 100 Ωm in high frequency. With frequency falling, the resistivity value is gradually growing, and then the transition zone, where the resistivity values have relatively small change, appears in about 200 Hz. Finally the frequency smaller, the resistivity bigger, the reisitivity-frequency curve is rising. The two points are on behalf of typically normal stratum; Figure 6(b) are the two maps of CSAMT apparent resistivity and phase curves on coal stratum. According to the resistivity-frequency curves, the resistivity values are still about 150 Ωm. However, with the frequency falling, the resistivity value has not bigger change till 100 Hz. At this point, the resistivity-frequency curve gradually begins to rise, and there is a summit in about 10 Hz, eventually the resistivity values start to decrease. There may be a distinct electric interface, so the conductive zone should be inferred. Based on the geological data and borehole information, the inferred result has been confirmed. In Figure 6, the red-hollow round represents electric field, and the blue triangle represents magnetic field.

3.4. Fitness Analysis

In order to give the reasonable geoelectric interpretation, the inverted result should be combined with original resistivity-frequency data, and fitness curve should be given to confirm the rationality of inverted result. So there are 22 frequency points listed in the Table 1. With the comparison of original resistivity in Table1 and inverted data in Figure 7, the fitness difference is relatively small, and the shape of inverted data is also similar to the shape of original data. Therefore we think that the inverted data is correct and believable, which could be used to have geological interpretation. In Figure 7, blue line expresses inverted data, and original data is respectively represented by red-hollow round.

Table 1: Fitness of original data and inverted data.

frequency	Ori-res	Inv-res	fitness	frequency	Ori-res	Inv-res	fitness
1.0	574.2	542.4	5.5%	128.0	288.6	278.7	3.4%
2.0	470.0	460.0	2.1%	189.5	218.4	214.3	1.9%
4.0	580.5	547.7	5.7%	256.0	184.0	181.5	1.4%
8.0	750.3	708.7	5.5%	369.2	152.2	150.3	1.2%
11.3	768.4	726.8	5.4%	512.0	149.6	146.2	2.3%
16.0	809.3	766.9	5.2%	724.5	138.4	135.0	2.5%
22.4	788.4	749.3	5.0%	1024.0	127.5	124.9	2.1%
32.0	686.2	655.8	4.4%	1440.0	144.5	140.5	2.8%
44.9	583.7	562.2	3.7%	2057.0	157.4	152.9	2.9%
64.0	454.3	443.4	2.4%	2880.0	162.8	158.5	2.7%
90.0	369.9	366.1	1.0%	4114.0	183.8	178.0	3.2%

4. RESULTS AND GEOLOGY EXPLORATION

We employed the inversion contour maps to display the qualitative results, which were three profile maps shown in Figure 9, Figure 11 and Figure 14 to interpret quantitatively and estimate coal

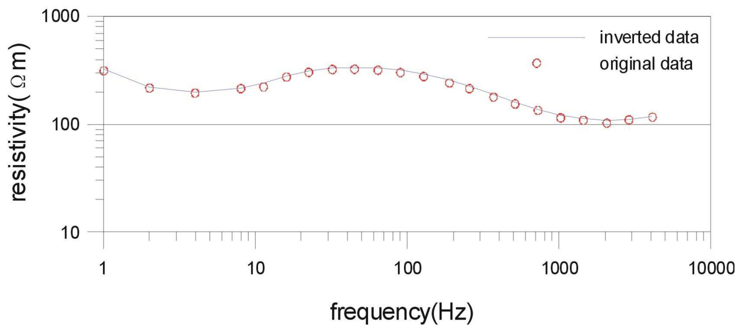


Figure 7: Curve comparison of original data and inverted data.

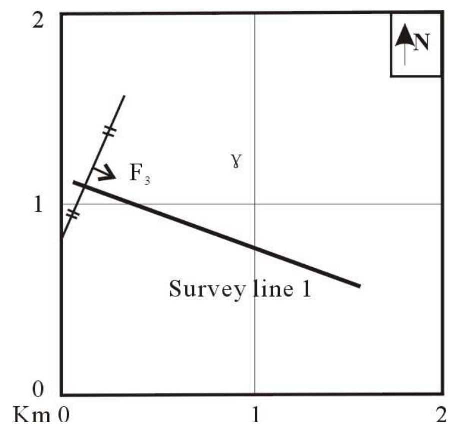
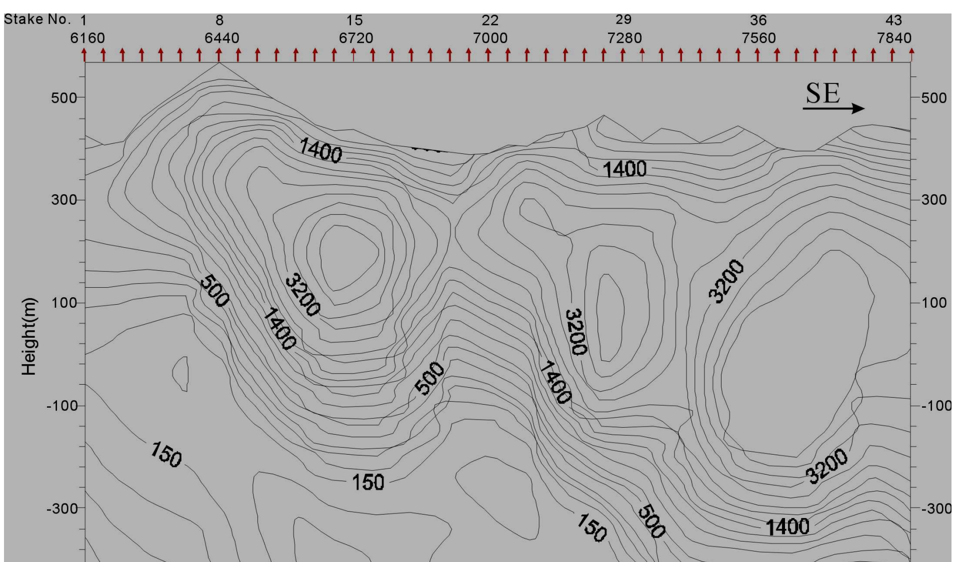


Figure 8: The trend of survey line 1.

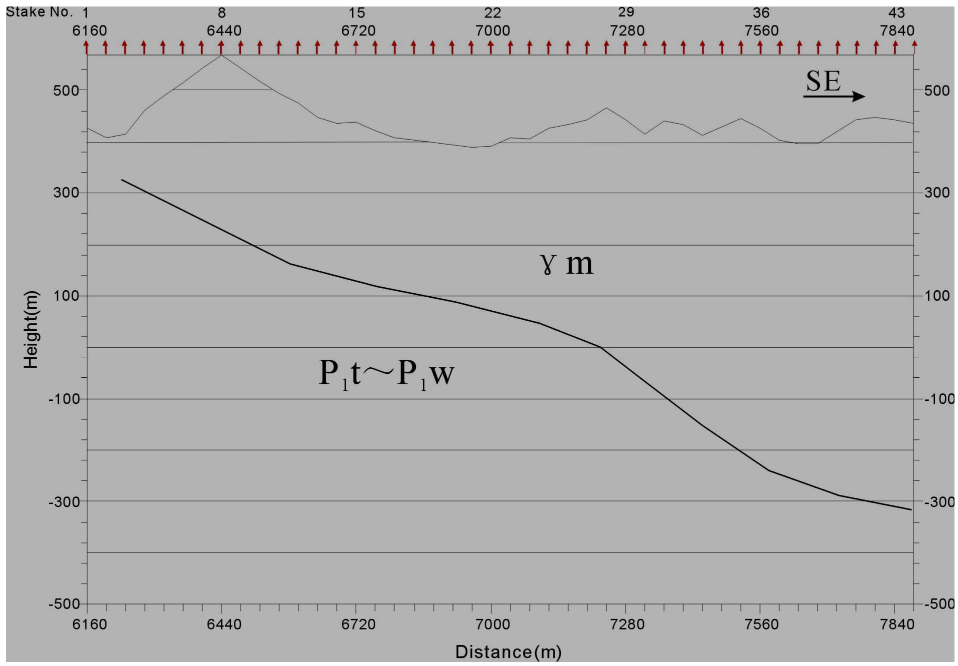
stratum in the survey areas.

To assess the subsurface qualitatively, three apparent resistivity contour maps with selected frequency respectively were drawn in Figure 9(a), Figure 11(a) and Figure 14(a). Figures 9(b), 11(b) and 14(b) depicted the final results along with the known geologic data also drawn to facilitate comparisons.

Three profiles were selected from the inversion results of the CSAMT data, which were prepared



(a)



(b)

Figure 9: Inverted contour map of line 1 and its geologic interpretation: (a) inverted counter map, (b) geologic interpretation. Horizontal coordinate indicates sounding points' position. Vertical coordinate depicts elevation. The interpreted positions of layers and their resistivity values in Ωm are also shown in this figure.

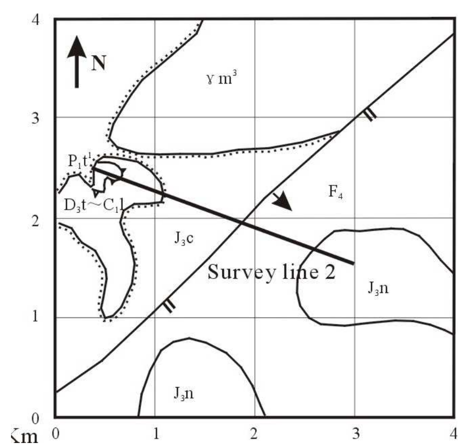


Figure 10: The trend of survey line 2.

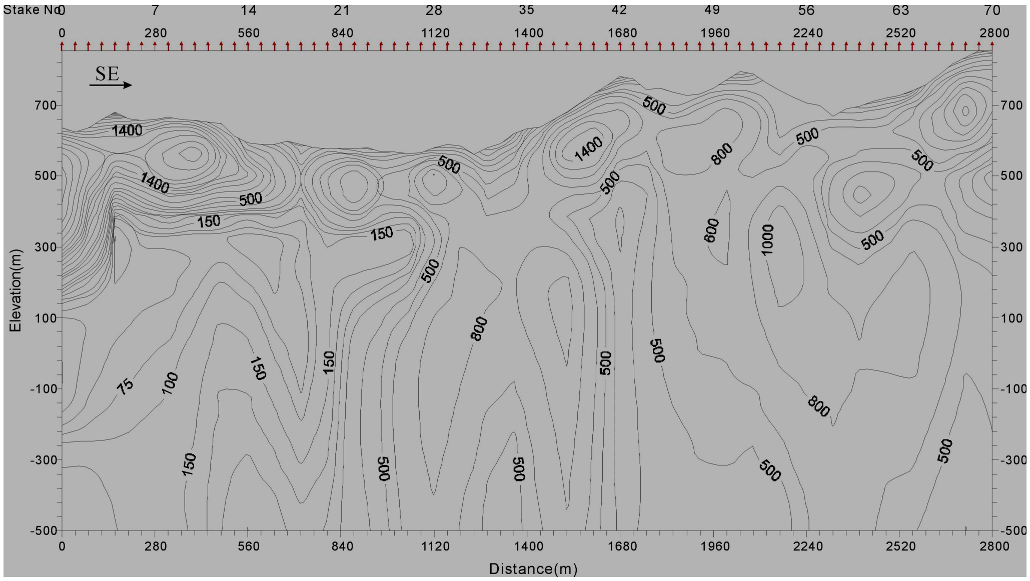
for interpretation of subsurface structure quantitatively, and to conform whether the difference exists between coal stratum and three other stratum including rhyolite, red-bed and old stratum or not, and whether the coal stratum exists in these stratums. The inversion profiles were shown as follows:

4.1. Profile A

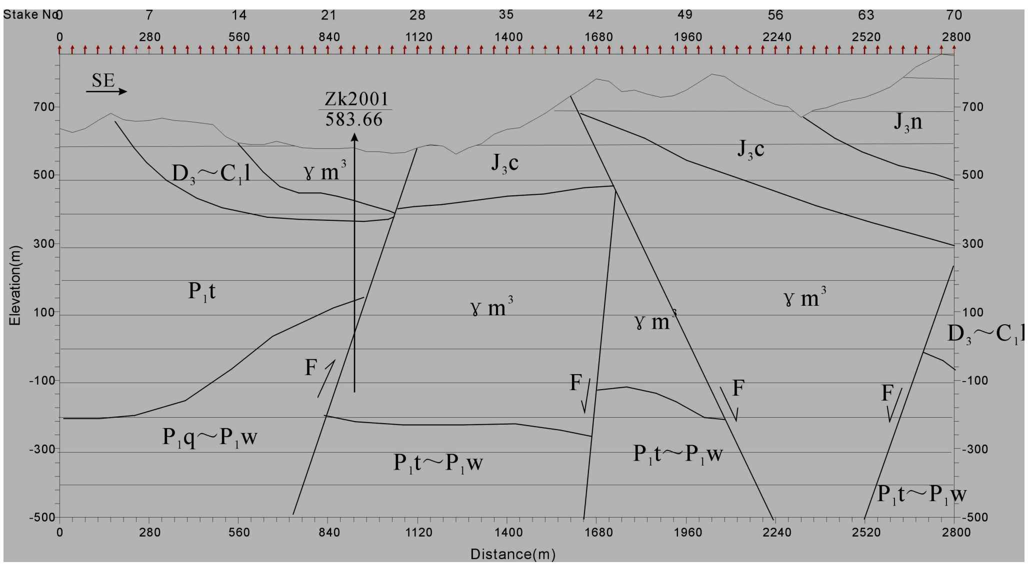
The 1.72 km long line selected, trending SE, was surveyed in CSAMT area one in Figure 8. To interpret the inversion result and compare geologic data well, we took the following map selected as an example in Figure 9. As for our goal, this was one of the relatively typical exploring coals under volcanic rock. This kind of exploring coals under the volcanic rock was relatively typical in Fujian province.

Based on geologic data, volcanic rock existed in this area, and there might be coal resource.

First of all, obviously, there was resistivity difference in the contour map shown in Figure 9(a). That was to say, there was a clear resistivity interface and conductive zone covered by resistive zone. Secondly, from the known geologic data map in Figure 9(b), an interface of rock properties was clearly visible in the geologic interpretation map, and upper stratum was volcanic (γm), while covered one was potential coal ($P_1t \sim P_1w$). Therefore, Figures 6 and 7 illustrated the results of CSAMT sounding on this profile well.



(a)



(b)

Figure 11: Inverted contour map of line 2 and its geologic interpretation: (a) inverted counter map, (b) geologic interpretation. Horizontal coordinate presents sounding points' position. Vertical coordinate expresses elevation. The interpreted positions of layers and their resistivity values in Ωm are also shown in this figure.

4.2. Profile B

This 2.8 km-long SE-trending profile in Figure 10 was the key profile to explore coal and verify the effect of CSAMT method as it ran through the outcrop of coal stratum in CSAMT area two. Figure 11 showed the inverted section derived from CSAMT sounding data. As follows:

Part 1: from 0 m to 240 m, there were clear properties, surface with higher resistivity and deeper rock with lower resistivity. Combined with outcrops and geologic data near the borehole, the lower resistivity zone was inferred as coal stratum ($P_1t \sim P_1w$).

Part 2 : from 240 m to 1200 m, resistivity properties: 200 m-thick higher-resistivity layer covered conductive layer, whose resistvity was lower than 200 Ωm. The resistive zone near the surface was interpreted as D_3-C_1l and γm. Their resistivity was above 1000 Ωm, however, there was obviously lateral discontinuity. Based on contact relationship, the resistive zone was interpreted as remainder

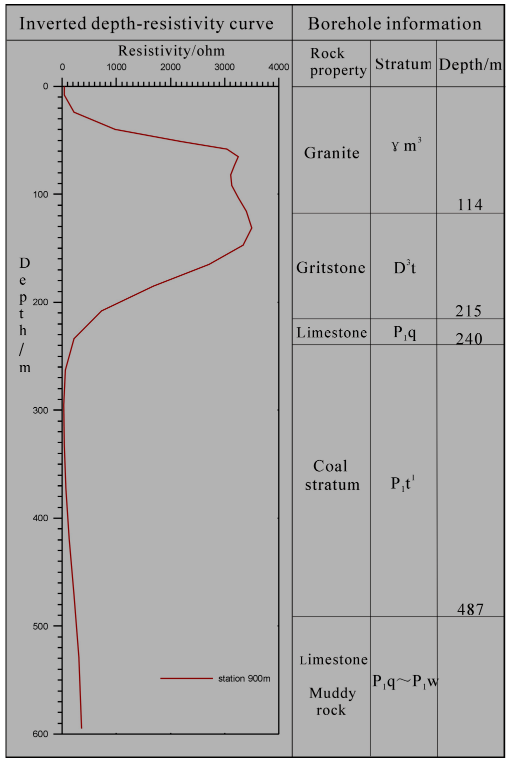


Figure 12: Comparison between inverted result and borehole information. The red line is the curve of inverted resistivity and depth in borehole ZK2001. The curve displays the geoelectric trend of actual underground components.

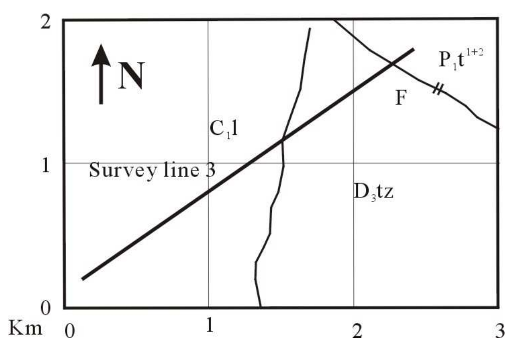


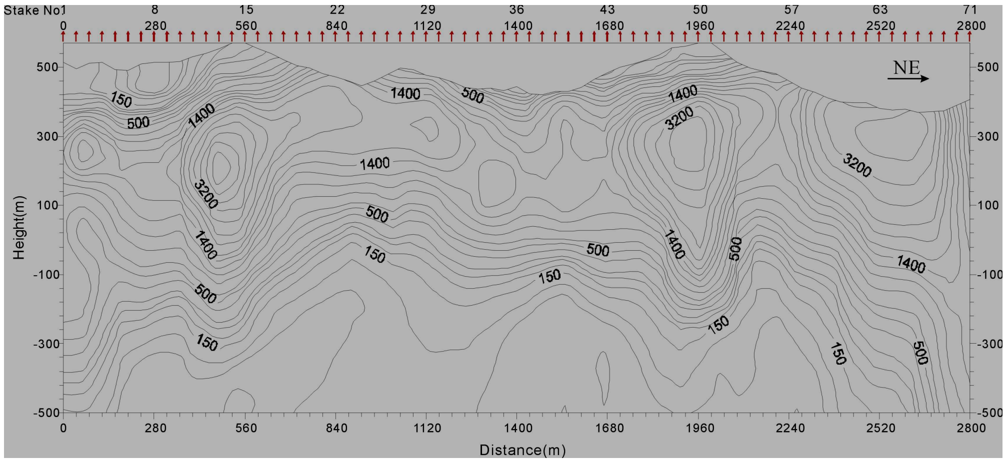
Figure 13: The trend of survey line 3.

cover, and the rock interface from 200 m to 240 m has been deduced fault. And deeper conductive zone was about 700 m thick, approximately sloping to small point with 40° angle. In combination with geologic data, the zone was inferred as coal stratum $(P_1t \sim P_1w)$. The resistivity of deeper layer was getting clearly high, and was inferred as P_1q. There was a borehole in about 900 m, our inverted result in Figure 12 conformed the geology information very well.

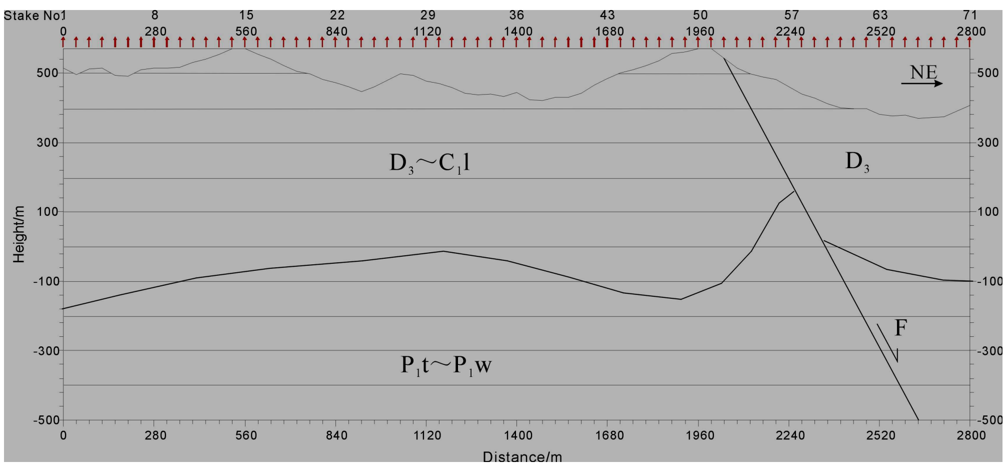
Part 3: from 1200 m to 2800 m, the main resistivity property as follows: upper middle-high layer (about 1000 Ωm), central thick zone of middle-high resistivity (about 800 Ωm), and then deeper stratum of lower resistivity. Near 1680 m point, there was clear low-resistivity strap, inferred as a fault. Based on relative geologic data, the main stratum series was interpreted as upper stratum J_3, center D_3-C_1l, and deeper stratum may be local remainder coal with bigger depth.

4.3. Profile C

The 2.8 km-long line, also trending NE, was surveyed in CSAMT area 3 shown in Figure 13. Figure 14 showed the results of CSAMT sounding on the profile. The inverted resistivity along the line had several prominent features. Near 2240 m point there was clear resistivity discontinuity, inferred as a fault. The whole profile had strong resistive zone in the central depth, then the deeper the elevation, the lower the resistivity. The wide conductive zone was interpreted as coal stratum ($P_1t \sim P_1w$), the low-resistivity structure could represent that coal stratum had clear difference with upper cover and confirmed availability of CSAMT method. However, the coal stratum interpreted should be approved by drilling.



(a)



(b)

Figure 14: Inverted contour map of line 3 and its geologic interpretation: (a) inverted counter map, (b) geologic interpretation. Horizontal coordinate indicates sounding points' position. Vertical coordinate depicts elevation .The interpreted positions of layers and their resistivity values in Ωm are also shown in this figure.

5. DISCUSSION AND CONCLUSION

The CSAMT method is a new technique for exploring resource. This method shows its unique advantages, and higher efficiency than other traditional electric sounding methods in certain areas, especially where other geophysical methods could not work efficiently. These places include the areas with thick vegetation, grave hypsography and high-depth prospecting target, which can be verified with the carried-out coal based on this research in Fujian.

On the results of the CSAMT soundings carried out in Fujian area, several low-resistivity anomalies are discovered, but our attention on this area is relatively small, so the conductive zones can not be connected together to the whole area.

The static correction and terrain correction are two types of difficulties for us in the course of data processing. We are confronted with great challenge in application because of surface electric discontinuity and complicated terrain. In processing some factors may be got over, but other features has still not eliminated. Therefore multi-resolution would occur. With an awful terrain like big cliff and deep valley, power lines and villages near the survey line, magnetic fields of part curves seriously have been influenced, so that they make it difficult for interpreting.

The results presented in the paper do show the potential of CSAMT method to be more useful for deep geologic mapping. Especially, it is useful for the detection of lateral boundaries, such as faults and fractures, where no detailed geological information is available, because not only the bedrock geology is concealed by overburden, but also borehole surveys have not yet begun.

ACKNOWLEDGMENT

The authors and their research group would like to acknowledge financial support from the 973 project 2002CB412702. The constructive review comments of Professor Ganquan Xie improved the quality of this manuscript.

REFERENCES

1. Basokur, A. T., T. M. Rasmussen, C. Kaya, Y. Altun, and K. Aktas, "Comparison of induced-polarisation and controlled-source audio-magnctotellurics methods for massive chalcopyrite exploration in a volcanic area," *Geophysics*, Vol. 62, No. 6, 1087–1096, 1997.
2. Batrel, L. C. and R. D. Jacobson, "Results of a controlled-source audiofrequency magnetotelluric survey at the Puhimau thermal area," *Geophysics*, Vol. 52, No. 4, 665–677, Kilauea Volcano, Hawaii, 1987.
3. Boerner, D. E., J. A. Wright, J. G. Thurlow, and L. E. Reed, "Tensor CSAMT studies at the Buchans Mine in central Newfoundland," *Geophysics*, Vol. 58, No. 1, 12–19, 1993.
4. Chen, C. S., "Application of CSAMT method for gold-copper deposits," *TAO*, Vol. 4, No. 4, 339–250, Chinkuashih Area, Northern Taiwan, 1993.
5. Di, Q. Y., M. Y. Wang, and K. F. S, "CSAMT research survey for preventing water flooding in mining," *Geophysica Sinica*, Vol. 45, No. 5, 1–5, 2002.
6. Eugene, A. B., E. E. Mark, G. A. Newman, and O. Biro, "Finite-element analysis of controlled-source electromagnetic induction using Coulomb-gauged potentials," *Geophysics*, Vol. 66, No. 3, 786–799, 2001.
7. He, J. S., *Control Source Audio Frequency Magnetotellurics Method*, Chinese Industry University Press, Changsha, 1990.
8. He, J. S., "The development and prospecting of electric method," *Geophysica Sinica*, Vol. 40, No. Supp. l, 308–316, 1997, (in Chinese).
9. Jiracek, G. R., "Near surface and topographic distortions in topographic induction," *Survey in Geophysics*, Vol. 11, No. 1, 163–203, 1990.
10. Liu, G. D., "The development of MT in China," *Geophysica Sinica*, Vol. 37, No. Supp. l, 301–310, 1994, (in Chinese).
11. Routh, P. S. and D. W. Oldenburg, "Inversion of controlled-source audio-frequency magnetotelluric data for a horizontally-layered earth," *Expanded Abstracts, 66th Ann. Internat. Mtg., Soc. Expl. Geophys.*, 1996.
12. Sandberg, S. K. and G. W. Hohmann, "Controlled-source audio-magnetotellurics in geothermal exploration," *Geophysics*, Vol. 47, No. 1, 100–116, 1982.
13. Sasaki, Y., Y. Yoshihiro, and K. Matsuo, "Resistivity imaging of controlled-source audiofrequency magnetotelluric dada," *Geophysics*, Vol. 57, No. 3, 952–955, 1992.
14. Shi, K. F, "New methods and technics of geophysical exploration, exploration technics and integrated evaluation method for ground water exploration system," Science Press, Beijing, 135–205, 1992.
15. Takeshhi, K., "Addenda to the paper: The topographic effect in resistivity prospecting," *Geophysical Exploration (Butsuri-Tansa)*, Vol. 6, No. 1, 51–54, 1953.
16. Wannamaker, P. E., "Tensor CSAMT survey over the sulphur springs thermal area," *Geophysics, Part I: Implications for Structure of the Western Equations*, Vol. 62, No. 4, 451–465, Valles Caldera, New Mexico, U.S.A., 1997.
17. Wu, L. P. and K. F. Shi, "Applied study of CSAMT method on ground water exploration," *Geophysica Sinica*, Vol. 39, No. 5, 712–717, 1996.
18. Xu, S. Z. and S. K. Zhao, "Topographic effect in MT exploration," *Seismology Journal of Northwest*, Vol. 7, No. 4, 422–427, 1985.
19. Yamashita, M. and P. G. Hallof, "CSAMT case histories with a multi-channel CSAMT system and discussion of near-field data correction," Phoenix Geophys., Ltd, 1985.
20. Zhang, S. Z., Q. Y. Wang, and Y. Z. Luo, "The development introduction of electric exploration in China," Vol. 37, No. Supp. l, 425–428, 1994, (in Chinese).

21. Zonge, K. L. and L. J. Hughes, "Controlled source audio-frequency magnetotellurics [A]," Nabighian, M. N., ed, "Electromagnetic methods in applied Geophysics [C]," *Soc. Expl. Geophys.*, Tulsa, Vol. 2, 713–809, 1991.
22. Zonge, K. L., A. G. Ostrander, and D. F. Emer, "Controlled-source audiofrequency magnetotelluric measurements," Vozoff. K., Ed., *Magnetotelluric Methods: Soc. Expl. Geophys Geophysics Reprint Series*, Vol. 5, 749–763, 1986.

Study of Bubble Size Distribution for Breaking Wave Propagates over a Submerged Dike

Tsung-Mo Tien, Chien-Hsun Lee, and Ching-Jer Huang

Cheng Kung University, Taiwan, China

Abstract— Breaking wave plays an important role in air-sea, wave-wave, wave-current, current-current interaction. Wave breaking not only disturbs communication of sound chiefly because of the transient noise of the air-entrained bubbles. A useful analytical method, Gabor transformation, is applied due to its advantages over nonsymmetric and abrupt transition signals in breaking events. Numbers of bubbles and size distribution in a specific wave condition is analyzed in laboratory experiments.

1. INTRODUCTION

Protective structures like sea walls, breakwaters, submerged dikes are popularly deployed in nearshore ocean engineering. Formation of propagating wave will deform due to the wind drafting force, topography variation and release the accumulated wave energy. Breaking of water waves plays an important factor in many ocean studies, such as air-sea interaction, remote sensing, wave dynamics, and gas transfer to the atmosphere. Breaking wave always results in severe destructions no matter if it happens on the wide open ocean or nearshore coast; vessels, mariners and creatures suffered for that.

Distortion of a periodic wave starts to collapse when the front of wave crest is onset to break beneath downstream, dubbed as an active phase [1]. The collapse wave front will entrain air volume into the surface water layer in a very short time, say whitcapping process and which will produce dense plumes of different sizes air bubbles. Breaking attracts more and more attentions with its specifically characteristics. Various methods and techniques such as acoustic methods, photography method [5, 7] and remote sensing method [4] are presented in order to detect bubble sizes at the breaking region.

Bubble clouds are distributed underlying propagating waves, the mainly disturb source of the underwater communication [2]. Estimation of the bubble size and numbers from the breaking event is normally very hard to obtain. Stokes (1993) used the photographic method to obtain the bubble size spectrum. As a result of fragmentation of turbulence, that dominates the production of larger bubble distribution. In terms of acoustic, Minnaert (1933) proved the noise produced by bubbles is related to the simple radial pulsation of bubbles and derived the resonance frequency of bubbles. Friedlander (1995) derived a representation lead for those signals with behavior of nonsymmetric and abrupt transitions, such as bubble emission noise. Application of Gabor coefficient is adapted in this study and got satisfied results of the distribution of bubble size when wave propagates through a rectangular submerged dike.

The resonant frequency of the bubble and the radius is expressed as

$$\omega_0 = \frac{1}{R_0}\sqrt{\frac{3\gamma P_0}{\rho}} \tag{1}$$

where P_0 is the steady pressure at infinity, R_0 is the equilibrium radius of a spherical bubble, γ is the ratio of specific heat (1.4 for air) and ρ is the density of the host liquid.

2. EXPERIMENTAL SETUP AND WAVE CONDITION

All experiments are carried out in a glass-walled wave tank with the scale of 25 m long, 0.6 m wide and 0.5 m high. Wave generator used in experiments is of piston type and controlled by an electromechanical hydraulic serve system. At the end of the tank is fitted with a beach which was constructed by anechoic armor units, to minimize the wave reflections. A submerged rectangular obstacle is mounted on the bottom of the tank and 12 m away from the paddle. The obstacle has the same width as the tank with which has the length of being $L = 0.15$ m. Schematics of the experiment is shown in Fig. 1, direction of wave propagation is from left to right and terminal at and anechoic beach.

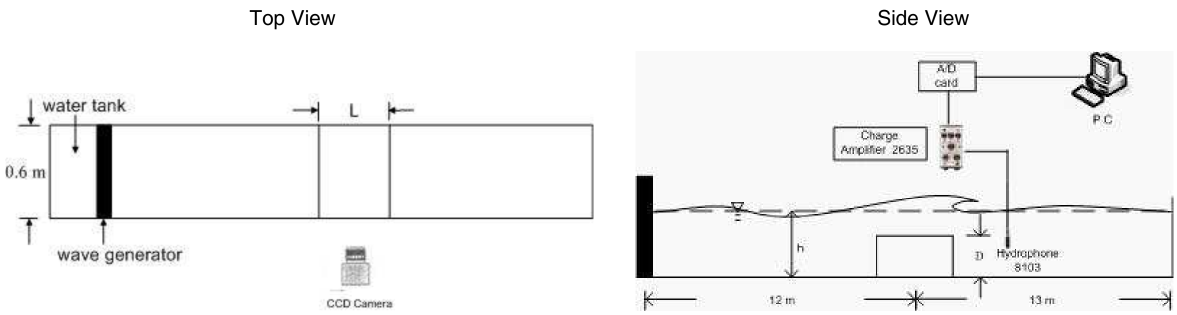


Figure 1: Schematics of the experiment.

The sound radiation from the entrained bubbles will be measured by the hydrophone mounted after the submerged dike. The sampling frequency of the hydrophone is 100 KHz, comes with one second recording time. A consistent water depth $h = 15\,\mathrm{cm}$, $T = 1.0\,\mathrm{sec}$, and the propagating wave height $H = 4\,\mathrm{cm}$ is adopted. Those measured signals will be first analyzed distribution on the frequency and go through the Gabor transform before the final employed. In the Fig. 2, it shows a series of results that measured and proceeded of the environmental noise. Fig. 2(a), it represents signal recorded in one second and which is also full of small amplitude fluctuations. From the frequency analysis can help to figure out the noise sources are composed by some frequency bands ($f > 30\,\mathrm{KHz}$, $f < 500\,\mathrm{Hz}$). The filtered result is shown in Figs. 2(b), and 2(c) represents the result in frequency and time domain that processed by the Gabor transform.

In the series results of Fig. 3, the short time existing and the pulse phenomenon of bubbles is obviously showed in the time domain like Fig. 3(b) shows. Compare the result in Fig. 3(c), several intense peaks occurred in the time domain, and via the work done by Gabor transform, distribution

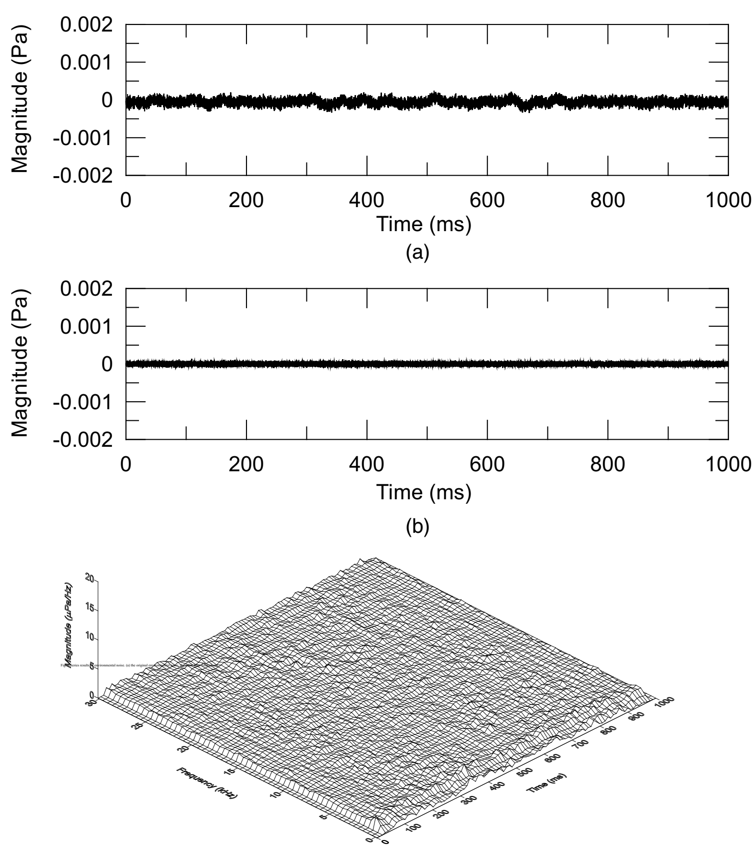


Figure 2: Series results of environmental noise. (a) the original environmental noise, (b) noise after filtered out high and low frequency ranges, (c) environmental noise in frequency and time domain.

on the frequency domain is conspicuous. The trapped picture of the breaking wave process is shown is Fig. 4, from the picture can chase down the quiescent phase [1] and dense plumes of bubbles formed.

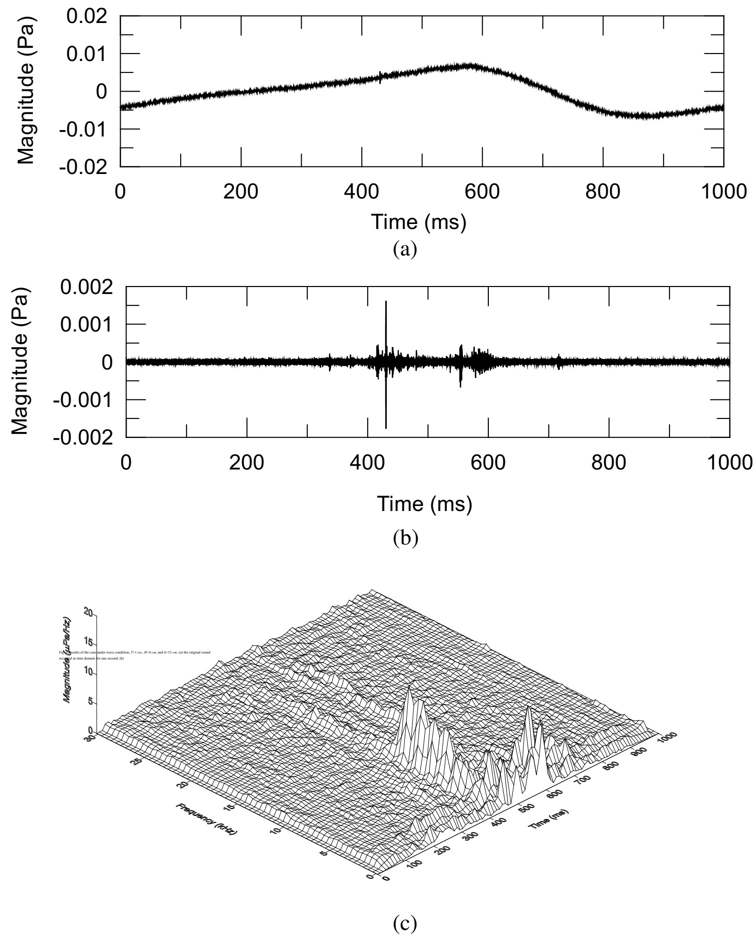


Figure 3: Results of the case under wave condition, $T = 1\,\text{sec}$, $H = 4\,\text{cm}$, and $h = 15\,\text{cm}$. (a) the original sound recorded in time domain for one second, (b) result of filtered out unwanted frequencies of (a), (c) result in time and frequency domain.

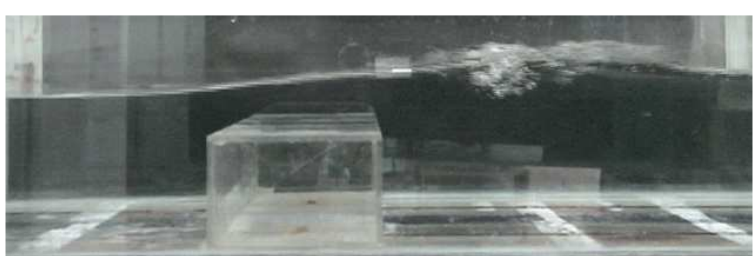


Figure 4: Breaking process occurred in the wave condition, time period $T = 1\,\text{sec}$, wave height $H = 4\,\text{cm}$, and water depth $h = 15\,\text{cm}$.

According to the characteristics of bubble size and frequency, take 25 Hz division to the frequency and 4 times of the averaged environmental energy intensity as a criterion level. Since that, the result of numbers and frequency will be derived out over the whole frequency. Due to the relation between frequency and the bubble radius, the histogram for averaged bubble numbers versus bubble radius is deployed as Fig. 5 shows.

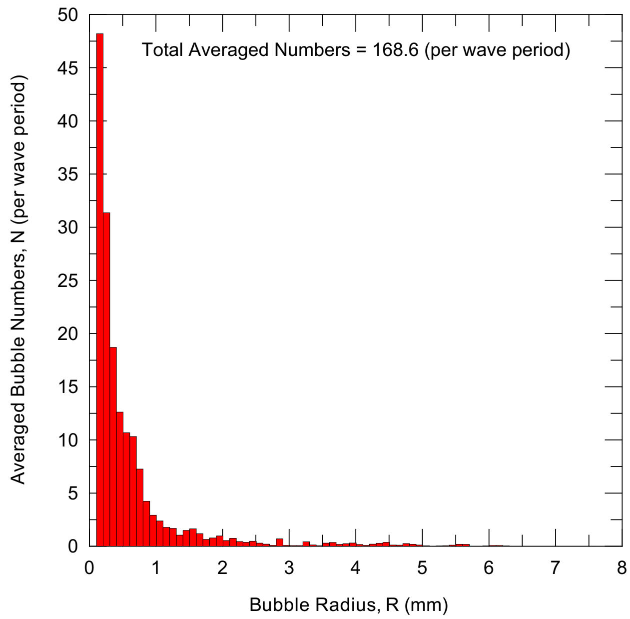


Figure 5: Histogram result for averaged bubble numbers versus bubble radius of wave propagate a designed submerged dike.

3. CONCLUSIONS

In this paper, an application of Gabor transform is deployed in estimating bubble distribution of a breaking event. From the results showed in the figures, one can realize there exist and relation between bubble size distributions and the wave height for propagating wave breaking over the under water obstacle. Many possible reasons will result in the bubble size distribution; further studies will follow up under different wave conditions. Bubble probability density for breaking events is expected to be accomplished when the whole measurements are done.

REFERENCES

1. Deane, G. B. and M. D. Stokes, "Scale dependence of bubble creation mechanisms in breaking waves," *Nature*, Vol. 418, No. 6900, 839–844, 2002.
2. Banner, M. L. and W. K. Melville, "On the separation of air flow over water waves," *J. Fluid Mech.*, Vol. 65, 647–656, 1985.
3. Friedlander, B. and A. Zeira, "Oversampled Gabor representation for transient signals," *IEEE Transactions on Signal Processing*, Vol. 43, No. 9, 2088–2094, 1995.
4. Loewen, M. R. and M. K. Melville, "Microwave backscatter and acoustic radiation from breaking waves," *J. Fluid Mech.*, Vol. 224, 601–623, 1991.
5. Medwin, H. and M. M. Beaky, "Bubble sources of the Knudsen sea noise spectra," *J. Acoust. Soc. Am.*, Vol. 86, No. 3, 1124–1130, 1989.
6. Minnaert, M., "On musical air bubbles and the sound of running water," *Philos. Mag.*, Vol. 16, 235–248, 1933.
7. Stokes, D., G. B. Deane, S. Vagle, and D. Farmer, "Measurements of large bubbles in open-ocean white caps," *Gas Transfer at Water Surfaces. Geophysical Monographs*, Vol. 127, (R. Wanninkhof, Ed. 2002), 279–284, 1993.

Least-squares Mixed Finite Element of Steady State Viscoelastic Fluid Flow

Yinghao Cai[3], **Hongwei Zhang**[1,2], **and Haitao Cai**[1]

[1]College of Mathmatics, Central South University, Changsha 410083, China
[2]College of Math & Computing Science, Changsha University of Science and Technology
Changsha 410077, China
[3]Institute of Automation, Chinese Academy of Sciences, 100080, China

Abstract— In the paper, we discuss least-squares mixed finite element method for steady state viscoelastic fluid flow. The approximate stress and velocity are respectively p_k discontinuous and p_{k+1} continuous. Existence and error estimates of the approximate solution were established.

1. INTRODUCTION

Lately, more and more attention have paid to least-squares mixed finite element method (see [1, 2]). The least-squares mixed finite element method has provided an optimal error estimates for solving many problems (see [3]). Bramble and Nitsche firstly used least-squares mixed finite element method to solve Dirichlet problem (see [4]). Raviart, Thomas, Carey and oden developed this method (see [5]). In this paper, we study the numerical analysis of the least-squares mixed finite element approximation for viscoelastic fluid flow obeying an Oldroyd B constitutive equation (see [6, 7]). At the same time, we discuss existence and error estimates of the approximate solution. The approximate stress, velocity and pressure are respectively p_k discontinuous, and p_{k+1} continuous (see [8]).

The outline of this paper is as follows: we begin by describing a model of viscoelastic fluid flow obeying an Oldroyd B type constitutive law in Section 1. In Section 2, we provide existence and uniqueness of least-squares mixed finite element approximation solution for steady viscoelastic fluid flow obeying an Oldroyd B type constitutive law. In Section 3, we give convergence analysis of least-squares mixed finite element.

Consider the model of time-dependent viscoelastic fluid flow obeying an Oldroyd B type constitutive law (see [6–8]):

$$\lambda\sigma_t + \sigma + \lambda(u\cdot\nabla)\sigma + \lambda g_\alpha(\sigma,\nabla u) - 2\alpha d(u) = 0, \text{ in } \Omega\times[0,T] \tag{1}$$

$$u_t - \nabla\cdot\sigma - 2(1-\alpha)\nabla\cdot d(u) + \nabla p = f, \text{ in } \Omega\times[0,T] \tag{2}$$

$$\nabla\cdot u = 0, \text{ in } \Omega\times[0,T] \tag{3}$$

$$u = 0, \text{ on } \Gamma\times[0,T] \tag{4}$$

If $t = 0$, then $u = u_0$, $\sigma = \sigma_0$. Here $\Omega \subset R^2$ is an open domain with boundary Γ, σ is the symmetrical stress tensor, u is velocity vector, p is the pressure (scalar), $d(u) = \frac{1}{2}\left(\nabla u + \nabla u^T\right)$ is the rate of strain tensor, $\lambda \geq 0$ is Weissenberg constant, α is a third dimensionless, which may be considered as the fraction of viscoelastic viscosity ($\alpha = 1$ for Max Well's model). For simplify, we only consider the case of $0 < \alpha < 1$. g_a: $R^4 \times R^4 \to R^4$ is a bilinear form:

$$\begin{aligned} g_a(\sigma,\nabla u) &= \sigma\omega(u) - \omega(u)\sigma - a(d(u)\sigma + \sigma d(u)) \\ &\quad + \frac{1-a}{2}(\sigma\nabla u + \nabla u^T\sigma) - \frac{1+a}{2}(\nabla u\sigma + \sigma\nabla u^T), \end{aligned}$$

Here $-1 < a < 1$ and $\omega(u) = \frac{1}{2}\left(\nabla u - \nabla u^T\right)$ is the vorticity tensor.

We use the standard Sobolev space $W^{m,p}(\Omega)$ with a norm $\|\cdot\|_{m,p}$ given by $\|\phi\|^p_{m,p} = \sum_{|\alpha|\leq m} \|D^\alpha \phi\|^p_{L^p(\Omega)}$. For $n = 2$, $H^2(\Omega) = W^{m,2}(\Omega)$ and $\|\cdot\|_m = \|\cdot\|_{m,2}$, $\|\cdot\| = \|\cdot\|_{0,2}$ (see [2]).

Definition 1 $\{\Gamma_h\}_{h>0}$ be a quasi-uniform family of finite element triangular partitions of Ω and any two neighbor units constitutes h^2-approximation patallelogram, there exists an independent of h constant C, such that $|P_1P_2 - P_3P_4| \leq Ch^2$, then the partitions family be denoted strong uniform triangular partitions [3].

Definition 2 If projection operator $\Pi_h : H^1(\Omega) \times H^1(\Omega) \to V_h$ such that for any $q \in H(div, \Omega)$, we have $(\nabla \cdot \Pi_h q, w_h) = (\nabla \cdot q, w_h)$, $\forall w_h \in W_h$, then Π_h be denoted standard finite element projection operator in W_h (see [6]).

2. EXISTENCE AND UNIQUENESS OF MIXED FINITE ELEMENT APPROXIMATION SOLUTION

For conveniently analysis the convergence property of approximation solution, we only consider steady viscoelastic fluid flow obeying an Oldroyd B type constitutive law:

$$\begin{aligned} \sigma - 2\alpha d(u) &= 0, \text{ in } \Omega && (5)\\ -\nabla \cdot \sigma &= f, \text{ in } \Omega && (6)\\ \nabla \cdot u &= 0, \text{ in } \Omega && (7)\\ u &= 0, \text{ on } \Gamma && (8) \end{aligned}$$

Let $p = -\sigma$, then the Equations (5)–(8) become:

$$\begin{aligned} p + 2\alpha d(u) &= 0, \text{ in } \Omega && (9)\\ \nabla \cdot p - f &= 0, \text{ in } \Omega && (10)\\ \nabla \cdot u &= 0, \text{ in } \Omega && (11)\\ u &= 0, \text{ on } \Gamma && (12) \end{aligned}$$

We define the energy space as follow

$$\begin{aligned} V &= \{v \in H^1(\Omega);\ v = 0 \text{ on } \Gamma,\quad \nabla \cdot v = 0 \text{ on } \Omega,\\ H(div;\Omega) &= \{q \in \left(L^2(\Omega)\right)^2; divq \in L^2(\Omega)\}. \end{aligned}$$

And the corresponding norm be

$$\begin{aligned} \|v\|_{1,\Omega} &= \left(\|v\|^2_{0,\Omega} + \|\nabla v\|^2_{0,\Omega}\right)^{1/2},\\ \|q\|_{H(div;\Omega)} &= \left(\|q\|^2_{0,\Omega} + \|divq\|^2_{0,\Omega}\right)^{1/2}. \end{aligned}$$

Now we define the least squares problem become: find $(u, p) \in V \times H$ such that

$$J(u,p) = \inf_{v\in V, q\in H} J(v,q) \tag{13}$$

where $J(v, q) = (divq - f,\ divq - f)_{0,\Omega} + (q + 2\alpha d(v),\ q + 2\alpha d(v))_{0,\Omega}$.

The corresponding variational formulation be: find $(u, p) \in V \times H$ such that

$$a(u,p;v,q) = (f, divq)_{0,\Omega},\ \forall v \in V,\ q \in H \tag{14}$$

where $a(u, p; v, q) = (divp, divq)_{0,\Omega} + (p + 2\alpha d(u), q + 2\alpha d(v))_{0,\Omega}$.

Lemma 1 There exists a positive constant C such that

$$a(v,q;v,q) \ge C\left(\|v\|^2_{1,\Omega} + \|q\|^2_H\right) \tag{15}$$

Proof According to the definition of $a(v, q; v, q)$, one obtains

$$\begin{aligned} a(v,q;v,q) &\ge \|divq\|^2_{0,\Omega},\\ a(v,q;v,q) &\ge \|q + 2\alpha d(v)\|^2_{0,\Omega}. \end{aligned}$$

So we have

$$\begin{aligned} a(v,q;v,q) &= (divq, divq)_{0,\Omega} + (q + 2\alpha d(v), q + 2\alpha d(v))_{0,\Omega}\\ &= (divq, divq)_{0,\Omega} + (q,q)_{0,\Omega} + 4\alpha(divq, v)_{0,\Omega} + 4\alpha^2(d(v), d(v))_{0,\Omega}\\ &\ge \|divq\|^2_{0,\Omega} - 4\alpha\|divq\|_{0,\Omega}\|v\|_{0,\Omega} + \|v\|^2_{0,\Omega} + (q,q)_{0,\Omega} + 4\alpha^2(d(v), d(v))_{0,\Omega} - \|v\|^2_{0,\Omega}\\ &\ge (\|divq\|_{0,\Omega} - 2\alpha\|v\|_{0,\Omega})^2 + (q,q)_{0,\Omega} + 4\alpha^2(d(v), d(v))_{0,\Omega} - \|v\|^2_{0,\Omega}\\ &\ge \|q\|^2_{0,\Omega}, \end{aligned}$$

Connection those inequalities, we obtain

$$C\|\nabla v\|_{0,\Omega}^2 \le C\left(\|q+2\alpha d(v)\|_{0,\Omega}^2 + \|q\|_{0,\Omega}^2\right) \le Ca(v,q;v,q).$$

By the Poincare inequality, we have $\|v\|_{0,\Omega}^2 \le C\|\nabla v\|_{0,\Omega}^2 \le a(v,q;v,q)$. The desired result then follows.

Theorem 1 Assume that $f \in L^2(\Omega)$, then the least-squares variational formulation (12) for steady viscoelastic fluid flow obeying an Oldroyd B type constitutive law exists uniqueness solution $(u,p) \in V \times H$.

Proof By lemma 2, then the bilinear form $a(v,q;v,q)$ is coercive and continuous. According to Lax-milgram Theorem, the desired result then follows.

3. CONVERGENCE ANALYSIS OF LEAST-SQUARES MIXED FINITE ELEMENT

Ω is supposed to be polygonal and equipped with uniform triangular partitions Γ_h, such that $\overline{\Omega} = \{\cup K, K \in \Gamma_h\}$. Let $P_k(K)$ denotes the space of polynomials of degree less orequal k on K. $V_h,\ H_h$ be any the usual mixed finite element approximating subspaces of $V,\ H$ (see [5]):

$$\begin{aligned} V_h &= \{v \in V;\ v|_K \in P_k(K),\ \forall K \in \Gamma_h\}, \\ H_h &= \{q \in H;\ q|_K \in P_r(K),\ \forall K \in \Gamma_h\}. \end{aligned}$$

The corresponding least squares mixed finite element approximation formulation become: find $(u_h,p_h) \in V_h \times H_h$ such that

$$a\left(u_h,p_h;v_h,q_h\right) = (f, divq_h)_{0,\Omega},\ \forall v_h \in V_h,\ q_h \in H_h.$$

Connection the two variational formulation, we have a error equation

$$a\left(u-u_h,p-p_h;v_h,q_h\right) = 0,\ \forall v_h \in V_h,\ q_h \in H_h.$$

For simplify, we introduce some properties in Sobolev space (see [9])

$$\|u-\Pi_h u\|_{0,\Omega} \le Ch^{k+1}\|u\|_{k+1,\Omega} \tag{16}$$

$$\|u-\Pi_h u\|_{1,\Omega} \le Ch^{k}\|u\|_{k+1,\Omega} \tag{17}$$

$$\|p-\Pi_h p\|_{0,\Omega} \le Ch^{r+1}\|p\|_{r+1,\Omega} \tag{18}$$

Theorem 2 Assume that $s = \min(k,r)$, then

$$\|u-u_h\|_{1,\Omega} + \|p-p_h\|_{H(div;\Omega)} \le Ch^s\left(\|u\|_{s+1,\Omega} + \|p\|_{s+1,\Omega}\right) \tag{19}$$

Proof By the coercive of the bilinear form $a(u,p;v,q)$ and the error equation , one obtains

$$\begin{aligned} & \|u_h-\Pi_h u\|_{1,\Omega}^2 + \|p_h-\Pi_h p\|_{H(div;\Omega)}^2 \\ \le\ & Ca\left(u_h-\Pi_h u, p_h-\Pi_h p; u_h-\Pi_h u, p_h-\Pi_h p\right) \\ =\ & Ca\left(u-\Pi_h u, p-\Pi_h p; u_h-\Pi_h u, p_h-\Pi_h p\right) \\ =\ & \left(div\left(p-\Pi_h p\right), div\left(p_h-\Pi_h p\right)\right)_{0,\Omega} + \left(p-\Pi_h p+2\alpha d\left(u-\Pi_h u\right), p_h-\Pi_h p+2\alpha d\left(u_h-\Pi_h u\right)\right)_{0,\Omega} \\ \le\ & C\Big[\left(\|div\left(p-\Pi_h p\right)\|_{0,\Omega}^2 + \|p-\Pi_h p+2\alpha d\left(u-\Pi_h u\right)\|_{0,\Omega}^2\right)^{1/2} \\ & + \left(\|div\left(p-\Pi_h p\right)\|_{0,\Omega}^2 + \|p_h-\Pi_h p+2\alpha d\left(u_h-\Pi_h u\right)\|_{0,\Omega}^2\right)^{1/2}\Big] \\ \le\ & C\left(\|u-\Pi_h u\|_{1,\Omega}^2 + \|p-\Pi_h p\|_{H(div;\Omega)}^2\right)^{1/2} \times \left(\|u_h-\Pi_h u\|_{1,\Omega}^2 + \|p_h-\Pi_h p\|_{H(div;\Omega)}^2\right)^{1/2} \\ \le\ & Ch^s\left(\|u\|_{s+1,\Omega}^2 + \|p\|_{s+1,\Omega}^2\right)^{1/2} \times \left(\|u_h-\Pi_h u\|_{1,\Omega}^2 + \|p_h-\Pi_h p\|_{H(div;\Omega)}^2\right)^{1/2}, \end{aligned}$$

So we have

$$\|u_h-\Pi_h u\|_{1,\Omega} + \|p_h-\Pi_h p\|_{H(div;\Omega)} \le Ch^s\left(\|u\|_{s+1,\Omega} + \|p\|_{s+1,\Omega}\right).$$

By triangle inequality, we obtain

$$\begin{aligned}
&\|u-u_h\|_{1,\Omega}+\|p-p_h\|_{H(div;\Omega)}\\
\le\ &\|u-\Pi_h u\|_{1,\Omega}+\|\Pi_h u-u_h\|_{1,\Omega}+\|p-\Pi_h p\|_{H(div;\Omega)}+\|\Pi_h p-p_h\|_{H(div;\Omega)}\\
\le\ &Ch^s\left(\|u\|_{s+1,\Omega}+\|p\|_{s+1,\Omega}\right).
\end{aligned}$$

The desired result then follows.

Theorem 3 Assume that $r=k+1$, then

$$\|u-u_h\|_{0,\Omega}+\|p-p_h\|_{H(div;\Omega)}\le Ch^k\left(\|u\|_{k+1,\Omega}+\|p\|_{k+2,\Omega}\right) \tag{20}$$

Proof Firstly, we define a projection operator $S_h : V\to V_h$ such that

$$(S_h u,\nabla\cdot v_h)=(u,\nabla\cdot v_h)\,,\quad \forall v_h\in V_h.$$

By the conclusion in [8], one obtains this posteriori estimate

$$\begin{aligned}
\|v-S_h v\|_{1,\Omega}\ &\le\ Ch^k\|v\|_{k+1,\Omega}.\\
\|v-S_h v\|_{0,\Omega}\ &\le\ Ch^{k+1}\|v\|_{k+1,\Omega}.
\end{aligned}$$

By lemma 2 and the error equation, one obtains

$$\begin{aligned}
&C\left(\|u_h-S_h u\|_{1,\Omega}^2+\|p_h-\Pi_h p\|_{H(div;\Omega)}^2\right)\\
\le\ &a\left(u_h-S_h u,p_h-\Pi_h p;u_h-S_h u,p_h-\Pi_h p\right)\\
=\ &a\left(u-S_h u,p-\Pi_h p;u_h-S_h u,p_h-\Pi_h p\right)\\
=\ &\left(div\left(p-\Pi_h p\right),div\left(p_h-\Pi_h p\right)\right)_{0,\Omega}+\left(p-\Pi_h p+2\alpha d\left(u-S_h u\right),p_h-\Pi_h p+2\alpha d\left(u_h-S_h u\right)\right)_{0,\Omega}\\
=\ &\left(div\left(p-\Pi_h p\right),div\left(p_h-\Pi_h p\right)\right)_{0,\Omega}+\left(p-\Pi_h p,p_h-\Pi_h p\right)_{0,\Omega}\\
&+2\alpha\left(p-\Pi_h p,d\left(u_h-S_h u\right)\right)_{0,\Omega}+2\alpha\left(d\left(u-S_h u\right),p_h-\Pi_h p\right)_{0,\Omega}+4\alpha^2\left(d\left(u-S_h u\right),d\left(u_h-S_h u\right)\right)_{0,\Omega}
\end{aligned}$$

By Cauchy-Schwarz inequality, we have

$$\begin{aligned}
&\left(div\left(p-\Pi_h p\right),div\left(p_h-\Pi_h p\right)\right)_{0,\Omega}+\left(p-\Pi_h p,p_h-\Pi_h p\right)_{0,\Omega}\\
\le\ &\|div\left(p-\Pi_h p\right)\|_{0,\Omega}\|div\left(p_h-\Pi_h p\right)\|_{0,\Omega}+\|p-\Pi_h p\|_{0,\Omega}\|p_h-\Pi_h p\|_{0,\Omega}\\
\le\ &Ch^{k+1}\|p\|_{k+2,\Omega}\|p_h-\Pi_h p\|_{H(div;\Omega)},\\
&2\alpha\left(p-\Pi_h p,d\left(u_h-S_h u\right)\right)_{0,\Omega}\\
\le\ &2\alpha\|p-\Pi_h p\|_{0,\Omega},\|d\left(u_h-S_h u\right)\|_{0,\Omega}\\
\le\ &Ch^{k+1}\|p\|_{k+2,\Omega}\|u_h-S_h u\|_{1,\Omega},\\
&2\alpha\left(p_h-\Pi_h p,d\left(u-S_h u\right)\right)_{0,\Omega}\\
\le\ &2\alpha\|p_h-\Pi_h p\|_{0,\Omega},\|d\left(u-S_h u\right)\|_{0,\Omega}\\
\le\ &Ch^k\|u\|_{k+1,\Omega}\|p_h-\Pi_h u\|_{H(div;\Omega)},\\
&4\alpha^2\left(d\left(u-S_h u\right),2\alpha d\left(u_h-S_h u\right)\right)_{0,\Omega}\\
\le\ &4\alpha^2\|d\left(u-S_h u\right)\|_{0,\Omega}\|d\left(u_h-S_h u\right)\|_{0,\Omega}\\
\le\ &C\|\left(u-S_h u\right)\|_{1,\Omega}\|\left(u_h-S_h u\right)\|_{1,\Omega}\\
\le\ &Ch^k\|u\|_{k+1,\Omega}\|u_h-S_h u\|_{1,\Omega}.
\end{aligned}$$

So, we obtain

$$\begin{aligned}
&C\left(\|u_h-S_h u\|_{1,\Omega}^2+\|p_h-\Pi_h p\|_{H(div;\Omega)}^2\right)\\
\le\ &Ch^{k+1}\left[\|p\|_{k+2,\Omega}\left(\|p_h-\Pi_h p\|_{H(div;\Omega)}+\|u_h-S_h u\|_{1,\Omega}\right)\right]\\
&+Ch^k\left[\|u\|_{k+1,\Omega}\left(\|p_h-\Pi_h p\|_{H(div;\Omega)}+\|u_h-S_h u\|_{1,\Omega}\right)\right]\\
\le\ &Ch^k\left(\|u\|_{k+1,\Omega}+\|p\|_{k+2,\Omega}\right)\left(\|p_h-\Pi_h p\|_{H(div;\Omega)}+\|u_h-S_h u\|_{1,\Omega}\right)
\end{aligned}$$

Furthermore, we have

$$\|u_h - S_h u\|_{1,\Omega} + \|p_h - \Pi_h p\|_{H(div;\Omega)} \leq Ch^k \left(\|u\|_{k+1,\Omega} + \|p\|_{k+2,\Omega}\right).$$

By triangle inequality, we have

$$\begin{aligned}
&\|u - u_h\|_{0,\Omega} + \|p - p_h\|_{H(div;\Omega)} \\
&\leq \|u - S_h u\|_{1,\Omega} + \|p - \Pi_h p\|_{H(div;\Omega)} + \|S_h u - u_h\|_{1,\Omega} + \|S_h p - p_h\|_{H(div;\Omega)} \\
&\leq Ch^k \left(\|u\|_{k+1,\Omega} + \|p\|_{k+2,\Omega}\right),
\end{aligned}$$

The desired result then follows.

REFERENCES

1. Bramble, J. H. and J. A. Nitsche, "A generalized Ritz-least-squares method for Dirichlet problems," *SIAM J. Number. Anal.*, Vol. 10, 81–93, 1973.
2. Pehlivanor, A. I., G. F. Carey, and R. D. Lazarov, "Leasr-squares finite elements for second-order elliptic problems," *SIAM J. Number. Anal.*, Vol. 31, 1368–1377, 1994.
3. Brezzi, F., "On the existence, uniqueness and approximation of saddle point problems arising from Lagrange multipliers," *RAIRO Ser. Anal.*, No. 8, 129–151, 1974.
4. Carey, G. F. and J. T. Oden, *Finite Elements*, Prentice-Hall, Englewood Cliffs, NJ, A second course, Vol. 11, 1289–1306, 1983.
5. Gu, H. and D. Yang, "Least-squares mixed finite elements method for a class of Stokes equation," *Applied Mathematics and Mechanics*, Vol. 21, 501–510, 2000.
6. Luo, Z. and J. Zhu, "A nonlinear Galerkin/Petrov least-squares mixed finite elements method for the stationary Navier-Stokes equations," *Applied Mathematics and Mechanics*, Vol. 23, 697–706, 2002.
7. Chen, Y. P. and J. Yang, "Superconvergence of least-squares mixed finite elements method for elliptic problems in triangulation," *Numerical Mathematics*, Vol. 12, 214–225, 2003.
8. Chen, Y. P., Y. Q. Huang, and Z. H. Shen, "Least-squares mixed finite elements method for degenerate elliptic problems," *Numer. Math. Sinica.*, Vol. 23, 87–94, 2001.
9. Chen, Y. P., Y. Q. Huang, and Z. H. Shen, "The least-squares mixed finite elements method for incompressible Miscible displacements," *Mathematical Numerical Sinica*, Vol. 3, 270–278, 2000.
10. Johnson, C. and V. Thomee, "Error estimates for some finite element methods for parabolic type problem [J]," *RAIRO Anal.*, No. 15, 41–78, 1981.
11. Baranger, J. and D. Sandri, "Finite elements approximation of viscoelastic fluid flow: Existence of approximate solutions and error bounds," *Numerische Mathematik*, Springer-verlag, Vol. 63, 13–27, 1992.
12. Sandri, D., "Finite elements approximation of viscoelastic fluid flow: Existence of approximate solutions and error bounds. II. continuous approximation of the stress," *SIAM Journal on Nuerical analysis*, Vol. 31, 362–377, 1994.

Frequency Response of Tri-axial Induction Logging Tool

Jing Li, Xuesheng Yu, Lili Zhong, and Richard C. Liu
Electrical and Computer Engineering Department, University of Houston
4800 Calhoun Rd, Houston, TX, 77004, USA

Abstract— A multi-component transmitter-receiver configuration in tri-axial induction tool can provide direct measurement to formation anisotropy. It is especially useful in deviated and horizontal drilling. Tri-axial induction too is composed of three mutually orthogonal transmitter-receiver coil pair oriented in three mutually perpendicular directions. The axis of the tool may intercept a formation with a constant dip angle, a constant azimuthal angle and a constant orientation angle. An algorithm and 1D code were developed based on rigorous analytic solution for the tri-axial logging tool response in 1-D formation, which neglects the borehole, and mandrel. Operating frequency plays an important part in the tool applications. In this paper, frequency responses for different formation contrast are investigated.

1. INTRODUCTION

In order to determine parameters in anisotropic formation with dipping angle more accurately, a tri-axial indirection tool configuration was developed. The tool is composed of three mutually orthogonal transmitter-receiver configurations, which can offer all the necessary data for formation resistivity measurements. Magnetic field generated by a magnetic dipole can be obtained by using Hertz potentials [1, 2]. After solving the H-fields, there are only five non-zero elements, which are Hxx, Hxz, Hyy, Hzx and Hzz. A 1D code based on this algorithm was developed, which neglect the borehole, the mandrel and coil size and the actual coils are replaced by point magnetic dipoles.

Figure 1 shows schematically a tri-axial logging tool containing both three orthogonal transmitter coils $M_{x'}$, $M_{y'}$, and $M_{z'}$, and three orthogonal receiver coils $R_{x'}$, $R_{y'}$, and $R_{z'}$, oriented in the coordinate system x', y', z' [1].

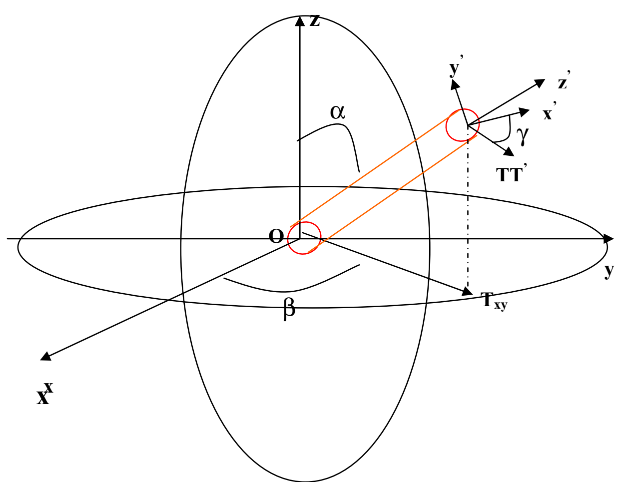


Figure 1: An induction sonde containing three orthogonal transmitter coils and three orthogonal receiver coils. α, β and γ are dip angle and azimuthal angle of borehole and tool rotation angle respectively.

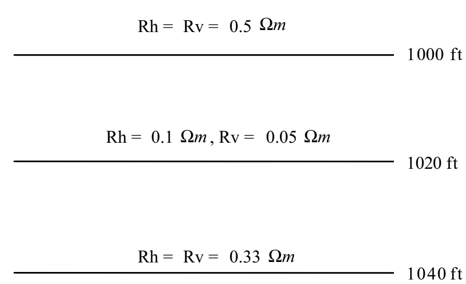
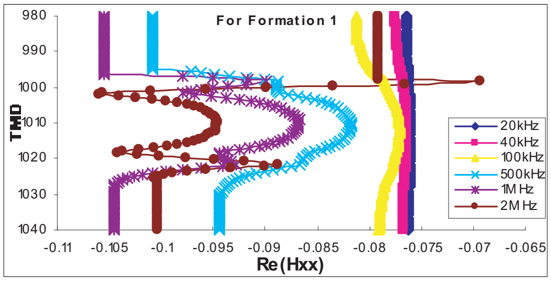


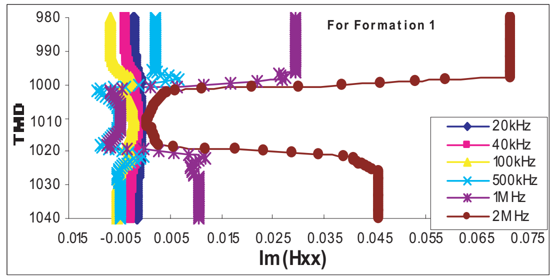
Figure 2: Formation 1.

2. FREQUENCY RESPONSE

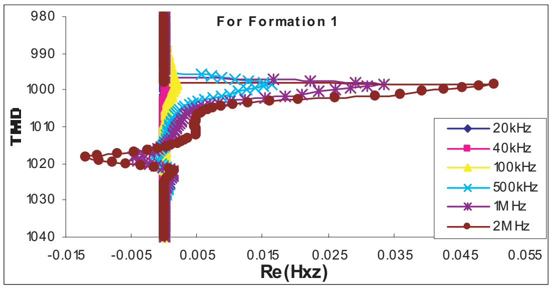
In this paper we will study the frequency response of tri-axial induction tool in different contrast formation models using the algorithm developed in [1]. The dipping angle is set to 60 degrees to cover both horizontal and vertical anisotropy. The frequencies used in the simulations are 20 kHz, 40 kHz, 100 kHz, 500 kHz, 1 MHz and 2 MHz. The formation model analyzed in this simulation is shown in Figure 2, with vertical resistivity of the second layer changes from 0.05 Ωm to 0.01 Ωm. The real and imaginary parts of non-zero H-field elements are plotted in Figure 3.



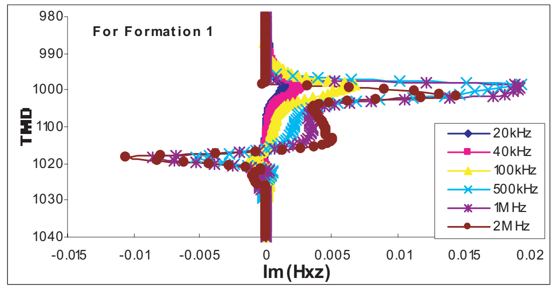
(a) Real part of Hxx



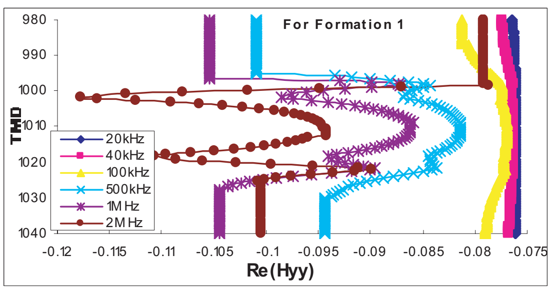
(b) Imaginary part of Hxx



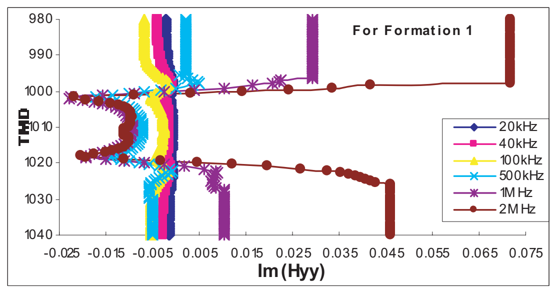
(c) Real part of Hxz



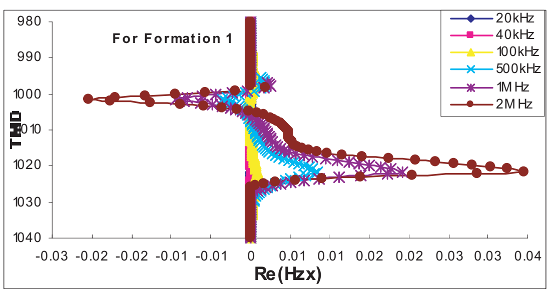
(d) Imaginary part of Hxz



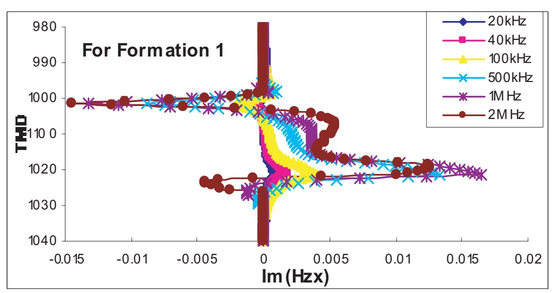
(e) Real part of Hyy



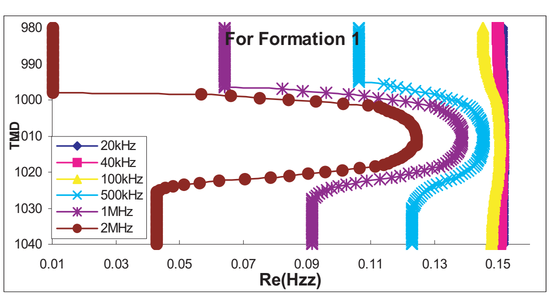
(f) Imaginary part of Hyy



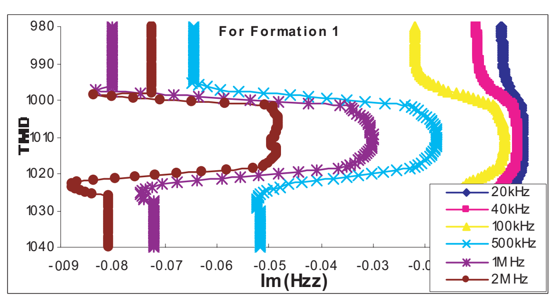
(g) Real part of Hzx



(h) Imaginary part of Hzx



(i) Real part of Hzz



(j) Imaginary part of Hzz

Figure 3: Frequency responses for Formation 1.

3. CONCLUSIONS

1. From the simulation results, it is seen that tri-axial induction tool response is frequency sensitive. The tool response begin to curve when the tool is entering into and out from the

reservoir. Moreover, the curvature of the tool response increases with increasing frequency.

2. The tool response is also sensitive to formation contrast. The greater formation contrast, the larger the tool response.

3. For sharper boundary response, higher frequency tool is preferred. However, the investigation depth may be affected, even though this is not simulated in this study.

4. To obtain deeper penetration and higher resolution, a multi frequency logging tool should be used.

ACKNOWLEDGMENT

This research is supported by a consortium of 11 oil and service companies including Saudi Aramco, Baker Hughes, BP, Chevron E & P Technology Company, ConocoPhillips, ExxonMobil Upstream Research Company, Halliburton Energy Service, Pathfinder, Shell E & P Technology Company, Schlumberger Oil Field Services, and Statoil.

REFERENCES

1. Zhong, L. L., "Simulation of triaxial induction tools in dipping anisotropic beds," *Dissertation*, University of Houston, 2004.
2. Moran, J. H. and S. Gianzero, "Effects of formation anisotropy on resistivity logging measurement," *Geophysics*, Vol. 44, 1266–1286, 1979.
3. Anderson, B., T. Barber, and T. M. Habashy, "The interpretation and inversion of fully triaxial induction data; a sensitivity study," *43rd Annual Logging Symposium, SPWLA*, 1–14, 2002.
4. Graciet, S. and L. C. Shen, "Theory and numerical simulation of induction and MWD resistivity tools in anisotropic dipping beds," *The Log Analyst*, Vol. 39, No. 1, 24–37, 1998.

Simultaneous Measurement of Capillary Pressure and Dielectric Constant in Porous Media

W. J. Plug, L. M. Moreno, J. Bruining, and E. C. Slob
Delft University of Technology, The Netherlands

Abstract— An experimental tool is presented with which the capillary pressure (P_c) and the dielectric constant (ε) of a porous medium are measured simultaneously. The equipment is designed to conduct measurements for the unconsolidated sand-distilled water-gas (CO_2) system for pressures up to 20 bar and constant temperature conditions. The bulk phase pressures of the gas and the water are measured at the top and bottom of the sample holder and we define the P_c as the averaged pressure difference. The sample holder is a parallel plate capacitor with stainless steel plates, which also serve as support for the sample. The plates are kept separated by a plastic ring. A precision component analyzer is connected to the sample holder and measures the impedance as a function of frequency (f). The total impedance is directly related to the effective value of the complex permittivity. The aim of this paper is to describe the details and the calibration of the tool. The accuracy and the validity of both $\varepsilon(S_w)$ and $P_c(S_w)$ are determined from reproducible data. Results are shown for a full capillary pressure cycle of drainage and imbibition at $f = 3\,\mathrm{MHz}$.

1. INTRODUCTION

Both the capillary pressure and the electrical properties, e.g., the resistivity and the complex dielectric constant, can be expressed as function of the water saturation (S_w). These properties are used in subsurface flow engineering applications such as hydrocarbon production and soil remediation techniques. The electrical and capillary characteristics are used to determine the connate water saturation, resistivity logs, the height of the transition zone, monitoring of the remediation and to assist in modeling purposes [1].

In literature, numerous examples can be found which describe the hysteresis in capillary pressure. The term "capillary hysteresis" can be defined as the difference in capillary pressure value between drainage (decreasing S_w) and imbibition (increasing S_w). Hysteresis depends on the saturation history and can be ascribed to contact angle hysteresis [2], irreversible changes in pore-scale fluid distributions [3] and the interfacial area [4].

Experimental studies have shown that pore scale mechanisms have direct influence on several geophysical properties, such as the electrical resistivity and the dielectric constant [1,5,6]. In practice electromagnetic techniques become very popular. Therefore it is important to understand the relation of the dielectric properties with respect to the saturation history, the fluid saturation and the phase distribution.

Knight and Nur [7] measured the dielectric constant of partially saturated sandstones in the frequency range of 60 kHz to 4 MHz. They found that a coated water layer around the grains is the predominant reason for the dielectric response. Pronounced hysteresis effects in electrical resistivity are found by Longeron et al. [8] and Knight [9]. Both studies reported that the values during drainage process were higher than for imbibition process. Nguyen et al., [6] have measured the complex dielectric constant with the FDR technique for a water-wet sample during the drainage and imbibition cycle and found dielectric hysteresis. At $S_w < 0.6$ the values of the permittivity for the imbibition is lower than those for the drainage. The opposite is found for $S_w > 0.6$.

The impact on surface interactions is explained by Knight and Abad [9]. In this study they concluded that the dielectric constant decreased when the sample altered from water-wet to oil-wet. Similar results were obtained by Nguyen [10]. Combined measurements of capillary pressure and the dielectric constant, as found in the work of Nguyen and Longeron, provide important information to improve the interpretation of capillary hysteresis.

Here we present a new method, with which the capillary pressure (P_c) and the dielectric constant (ε) of a porous medium are measured simultaneously. The capillary pressure is measured under quasi-static conditions using the set-up discussed by Plug et al., [11]. The sample holder is designed as a parallel plate capacitor, where the two stainless steel end-pieces act as electrodes. The advantage of this technique is that the sample remains intact during the measurements and the complex permittivity can be measured continuously as a function of the frequency. Calibration of the impedance tool is done using materials with known dielectric constants and the experimental technique is validated with reproducible data.

2. EXPERIMENTAL SET-UP

2.1. The Capillary Pressure Tool

In this section we briefly describe the set-up that measures the capillary pressure as function of the water saturation under quasi-static conditions. The equipment, presented by Plug et al., [11], is based on the porous plate technique combined with the micro-pore membrane technique, discussed by Jennings et al., [12] and Longeron et al., [13]. A schematic diagram of the set-up is shown in Figure 1.

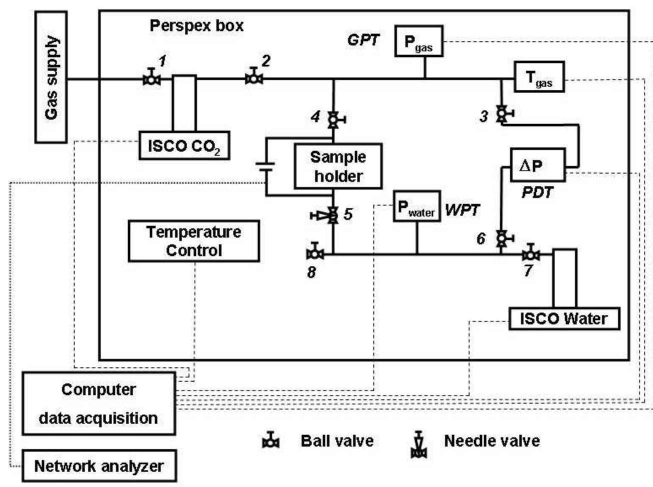


Figure 1: Schematic lay-out of the experimental set-up.

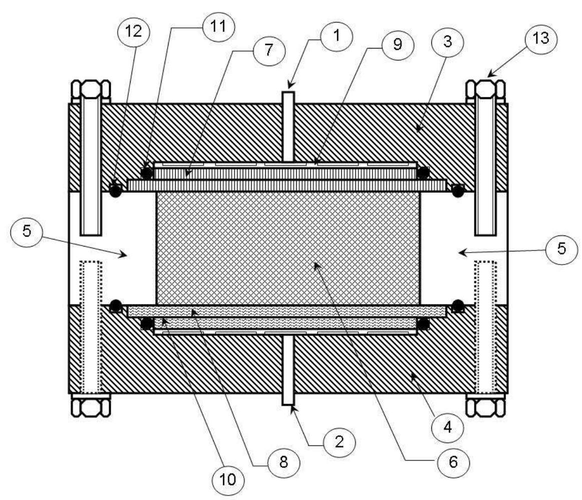


Figure 2: The sample-holder: 1. gas-inlet; 2. water-inlet; 3. stainless steel end piece 1; 4. stainless steel end piece 2; 5. PEEK ring; 6. porous medium; 7. Perforated plates; 8. SIPERM plates; 9. concentric grooves; 10. Water-wet filter; 11. O-rings (2.1 mm); 12. O-rings (4 mm); 13. Stainless steel bolts.

Two syringe pumps (ISCO pump, 260D) are connected to the in- and outlet of the sample holder (see Figure 2) and can be set to a constant injection rate (accuracy of 0.005 ml/h) or a constant pressure (accuracy ± 0.1 bar). The gas phase is injected (drainage) or produced (imbibition) at the top of the sample holder and the water is collected or injected at the bottom using the second (water) syringe pump. The gas pressure transducer (GPT) and the water pressure transducer (WPT) record the single phase pressures (range 0-100 bar, accuracy ± 0.01 bar).

The differential pressure between the gas and the water phase is measured by the pressure difference transducer (PDT, 0–500 mbar, accuracy ± 0.1 mbar), which is placed at the middle of the sample, such that no correction for gravity effects is required. To maintain a constant temperature we cover the entire set-up with a Perspex box, sealed by polystyrene. We allow temperature equilibration for at least two days to ensure that the total set up is at equal temperature.

The sample holder, see Figure 2, consists of a PEEK (Polyetheretherketone) ring which contains the unconsolidated sand sample. The height, H, of the sample is 27 mm and the diameter, D, is 84 mm. The grains are kept in place using a combination of plates at the top and bottom of the sample. At the bottom, two porous plates (SIPERM R, Cr-Ni-Steel basis) with a diameter of, $D_{s,1} = 90$ mm and $D_{s,2} = 84$ mm respectively, a permeability of $2 \times 10^{-12}\,\mathrm{m}^2$ and a porosity of 0.32, support the sample and protect the hydrophilic membrane (0.1 µm). Two stainless steel plates ($D_{ss,1} = 90$ mm and $D_{ss,2} = 84$ mm) both with 32 single perforations ($D_p = 5$ mm) are used at the top directly above the sample in combination with a nylon filter with a pore size of 210 µm. To avoid leakage of gas or water over the hydrophilic membrane, we seal the outer perimeter with a rubber O-ring (thickness of 2.1 mm). Concentric flow grooves in the end-pieces redistribute the injected and produced phase over the total sample area to avoid preferential flow and fast breakthrough of the injected phase.

2.2. The Impedance Tool

A precision component analyze (Precision Component Analyzer, 6640A) is connected to the sample holder and measures the impedance as a function of frequency (f). The analyzer supplies an AC potential difference of 990 mVac and the impedance and phase angle are measured. The total impedance is directly related to the effective value of the effective complex dielectric constant (ε) of the mixture of gas, water and grains. To measure correct values for the dielectric constant, the PEEK ring must separate the electrodes. In our design, the electrodes are the two end pieces (3 and 4, Figure 2), including the support plates (7 and 8, Figure 2). The PEEK material is non-conductive and the total sample holder will act as a parallel plate capacitor. As a result of this configuration a parallel circuit is established for the sample inside the PEEK ring and the PEEK ring itself. To keep the different parts of the sample holder together, 4 stainless steel bolts are used at the top and bottom of the sample holder. The placement of the bolts is such that no short-circuit for the electrical current occurs. To ensure total isolation of the sample holder, Teflon tubing are used for the in- and outlet.

3. CALIBRATION OF IMPEDANCE TOOL

Different configurations of the sample holder are investigated to obtain the most accurate data for the dielectric constant and capillary pressure. It appears that the type and combination of the support plates, the presence and the number of the stainless steel bolts do not have an effect on the measurements. The use of the stainless steel bolts, the bolt holes (in the PEEK ring and the end-pieces) and the presence of electronic devices in the set-up, will introduce background noise. To account for this, we obtain the capacitance of the PEEK ring, where the sample holder is filled with air. Since ε of air is 1, we can derive the capacitance of the PEEK ring using

$$Z^* = \frac{1}{i\omega\left(C_{\text{sample}} + C_{\text{PEEK\_RING}}\right)} \tag{1}$$

where $Z^*[\Omega]$ is the complex impedance, $i\omega$ is the complex frequency and C is the capacitance [F] expressed as $C = \varepsilon_0 \varepsilon A/d$. Eq. (1) and the value for $C_{\text{PEEK\_RING}}$ is used in all of the derivations of the dielectric constant of the material inside the PEEK ring. Calibration is done at room temperature using materials with known dielectric properties. Figure 3 shows the straight calibration line (dashed line) and the values obtained for the different materials at $f = 3$ MHz. Good agreement is found between the experimental results and the theoretical values and the maximum relative error of 7% is measured for n-Butanol.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this section we describe the validity of the experimental technique. Two different experiments conducted for the CO_2-sand-water system at 8 bar and 28°C. We use an unconsolidated sand sample with a grain-size fraction of $360 < D < 410$ µm and a porosity of 0.38. Reproducible capillary pressure curves are obtained for both the drainage and imbibition process (see Figure 4). High precision is obtained for the saturation range between 0.15 and 0.9. Due to a power failure at

A (Figure 4), the primary drainage data are missing near $S_w = 0.9$. Non-monotonic behavior of the imbibition curve is observed at point B. These irregularities are attributed to summer temperatures in the laboratory, exceeding the upper limit of the temperature control system. Capillary pressure hysteresis is measured and is similar for both experiments.

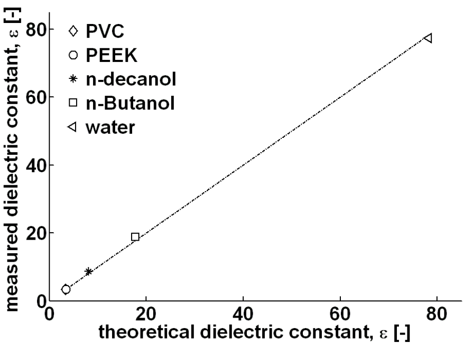


Figure 3: Calibration line for different materials obtained for the impedance tool for $f = 3\,\mathrm{MHz}$.

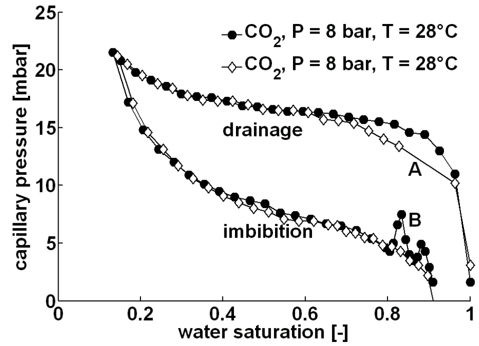


Figure 4: Drainage and imbibition capillary pressure curves for the sand-CO_2-water system. In this case the capillary hysteresis is very pronounced.

The corresponding dielectric constants as function of the water saturation for the drainage and imbibition experiments are presented in Figure 5 and Figure 6. As expected, the dielectric constant decreases during drainage and increases for imbibition. Comparison of the two experiments shows a small deviation, $\varepsilon = \varepsilon \pm 0.5$, in dielectric constant for both drainage and imbibition. We ascribe these small differences to different sample packing.

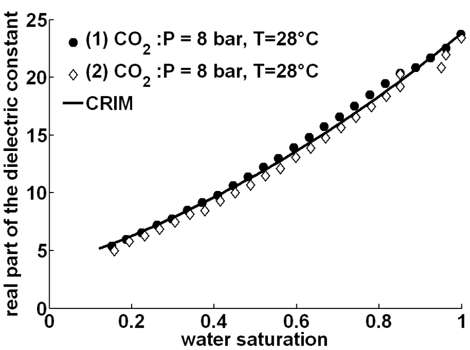


Figure 5: The real part of the dielectric constant as function of the water saturation during drainage for $f = 3\,\mathrm{MHz}$.

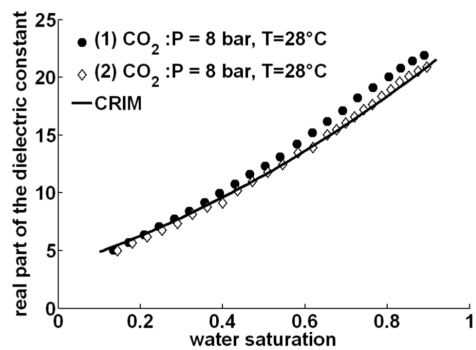


Figure 6: The real part of the dielectric constant as function of the water saturation during imbibition for $f = 3\,\mathrm{MHz}$.

As can be seen from our experimental results, the data can be modeled using the CRIM model [9]. As input model we use $\varepsilon_{\text{gas}} = 1$ for gas, $\varepsilon_{\text{water}} = 77.5$ for water and $\varepsilon_{\text{grains}} = 6.1$ for the sample. For the drainage results the CRIM estimates the experimental data over the entire saturation range with an acceptable deviation. Figure 6 shows that the CRIM fits well for the experimental results for the imbibition experiment 2 and underestimates the results of experiment 1. Comparison of the dielectric constant for drainage and imbibition shows that dielectric hysteresis is measured. In contrast to resistivity measurements [5] the imbibition values for the dielectric constant are higher than the drainage values. The reasons of the dielectric hysteresis are found in the distribution of the water and gas phase as well as the surface area [7]. For a better understanding of this phenomenon, low frequency data must assist in the interpretation.

5. CONCLUSIONS

- The experimental tool has proven to measure adequate dielectric constants for different fluids and materials.

- We have developed an experimental tool with which reproducible data for both the capillary pressure and the complex dielectric constant are measured.
- The dielectric values for drainage and imbibition are validated with CRIM, which indicates the correct measurement of the water saturation.
- Hysteresis is found for both the capillary pressure and the dielectric constant.
- The imbibition dielectric values are higher than the drainage values.
- Low frequency data is necessary for a better interpretation of the dielectric measurements.

ACKNOWLEDGMENT

The research presented in this work was carried out as part of the CATO program: CO_2 Capture, Transport and Storage in the Netherlands (http://www.CO2-cato.nl/) and the DIOC water project. The financial support is gratefully acknowledged. We thank L. Vogt, P. S. A. de Vreede and H. G van Asten for technical support.

REFERENCES

1. Fleury, M. and D. Longeron, "Combined resistivity and capillary pressure measurements using micropore membrane technique," *J. Pet. Sci. Eng.*, Vol. 19, 73–79, 1998.
2. Anderson, W. G., "Wettability literature survey-part 2: wettability measurement," *J. Pet. Sci. Technol.*, SPE 13933, 1246–1262, 1986.
3. Morrow, N., "Physics ad thermodynamics of capillary action in porous media," *Ind. Eng. Chem.*, Vol. 62, No. 6, 32–56, 1970.
4. Reeves, P. and M. A. Celia, "Functional relationship between capillary pressure, saturation and interfacial area as revealed by a pore-scale network model," *Water Resour. Res.*, Vol. 32, No. 8, 2345–2358, 1996.
5. Knight, R., "Hysteresis in the electrical resistivity of partially saturated sandstones," *Geophysics*, Vol. 56, No. 12, 2139–2147, 1991.
6. Ngyuen, B.-L., J. Bruining, and E. C. Slob, "Hysteresis in dielectric properties of fluid-saturated porous media," *Proceedings of SPE Asia Pacific Improved Oil Recovery Conference*, Kuala Lumpur, Malaysia, October 1999.
7. Knight, R. J. and A. Nur, "The dielectric constant of sandstones, 60 kHz to 4 MHz," *Geophysics*, Vol. 52, No. 5, 644–654, 1987.
8. Longeron, D. G., M. Argaurd, and J. P. Feraud, "Effects of overburden pressure and the nature and microscopic distribution of fluids on electrical properties of rock samples," *SPE Format. Eval.*, Vol. 4, 194–202, 1989.
9. Knight, R. and A. Abad, "Rock/water interaction in dielectric properties: experiments with hydrophobic sandstones," *Geophysics*, Vol. 60, No. 2, 431–436, 1995.
10. Ngyuen, B-L., J. Bruining, and E. C. Slob, "Effects of Wettability on dielectric properties of porous media," *Proceedings of SPE Annual Technical Conference and Exhibition*, Houston, Texas, October 1999.
11. Plug, W. J., S. Mazumder, J. Bruining, N. Siemons, and K. H. Wolf, "Capillary pressure and wettability behavior of the coal-water-carbon dioxide system at high pressures," *Proceedings of International Coalbed Methane Symposium*, Tusculoosa, Alabama, May 2006.
12. Jennings, J. W., D. S. McGregor, and R. A. Morse, "Simultaneous determination of capillary pressure and relative permeability by automatic history matching," *SPE Format. Eval.*, 322–328, 1988.
13. Longeron, D., W. L. Hammervold, and S. M. Skjaeveland, "Water-oil capillary pressure and wettability measurements using micropore membrane technique," *Proceedings of International Meeting on Petroleum Engineering*, Bejing, China, November 1995.

Numerical Simulation for the Effective Conductivity of Composite Medium in High Frequency

D. K. Yang, **X. Y. Hu**, **S. Y. Zhang**, and **M. Dai**

China University of Geosciences, China

Abstract— Because of the disunity of resolution and exploration range in applied geophysics, the effective medium theory (EMT) should be developed to help us to understand the geological microstructure. We extended the EMT to complex permittivity in high frequency, and calculated the imaginary part of effective complex permittivity of composite as the effective conductivity using Finite-Difference Time-Domain (FDTD) numerical method. The result we obtained is the intrinsic property of the equivalent medium, which has explicit geological signification to satisfy the requirement of interpretation.

1. INTRODUCTION

The electromagnetic (EM) method in applied geophysics can map the distribution and property of geological object underground efficiently, but there are always some important parameters referring to microstructure we cannot detect directly. For example, petroleum geologists have strong interest in the porosity and saturation of oil-bearing rocks. Unfortunately, most geophysical methods do not have enough resolution to study microscale anomaly, so balance between the exploring range and investigating resolution become a problem.

Effective medium theory (EMT) is a good solution to simplify the composite problem to equivalent homogeneous medium [3]. So researchers can concentrate on the relationship between heterogeneousness in microscale and effect in macroscale. Maxwell, Bruggeman, Rayleigh, Pauly and Schwan all contributed their distinguished works on how to predict the effective electric properties of composite exactly. Archie formula is a famous empirical mixing law in geophysics, by which the stratum physical parameters can be inferred from resistivity data after inversion.

However, most researches paid attention to static or low frequency field (under 1 MHz) considering conductivity or permittivity singly, and achievements of microwave mainly came from sample test in laboratory. The propagation of HF EM wave is governed by Helmholtz Equation ($\boldsymbol{H}$ field has the same format)

$$\nabla^2 E - \gamma^2 E = 0, \tag{1}$$

where $\gamma = j\omega\sqrt{\mu\tilde{\varepsilon}}$ is propagating constant, and complex permittivity $\tilde{\varepsilon} = \varepsilon - j\frac{\sigma}{\omega}$ containing both permittivity and conductivity characters the intrinsic electric properties of medium. Therefore, the EMT of complex permittivity is needed.

But the complex permittivity is a frequency-dependent parameter without any geological information; it also cannot satisfy the customary requirement of conductivity mapping by geophysicists. We noticed that signal attenuation (skin depth) due to conductive medium is the main effect during wave propagation, so it is rational to substitute the multi-phase mixture with homogeneous medium having equivalent attenuation ability. There are two main factors causing EM energy decay in earth medium: conductive loss and polarization loss. In our research, to meet the need of geological interpretation, we aimed at calculating the imaginary part (multiplied by circular frequency) of effective complex permittivity of the composite, which is defined as the effective conductivity in HF. This effective conductivity $\boldsymbol{\sigma}_{\mathit{eff}}$ is the intrinsic property of hypothetical equivalent medium following Ohm's Law

$$J = \sigma_{\mathit{eff}} E \tag{2}$$

and related only with conductive current by the transmission of electrons or ions in rocks.

In this paper, we adopt "black box" method applying numerical simulation to infer the medium properties from surveyed EM field behavior (as output). Our approach avoided some impractical preconditions in analytical modeling before, and would not be disturbed by irrelevant factors as in sample test. We can also simulate the extreme case using difference gridding while Maxwell Equation is still obeyed.

2. MODELING

To be easily-computed and comparable, our research began with a simplified two-dimensional mixing system abstracted from real-world rock background. We suppose that:

(1) There are two phases in the system, namely matrix and impurity; the grains of impurity are embedded in the matrix randomly.

Although high frequency EM wave has very short wavelength, the geological heterogeneousness is always much tinier than the exploring wavelength. For example, pore size in coarse grain sandstone of Woodbine C, Kurten Oilfield, Southeastern Texas, is scaled in hundreds μm [4], and if we use EM wave at frequency 1 GHz, the wavelength will be several decimeters (general conductive ground) at least. It is Rayleigh region according to radar theory. Because of such striking scaling contrast, we do not need to care about the shape of impurity grain. So the simplest spherical scatter is chosen to describe the heterogeneousness in rocks. Figure 1 shows grid simulation of random distribution in disseminated sandstone. Also, the grain size and sorting grade are adjustable.

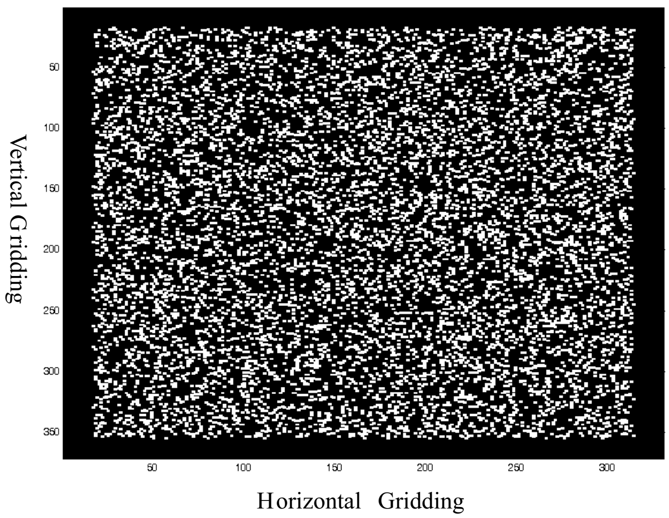


Figure 1: Analog of two-phase mixture (The white spots are embedded scatters; the matrix medium is black; the dotted area is about 10 cm^2 divided into 340 × 300 cells).

(2) What we simulated is a differential area $\boldsymbol{dS}$ of 2D EM field in the earth. It is small enough to not take the spatial asymmetry of EM wave into account; therefore, the hypothesis of incident plane wave can give good approximation. Farther, we think that the current density $\boldsymbol{J}$ and electric intensity $\boldsymbol{E}$ are not strongly-varied in such a small area, so the effective conductivity of $\boldsymbol{dS}$ can be obtained by average Ohm's Law [5]

$$\sigma_{\mathit{eff}} = \frac{\langle J \rangle}{\langle E \rangle}, \tag{3}$$

where $\langle J \rangle$ is average current density and $\langle E \rangle$ is average electric intensity in $\boldsymbol{dS}$. In order to approach better accuracy, the size of $\boldsymbol{dS}$ is restricted within limit of centimeters.

(3) The intrinsic properties of both materials in our system, including conductivity, dielectric constant, magnetic permeability, magnetic conductivity, are evaluated beforehand. If they are frequency-dependent or temperature/pressure-dependent, supposedly, there have been precorrection before computation. The geochemical interaction is also not considered. It means that we only want to know how dose the media distribution and electric properties affect the effective conduc-

Table 1: Experimental parameters.

Parameters (Unit)	Typical Values
Conductivity (S/m)	1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}
Relative Dielectric Constant	1, 5, 20, 80

tivity. Some typical parameters were determined by near-surface petrophysics as shown in Table 1 [6].

3. NUMERICAL METHOD

Finite-Difference Time-Domain (FDTD) is a well-developed finite difference method operating in time domain, by which we can obtain the numerical solution of Maxwell equation with good speed and enough accuracy. And we feel very free to put arbitrarily-shaped impurity into the difference grid.

In this research, we used centered difference with two-order accuracy and perfectly matched layer (PML) absorbing boundary condition (Berenger, 1994). The wave was 2D TM mode concerning three components ($\boldsymbol{E_Z}$, $\boldsymbol{H_X}$, and $\boldsymbol{H_Y}$. The time series of sinusoidal $\boldsymbol{E_Z}$ source was generated by 1D FDTD, and added into the impurity area via connective boundary surrounding the white-dotted rectangle in Figure 1. Totally, there were three boundaries from outside to inside: absorbing boundary, connective boundary and target boundary.

The grid size $\boldsymbol{dx}$ (or $\boldsymbol{dy}$) in our program is constantly 10^{-4} m to give enough fitting to the grain size in rocks, especially sedimentary rocks. We designed an even gridding system of 300×300 size following staggered cell style (Yee, 1966). The impurity distribution was realized by designating different electric parameters at the gridding nods randomly. According to Courant's condition and general geological background, we conservatively chose $\boldsymbol{dt} = 1.67 \times 10^{-13}$ s as time separation between two iterations. The iteration would not stop until the field intensity in each nod become stable.

Repeated computations and averaging is helpful to improve the final accuracy.

All computation has done on the platform of Mathworks MatLab 7.0, by which we were able to monitor the frame movie of wave propagation while running.

4. VALIDITY TEST

We tested a set of data from real world, and compare the computational result with traditional formulaic solution to verify the validity of our numerical method.

The Tengger Desert in Northeastern China is famed for silver sand having very small grain diameter. A sand-air mixing sample from there was measured in wave-guide at frequency 33.5 GHz. The lab investigation showed the complex relative permittivity of that sample was $5.43 - \boldsymbol{j}0.074$, and volume fraction of air was 0.08 [9].

Because of the EM wavelength forty times greater than the pore size in the sample, it is reasonable to refer this composite problem in dynamic EM field to Bruggeman's formula [10]

$$f_1 \frac{\varepsilon_1 - \varepsilon_{eff}}{\varepsilon_1 + 2\varepsilon_{eff}} + f_2 \frac{\varepsilon_2 - \varepsilon_{eff}}{\varepsilon_2 + 2\varepsilon_{eff}} = 0 \tag{4}$$

where f_1, f_2, ε_1, ε_2 are volume fractions and complex permittivities of matrix and impurity respectively, and ε_{eff} is the effective complex permittivity. After substituting variables: $f_1 = 0.92$, $f_2 = 0.08$, $\varepsilon_1 = 5.43 - j0.074$, $\varepsilon_2 = 1$ (pure air) into Bruggeman's formula, we obtained $4.947 - \boldsymbol{j}0.0654$ as the effective complex permittivity of sand-air composite, so the effective conductivity of analytical solution was 0.122 S/m.

We also ran our program with the same parameters above, and then got a group of simulating results by five repeating computations as listed in Table 2. Statistic analysis demonstrated that the RMS was in allowable tolerance; therefore the arithmetical average value was accepted. The numerical solution showed good agreement to analytical solution within the error of 0.24%.

Table 2: Five simulations of sand-air composite and the statistic analysis (Unit: S/m).

Simu. 1	Simu. 2	Simu. 3	Simu. 4	Simu. 5	Average	RMS
0.126	0.121	0.123	0.125	0.123	0.124	0.17%

More numerical tests have proved that the initial amplitude of incident wave and gridding number did not impact the final result significantly.

5. COMPUTATIONAL RESULTS

Several groups of results from numerical experiments are summarized in Table 3.

Table 3: Simulation results.

Freq.	Matr. Cond.	Matr. Perm.	Impu. Cond.	Impu. Perm.	Volume Ratio	Eff. Cond.
10^8 Hz	0.0001 S/m	1	0.001 S/m	1	80/20	0.00028 S/m
10^8 Hz	0.0001 S/m	1	1 S/m	1	20/80	0.7968 S/m
10^8 Hz	0.0001 S/m	1	0.001 S/m	1	60/40	0.00046 S/m
10^9 Hz	0.0596 S/m	5.43	0	1	90/10	0.0536 S/m
10^9 Hz	0.01 S/m	5.0	1 S/m	80	85/15	0.1679 S/m
10^9 Hz	0.0596 S/m	5.43	0	1	40/60	0.0238 S/m
10^9 Hz	0.1 S/m	1	1 S/m	1	80/20	0.2804 S/m
10^{10} Hz	0.0001 S/m	1	1 S/m	1	60/40	0.3956 S/m
10^{10} Hz	0.0001 S/m	1	1 S/m	10	60/40	0.4010 S/m

6. DISCUSSION & CONCLUSION

(1) The EMT can be studied by FDTD methods, but the numerical simulation is too time-consuming. We have to suffer much more iterations if the wavelength is much greater than the heterogeneousness size. So we are looking forward to the advance of computational electromagnetics.

(2) Although our program agreed Bruggeman's formula well, we had batter conclude after all band verification in further work. If possible, laboratory experiment or real rock test must be carried out.

(3) In conductivity contrasted two-phase medium, there will be a "dominant" phase having much greater conductivity relatively. The effective conductivity is controlled by the conductivity and volume fraction of the dominant phase.

(4) If the conductivities of two phases are close, the effective conductivity is contributed by both phases with their volume fractions as weighting. In this case, the difference of permittivity should be seriously concerned, if any.

(5) The grain size of impurity doesn't affect effective conductivity apparently, but we have observed slight decrease of effective conductivity with the grain growth. Absolutely, the simulating accuracy becomes worse when the grain size is larger.

ACKNOWLEDGMENT

We sincerely thank the funding support of "Academic Research and Innovation Fund for Graduate Students" of China University of Geosciences, Wuhan China.

REFERENCES

1. Maxwell, J. C., *Treatise on Electricity and Magnetism*, Oxford Univ. Press, U.K., 1873.
2. Pavlin, M., T. Slivnik, and D. Miklavcic, "Effective conductivity of cell suspensions," *IEEE Trans. on Biom. Eng.*, Vol. 49, No. 1, 77–80, 2002.
3. Bruggeman, D. A. G., "Berechnung verschiedener physikalischer Konstantenvon heterogenen Substanzen," *Ann. Phys.*, Vol. 24, 636–641, 1935.
4. Robert, R. B., *Reservior Sandstone*, Prentice-Hall Inc., Englewood cliffs, New Jersey, 1986.
5. Ma, Y. Q. and Z. Y. Li, "Effective-medium theory for two-phase random composite/system," *Acta Physica Sinica*, Vol. 39, No. 3, 457–462, 1990.
6. Touloukian, Y. S., W. R. Judd, and R. F. Roy, *Physical Properties of Rocks and Minerals*, McGraw-Hill, USA, 1981.
7. Taflove, A., *Computational Electromagnetics — The Finite-Difference Time-Domain Method*, Artech House, Boston MA, 1995.

8. Ge, D. B. and Y. B. Yan, *FDTD Method of Electromagnetic Wave (2nd Edition)*, Xidian Univ. Press, Xi'an China, 2003.
9. Dong, Q. S., "Physical characteristics of sand and dust in different deserts of china," *Chin. J. Rad. Sci.*, Vol. 12, No. 1, 15–25, 1997.
10. Itzhak, W., J. Joshua, and M. H. Cohen, "Theory of optical and microwave properties of microscopically inhomogeneous materials," *Phys. Rev.*, Vol. B15, No. 12, 5712–5723, 1977.

Stitching a Reference Plane Split Using Routing Layer Traces to Improve I/O Bus Signal Integrity

Helen K. Pan and Christopher Y. Pan
Intel Corporation, Hillsboro, OR, USA

Abstract— While on-board interconnects are operating at much higher frequencies, the cost of the overall computing system has been decreasing. Computer manufacturers are therefore continuously looking for ways to lower the production cost. In a mobile computing system, where size is a main constraint and expensive high layer count printed circuit board (PCB) technology has been used to control routing congestion, using cheaper technologies with a smaller number of routing layers is an attractive mean of controlling manufacturing cost. On the other hand, routing on a severely size-constrained mobile platform with a smaller number of routing layers can inadvertently expose wide high speed digital interfaces to non-ideal conditions such as reference plane gaps produced by various power and ground islands. Multi-GHz operation of these high speed signals also means that the traditional mitigation method of adding bypass capacitors across the reference gaps is becoming ineffective. In this paper, we propose to use stitching patch underneath the splitting power and ground plane that can be used in conjunction with bypass capacitors to mitigate the effect of these reference plane gaps on signal quality, as well as system EMI and ESD.

1. INTRODUCTION

In the past few decades, microprocessor operating frequency has increased from a few megahertz to multi gigahertz. Board-level high speed digital interfaces, such as system bus, system memory, and system I/O, that carry information between microprocessor, memory controller, I/O controller, and other components, however, have been operating at a lower frequency. This is in part because transmission line and other higher order electromagnetic effects can dominate board-level interconnect designs due to electrically significant interconnect sizes. Yet, due to the ever increasing demand for more bandwidth, operating frequencies of these board-level high speed interfaces have also been doubling in roughly two-year cycles. Parallel high speed interfaces such as Front-Side Bus (FSB) and Double Data Rate-III (DDR-3) SDRAM, for example, are fast approaching 1 GT/s, where significant frequency content can extend well into multi-GHz. In the mean time, PCI Express-like serial differential high speed interfaces, already operating at speeds over 1 GT/s, are designed into new cross-chip interconnects such as the Direct Media Interface (DMI). Routing a system with such a large number of high speed digital buses while insuring signal integrity, EMI compatibility, and ESD susceptibility, is a challenge.

The power and ground planes are generally used as reference planes for high speed I/O buses. As the number of power supply rails increased dramatically in desktop and mobile systems, multiple power planes on a real estate constraint platform often forces high speed interfaces to cross different reference planes. It is a known fact that when high speed transmission lines cross a reference plan split, cross coupling and radiated emission are greatly increased [1–5]. Increased cross coupling leads to greater-than-expected signal quality degradation, and on critical interfaces, have been known to cause signaling failures. Added radiated emission, on the other hand, may lead to system Electromagnetic Interference (EMI) failure. We have known cases of such EMI failures such as when DDR2 signals crossed a reference plane split and coupling large amounts of common mode noise to nearby LVDS signals, which in turn lead to platform EMI failure.

In most cases when reference plane crossing is unavoidable, discrete decoupling capacitors are used to reduce its impact to signal integrity, as show in Figure 1. However, discrete decoupling capacitors have finite parasitic inductance, making them less effectiveness at high frequencies. Typical on-board decoupling capacitors have a resonance frequency on the order of a few hundreds Mega Hertz (MHz). On the other hand, however, typical high speed interfaces have frequency contents greater than 1 GHz. Discrete capacitors also add to board cost.

In this paper, we "stitch" a reference plane split below with a conductor patch on the second routing layer, and use the parasitic capacitance between the patch and the second reference plane to reduce the effect of reference plane crossing.

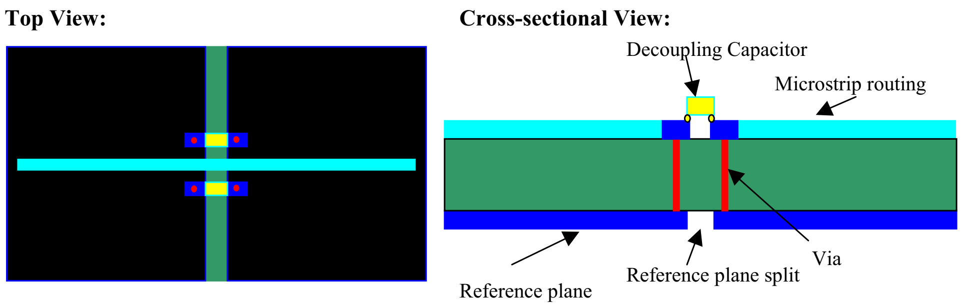
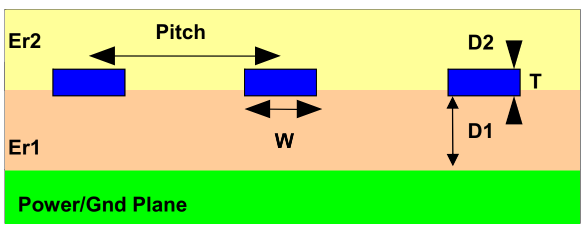


Figure 1: Decoupling capacitors are used to reduce the effect of crossing a reference plane split.

2. REFERENCE PLANE DISCONTINUITY STUDY

We studied the microstrip line performance due to the reference plan gap based on the typical microstrip geometries on computer mother board as shown in Figure 2. All microstrip transmission line study in this paper is based on the geometries descript in Figure 2 with 5000 mil length.



Microstrip	Typical value
Er1	4
Er2	3.4
D1	4 mil
D2	1 mil
T	2 mil
W	5 mil

Figure 2: Typical microstrip transmission line geometry for computer mother board.

Interconnects over a split power/ground plane are evaluated using 3D EM simulator HFSS. Based on the HFSS simulation, we observed the E-field discontinuity when the reference plane has slot or complete gap as shown in Figure 3. The splitting reference plane effects to interconnect transmission performance were plotted in Figure 4 with various gap widthes comparing with ideal transmission line.

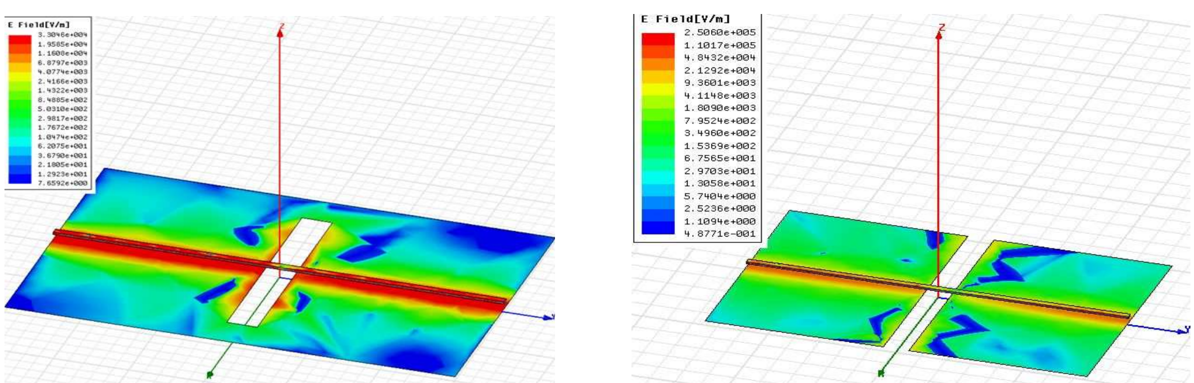
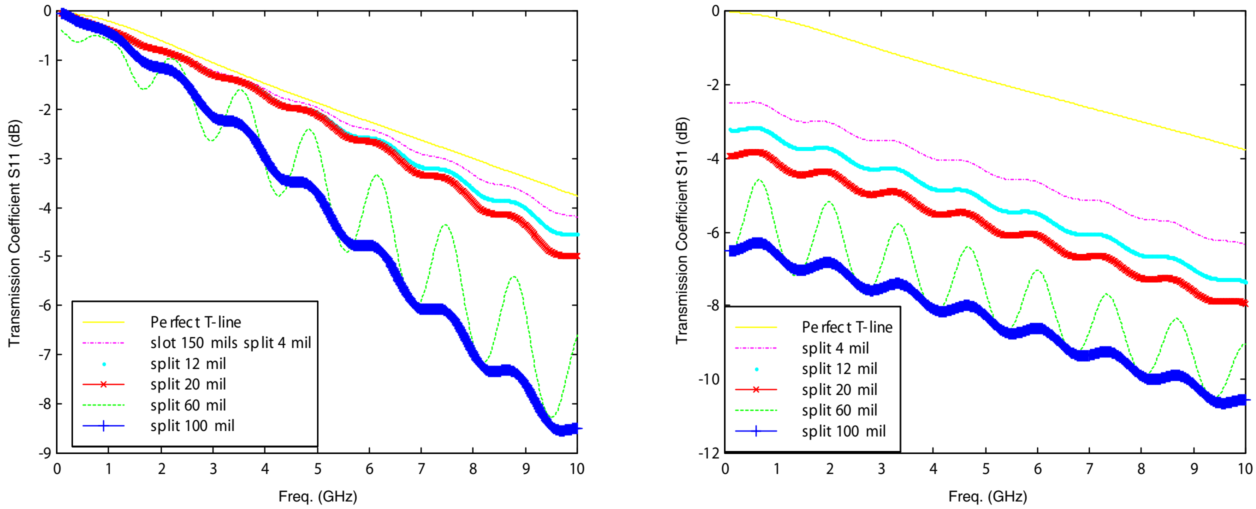


Figure 3: E-field distribution plot for reference plane slot (left) and complete split (right).

From the Figure 4, we observed that the gaps in reference planes cause discontinuities on microstrip lines to increase transmission loss as frequency goes up. Splitting reference plane causes more severe transmission line performance degradation comparing with slot opening in reference plane. As the width of the slot/split increase, the trace transmission loss increase further.

3. MITIGATION METHOD FOR SPLITTING REFERENCE PLANE

Our approach to solve splitting reference planes is to stitch a conductor patch on the second routing layer below the reference plane split, as show in Figure 5. Here we use vias to connect the patch



(a) (b)

Figure 4: Microstrip line transmission coefficient with different split gap width 4, 12, 20, 60, 100 mil for 150 mil long slot gap on reference plane (left) and complete reference plane split (right).

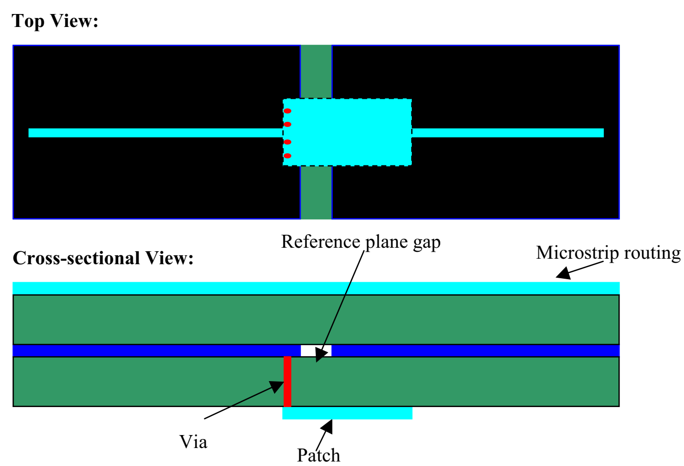


Figure 5: A patch on the routing layer below the reference plane is used to reduce the effect of reference plane split.

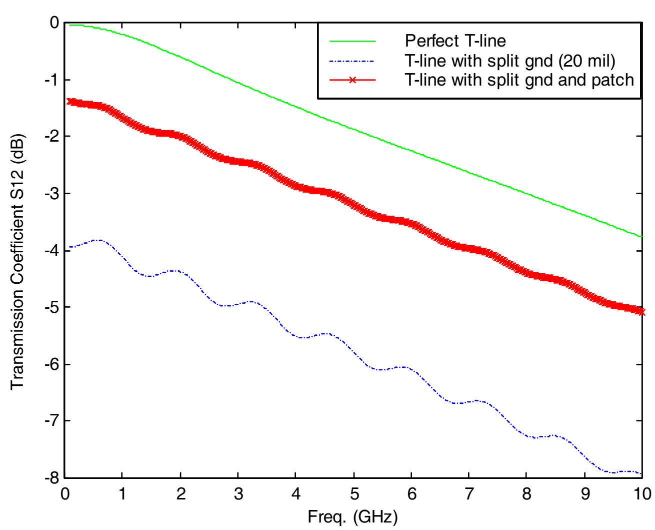


Figure 6: Input reflection coefficient comparison of perfect T-line, T-line with split ground (20 mil gap) and patch (40 mil wide and center via) with split ground T-line.

to one reference plane and use parasitic capacitance between the patch and the second reference plane to reduce the effect of reference plane crossing. Using the parasitic capacitance between two parallel plates (patch and the second reference plane) allows a continuous return path beneath the high speed I/O signal that must cross the reference plane split. The transmission loss plot in Figure 6 is shown the conducting patch beneath the splitting reference plane reduce reference plane split effects by $\sim 3\,\text{dB}$. Compare to the prior art of using discrete decoupling capacitors to provide a return path, this approach is much more effective at high frequencies and is more cost effective.

The main advantage of using a conductor patch on the routing layer below the reference plane to reduce the effect of reference plane split is its effectiveness at high frequencies. Past its self resonance frequency, a discrete decoupling capacitor is no long effective. Given most of the high speed I/O signals have significant frequency content in Giga-Hertz range, using decoupling capacitors is not an effective mean to reduce the impact of reference plane splits. The conductor patch, on the other hand, is effective for frequencies up to 100 GHz. The BOM cost is also reduced because discrete components are not needed.

4. CONCLUSION

Impact of reference plane split to transmission line performance was presented in this paper. A stitching conducting patch beneath splitting power/ground plane to reduce the reference plane split discontinuities effect was proposed. Based on the HFSS simulation results, we observed stitching conducting patch effectively reduce the reference plan split effects.

REFERENCES

1. Senthinathan, R., et al., "Reference plane parasitics modeling and their contributions to power and ground path effective inductance as seen by the output drivers," *IEEE Trans. Microwave Theory Tech.*, Vol. 42, 1765–1773, Sept. 1994.
2. Chen, H., et al., "Effects of gaps and bypass capacitors on interconnect of PCB with multilayered geometry," *IEE Proc. Microw. Antenna Propag.*, Vol. 148, No. 3, 167–173, 2001.
3. Cangellaris, A. C., "Electromagnetic characterization of high-speed VLSI interconnects with perforated reference planes," in *Proc. of IEEE AP-S*, Chicago, IL, USA, 98, July 1992.
4. Xue, Z., et al., "Electrical characteristics of multiconductor interconnects with perforated reference planes," in *Proc. of IEEE 3rd Topic Meeting on EPEP*, Monterey, CA, USA, 98–100, Nov. 1994.
5. Yuan, F., "Electromagnetic modeling and signal integrity simulation of power/ground networks in high speed digital packages and printed circuit boards," in *Proc. of Design Automation Conference*, San Francisco, CA, USA, 421–426, June 1998.

Signal Waveform Distortion on Terminatorless Transmission Line of UART-CSMA/CD Control Network

C. Ninagawa[1], **K. Yokohama**[1], **F. Aoi**[1], **and Y. Miyazaki**[2]

[1]Mitsubishi Heavy Industries, Ltd., Japan
[2]Aichi University of Technology, Japan

Abstract— Regarding a control network for a building air-conditioning system, signal waveforms on the transmission lines without terminators were calculated and compared with experiments. The calculation predicated the severe distortion in the signal waveform due to line end reflections in the cases of improper line connections such as loops or long branches.

1. INTRODUCTION

Cost is one of the most important factors in control networks for building facilities such as airconditioning or lighting systems. As the total length of a control network is very long such as a thousand meters, reducing installation cost is strongly demanded. We have proposed a low-cost media access control method UART-CSMA/CD (Carrier Sense Multiple Access with Collision Detection using Universal Asynchronous Receiver Transmitter) which enables to omit terminators for impedance matching on the transmission line [1]. However, the affection of signal reflections has been studied only in the case of a point-to-point connection using a lossless transmission line model [2].

In this research, we use our new model for waveform calculations assuming lossy transmission line and complex line connections. The calculation results have been verified by comparing with the experimental results. Our new model predicts the level of the returning reflection signal for a variety of terminatorless line connections. In the case of an improper connection such as a loop connection, our calculation and experiment indicate a wide distortion region in the front part of one bit pulse resulting in wrong bit sampling. We confirmed that our calculation is useful for prior evaluation of actual line connections in the building.

2. UART-CSMA/CD CONTROL NETWORK

Figure 1 shows the concept of a control network for air-conditioning facility. In a typical building airconditioning system, there are hundreds of controllers distributed throughout the building. Fig. 2 shows the structure of an UART-CSMA/CD controller connected to the transmission line. The total line length often reaches as long as approximately 1000 m. Therefore, not only the costs of each controller but also cost of network installation should be decreased.

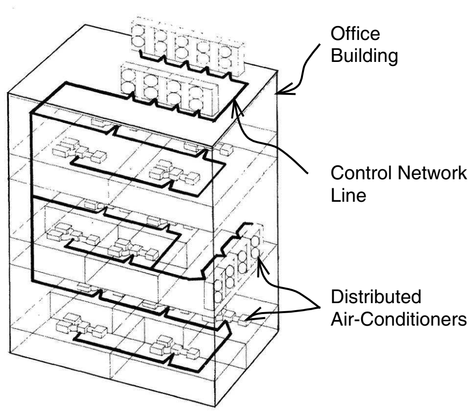


Figure 1: Air-conditioner control network.

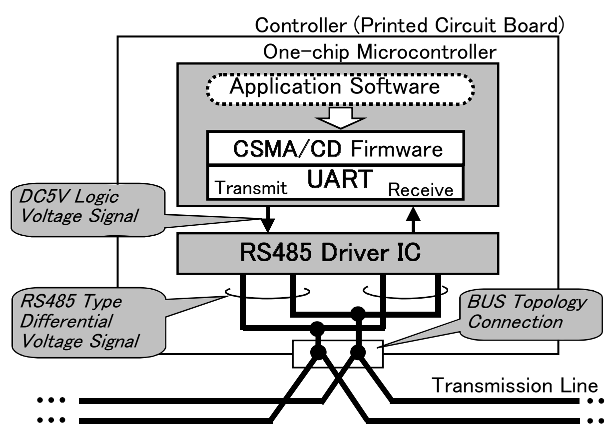


Figure 2: UART-CSMA/CD controller.

Since this network has to be installed in the most inexpensive way, any network wiring devices such as line terminators are not allowed to be used. It is well known that impedance mismatching at the terminal points adversely affects on the signal waveform in the case of a significantly long line.

As shown in Fig. 1, the routes of transmission lines in the building are very complicated including many branches, stubs, and very long lines. Therefore, it is extremely important to predict signal waveform distortions on the terminatorless lines in the particular case of the building before the construction of the control network.

3. SIGNAL TRANSMISSION WAVEFORMS

3.1. Calculation Model

In this research, we use a mathematical model for a lossy transmission line. The fundamental differential equations of signal voltage $v(x, t)$ and current $i(x, t)$ for a distributed constant circuit are

$$\begin{cases} -\dfrac{\partial v(x,t)}{\partial x} = L_0 \dfrac{\partial i(x,t)}{\partial t} + R_0 i(x,t) \\ -\dfrac{\partial i(x,t)}{\partial x} = C_0 \dfrac{\partial v(x,t)}{\partial t} + G_0 v(x,t) \end{cases} \tag{1}$$

where L_0, R_0, C_0, and G_0, are first order circuit parameters of the transmission line. Since it is extremely difficult to obtain an analytical solution in cases of complex line connections that make the boundary conditions very complicated, we have made a discrete model in which each line consists a number of small segments with lumped circuit parameters. Therefore, the differential equations (1) were converted to the difference equations for a segment of the transmission line as shown in Fig. 3. Using Kirchihoff's law, the difference equations for a segment are expressed as follows,

$$\begin{cases} \Delta v = -L_0 \Delta x \dfrac{di(t)}{dt} - R_0 \Delta x i(t) \\ \Delta i = -C_0 \Delta x \dfrac{dv(t)}{dt} - G_0 \Delta x v(t) \end{cases} \tag{2}$$

By giving boundary conditions, initial conditions, and a source voltage function, the difference equations were solved using the numerical calculation algorithm. Since the bit rate of our signal is very low such as 19.2 or 9.6 kbps, the length of one segment of the model is set as 50 m, that is, the number of segments is 20 in the case of full line length.

Figure 3 shows an example of a complex connection of the transmission lines that consist of many segments. The example is called a loop connection. Although the Fig. 3 shows only an example of a loop connection, our model can support a variety of connection topologies such as basic point-to-point, long branch, and short stub connections.

3.2. Numerical Calculations and Observations

We have compared signal waveforms calculated using the above-mentioned model with those observed by experiments. The bit rate has been chosen as $R_{BIT} = 19.6$ Kbps, that is, the bit width is $W_{BIT} = 52\,\mu$s. The source voltage is an alternating pulse series of $E_1 = \pm 3.0$ V, the source

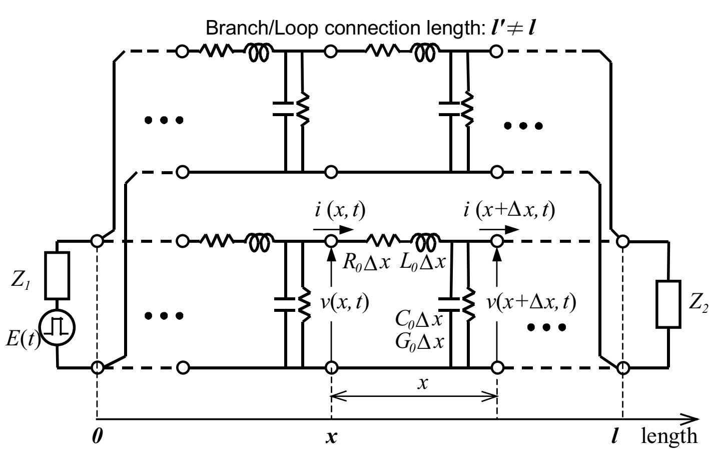


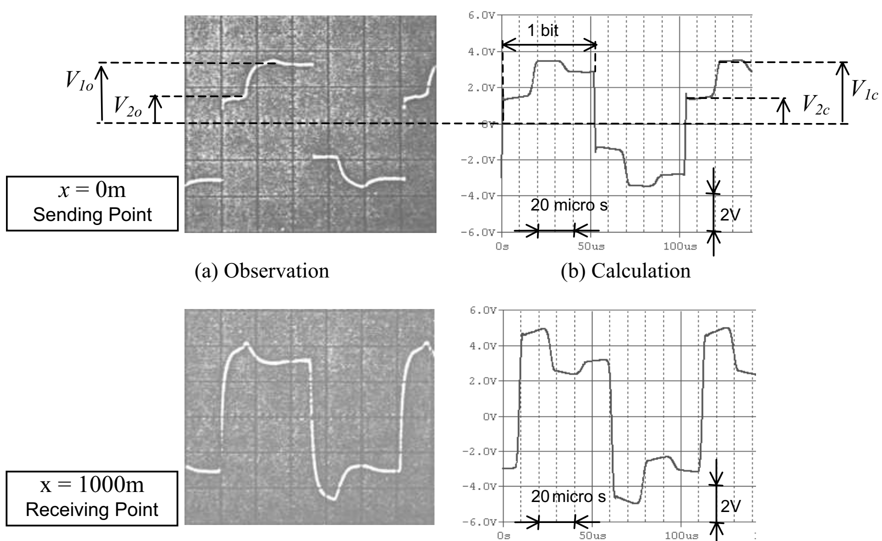
Figure 3: Example of calculation model for improper line connection.

Table 1: Measured cable parameters.

Parameter	Value
R_0	1.85×10^{-2} [Ω/m]
L_0	0.59×10^{-6} [H/m]
C_0	0.12×10^{-9} [F/m]
G_0	6.07×10^{-8} [S/m]

impedance is $Z_1 = 28\,\Omega$, and the receiving impedance $Z_2 = 12\,\mathrm{K}\Omega$, where those are of an RS-485 driver/receiver IC. The type of the cable was a polyvinyl chloride (PVC) twisted cable with $2\,\mathrm{mm}^2$ (AWG14) wires. The values of cable parameters have been measured as shown in Table 1. As the characteristic impedance of the cable is $Z_0 = 69\,\Omega$ and the receiving impedance of the RS-485 receiver is $Z_2 = 12\,\mathrm{K}\Omega$, severe impedance mismatching will take place at the end if any terminator is not connected.

Figure 4 shows the calculation and observation of the signal waveforms for the case of a basic point-to-point connection without terminators. At the sending point, $x = 0\,\mathrm{m}$, a typical kink shape is seen on the rising part of every bit pulse due to returning reflection signal form the far end. The reflection signal level ratio at the sending point was $L_{Ro} = V_{1o}/V_{2o} = 2.0$ by the observation and $L_{Rc} = V_{1c}/V_{2c} = 2.2$ by the calculation. The shape of the signal waveforms by the calculations are similar to those observed on the actual transmission line. At the receiving point, $x = 1000\,\mathrm{m}$, the waveform shape of the calculation is less similar to that of the observation than the case of the sending point. However, the calculated waveform correctly indicates a feature of the reflection in the case that the receiving impedance Z_2 is very large as compared to the cable impedance Z_0.



(a) Observation (b) Calculation

(c) Observation (d) Calculation

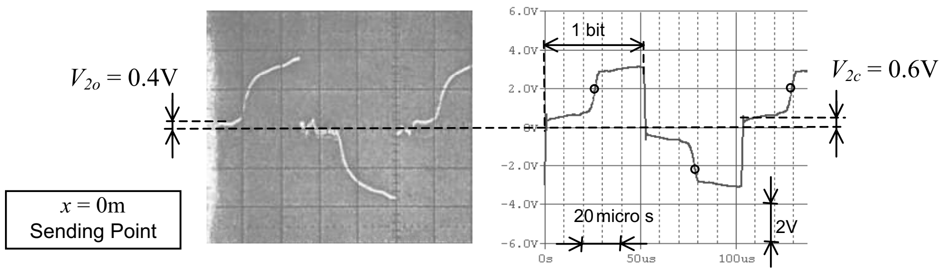
Figure 4: Signal waveforms for the transmission line of 1000 m.

4. COMPLEX LINE CONNECTIONS

We calculated the signal waveforms in many cases of transmission line connections such as extremely long point-to-point, long branch, short stub, and loop connections, which might appear in the actual installations. Our model indicates that the most severe distortion occurs in the case of loop connections.

Figure 5 shows signal waveforms by observation and calculation in the case of a loop connection. In this case, the sending point and receiving point are connected by two lines consequently the lines form a loop connection as shown in Fig. 3. It was assumed that the length of one line is 1000 m and that of the other line is 2000 m. Since there are two returning reflected signal traveling for different periods, the rising part of the waveform at the sending point is heavily distorted. The signal level of the distorted region was $V_{2o} = 0.4\,\mathrm{V}$ in the case of observation and $V_{2c} = 0.6\,\mathrm{V}$ in the case of calculation. Since the UART reads the signal level at the middle of the bit width as shown in small circles in the Fig. 5(b), there is a possibility for the UART to read the signal level in the distorted region. Our calculation successfully predicts the possibility of misjudge of the digital value "0" or "1" in the case of such unstable samplings. Although this is a simple case of packet sending by one controller, the voltage level sampling will be extremely difficult in the case of packet collision because the waveforms of the superimposed bit signals that will suffered form the more severe distortions.

Figure 6 shows the relationship between the signal waveforms discussed above and our previously proposed communication protocol UART-CSMA/CD. Fig. 7 shows the logical performance of our



(a) Observation (b) Calculation

Figure 5: Signal waveforms at sending point for 1000 m/2000 m loop connection.

protocol for the bit rate of 9.6 kbps, that is, the throughput versus the offered communication load evaluated by computer simulations and experiments. In the case of simulation, it was assumed the signal on the transmission line is perfect and never contains any distortion in the waveform. In the region of light offered load, the simulation and experiment show the same characteristics. However, in the region of extremely heavy offered load where packet collisions occur intensively, there are differences between simulation and experiment data. There is a possibility that the difference might be related to the signal waveform distortions due to impedance mismatching on the terminatorless transmission line.

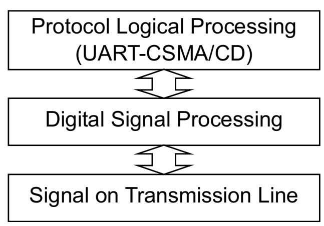


Figure 6: Relation between signal waveform and logical protocol.

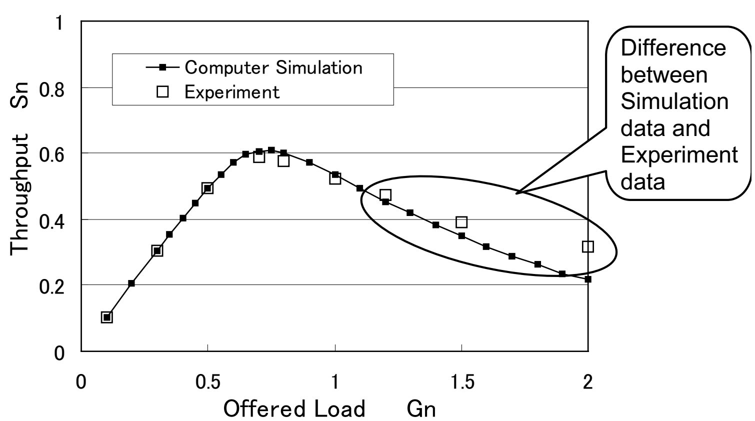


Figure 7: Communication Performance of UART-CSMA/CD.

5. CONCLUSIONS

In this paper, we have shown that our calculation model of lossy transmission line is effective to predict waveform distortions by reflections in the cases of a variety of line connection topologies. We confirmed that our calculation is useful for prior evaluation of actual line connections in the building. In future, we will study the relationship between the waveform distortions and the collision detection failures.

ACKNOWLEDGMENT

The authors would like to thank Kazuhiro Omura of Suntechno Co., Ltd. for his assistance in numerical calculations.

REFERENCES

1. Ninagawa, C., K. Yokohama, F. Aoi, and Y. Miyazaki, "Performance analysis on air-conditioner control CSMA/CD protocol by embedded microcontroller UART collision detection," *IEEJ Trans. IA*, Vol. 126, No. 5, 665–672, 2006 (in Japanese).
2. Ninagawa, C., K. Yokohama, F. Aoi, H. Otake, and Y. Miyazaki, "Transmission Characteristics of UART-CSMA/CD control network with one-chip microcontroller and RS-485 driver IC," *Proceedings of Progress in Electromagnetic Research Symposium, PIERS 2006-TOKYO*, 271–275, Aug. 2–5, 2006.

Input Impedances and Current Distributions for Meander Line Antennas with Planar Coupled Parasitic Meander Element

K. Taki[1] **and Y. Miyazaki**[2]

[1]NID Research and Development Department, Brother Industries, Japan
[2]Department of Electronics and Information Eng., Aichi University of Technology, Japan

Abstract— Meander line antennas with planar coupled parasitic meander elements have been proposed in order to increase the antenna input resistances for RFID tags. The input impedances, current distributions and the bandwidths have been investigated by using the method of moment. At the second resonant frequency in the UHF band, the current flow along the higher current concentration on each element were in the same direction and higher input resistance were obtained.

1. INTRODUCTION

Radio frequency identification (RFID) technology using UHF bands is rapidly developing for various applications and small tag antennas are required for the purpose of reducing the tag sizes [1, 2]. An RFID tag consists of an antenna and an IC chip. Since passive RFID tags get all the energy needed to operate from the electromagnetic wave transmitted by readers, proper impedance match between the antenna and the IC is very important. The antenna should be directly matched to the IC without adding any external matching networks because of lower cost and easier fabrication. In the UHF band, especially below 1 GHz, meander line antennas are very attractive to reduce tag sizes [3]. However, the input resistance of a small meander line antenna is too low and it is difficult to achieve direct impedance match.

We have proposed meander line antennas with planar coupled parasitic meander elements in order to increase the antenna input resistance [4]. In this paper, the relation between input impedance and current distribution is investigated by using the method of moment. The bandwidth is presented for modified meander configurations.

2. CONFIGURATION OF A MEANDER LINE ANTENNA WITH A PARASITIC MEANDER ELEMENT

The antenna configuration is shown in Fig. 1. A parasitic meander element is placed close to a driven element so that the each narrower meander section should be inserted into the meander section of the driven element. This planar coupled configuration is convenient to RFID tags because the total occupied area is small. The antenna input resistance and the resonant frequencies vary with the width w_2 and the spacing w_3.

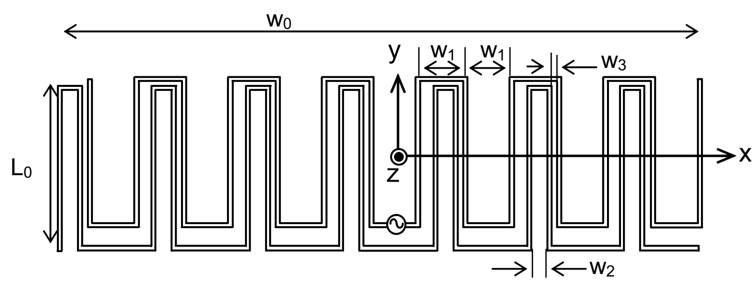


Figure 1: A meander line antenna with a parasitic meander element.

3. INPUT IMPEDANCE AND CURRENT DISTRIBUTION

The length and width of the total antenna were chosen to be $w_0 = 67$ mm and $L_0 = 18.5$ mm, respectively. This length L_0 is about 0.2λ at 915 MHz. The line width of each element was set to be 0.5 mm. Each element is assumed to be lossless and the input impedance was calculated with the method of moment. The input impedance is plotted on the Smith chart as shown in Fig. 2 where $w_1 = 0.5$ mm and $w_2 = w_3 = 0.2$ mm.

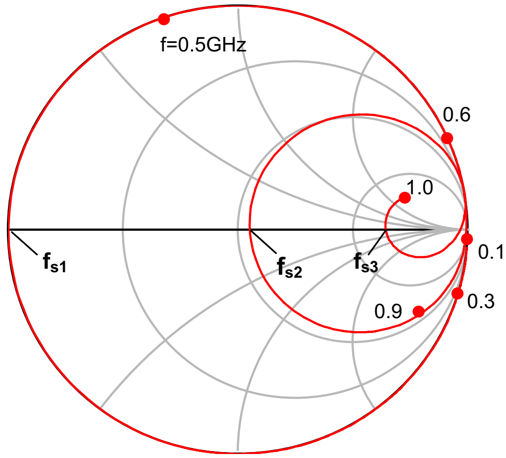


Figure 2: Input impedance for 0.1 to 1 GHz.

The first resonant frequency f_{s1} was obtained at 475MHz but the radiation resistance was 0. As shown in Fig. 3(a), the current distribution in the driven element is similar to the parasitic element but the directions of the current flows are opposite each other. This antiparallel current flow shows the transmission line mode with no radiation [5].

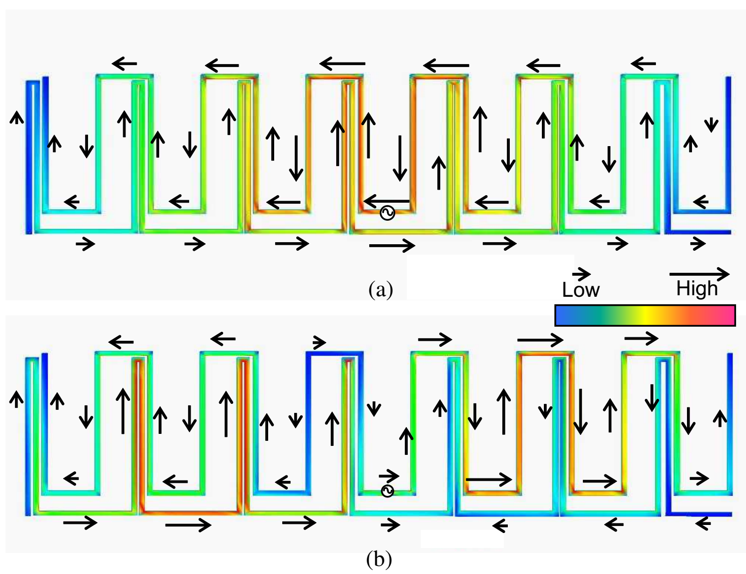


Figure 3: Current distributions at f_{s1} and f_{s2}. (a) $f_{s1} = 475$ MHz, (b) $f_{s2} = 929$ MHz

At the second resonant frequency of $f_{s2} = 929$ MHz, the current concentrates on the parasitic element at the left side and on the driven element at the right side as shown in Fig. 3(b). Since the current flows along high current concentration denoted by longer arrows in Fig. 3(b) are in the same direction, the efficient radiation is achieved. The current density is relatively low at feed point. Therefore, the input resistance of 55 Ω which is five times as large as the input resistance of a conventional meander line antenna without parasitic elements, is realized with this antenna configuration.

4. BANDWIDTH FOR MODIFIED ANTENNA CONFIGURATION

The antenna configuration is modified so that the feed point should be positioned on the center of the driven element as shown in Fig. 4. The width and spacing was chosen to be $w_2 = 2.5$ mm and $w_3 = 0.5$ mm, respectively. Since the constant input resistance is obtained around the second resonant frequency for this antenna configuration, the proper impedance matching is obtained with ease. The equivalent input circuit of an IC chip for RFID is represented with the combination of resistance and capacitance [6]. When the input impedance of an IC chip is assumed to be 65-j150, the reflection coefficient S11 is shown in Fig. 5. The bandwidth defined with $|S11| < 9.45\,\text{dB}$, that is VSWR<2, was 18 MHz. Since the input impedance of an IC chip has capacitive component, the antenna must operate at the higher frequency than the second resonant frequency. Therefore, the antenna configuration should be modified to achieve the moderate change in input reactance around the frequency where the inductive reactance of the antenna input impedance is conjugate matched to the capacitive reactance of the IC.

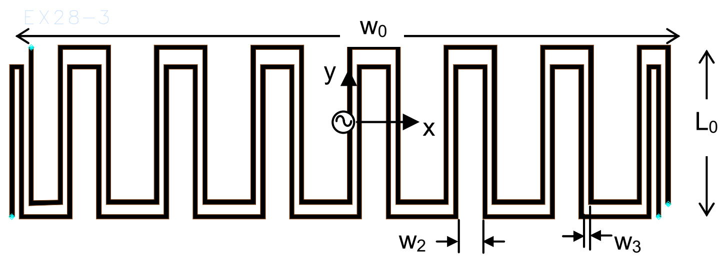


Figure 4: Modified antenna configuration.

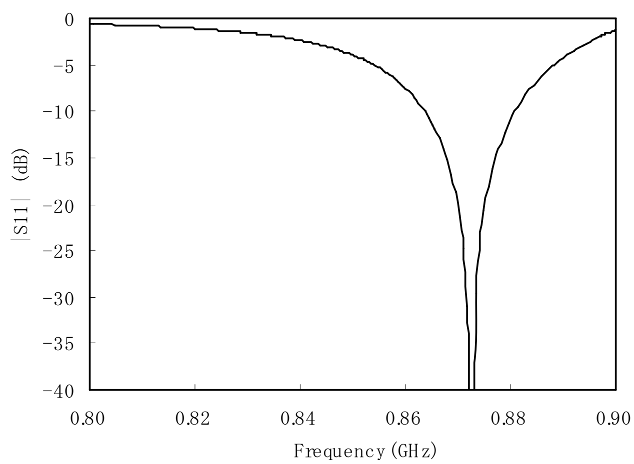


Figure 5: $|S_{11}|$ for the modified antenna configuration.

5. CONCLUSIONS

The input impedance and current distribution have been investigated for meander line antennas with planar coupled parasitic meander elements. The large input resistance has been obtained at the second resonant frequency. The bandwidth has been presented for modified meander configurations. The antenna configuration will be optimized to obtain the wider bandwidth in future.

REFERENCES

1. Seshagiri Rao, K. V., P. V. Nikitin, and S. F. Lam, "Antenna design for UHF RFID tags: A review and a practical application," *IEEE Trans. Antennas Propagat.*, Vol. 53, No. 12, 3870–3876, 2005.
2. Vita, G. D. and G. Iannaccone, "Design criteria for the RF section of UHF and microwave passive transponders," *IEEE Trans Microwave Theory and Tech.*, Vol. 53, No. 9, 2978–2990, 2005.
3. Marrocco, G, "Gain-optimized self-resonant meander line antennas for RFID applications," *IEEE Antennas Prapagat. Lett.*, Vol. 2, 302–305, 2003.
4. Taki, K. and Y. Miyazaki, "Radiation characteristics of meander line antennas with planar coupled parasitic meander elements for RFID tags," *Proceedings of Progress in Electromagnetics Research Symposium*, 256, Tokyo, Japan, August 2006.
5. Nogchi, K., M. Mizusawa, T. Yamaguchi, and Y. Okumura, "Increasing the bandwidth of a meander line antenna consisting of two strips," *1997 IEEE Antennas and Popagat Soc. Int. Symp. Digest*, Vol. 4, 2198–2201, 2003.

6. Joe, J. and M. Y. W. Chia, "Voltage, efficiency calculation and measurement of low power rectenna rectifying circuit," *1998 IEEE Antennas and Popagat Soc. Int. Symp. Digest*, Vol. 4, 1854–1857, 1998.

Conceptual Design of A High Resolution, Low Cost X-Band Airborne Synthetic Aperture Radar System

Y. K. Chan, V. C. Koo, and T. S. Lim

Multimedia University, Malaysia

Abstract— This paper describes the conceptual design of a low cost airborne Synthetic Aperture Radar (SAR) capable to obtain high-resolution image. The proposed system is an X-band, single polarization, high bandwidth linear FM radar and high resolution airborne SAR system. The system can be used as monitoring and management of earth resources such as paddy field, oil palm plantation and soil surface. First of all, the high level design will be discussed and detail design parameters are presented. It followed by radar electronics design, which outlined the detail in radar transmitter and receiver.

1. INTRODUCTION

Radar has long been used for military and non-military purposes in a wide variety of applications such as imaging, guidance, remote sensing and global positioning [6]. Development of radar as a tool for ship and aircraft detection was started during 1920s. The first imaging radar, developed during World War II, used the *B-Scan*, which produced an image in a rectangular format. In the 1950s, the *Side Looking Airborne Radar* (SLAR) was developed. Scanning had been achieved with the SLAR by fixed beam pointed to the side with aircraft's motion moving the beam across the land.

However, the image formed by SLAR is poor in azimuth resolution. For SLAR the smaller the azimuth beamwidth, the finer the azimuth resolution. In order to obtain high-resolution image one has to resort either to an impractically long antenna or to employ wavelengths so short that the radar must contend with severe attenuation in the atmosphere. In airborne application particularly the antenna size and weight are restricted. Another way of achieving better resolution from radar is signal processing. Synthetic Aperture Radar (SAR) is a technique which uses signal processing to improve the resolution beyond the limitation of physical antenna aperture [8]. In SAR, forward motion of actual antenna is used to 'synthesise' a very long antenna. SAR allows the possibility of using longer wavelengths and still achieving good resolution with antenna structures of reasonable size.

There are many applications or potential applications such as biomass estimation, crops monitoring, vegetation cover mapping, mineral exploration, ice dynamics modelling, forest fire, oil spill and biological water monitoring. Some of these have not yet been adequately explored because lower cost electronics are just beginning to make SAR technology economical for smaller scale uses. This paper describes the conceptual design and the proposed X-band airborne SAR system.

2. DESIGN CONSIDERATION

The primary goal of this project is to develop a low cost airborne X-band SAR system capable to illuminate a small size terrain and construct the high-resolution image of scanned area. High-level system design and subsystem level requirements have been carefully considered. High-level design consideration include:

2.1. Operating Frequency and Polarization

For remote sensing applications, frequency range from 1 to 30 GHz is normally used. In the 1–10 GHz range, the transmissivity through air approaches 100%. Thus, a SAR operating in this frequency range is always able to image the earth's surface independent of the cloud cover or precipitation. Our system is designed to operate at X-band (9.6 GHz or 3 cm wavelength), which is within the allowable spectrum (9.5G–9.8 GHz) defined by International Telecommunication Union (ITU) for Earth Exploration Satellite System (EESS) [4].

The size of an X-band antenna is considerably small and suitable to be used in airborne platform. In this frequency band, incident wave tends to be reflected by the surface layer. Potential application of X-band SAR included high-resolution land imaging, and ocean observation. In our design, single polarisation will be ultilized. The VV polarization is chosen because it is more suitable for remote sensing of earth terrain. Change detection and simple classification can be done by VV polarization.

2.2. Modulation and Mode of Operation

Modern radar uses Linear Frequency Modulation (LFM) waveform or Chirp to increase range resolution when long pulses are required to get reasonable signal to noise ratio (S/N). The same average transmitting power as in a pulse system can be achieved with lower peak amplitude. The LFM configuration is employed in this project since it gives better sensitivity without sacrificing range resolution and ease of implementation. The lower peak power allows for the use of commercially available microwave components that have moderate peak power handling capability.

Stripmap is standard mode of SAR operation, widely used by airborne SAR sensors where a strip (swath) to the side of the aircraft is imaged. The radar antenna pattern is oriented towards the ground, orthogonal to the flight track and to one side of aircraft. As the aircraft moves, a swath is mapped out on the ground by antenna footprint. In our design, stripmap mode will be used.

2.3. Resolution and Swath Width

Typical resolution of airborne SAR range from 1 m to 20 m [1]. It depends mostly on the application requirements. Since our main objective is to establish high-resolution airborne SAR system, resolution of 1×1 m for both range and azimuth direction is chosen. A swath width of 1 Km is used in our design for small terrain coverage.

2.4. Operation Platform and Antenna

A small aircraft flying at low altitude will be used in the design in order to achieve the low cost SAR design. The SAR system should support true ground speed at 100 m/s and operating altitude about 1000 m. Typical airborne SAR antenna has the gain of 17 dB to 28 dB. The 25 dBi gain microstrip patch antenna will be used in this system design.

2.5. Range of Incident Angle

From the open literatures, the incident angle from 0°–80° is utilised by present airborne SARs. The backscattering coefficient of nature targets such as soil, grass and vegetable are maintained almost constant over the incident angle of 40° to 60° [6, 7]. Base on the swath width requirement and operating altitude, 50° incident angle with 24° elevation angle is chosen in our SAR system.

2.6. Dynamic Range of Backscattering Coefficient σ°

The required system sensitivity is determined based on the various categories of earth terrain to be mapped such as man made target, ocean, sea-ice, forest, natural vegetation and agriculture, geological targets, mountain, land and sea boundary. From the open literatures, the typical value of σ° falls in the range of +20 dB to −40 dB [3, 5]. For vegetation the typical value of σ° vary from +0 dB to −20 dB. In our system, a dynamic range of 50 dB is targeted from +20 dB to −30 dB in order to facilitate the measurement of various types of earth terrain.

3. DESIGN PARAMETER

In this Section, the calculation associated with the design and hardware implementation of airborne SAR system is presented. The analysis of data rate and data volume are also illustrated. Fig. 1 shows a sketch of the operating condition.

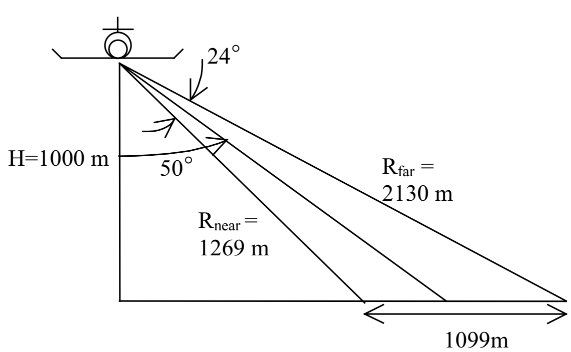


Figure 1: Operating condition of airborne SAR system.

This X-band SAR system is proposed to operate at 9.6 GHz. The LFM waveform with bandwidth of 200 MHz is selected in our design. The theoretical range resolution of this system is 0.75 m.

Assuming a 0.8 m length of antenna will be used in this system, the theoretical azimuth resolution of 0.4 m can be achieved.

The upper limit of Pulse Repetition Frequency (PRF) is attained from a consideration of the maximum mapping range and the fact that the return pulse from this range should come within the interpulse period [2]. In order to adequately sample the Doppler bandwidth, the radar must be pulse with a PRF greater or equal to this bandwidth. Combination of lower and upper limit, and assuming antenna length of 0.8 m is used, a medium PRF of 1000 Hz is chosen.

Since the radar system is planned to operate at altitude of 1000 m, the first return from the target will be at 8.46 μs. Therefore the transmitted pulse-width is pre-selected as 8 μs. Assuming a S/N of 10 dB is sufficient for mapping of earth terrain, for the lowest value of backscattering coefficient, $\sigma_o = -30$ dB and the system losses is assume as 6 dB, the minimum average power required to be transmitted will be 0.208 watt. Thus the peak power requirement is 26 watts. Therefore a high power amplifier that has a 1-dB compression level larger than 26 watts (+44 dBm) can be employed.

System design and performance can be impacted by limitation on data rate and data volume. The impacts of limitation include smaller swath, fewer receiver channels, reduce range resolution, fewer bits per sample and reduce pulse duration. Refer to Fig. 1, the far range and near range is 1269 m and 2130 m respectively and the swath width of 1099 m can be obtained. The time delay for far range and near range can be calculated, there are 8.46 μs and 14.2 μs respectively.

The Nyquist Criterion states that in order to construct a band-limited signal from its samples, the signal must be sampled at least twice the highest frequency. In practice, the signal is over-sampled at a rate higher than the Nyquist by 25% in order to account for non-ideal filter behaviour. In our design, sampling rate of 500 MHz is chosen. In order to capture all the data during the flight mission for geometry given in Fig. 1, the Data Acquisition Unit (DAU) needs to start range sampling at time of return from near-range of swath and stop sampling at end of return from edge of swath.

Assuming the aircraft flies at the constant height of 1000 m, the data window is given by *time of flight far range* minus *time of flight near range plus pulse duration,* and is equal to 13.74 μs. Thus the data rate for single ADC channel (assuming 8 bits per sample) will be 54.96 M bits/second. Therefore the total data rate can be calculated by multiplying the data rate for single channel with the number of receive channels which is equal to 6.87 M byte/second. Assuming an image length of 100 km will be taken for each flight mission, the data volume required for each data take is about 6.87 GB.

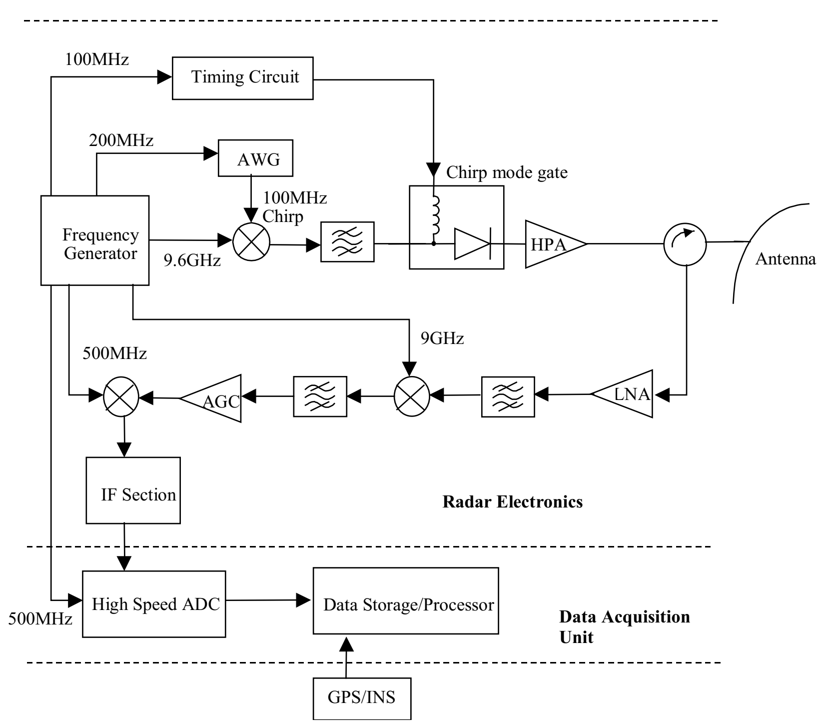


Figure 2: Proposed block diagram of X-band SAR system.

The data rate of at least 6.87 M bytes per second is required to record all the information captured during the data window. In addition, the data storage of more than 6.87 G bytes are needed for recording an image length of 100 km. Thus for a flight time of one hour, the storage required for raw data storage is 24.73 G bytes.

4. PROPOSED AIRBORNE X-BAND SAR

The proposed block diagram of the X-band SAR system is shown in Fig. 2. The system is based on a superheterodyne design. It consists of a microstrip antenna, a radar electronics subsystem and a data acquisition system.

The microwave source is generated from an ovenized Stable Local Oscillator (STALO). The frequency plan of this X-band SAR is shown in Fig. 3. The entire reference signal phase-locked to 100 MHz STALO to preserve the coherency of received signal. An arbitrary waveform generator (AWG) is used to generate the required LFM chirp signal. The timing circuit provides the control signal to switch the chirp mode gate so that the chirp pulse width is properly control. The output of the chirp mode gate is routed to a solid-state high power amplifier with 40 dB gain. The amplified signal is then radiated through the antenna via a circulator. The transmitted waveform is centered at 9.6 GHz with 200 MHz bandwidth. The first stage of the receiver is a low noise amplifier (LNA) and followed by a band-pass filter.

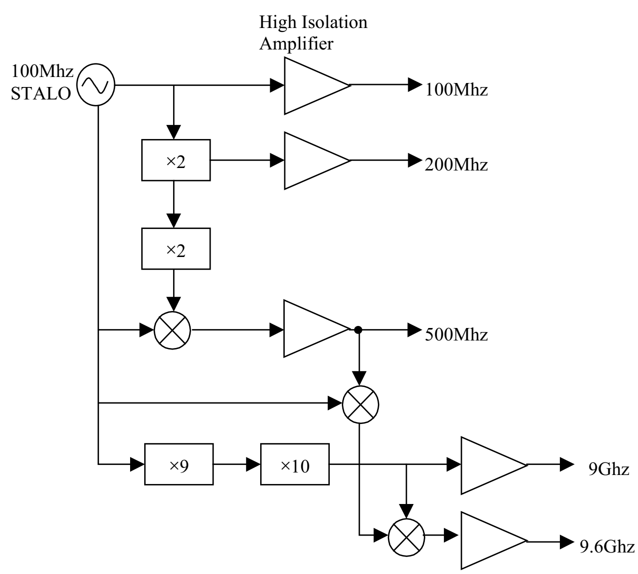


Figure 3: X-Band SAR frequency plan.

The first down-converter mixer is used to convert the received signal to an intermediate frequency (IF) centered at 600 MHz. This IF signal is filtered and amplified by Auto-Gain Control (AGC) Amplifier before feed into second mixer. The signal from second mixer will be routed to IF section which consists of IF filter and amplifier. The X-band SAR system is proposed to employ a PC-based digital signal processing system for data acquisition. It consists of a high-speed 8-bits 500MHz analogue-to-digital converter (ADC). The ADC is capable of converting the down-converted SAR echoes into digital signals and stores them into high-density digital disk for future processing.

5. CONCLUSIONS

The conceptual design of a low cost X-band airborne SAR system has been presented. This airborne SAR system can be used as a tool for monitoring and management of earth resources. This low cost system can be achieved by using simple RF subsystem, commercial component for chirp generation, PC based data acquisition and processing system.

REFERENCES

1. Birk, R., W. Camus, E. Valenti, and W. J. McCandles, "Synthetic aperture radar imaging systems," *IEEE Aerospace and Electronic Systems Magazine*, Vol. 10, No. 11, 15–23, 1995.
2. Hovanessian, S. A., *Introduction to Synthetic Array and Imaging Radars*, Artech House, Dedham, 1980.

3. Hyyppa, J., J. Pulliainen, K. Heiska, and M. T. Hallikainen, "Statistics of backscattering source distribution of boreal coniferous forests at C- and X-band," *Proceeding of the 1986 International Geoscience and Remote Sensing Symposium*, Vol. 1, 241–242, 1986.
4. International Telecommunication Union's (ITU) World Radiocommunication Conference 1997 (WRC-97).
5. Pulliainen, J. T., K. Heiska, J. Hyyppa, and M. T. Hallikainen, "Backscattering Properties of boreal forests at the C- and X-bands," *IEEE Trans on Geosc. and Remote Sensing*, Vol. 32, No. 5, 1041–1050, 1994.
6. Skolnik, M. I., *Radar Handbook*, McGraw-Hill, New York, 1970.
7. Ulaby, F. T., R. K. Moore, and A. K. Fung, *Microwave Remote Sensing: Active and Passive*, I, Artech House, Norwood, 1981.
8. Wiley, C. A., "Synthetic aperture radar — A paradigm for technology evolution," *IEEE Trans. Aerospace Elec. Sys.*, Vol. AES-21, 440–443, 1985.

A Real-time Hybrid Correlator for Synthetic Aperture Radar Signal Processing

V. C. Koo, Y. K. Chan, and T. S. Lim

Multimedia University, Malaysia

Abstract— Synthetic Aperture Radar (SAR) image formation is based on a coherent processing approach to build a long azimuth synthetic aperture. Due to the curvature and the range variance of the azimuth modulation, SAR image formation is inherently a two dimensional process. Two key elements in designing a digital SAR processor are the processing speed and memory requirements. For real-time SAR processing, the high computational requirements cannot be provided by a sequential processor system. In this paper, a new real-time SAR processor based on sub-aperture processing approach is presented. This scheme, called the Real-Time Hybrid Correlator (RTHC), achieves range correlation in frequency domain, followed by azimuth correlation in time domain. The proposed RTHC is a computationally efficient scheme without the need of corner-turn memory. Due to its distributed architecture, dedicated parallel processors can be employed for real-time processing.

1. INTRODUCTION

In essence, the processing of SAR data is a two-dimensional space-variant convolution. Two key elements in designing a digital SAR processor are the processing speed and memory requirements. For ground processing where the recorded SAR data are processed after the flight, these two requirements can easily be met by today's computer technology. However, for real-time SAR processing, the high computational requirements cannot be provided by a sequential processor system. It is therefore mandatory to partition the SAR algorithm into independent sub-tasks to run on multiple processors. In addition, for real-time SAR processing, it is necessary to efficiently access the memory array in both the range and azimuth domains. This requires transpose operation that is commonly known as corner turn operation. The *corner turn* is a potential source of problems even for solid-state memory since transpose operations on large arrays are inefficient in many memory architectures that rely on fast DRAMs.

Depending on the nature of the processing, SAR architecture can be categorized into three types: time domain, frequency domain, and hybrid domain architectures. Time domain architecture is based on the correlation or the convolution integral for range and azimuth compressions. Time domain convolution is computational intensive since its implementation involves multiply, add, and shift operations. However, the repetitive nature of the correlation algorithm is best suited in the implementation of high-speed computer chips. The range correlation is usually done before the azimuth correlation. This architecture requires a large memory to store all the compressed range lines before the azimuth compression can begin.

In the frequency domain architecture, Fast Fourier Transform (FFT) techniques are employed to perform range and azimuth compressions. For a large data length, this can save the computation time since the convolution is now reduced to only multiply operations. However, a large memory is required to store the intermediate data before and after Fourier transformation. The hybrid domain architecture uses both time and frequency domain processing. Generally, this architecture is more flexible and the order in which the range and azimuth compressions are done give rise to different performances.

Figure 1 depicts the basic elements in a typical SAR processor. The analog-to-digital converter (A/D converter) in the first stage digitizes the SAR echoes received by the radar as it moves along the flight track. The raw SAR data are complex values that can be recorded in one of the following two formats. If the received signal has been heterodyned to baseband, then each sample is a complex pair designated as I and Q channels (in-phase and quadrature-phase). If the receiver output spectrum is on a carrier offset from DC, then only one A/D converter is required and the data samples are real valued in offset video format. The net number of data is the same for the two configurations.

Since the range and azimuth time scales are so different, the range and azimuth domains are usually treated separately during processing. The range correlator compressed the spread energy in the range direction with a reference signal, which is an exact replica of the transmitted signal. An

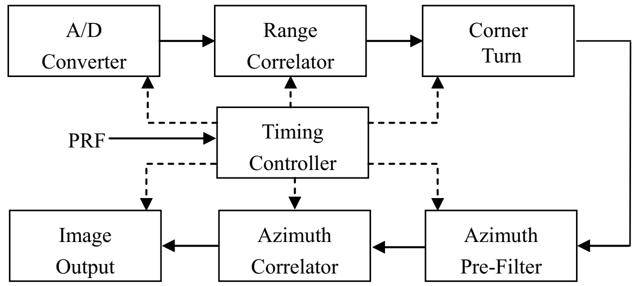


Figure 1: Basic elements of a SAR processor.

important parameter for determining the sidelobe level and resolution of the range-compressed data is the window applied to the range reference function. Typical windows used include the Hamming, Kaiser and Bessel weighting functions. The output of the range-compressed data is typically stored in a corner-turn memory (CTM). The CTM derives its name from the fact that the in-coming data is placed in the memory by columns and the out-going data is read out by rows before feeding it to the azimuth correlator.

Most of the SAR processors have the option of decimating the data in azimuth by pre-filtering around the Doppler centroid prior to azimuth correlator. This is desirable since the azimuth bandwidth of the in-coming data is usually much higher than what is needed to achieve the required azimuth resolution. The azimuth correlator is similar to the range correlator except that it uses a different reference function to compress the spread in the azimuth direction. Finally, the timing controller serves as an overall coordinator of all the processing and memory operations.

2. PROPOSED RTHC ARCHITECTURE

As discussed in previous section, SAR image formation can be viewed as the process of compressing a return signal by correlating it with a reference point target response. Correlator implementation is affected by a number of factors including complexity of the reference function, processing-time requirements, and availability of radar and signal-processing hardware. No single image formation technique is appropriate for all applications.

The implementation scheme proposed here is designed for real-time SAR operation with integrated autofocus capability. This scheme, called the Real-Time Hybrid Correlator (RTHC), achieves range correlation in frequency domain, followed by azimuth correlation in time domain. Its implementation scheme is illustrated in Figure 2.

Table 1: Summary of processing requirements for RTHC Scheme.

Process	**Computational Load (FLOPS) (1 complex multiply = 6 flops, 1 complex add = 2 flops)**	**Memory Requirement (samples)**
Range correlation	$10N_r \log_2 N_r + 6N_r$ (FFT/IFFT: 1 complex multiply, 2 complex adds; Mixing: 1 complex multiply)	$6N_r$ real ($2N_r$ input buffer, $2N_r$reference function, $2N_r$ output buffer)
Azimuth pre-filter	$4N_r$ (FIR: 1 complex-real multiply, 1 complex add)	$2N_r$ real ($2N_r$ input/output buffer)
Azimuth correlation	$6N_{sub}^2$ per sub-aperture (N_{sub} complex multiply)	$4N_r + 2K_rN_{syn}$ real ($2N_r$ input buffer, $K_r \times 2N_{syn}$ reference function, $2N_r$ output buffer)
Square-law detector	$4N_r$ (2 real multiplies, 1 real add, 1 real divide)	$2N_r$ real ($2N_r$ input/output buffer)

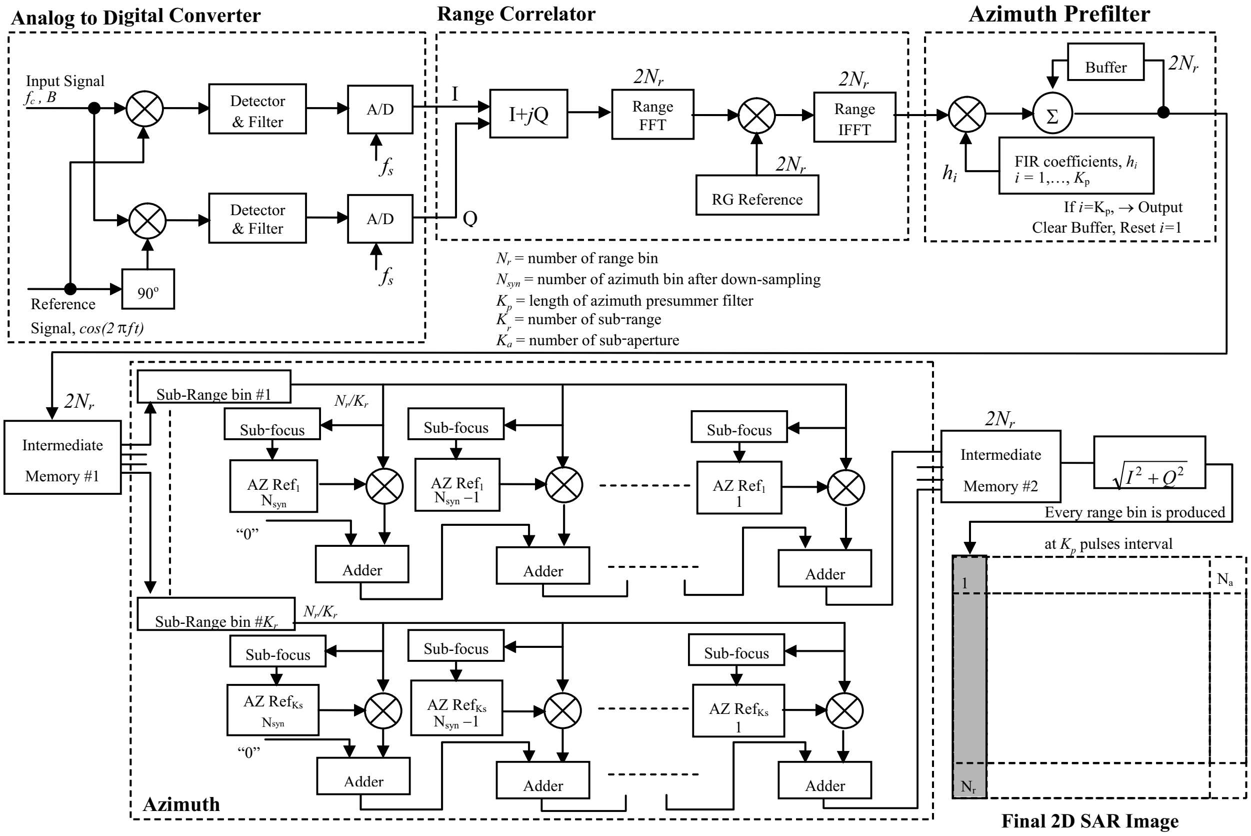


Figure 2: Proposed RTHC implementation scheme.

The A/D converter converts the input data into digital samples and filters out the unwanted signals. Each sample is read into the memory as an $N_r \times 1$ row vectors. The real-to-complex converter combines the I-Q samples into complex data of size $2N_r$. Range correlation is carried out in frequency domain by first performing FFT, followed by complex multiplication with a reference function, and finally performing inverse-FFT. The range reference function is basically the replica of linear FM pulse. In terms of processing complexity, this operation requires

$$\underbrace{10 \times \frac{N_r}{2} \log_2 N_r}_{FFT} + \underbrace{6N_r}_{mixing} + \underbrace{10 \times \frac{N_r}{2} \log_2 N_r}_{IFFT} = 10 N_r \log_2 N_r + 6N_r \text{flops}$$

where flops is floating point operations per second.

A pre-summer filter is employed to reduce the azimuth bandwidth and hence the incoming data rate. The output of the range-compressed data is stored in a temporary memory of size $2N_r$. The implementation of this hybrid approach to divide the processing into separate one-dimensional range and azimuth processing sequences is based on the assumption that the range migration of a point target is less than a resolution cell.

The conventional azimuth processing requires waiting for N_{syn} azimuth samples (that comprises of a full synthetic length) before the azimuth correlation is carried out. This process usually involves large memory storage (at least $N_r \times N_{syn}$). In the RTHC implementation, however, the azimuth correlation employs a time domain convolution against an azimuth reference. This process is performed by sliding the input data over the referenced function and accumulating the product of the overlapping points. For short reference function where $N_{sub} = N_{syn}/K_a$, the required complexity is $6N_{sub}$ flops, as compared to $(10 \log N_{syn} + 6)$ flops in the frequency domain compression technique. Thus, the time domain convolution is more efficient in the case of sub-aperture processing. Moreover, this scheme does not require storing large amount of data for azimuth processing.

3. CONCLUSION

Table 1 presents the summary of the processing requirements for the major components of the RTHC. In principle, the proposed RTHC implementation scheme has the followings advantages:

1. It is computationally efficient and well-suited for real-time processing
2. All the processes are one-dimensional, which greatly reduces the complexity of the actual hardware design
3. Due to its distributed architecture, dedicated parallel processors can be employed for various processes
4. It does not require corner turn memory
5. Autofocus algorithms that are based on sub-aperture approach can be easily integrated within the basic structure of RTHC [2]

REFERENCES

1. Franceschetti, G. and R. Lanari, *Synthetic Aperture Radar Processing*, Crc Press, 1999.
2. Koo, V. C., Y. K. Chan, and H. T. Chuah, "Multiple phase difference method for real-time SAR autofocus," *Journal of Electromagnetic Waves and Applications (JEMWA)*, Vol. 20, No. 3, 375–388, 2006.

Analysis of Electromagnetic Wave Propagation in Out-door Active RFID System Using FD-TD Method

Tadahiro Masuda[1], **Yasumitsu Miyazaki**[2], **and Yoshitaka Kashiwagi**[1]

[1]System Engineering Department, Synclayer, Inc.
1-20 Himegaoka, Kani 509-0249, Gifu, Japan
[2]Department of Electronic & Information Engineering, Aichi University of Technology
50-2 Umanori, Nishihasama-cho, Gamagori 443-0047, Aichi, Japan

Abstract— In recent years, RFID systems have received much attention in security, logistics and medical fields. However, most of these systems are used at in-door and the RF tags are passive tags that are controlled by reader. We consider to use active RFID system of very weak UHF electromagnetic waves at out-door. In this paper, we describe the characteristics of electromagnetic wave propagation by RF tags using FD-TD method. In transmit and receive points of the weak electromagnetic wave, receive characteristics are influenced by the established environment. We consider the road models as this environment. RF tags are carried by the human on the road, and the readers are installed on the electric pole and connected to the cable network. As out-door models, we consider road models of three types, the straight road, the T-type road and the cross road.

1. INTRODUCTION

The reader of active RFID system receives periodic weak electromagnetic wave pulses of UHF carrier band from multiple tags and recognizes multiple tags. We considered to build human identification system for elementary school students by active RFID using UHF band. In transmitting and receiving points of the weak electromagnetic wave, receive characteristics are influenced by the established environment. We consider the road models as this environment. RF tags are carried by the human on the road, and the readers are installed on the electric pole and connected to the cable network. Therefore, understanding of out-door propagation characteristics of RFID is necessary to develop high performance antenna of reader. In this paper, FD-TD method is applied to show the propagation characteristics of RFID tags. The radiation field generated at tag locations on the road, which leads to weak intensity of electric fields at the position of reader are investigated [1–3].

2. E-M ANALYSIS OF RF PROPAGATION IN RFID SYSTEM

Electro-magnetic fields of RF propagation in RFID system of many radiation sources of RF tags can be studied by Maxwell equations and boundary conditions. For RFID system, we consider electromagnetic waves radiated at m current sources of positions $\mathbf{r}_s$ and radiation time at t_s.

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{H}(\mathbf{r}, t)}{\partial t} \tag{1}$$

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = -\frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} + \sum_{s=1}^{m} \mathbf{J}_s(\mathbf{r}, t) \tag{2}$$

where, position vector is $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$, and $\mathbf{J}_s$ are source currents corresponding to RF tags of each user at $\mathbf{r} = \mathbf{r}_s$. The source currents of Equation (2) are represented by the following equations for spatial and temporal functions.

$$\mathbf{J}_s(\mathbf{r}, t) = \mathbf{J}(\mathbf{r} - \mathbf{r}_s, t - t_s) = \mathbf{J}(\mathbf{r} - \mathbf{r}_s) f_0(t - t_s) \tag{3}$$

$$\mathbf{J}(\mathbf{r} - \mathbf{r}_s) = \mathbf{J}_0 \delta(\mathbf{r} - \mathbf{r}_s) \tag{4}$$

$$f_0(t - t_s) = \cos 2\pi f_0 (t - t_s) e^{-\alpha(t - t_s)^2} \tag{5}$$

The signal wave $f_0(t)$ for source current of Equation (3) and wave packet pulse of (5) is represented in Figure 1, where f_0 is carrier frequency and signal wave envelop is Gaussian at radiation time t_s.

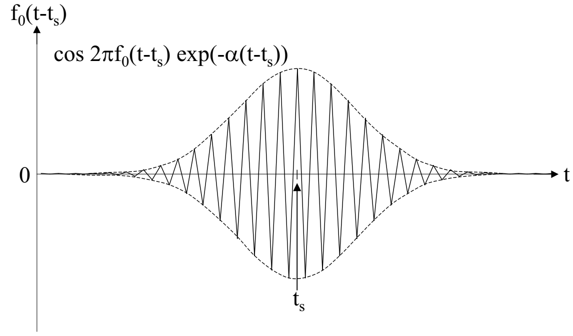


Figure 1: The transmit signal wave of tag.

3. NUMERICAL ANALYSIS OF PROPAGATION BY FD-TD METHOD

In this paper, propagation characteristics are investigated by numerical computation method of FDTD. Electromagnetic fields of RFID system for complicated environment as urban status containing many houses, trees and cars on the roads, can not be easily studied by analytic methods. The analysis models for several types of roads are shown in Figures 2(a)–(c). These models are consisting of the straight road (a), the T-type road (b) and the cross road (c). The parameters of the straight road are length L and width W, and the parameters of the T-type and cross roads are horizontal length L_1 and width W_1 and vertical length L_2 and width W_2. In Figure 2 , the points A,B and C indicate the positions of receiving reader antennas and the point T_s indicates the position of one tag. Ni and Nj are number of divisions in x and y direction. The walls of the road are considered to be perfect conducting plane or concrete.

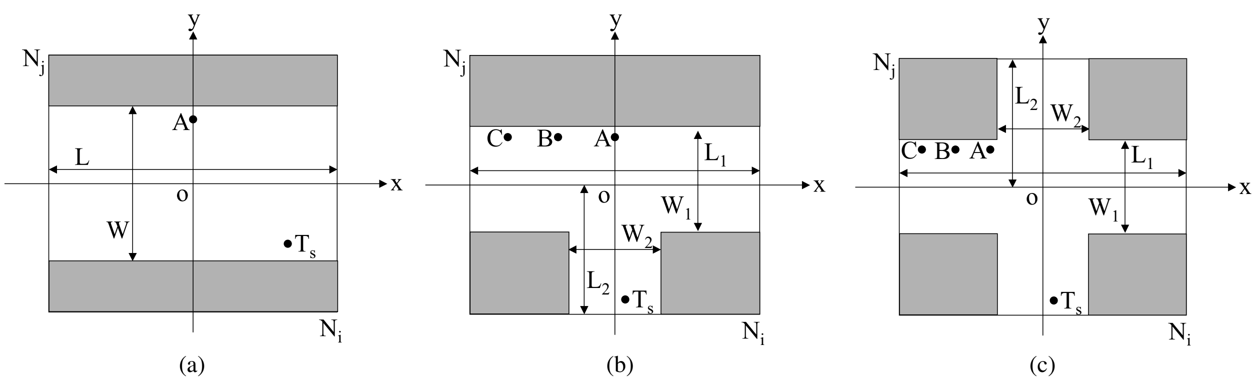


Figure 2: Analysis Models for FD-TD method. (a) straight road model, (b) T-type road model, (c) cross road model.

The FD-TD method is formulated by discretization of Maxwell equations, if we define space coordinates and time parameters as $x = \Delta si$, $y = \Delta sj$, $t = \Delta tn$ in Equations (1) and (2), as follows.

$$H_x^{n+\frac{1}{2}}\left(i, j+\frac{1}{2}\right) = H_x^{n-\frac{1}{2}}\left(i, j+\frac{1}{2}\right) - C_1\left\{E_z^n\left(i, j+1\right) - E_z^n\left(i, j\right)\right\} \tag{6}$$

$$H_y^{n+\frac{1}{2}}\left(i+\frac{1}{2}, j\right) = H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2}, j\right) + C_1\left\{E_z^n\left(i+1, j\right) - E_z^n\left(i, j\right)\right\} \tag{7}$$

$$\begin{aligned} E_z^n\left(i, j\right) = {} & C_2\, E_z^{n-1}(i, j) - C_3 \sum_{s=1}^{m} J_s^{n-\frac{1}{2}}\left(i_s, j_s\right) \\ & + C_4\left\{H_x^{n-\frac{1}{2}}\left(i, j+\frac{1}{2}\right) - H_x^{n-\frac{1}{2}}\left(i, j-\frac{1}{2}\right) - H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2}, j\right) + H_y^{n-\frac{1}{2}}\left(i+\frac{1}{2}, j\right)\right\} \end{aligned} \tag{8}$$

where,

$$C_1=\frac{\Delta t}{\Delta s\mu\,(i,j)}\ C_2=\frac{1-\Delta t\sigma\,(i,j)\,/2\varepsilon\,(i,j)}{1+\Delta t\sigma(i,j)/2\varepsilon\,(i,j)}\ C_3=\frac{\Delta t/\varepsilon\,(i,j)}{1+\Delta t\sigma\,(i,j)\,/2\varepsilon\,(i,j)}\ C_4=\frac{\Delta t/\Delta s\varepsilon\,(i,j)}{1+\Delta t\sigma\,(i,j)\,/2\varepsilon\,(i,j)}$$

i, j indicate the position in x-y plane, and n is number of time steps. $\Delta s = \Delta x = \Delta y$ is space increment, and Δt is time increment. For stability of the FD-TD solution, Δs and Δt should satisfy the condition, $\Delta t \leq \Delta s/\sqrt{2}c$. c is propagation velocity in the free space. $\sigma\,(i,j)$ and $\varepsilon\,(i,j)$ are conductivity and dielectric constant of air space and road wall at the coordinate (i,j),respectively.

J_s is equivalence source and corresponds to a transmitting antenna of tag. (i_s, j_s) is the position coordinates of tag source current that represent by T_s in Figure 2.

4. CONCLUSIONS

For the optimum design of RFID system, we analyzed the characteristics of electromagnetic wave propagation in out-door active RFID system consisting of multiple tags and reader using FD-TD method. In this paper, we formulated FD-TD method for electromagnetic fields radiated from source current of RFID Tag. We study computer simulation using the above FD-TD formulation, and make clear the characteristics of propagation. As a next step, we perform field experiments using actual tags and compact reader in out-door and compare numerical results with the experimental results. These results yield basic foundation for the design of RFID system, and system design of resolving collision problems for multiple pulses.

REFERENCES

1. Rodriguez, G., Y. Miyazaki, and N. Goto, *IEEE Trans. on Antennas and Propagation*, Vol. 54, No. 3, 785–796, 2006.
2. Masuda, T. and Y. Miyazaki, *Trans. IEE Japan*, Vol. 118-C, No. 1, 112–117, 1998.
3. Taki, K. and Y. Miyazaki, *Proc. Progress in Electromagnetic Research Symposium*, Tokyo, Japan, 256, 2006.

Arbitrary Optical Waveform Generation Using Planar Lightwave Circuits

Lawrence R. Chen and Bing Xia
Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 2A7, Canada

Abstract— We provide an overview of the direct temporal domain approach for designing optical filters for ultrafast photonic signal processing applications. We discuss different filter designs, including lattice-form Mach-Zehnder interferometers and two-dimensional ring resonator arrays, to perform pulse repetition rate multiplication and arbitrary optical waveform generation. Simulation and experimental results are presented.

1. INTRODUCTION

Techniques for the generation, control, and manipulation of optical pulses attract considerable interest for numerous applications and have become increasingly important in many scientific areas. Of specific interest are techniques for pulse repetition rate multiplication (PRRM), which are used to obtain ulrafast optical pulse trains from a lower repetition rate input pulse train, as well as for arbitrary optical waveform generation (AOWG). Traditional pulse shaping methods are based on frequency domain processing (i.e., spectral filtering) in which we specifically manipulate the different spectral components of the input pulse in amplitude and/or phase. However, the relationship between the input pulse spectrum and the target temporal waveform is not always straightforward, especially for phase-only filtering processes. In this paper, we present an overview of our recently developed direct temporal domain approach for designing optical filters to achieve PRRM and AOWG.

2. THEORY

The use of spectrally-periodic (SP) amplitude and/or phase filters for performing PRRM is well known and there have been several demonstrations of generating ultrahigh repetition rate pulse bursts or trains [1]. The principle of PRRM using SP amplitude filters is based on mode-selection theory: a uniform input pulse train with a repetition rate R will have a series of modes in the frequency domain which are also periodic with the same spacing R (based on simple Fourier relations). If the amplitude of an SP filter that is used to process the input train has a (larger) mode spacing equal to a multiple of R, then the output will have the same increased mode spacing, thereby increasing the repetition rate of the output train (the output train is otherwise uniform). More generally, based on spectral analysis, it is not always straightforward to see how the amplitudes of the individual pulses in the newly generated output pulse train can be controlled and modulated, especially when using phase-only filtering approaches, to achieve envelope shaping or AOWG.

Recently, we have proposed the direct temporal domain approach as a means for designing SP filters to perform PRRM with arbitrary envelope shaping [2]. In particular, when the free spectral range (FSR) of the SP filter satisfies certain conditions with respect to the new (output) repetition rate, not only is PRRM achieved, we can also specify the amplitudes of the output pulses (within the original period). The only requirement is that the input pulses are sufficiently "short". The principle is based on the double convolution of the input signal and the filter impulse response and requires a spectrally-periodic filter, as shown below in Eq. (1) [1]:

$$a_2(t) = a_0(t) \otimes \sum_{n=-\infty}^{\infty} \sum_{n'=-\infty}^{\infty} h_0(n'T_{FSR})\delta(t - nT - n'T_{FSR}) \quad (1)$$

where $a_0(t)$ is the complex envelope of an individual input pulse, $T = 1/R$ with R being the input repetition rate, $a_2(t)$ is the output signal, T_{FSR} is the unit delay of the SP filter and is given by the inverse of the FSR, and $h_0(t)$ is the filter impulse response. Eq. (1) shows that the output signal is periodic with a new repetition rate which is determined by the original pulse repetition period T and the filter unit delay T_{FSR}. For applications where only the intensity of the output signal is of importance, we can ignore the phase response of the SP filter. In this case, Eq. (1) shows that we can control the amplitudes of the pulses in the output train by controlling the impulse response of the filter at the discrete instants $t = nT_{FSR}$. The required impulse response is obtained by optimizing the filter parameters.

Figure 1(a) illustrates the principle of performing PRRM using the SP filter. The same approach can also be used to perform AOWG as follows. When an input pulse train comprising "broader" pulses is launched into an SP filter, PRRM still takes place, but there will now be an interference among the output pulses. By optimizing $h_0(t)$, the amplitude and phase of each individual output pulse in the multiplied train can be manipulated, thereby forming a specified waveform (which repeats at the input repetition rate). This process is illustrated in Fig. 1(b). Note that with this approach, traditional frequency domain processing is not used (no attention is paid to the frequency response of the SP filter). In either of the above cases, the key is to find an SP filter which can simultaneously perform PRRM and manipulate the amplitude and/or phase of each individual output pulse.

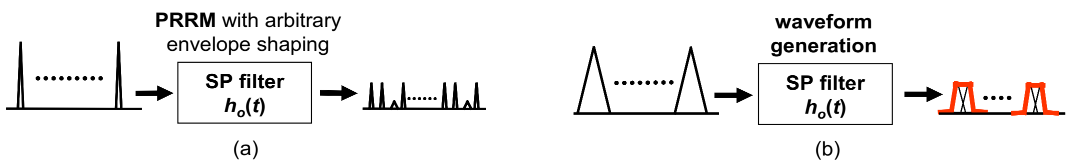


Figure 1: Schematic of using the direct temporal domain approach for photonic signal processing and pulse shaping: (a) PRRM with arbitrary envelope shaping from an input pulse train with narrow pulse widths; (b) waveform generation from an input pulse train with a wide pulse width.

3. SIMULATION RESULTS

Two possible filters which can be used are lattice-form Mach-Zehnder interferometers (LF-MZIs) [3] and two-dimensional ring resonator arrays (2D RRAs) [4]. Both have attracted considerable interest for their applications in dispersion compensation and signal processing. The LF-MZI is a finite impulse response filter which does not incorporate any feedback loops and performs both amplitude and phase filtering, see Fig. 2(a). It comprises m MZIs where κ_m is the coupling coefficient of the mth directional coupler, ΔL is the length difference of the two arms in an MZI, and the FSR is given by $FSR = c_0/n_e\Delta L$ where n_e is the effective index of the waveguide medium. Φ_m is the relative phase shift added to each MZI stage. Fig. 2(b) shows a typical configuration of an $M \times N$ 2D RRA, where M represents the number of rings in the vertical direction and N represents the number in the horizontal direction. Details of the individual ring elements, which incorporate a directional coupler with splitting ratio κ and a phase shifter with an additional phase shift ϕ, are also shown. In our simulations, we assume that the rings are identical in size and are coupled in the vertical direction only. While there is no coupling between the rings in the horizontal direction, the signals propagate in the horizontal direction through the two waveguide buses that are placed at the top and the bottom of the RRA. The responses of both filtering devices can be calculated by using suitable transfer matrices [5, 6].

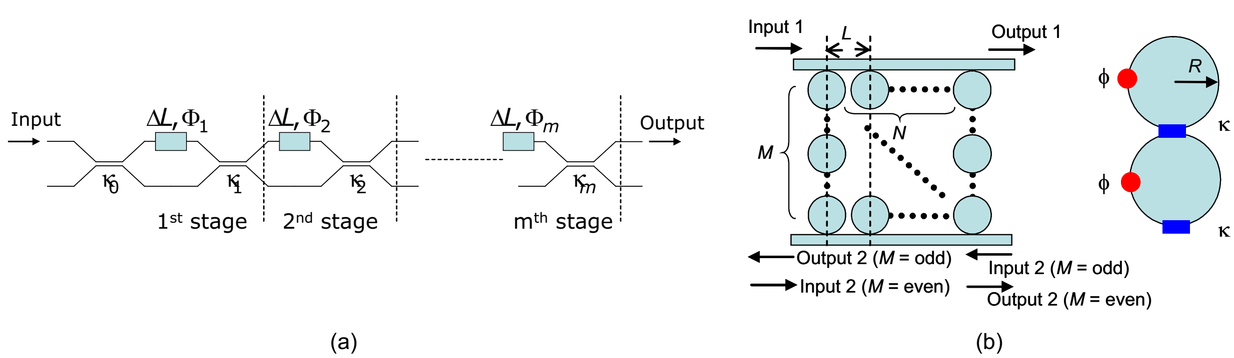


Figure 2: Schematic of (a) LF-MZI and (b) 2D RRA.

As a first example, we consider the use of a LF-MZI for generating an output pulse train at 80 GHz with a triangular envelope from a 10 GHz train of sech^2 pulses. We choose a filter $FSR = 80/3$ GHz to demonstrate the power of the direct temporal domain approach, namely that unlike conventional spectral filtering, we do not require the filter FSR to correspond to the output repetition rate. The number of MZI stages is 7. The results are shown in Fig. 3. Indeed, we are able to obtain PRRM and multi-level amplitudes in the output profile.

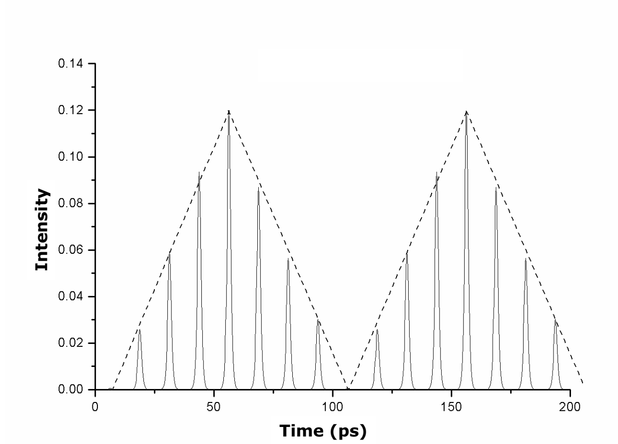


Figure 3: Generation of an 80 GHz pulse train with a triangular profile from a 10 GHz pulse train using LF-MZIs.

Next, we show how the 2D RRAs can be used to perform PRRM with binary code profiles. We design the RRAs to generate a pair of 40 GHz pulse trains with binary code profiles '1101' at output 1 and '1010' at output 2 from a uniform 10 GHz input pulse train comprising 1 ps sech^2 pulses. We consider a 3×3 RRA where the FSR is equal to 40 GHz and the waveguide loss is 1 dB/cm. We optimize the through-amplitude coefficient $t = \sqrt{1-\kappa}$ and phase shifts φ. The results are shown in Fig. 4(a). Note that a pair of 40 GHz pulse trains with binary codes '1101' and '1010' are produced *simultaneously* at outputs 1 and 2.

Finally, we demonstrate AOWG using a 2D RRA. We consider a 5×5 RRA with an FSR of 160 GHz. A 10 GHz train of sech^2 pulses with 8 ps FWHM is launched at input 1. We optimize the filter parameters to transform each input pulse into either (1) a square waveform with 5 ps rise and fall times, and a 40 ps flat-top or (2) a triangular waveform with 30 ps rise and fall times. The generated waveforms are shown in Fig. 4(b) and are well matched to their targets. This demonstrates the power of the direct temporal domain approach for designing SP filters to perform AOWG.

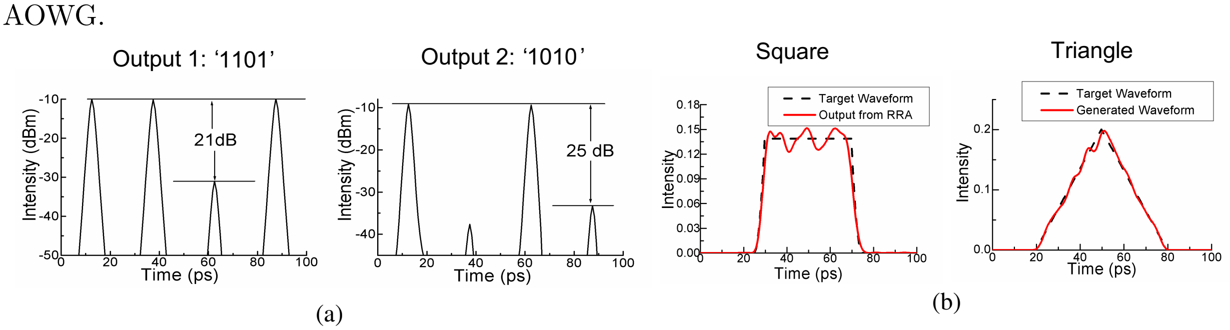


Figure 4: (a) Simultaneous generation of 40 GHz pulse trains with binary code profiles from a 3×3 RRA. (b) AOWG using a 5×5 RRA.

4. EXPERIMENTAL RESULTS

To demonstrate experimentally the temporal domain approach, we have designed and fabricated LF-MZIs on silica-based planar lightwave circuits (PLCs) [7]. In particular, we design the LF-MZIs to perform 10 GHz to 40 GHz PRRM with binary code profiles. The device consists of 4 MZI stages and each stage has a pair of asymmetrical arms. The filter FSR is 40 GHz. We use a large bending radius ($R = 8$ mm) to minimize bending losses in the design. There are 5 couplers ($\kappa 0$-$\kappa 4$) and 4 phase shifters ($\varphi 1$-$\varphi 4$). A single 3 dB MMI coupler design is used for all 5 couplers in the device, which greatly reduces the complexity of the design and increases the stability and tolerance to fabrication errors. The phase shifts provided by the shifters are optimized for the desired output code profile. The waveguide core is made of Ge-doped SiO_2 with a dimension of $3.5 \times 3.5\,\mu m^2$ and is produced by plasma enhanced chemical vapor deposition using silane (SiH_4), nitrous oxide (N_2O) and germane (GeH_4) as precursors. The top cladding is a $12\,\mu$m borophosphosilicate glass (BPSG) with dopants introduced by using diborane (B_2H_6) and phosphine (PH_3), respectively. The waveguides are defined using refractive ion etching.

Figure 5 summarizes the results for generating binary codes '1011' and '1101'. The output repetition rate is ≈ 39 GHz, which is 4 times that of the original repetition rate. The four pulses within the original repetition period exhibit the desired binary code patterns although there are some intensity variations among the '1' bits which arise from imperfections in the device fabrication. Fig. 5(d) shows the simulated output pulse train for '1011' using parameters based on the experimental conditions: the waveguide loss = 0.8 dB/cm, the input train at 9.733 GHz comprises 3 ps Gaussian pulses, and the output signal is convolved with the 16 ps impulse response time of the detection system. The simulated output pulse train is in very good agreement with the measurements (note that the background in the measured output is due to unfiltered amplified spontaneous emission from the amplifiers used in the experiments).

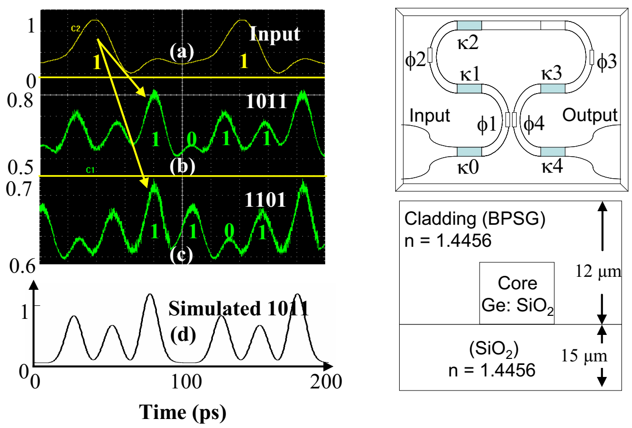


Figure 5: Measured (a) input pulse train from the mode-locked laser; (b) output binary code pattern "1011"; (c) output binary code pattern "1101"; (d) simulated output signal. The device geometry is also shown.

5. CONCLUSION

In summary, we have provided an overview of the direct temporal domain approach for PRRM and AOWG. We have provided simulations to demonstrate the principle using two very different types of filter structures: LF-MZI and 2D RRAs. We have also experimentally demonstrated the use of LF-MZIs for performing PRRM with binary code profiles. The use of tunable PLCs will allow for reconfigurable operation, thereby enhancing the capabilities even further.

ACKNOWLEDGMENT

This research was supported in part by the NSERC Canada and industrial partners, through the Agile All-Photonic Networks (AAPN) Research Program. We thank P. Dumais, C. Callender, and C. Ledderhof for fabricating the PLCs used in the experiments.

REFERENCES

1. Weiner, A. M., "Femtosecond optical pulse shaping and processing," *Prog. Quantum Electron.*, Vol. 19, 161–237, 1995.
2. Xia, B. and L. R. Chen, "A direct temporal domain approach for pulse repetition rate multiplication with arbitrary envelope shaping," *IEEE J. Sel. Topics in Quantum Electronics*, Vol. 11, No. 1, 165–172, 2005.
3. Takiguchi, K., K. Jinguji, K. Okamoto, and Y. Ohmori, "Variable group-delay dispersion equalizer using lattice-form programmable optical filter on planar lightwave circuit," *IEEE J. Sel. Topics Quantum Electron.*, Vol. 2, No. 2, 270–276, 1996.
4. Agarwal, A., P. Toliver, R. Menendez, S. Etemad, J. Jackel, J. Young, T. Banwell, B. E. Little, S. T. Chu, W. Chen, W. Chen, J. Hryniewicz, F. Johnson, D. Gill, O. King, R. Davidson,

K. Donovan, and P. J. Delfyett, "Fully programmable ring-resonator-based integrated photonic circuit for phase coherent applications," *IEEE/OSA J. Lightwave Technol.*, Vol. 24, No. 1, 77–87, 2006.

5. Xia, B. and L. R. Chen, "Arbitrary optical waveform generation using 2D ring resonator arrays," *Optics Express*, Vol. 14, No. 15, 6619–6627, 2006.
6. Xia, B. and L. R. Chen, "Ring resonator arrays for pulse repetition rate multiplication and shaping," *IEEE Photonics Technol. Lett.*, Vol. 18, No. 19, 1999–2001, 2006.
7. Xia, B., L. R. Chen, P. Dumais, and C. L. Callender, "Ultrafast pulse train generation with binary code patterns using planar lightwave circuits," *IEE Electron. Lett.*, Vol. 42, No. 19, 1119–1120, 2006.

Investigation of Crosstalk Effects due to Optical Fiber Nonlinearities in WDM CATV Networks

Rakkappan Balasubramanian[1], **Yasumitsu Miyazaki**[3], **and Masayoshi Kondo**[2]

[1]System Engineering Department, Synclayer, Inc., 1-20 Himegaoka, Kani 509-0249, Gifu, Japan
[2]Engineering Department, Synclayer, Inc., 1-20 Himegaoka, Kani 509-0249, Gifu, Japan
[3]Department of Electronic & Information Engineering, Aichi University of Technology
50-2 Umanori, Nishihasama-cho, Gamagori 443-0047, Aichi, Japan

Abstract— Crosstalk due to optical fiber nonlinearities is an important parameter to be considered to determine the performance of WDM based CATV optical networks. In this paper we present the results of our investigation on the crosstalk due to stimulated Raman scattering and cross phase modulation. Due to the limited bandwidth availability for the upstream data communication in the CATV network, WDM technology is widely employed. We considered different WDM configurations of CATV network and investigated, both numerically and experimentally, the effect of these nonlinear effects with respect to optical power, wavelength spacing, transmission distance and carrier frequency.

1. INTRODUCTION

Wavelength division multiplexing (WDM) technology has been given wide attention in the broadband optical networks in order to utilize the enormous bandwidth of the optical fiber. However, when several optical signals are multiplexed, there arise coupling and crosstalk between the multiplexed signals due to different fiber nonlinear effects, and these phenomena limit the ultimate WDM network performance [1–3]. Typical wavelength band used in these networks is between 1310 nm and 1550nm. Subcarrier multiplexed optical systems have traditionally been used for cable television (CATV) distribution, and system operators have deployed video narrow casting systems to serve different analog contents to different subscribers. More recently, in the access networks, systems are upgraded to handle two-way communication of voice, data and analog video overlay through WDM as in the case of Passive Optical Network (PON) system. WDM techniques have also been employed to effectively use the limited upstream bandwidth of 5 MHz-55 MHz in the CATV data communications systems. The optical nodes serving the subscribers are segmented and each node is assigned an unique CWDM wavelength. The multiplexed CWDM signal is then transmitted back to the head-end. The dominant fiber nonlinearities are stimulated Raman scattering (SRS) and cross phase modulation (XPM). The crosstalk effects in such system are investigated.

2. THEORETICAL BACKGROUND

SRS in WDM systems is due to nonlinear coupling effect through Raman gain and thus the modulated optical signal causes modulation of all other co-propagating signals. XPM effect arises from the nonlinear refractive index of the fiber material, and thus the power of the modulated channel induces small change in the refractive index and in turn phase modulation of all other wavelength channels. In a dual wavelength WDM system, the SRS and XPM induced crosstalk can be investigated by studying the effect of a modulated signal on a CW signal. The interacting signal in this case is described by the following equations.

$$\frac{\partial P_{CW}(z,t)}{\partial z}+\frac{1}{v_{CW}}\frac{\partial P_{CW}(z,t)}{\partial t}=-\alpha_{CW}P_{CW}(z,t)+\left(\frac{g_R}{K\,A_{eff}}\right)P_{MOD}(z,t)P_{CW}(z,t) \quad (1)$$

$$\frac{\partial P_{MOD}(z,t)}{\partial z}+\frac{1}{v_{MOD}}\frac{\partial P_{MOD}(z,t)}{\partial t}=-\alpha_{MOD}P_{MOD}(z,t)-\left(\frac{g_R}{K\,A_{eff}}\right)P_{MOD}(z,t)P_{CW}(z,t) \quad (2)$$

Here, the subscripts CW and MOD refers to CW and modulated signal respectively, v is the group velocity, α is the propagation loss, P is the optical power, g_R is the Raman gain coefficient, A_{eff} is the effective fiber core area. For copropagating signals, the optical power of the CW signal can be derived as,

$$P_{CW}(z)=P_{CW}(0)e^{-\alpha z}\left[1+\frac{Kg_RP_{MOD}L_{eff}}{A_{eff}}+\frac{Kg_RP_{MOD}m}{A_{eff}}\cdot\frac{\sqrt{1+e^{-2\alpha z}-2e^{-\alpha z}\cos(\Omega\phi_{12}z)}}{\sqrt{\alpha^2+(\phi_{12}\Omega z)^2}}\right] \quad (3)$$

The first and second terms within the bracket correspond to the amplification and the third term corresponds to the modulation transfer. Eq. (3) is used to calculate the crosstalk due to SRS. Similar derivation will be presented for the case of XPM induced crosstalk.

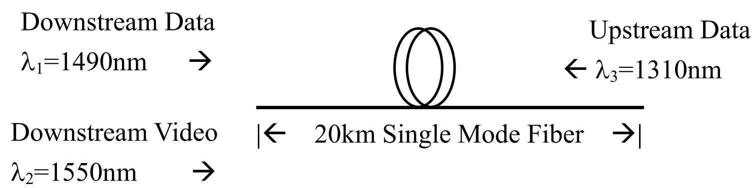


Figure 1: Typical CATV network using WDM system.

3. NUMERICAL CALCULATIONS AND EXPERIMENTS

Figure 1 is a typical CATV network used in the CATV network. Fig. 2 shows the Raman amplification of optical signal due to strong pump. The pump signal of 1450 nm induces Raman amplification on a copropagating signal (tunable wavelength signal in the range of 1300 nm–1600 nm is used in this experiment). Though the peak amplification is at 100 nm on the higher side of pump wavelength, amplification can be observed over a bandwidth of 1450 nm–1600 nm.

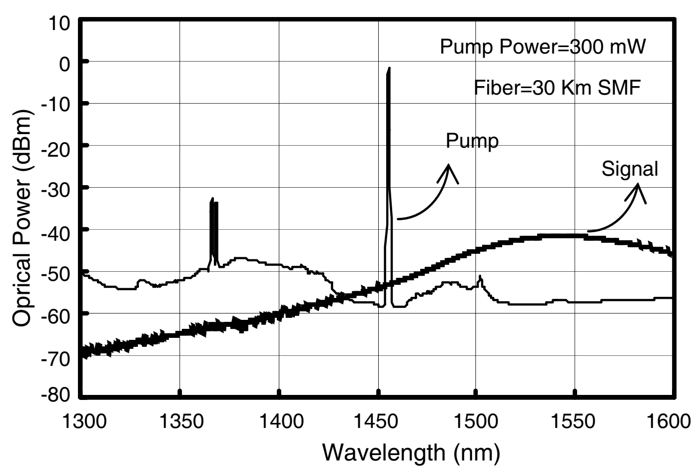


Figure 2: Signal amplification due to SRS.

We investigated a dual wavelength WDM optical system with transmission distance of 20 km, where 1490 nm and 1550 nm are used for the downstream data and analog video transmission. The data signal of wavelength at 1490 nm will act as the pump wave, and when modulated, optical wave of this wavelength induces crosstalk on the video channel.

We experimentally measured the crosstalk levels at 1550 nm video channel by modulating 1490 nm data channel with the frequency range between 70 to 770 MHz, which is typically used for CATV video distribution. We compared the measured data with the theoretical prediction and find good agreement between the results. Different wavelength spacing used in the CATV network, 20 nm and 0.8 nm wavelength spacing for CWDM and DWDM systems respectively. Crosstalk effects are investigated against the system parameters such as RF frequency bandwidth, optical wavelength spacing, transmission distance and transmission direction.

The SRS and XPM induced crosstalks were observed to affect the low frequency and high frequency RF channels respectively, and the total crosstalk will be the sum of these individual crosstalk mechanisms.

4. CONCLUSIONS

In this paper, we presented the results of our investigation on the crosstalk mechanism and the effect of crosstalk on the system performance in the context of WDM systems employed in the CATV network. Different system architectures are considered for the present investigation and against system parameters such as RF frequency bandwidth, optical wavelength spacing, transmission distance and transmission direction.

REFERENCES

1. Agrawal, G. P., *Nonlinear Fiber Optics*, 3rd ed., Academic Press, San Diego, 2001.
2. Phillips, M. R. and D. M. Ott, "Crosstalk due to optical fiber nonlinearities in WDM lightwave systems," *J. Lightwave Technol.*, Vol. 17, 1782–1791, 1999.
3. Balasubramanian, R., Y. Miyazaki, and M. Kondo, "Performance limitations in CATV networks employing WDM system due to optical fiber nonlinearities," *Proc. Progress in Electromagnetic Research Symposium*, 253, Tokyo, Japan, 2006.

Code Recognition for Optical Time-series WDM Coded Label Using Acoustooptic Switch Array

Nobuo Goto[1] **and Yasumitsu Miyazaki**[2]

[1]Toyohashi University of Technology, Japan
[2]Aichi University of Technology, Japan

Abstract— Optical label processing has been expected to increase the throughput of packet routing in photonic networks. We have studied label recognition using collinear acoustooptic (AO) devices, where the label is encoded in time and spectral domains. In the label recognition system in time domain, time gating devices were required to improve the recognition characteristics. In this report, we investigate a label recognition system for optical codes encoded in time and spectral domains where time gating function is not required. We discuss the recognition characteristics using numerical analysis. The proposed system will be useful to increase the number of distinguishable codes and to improve the flexibility of routing control in the networks.

1. INTRODUCTION

Routing processing of optical packets will be a bottleneck in large-capacity photonic networks. Label routing system has been expected to improve the processing of optical packets. A variety of label processing methods have been proposed to use effectively the potential of optical signal processing. [1, 2] The authors have studied on collinear acoustooptic (AO) devices [3, 4] and applications to optical label recognition [5, 6]. Since the AO devices can handle wavelength-division-multiplexed (WDM) optical signals, code recognition for optical codes encoded in spectral domain can be realized with a simple processing system. However, encoding in spectral domain limits the WDM number for optical packets. So, we consider combination of encoding in spectral domain and in time domain. Previously proposed systems for recognition in time domain required time gating to generate output signal. In this report, we propose a system that does not require time gating.

2. STRUCTURE OF OPTICAL LABEL RECOGNITION PROCESSOR

The proposed processor consists of optical power divider, optical delay lines, parallel array of collinear AO switches, photodetectors and an electric multiplier as shown in Fig. 1. Optical label is encoded in time and spectral domains. The optical code m consists of a pulse train of N_t pulses, $\boldsymbol{C}^m = (\boldsymbol{c}_1^m, \boldsymbol{c}_2^m, \ldots, \boldsymbol{c}_{N_t}^m)$. Each pulse is wavelength multiplexed with N_λ components as given by $\boldsymbol{c}_i^m = (c_{i1}^m, c_{i2}^m, \ldots, c_{N_\lambda}^m)^t$, $(i = 1, \ldots, N_t)$. Thus, the optical code m is represented as

$$\boldsymbol{C}^m = \begin{bmatrix} c_{11}^m & c_{21}^m & \cdots & c_{N_t 1}^m \\ c_{12}^m & c_{22}^m & \cdots & c_{N_t 2}^m \\ \vdots & \vdots & \ddots & \vdots \\ c_{1N_\lambda}^m & c_{2N_\lambda}^m & \cdots & c_{N_t N_\lambda}^m \end{bmatrix} \tag{1}$$

We consider On-Off-Keying, OOK, orthogonal codes for each WDM encoded pulse. The number of codes represented by a single WDM pulse is about N_λ. For pulse trains of N_t pulses, the total number of codes is $(N_\lambda)^{N_t}$.

When an optical pulse train representing label $\boldsymbol{C}^m$ is incident, the pulse train is divided into N_t pulse trains, denoted by $\boldsymbol{C}_l^m, (l = 1, \ldots, N_t)$ with the optical power divider. The lth pulse train $\boldsymbol{C}_l^m$ is delayed with an amount of $(N_t - l)\Delta t_p$ through the delay line, and incident into the lth AO switch, where frequency-multiplexed surface acoustic waves (SAWs) representing a code $\boldsymbol{c}_l^k$ are propagated along the interaction region. The outputs from the parallel AO switches are balanced-detected with PDs. The electric pulse trains are multiplied with the multiplier, generating the matched pulse signal.

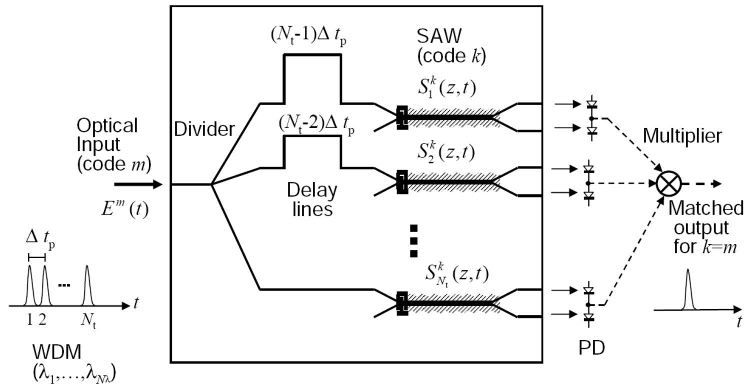


Figure 1: Architecture of label recognition processor.

It is noted that the SAWs are not required to change during the label recognition for code $\boldsymbol{C}^k$. Therefore, even if it takes a few µs for SAWs to propagate over the interaction region, high-speed code recognition can be performed by using short optical pulses. The electrical processing is used only for multiplication, high-speed processing will be expected.

3. PRINCIPLE OF CODE RECOGNITION

The electric field of optical incident pulses representing code C^m is written as

$$E^m(t) = E_0 \sum_{l=1}^{N_t} \sum_{i=1}^{N_\lambda} c_{li}^m g_{in,0}(t-(l-1)\Delta t_p) \exp\left[j\omega_i(t-(l-1)\Delta t_p)\right] \tag{2}$$

where $g_{in,0}(t)$ denotes the envelope of a single pulse, and ω_i is the angular frequency corresponding to the wavelength λ_i.

The incident pulse train is equally divided into N_t pulse trains and delayed with the delay lines. The incident pulses into the l'th AO switch is given by

$$\begin{aligned} E_{l'}^m(t) &= \frac{E_0}{\sqrt{N_t}} E^m(t-(N_t-l')\Delta t_p) \\ &= \frac{E_0}{\sqrt{N_t}} \sum_{l=1}^{N_t} \sum_{i=1}^{N_\lambda} c_{li}^m g_{in,0}(t-(N_t-l'+l-1)\Delta t_p) \\ &\quad \times \exp\left\{j\left[\omega_i(t-(N_t-l'+l-1)\Delta t_p) - \beta_i(N_t-l')\Delta t_p - c/N_i\right]\right\} \end{aligned} \tag{3}$$

where c is the optical velocity in free space. β_i and N_i are the propagation constant and its effective index of optical guided waves at wavelength λ_i in the optical waveguides.

The code to be recognized in this system is supposed to be $\boldsymbol{C}^k$. Then the SAW strain in the l'th AO switch is written as

$$S_{l'}^k(t,z) = \sum_{i=1}^{N_\lambda} c_{il'}^k s_0 \cos(\Omega_i t - K_i z) \tag{4}$$

where s_0 is the strain component that contributes to the collinear AO switching. Ω_i and K_i are the angular frequency and the propagation constant of SAW propagated in z direction. The electric field of the optical output from the switched port of the AO switch is given by

$$\begin{aligned} E_{out1,l',k}^m(t) = &\frac{E_0}{\sqrt{N_t}} \sum_{l=1}^{N_t} \sum_{i=1}^{N_\lambda} c_{li}^m c_{il'}^k g_{out,0}(t-(N_t-l'+l-1)\Delta t_p - l_{SW}N_i/c) \\ &\times \exp\left\{j\left\{(\omega_i-\Omega_i)\left[t-(N_t-l'+l-1)\Delta t_p - l_{SW}N_i/c\right] - (\beta_i-K_i)\left[l_{SW}+(N_t-l')\Delta t_p c/N_i\right]\right\}\right\} \end{aligned} \tag{5}$$

where $g_{out,0}(t)$ denotes the envelope of the switched single pulse. l_{SW} is the interaction length of the collinear AO switches. It is noted that the optical frequency of the switched pulse is shifted by an amount of the SAW frequency. The unswitched fields from the other output ports are written as

$$\begin{aligned} E^m_{out2,l',k}(t) &= \frac{E_0}{\sqrt{N_t}} \sum_{l=1}^{N_t} \sum_{i=1}^{N_\lambda} c^m_{li} \Big[\bar{c}^k_{il'} g_{in,0}\left(t-(N_t-l'+l-1)\Delta t_p - l_{SW}N_i/c\right) \\ &+ c^k_{il'} g^{res}_{out,0}\left(t-(N_t-l'+l-1)\Delta t_p - l_{SW}N_i/c\right)\Big] \\ &\times \exp\left\{j\left\{\omega_i\left[t-(N_t-l'+l-1)\Delta t_p - l_{SW}N_i/c\right] - \beta_i\left[l_{SW}+(N_t-l')\Delta t_p c/N_i\right]\right\}\right\} \end{aligned} \tag{6}$$

where $\bar{c}^k_{il'} = 1 - c^k_{il'}$, and the pulse shape $g^{res}_{out,0}(t)$ corresponds to the unswitched output due to incomplete switching.

By converting these optical outputs with PDs, the photo current is given by

$$\begin{aligned} I^m_{out1,l',k}(t) &= \frac{I_0}{|E_0|^2}\overline{|E^m_{out1,l',k}|^2} \\ &\simeq \frac{I_0}{N_t}\sum_{l=1}^{N_t}\sum_{i=1}^{N_\lambda}\left(c^m_{li}c^k_{il'}\right)^2 g^2_{out,0}\left(t-\left(N_t-l'+l-1\right)\Delta t_p - l_{SW}N_i/c\right) \end{aligned} \tag{7}$$

and

$$\begin{aligned} I^m_{out2,l',k}(t) &= \frac{I_0}{|E_0|^2}\overline{|E^m_{out2,l',k}|^2} \\ &\simeq \frac{I_0}{N_t}\sum_{l=1}^{N_t}\sum_{i=1}^{N_\lambda}\Big[\left(c^m_{li}\bar{c}^k_{il'}\right)^2 g^2_{in,0}\left(t-\left(N_t-l'+l-1\right)\Delta t_p - l_{SW}N_i/c\right) \\ &+ \left(c^m_{li}c^k_{il'}\right)^2 g^{res}_{out,0}\left(t-(N_t-l'+l-1)\Delta t_p - l_{SW}N_i/c\right)\Big] \end{aligned} \tag{8}$$

Here, we assume the collinear AO switching is ideal, that is, the output pulse envelopes are approximated as

$$g_{out,0} \simeq g_{in,0}, \quad g^{res}_{out,0} \simeq 0 \tag{9}$$

By subtracting Eq. (8) from (7), we obtain the balanced output as

$$\begin{aligned} I^m_{l',k}(t) &= I^m_{out1,l',k}(t) - I^m_{out2,l',k}(t) \\ &\simeq \frac{I_0}{N_t}\sum_{l=1}^{N_t}\sum_{i=1}^{N_\lambda}\left[\left(c^m_{li}c^k_{il'}\right)^2 - \left(c^m_{li}\bar{c}^k_{il'}\right)^2\right] g^2_{in,0}\left(t-\left(N_t-l'+l-1\right)\Delta t_p - l_{SW}N_i/c\right) \end{aligned} \tag{10}$$

Here, we assume the wavelength dispersion of the waveguides can be ignored. By assuming $N_i \simeq N_1$, we have

$$I^m_{l',k}(t) \simeq \frac{I_0}{N_t}\sum_{l=1}^{N_t}\left\{\sum_{i=1}^{N_\lambda}\left[\left(c^m_{li}c^k_{il'}\right)^2 - \left(c^m_{li}\bar{c}^k_{il'}\right)^2\right]\right\} g^2_{in,0}\left(t-\left(N_t-l'+l-1\right)\Delta t_p - l_{SW}N_1/c\right) \tag{11}$$

Multiplication of N_t pulse trains results in

$$I_{out} = \prod_{l'=1}^{N_t} I^m_{l'k} \tag{12}$$

Since the delayed pulse trains are multiplied, only the pulse at $l' = l$ remains in the multiplied output as

$$\begin{aligned} I_{out} &\simeq \prod_{l'=1}^{N_t}\frac{I_0}{N_t}\left\{\sum_{i=1}^{N_\lambda}\left[\left(c^m_{l'i}c^k_{il'}\right)^2 - \left(c^m_{l'i}\bar{c}^k_{il'}\right)^2\right]\right\} g^2_{in,0}\left(t-(N_t-1)\Delta t_p - l_{SW}N_1/c\right) \\ &= \left\{\prod_{l'=1}^{N_t}\sum_{i=1}^{N_\lambda}\left[\left(c^m_{l'i}c^k_{il'}\right)^2 - \left(c^m_{l'i}\bar{c}^k_{il'}\right)^2\right]\right\}\left[\frac{I_0}{N_t}g^2_{in,0}\left(t-(N_t-1)\Delta t_p - l_{SW}N_1/c\right)\right]^{N_t} \end{aligned} \tag{13}$$

4. CODE RECOGNITION IN CASE OF IDEAL SWITCHING

We consider each pulse of the pulse train is wavelength multiplexed with OOK orthogonal codes. As the orthogonal codes, we employ codes generated from a Hadamard matrix by deleting row and column having all zeros [6]. For example, the codes in case of $N_\lambda = 3$ are given by

$$\boldsymbol{c}^m_i = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\} = \left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}\right\} \tag{14}$$

In these codes, the number of codes represented by a single WDM pulse is equal to the number of WDM, N_λ. Since the pulse train consists of N_t pulses, the total number of codes encoded in time and spectral domains N_{code} is given by

$$N_{code} = N_\lambda^{N_t} \tag{15}$$

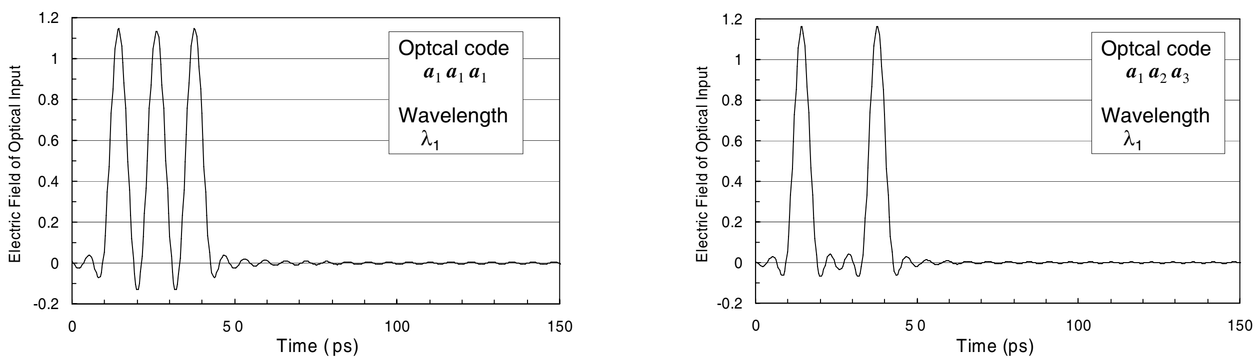
The value $I_{out,code}$ corresponding to the parenthesis $\{\,\cdot\,\}$ in Eq. (13) has non-zero value only when the codes are matched, that is $m = k$, resulting in

$$I_{out,code} = \begin{cases} [(N_\lambda + 1)/2]^{N_t} & (m = k) \\ 0 & (m \neq k) \end{cases} \tag{16}$$

where $(N_\lambda + 1)/2$ is equal to the number of wavelength components in each pulse.

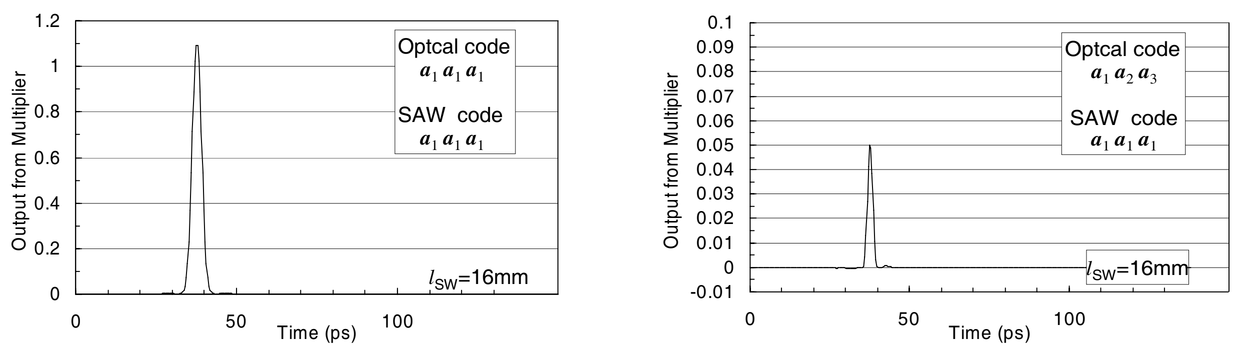
5. EVALUATION OF CODE RECOGNITION CONSIDERING AO SWITCHING CHARACTERISTICS

Next, we evaluate the code recognition characteristics by considering wavelength selectivity of the collinear AO switches. As a numerical example, we assume a bandwidth limited optical pulse train with pulse period 12 ps, pulse width 5.9 ps, bandwidth 160 GHz and $N_t = 3$. The interaction length of collinear AO switch is assumed to be $l_{SW} = 16$ mm.



(a) (b)

Figure 2: Electric fields of optical pulse trains incident at 3rd AO switch at wavelength λ_1. (a) Optical code $\boldsymbol{C}^m = (\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1)$, (b) Optical code $\boldsymbol{C}^m = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$.



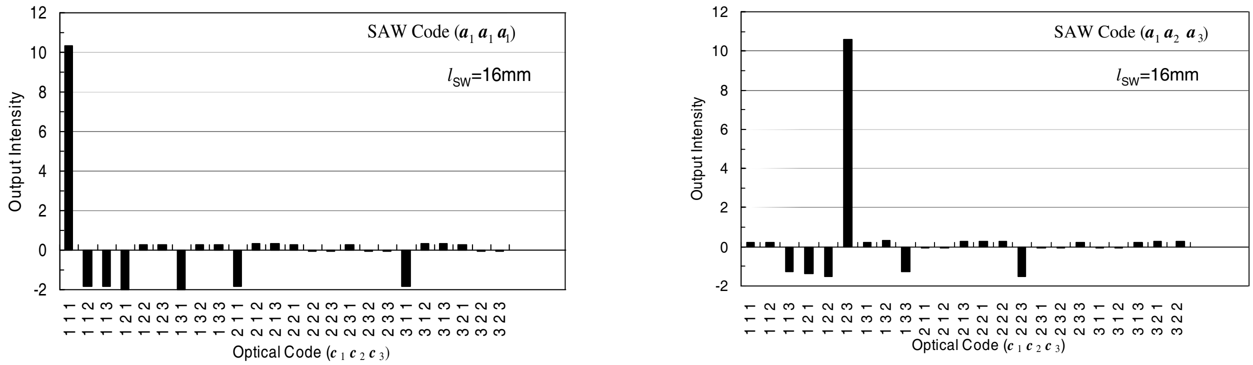
(a) (b)

Figure 3: Output current from multiplier. (a) Optical code $\boldsymbol{C}^m = (\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1)$, (b) Optical code $\boldsymbol{C}^m = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$.

Optical incident pulses $E_1^m(t)$ at wavelength λ_1 are shown for the cases of $\boldsymbol{C}^m = (\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1)$ and $\boldsymbol{C}^m = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$ in Fig. 2.

We assume the SAWs in collinear AO switches are set for recognizing codes $\boldsymbol{C}^k = (\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1)$. The optical output is evaluated for the cases of optical inputs $\boldsymbol{C}^m = (\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1)$ and $\boldsymbol{C}^m = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$.

Time shifted pulse trains are obtained from the parallel AO switches and multiplication of these pulses results in the output as shown in Fig. 3. The output power for all the optical codes are calculated as shown in Fig. 4. It is found that only the matched code results in large output. Fig. 5 shows the outputs corresponding to the matched code and minimum value of unmatched output, that is, the maximum absolute value, for all the codes. It is concluded that all the codes can be distinguished.



(a) (b)

Figure 4: Output power from multiplier for all optical codes. (a) SAW code $\boldsymbol{C}^k = (\boldsymbol{a}_1, \boldsymbol{a}_1, \boldsymbol{a}_1)$, (b) SAW code $\boldsymbol{C}^K = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3)$.

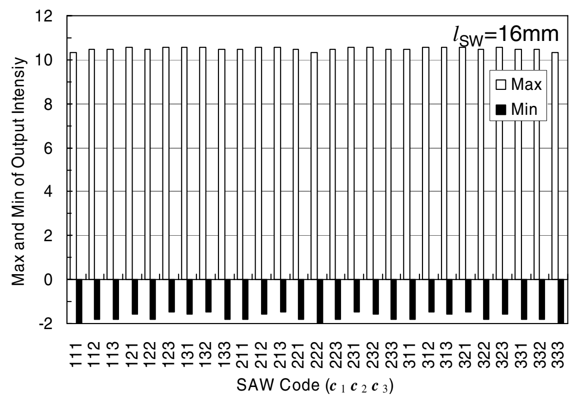


Figure 5: Maximum and minimum value of output from multiplier for all SAW codes.

6. CONCLUSION

Optical code recognition system consisting of collinear AO switch array was investigated. By using electrical multiplication, a simple processing which does not require time gating is performed. The code recognition characteristics were evaluated for the case of ideal switching and that of actual AO switching having wavelength selectivity. In future plan, the total system for routing will be discussed. We also consider experimental verification of the proposed code recognition.

REFERENCES

1. Blumenthal, D. J., B. E. Olsson, G. Rossi, T. E. Dimmick, L. Rau, M. Masanovic, O. Lavrova, R. Doshi, O. Jerphagnon, J. E. Bowers, V. Kaman, L. A. Coldren, and J. Barton, "All-optical label swapping networks and technologies", *J. Lightwave Technol.*, Vol. 18, No. 12, 2058–2075, 2000.
2. Kitayama, K., N. Wada, and H. Sotobayashi, "Architectural considerations for photonic IP router based upon optical code correlation", *J. Lightwave Technol.*, Vol. 18, No. 12, 1834–1844, 2000.
3. Kondo, T., Y. Miyazaki, and Y. Akao, *Jpn. J. Appl. Phys.*, Vol. 17, 1978, 1231.
4. Goto, N. and Y. Miyazaki, "FDTD analysis of wavelength-selective optical switch using collinear acoustooptic interaction for WDM routing", *Trans. IEICE Electron.*, Vol. J86-C, No. 12, 1272–1281, 2003.
5. Goto, N. and Y. Miyazaki, "Label recognition with wavelength-selective collinear acoustooptic switches for photonic label routers", *Jpn. J. Appl. Phys.*, Vol. 44, No. 6B, 4449–4454, 2005.
6. Goto, N. and Y. Miyazaki, "Wavelength-division-multiplex photonic routers with collinear acoustooptic processing for optical label matching in spectral and time domains", *Jpn. J. Appl. Phys.*, Vol. 45, 2006.
7. Goto, N. and Y. Miyazaki, "Acoustooptic processing for recognition of optical binary-phase-shift-keying codes for photonic routers", *IEEE International Ultrasonics Symposium*, Vancouver, P2K-1, Oct. 2006.

Spatial Filtering Characteristics of Transmitted and Scattered Nano-meter Electromagnetic Waves and X-rays in Bio-medical Media by Waveguide-type Grids

Yasumitsu Miyazaki
Department of Electronics and Information Engineering, Aichi University of Technology
50-2 Manori, Nishihazama-cho, Gamagori 443-0047, Japan

Abstract— Medical image diagnosis and computer aided diagnosis are modern important medical techniques developed with computer technology. Particularly, medical image diagnosis using X-rays is very important technical tools for physiological examination of human body. Image responses of X-ray transmitted projection include X-ray scattering characteristics that disturb X-ray transmission properties through biological structures concerned with X-ray absorption effects due to biological characteristics consisting of atomic structure. In this paper, technical methods of spatial filtering for X-ray scattering superposed on transmitted and attenuated waves are discussed to improve image diagnosis. Statistical theory of X-ray is described for X-ray propagation, attenuation and scattering in random inhomogeneous biological media. Spatial filtering characteristics of grid structure are shown for exact image X-ray projection excluding scattering effects through physiological media. Spatial filtering characteristics for off-axial scattering X-rays are discussed by mode propagation properties with large attenuation for higher order modes in waveguide-type grids.

1. INTRODUCTION

Medical image diagnosis and computer aided diagnosis are very important medical techniques using X-ray. X-ray, that is, nano-meter electromagnetic wave is important physical tool for medical diagnosis and recently automatic image diagnosis using X-ray with computers is rapidly developed [1]. However, in image processing for medical diagnosis, based on photo-electric absorption, interactions between electrons and photons of X-ray, Thomson scatterings and Compton scatterings, physiological and physical phenomena of X-rays in biomedical media are not so much studied [2–5]. The spatial characteristic in the received X-ray image is determined by the intensity of transmitted and attenuated X-ray waves superposed with scattered waves. The characteristics of transmitted waves depend on both the absorption and scattering characteristics. Therefore, for the identification characteristics of biomedical media by X-ray transmission, signal to noise ratio of primitive X-ray diagnosis without scattering filtering is not so high. Although the use of the X-ray grid is efficient to remove the scattering wave from the transmitting wave, the grid has not been studied sufficiently. Also, by using spatial grid filter and the characteristic investigation of scattering, absorption, dispersion and spectroscopy properly, it may be possible to find more accurate new method of X-ray image diagnosis. In this paper, in order to find excellent computer aided diagnosis system based on the electromagnetic wave characteristics, primarily, the scattering filter characteristics of X-ray beam using the grid consisting of lossy waveguide arrays are studied. Based on this analysis, the optimum scattering wave filter may be found.

Statistical theory of X-ray propagation in random media consisting of biological tissues is discussed by integral equations with Green's function, using correlation functions of random media. Scattering characteristics of incident Gaussian beam in random physiological media consisting of biological materials such as proteins are discussed. Electromagnetic filtering properties by X-ray waveguides with lossy wall for off-axial scattered fields are discussed using impedance characteristics of waveguides consisting of lossy metal walls.

2. ELECTROMAGNETIC X-RAY SCATTERING FIELD IN BIOLOGICAL RANDOM MEDIA

In the human biological body region, using Maxwell's equations X-ray transmitted and scattered waves through biomedical media as random media are studied by statistical theory of electromagnetic waves. X-rays, nanometer electromagnetic waves are incident on random media (I) ($0 \leq z \leq \ell$) from left side in Fig. 2. Incident X-ray has y-direction linearly polarization $\mathbf{E} = \phi(x, z)\mathbf{i}_y$. Biological random media of region (I) have dielectric constants as

$$\varepsilon_t = \varepsilon + \varepsilon \Delta\eta(\mathbf{r}_t) \tag{1}$$

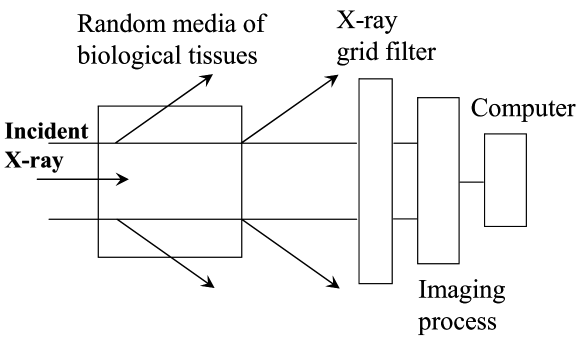


Figure 1: Computer aided diagnosis and scattering X-ray filtering system.

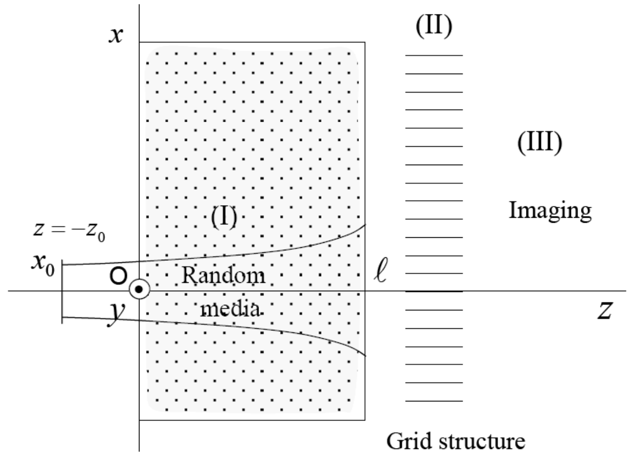


Figure 2: X-ray scattering and X-ray grid arrays.

where $\Delta\eta$ is random function and $\varepsilon = \varepsilon' - j\varepsilon''$. Here, transverse vector of 2 dimensional space is $\mathbf{r}_t = x\mathbf{i}_x + z\mathbf{i}_z$, and we consider two dimensional scattering fields. The field function $E(x,z)$ concerned with electric field of y polarization satisfies the following wave equations, using $k^2 = \omega^2\varepsilon\mu = (k_r - jk_i)^2$

$$\nabla^2_{xy}E(\mathbf{r}_t) + k^2E(\mathbf{r}_t) = -\omega^2\varepsilon\Delta\eta E(\mathbf{r}_t) \tag{2}$$

Incident Gaussian X-ray of $\mathbf{E}^{inc} = E_{inc}(\mathbf{r}_t)\mathbf{i}_y$ is written as

$$\mathbf{E}_{inc}(\mathbf{r}_t) = \mathbf{i}_y \frac{A}{\sqrt{1-j\varsigma}} e^{-jk(z+z_0)} e^{-\frac{x^2}{x_0^2(1-j\varsigma)}} \tag{3}$$

where beam parameters are $\varsigma = \frac{2(z+z_0)}{kx_0^2}$, beam waist is $z = -z_0$, and beam spot size is x_0.

Green's function $G(\mathbf{r}_t, \mathbf{r}'_t)$ for free space, is satisfying the following wave equation $\nabla^2_t G + k^2 G = -\delta(\mathbf{r}_t - \mathbf{r}'_t)$

$$G(\mathbf{r}_t, \mathbf{r}'_t) = -\frac{j}{4} H_0^{(2)}\left(k|\mathbf{r}_t - \mathbf{r}'_t|\right) \tag{4}$$

Applying Green's formula in random biological media (I) for total field $E_{totl} = E_{scatt} + E_{inc}$, using Green's function G, we have when $\underline{\mathbf{\Gamma}} = G\mathbf{I}$, using unit dyad $\mathbf{I}$,

$$\begin{aligned} E_{totl} &= E_{inc} - \frac{j}{4}\int_{S_I} dS' \left(\omega^2\varepsilon\Delta\eta\right) E_{totl}(\mathbf{r}'_t) H_0^{(2)}\left(k|\mathbf{r}_t - \mathbf{r}'_t|\right) \\ \mathbf{E}_{totl} &= \mathbf{E}_{inc} + \int k^2\Delta\eta\underline{\mathbf{\Gamma}}\left(\mathbf{r}_t, \mathbf{r}'_t\right)\cdot\mathbf{E}(\mathbf{r}'_t)dS' \end{aligned} \tag{5}$$

From Maxwell's equation, magnetic field is derived as, $\mathbf{H}_{inc} = \frac{j}{\omega\mu}\nabla\times\mathbf{E}_{inc}$

$$\mathbf{H}_{totl} = \mathbf{H}_{inc} - \frac{1}{j\omega\mu}\int k^2\Delta\eta\nabla\times\underline{\mathbf{\Gamma}}\cdot\mathbf{E}dS' \tag{6}$$

when $\mathbf{E} = E\mathbf{i}_y$, $\mathbf{H} = \frac{1}{j\omega\mu}\mathbf{i}_y\times\nabla\mathbf{E}_y$, we have $\frac{1}{-j\omega\mu}\nabla\times\underline{\mathbf{\Gamma}} = \frac{1}{j\omega\mu}(\mathbf{i}_y\times\nabla G)\mathbf{i}_y$

Based on the fundamental Equations (5) and (6) in the previous section, we calculate the scattered field by iterative method in this section. Employing the representation of functional analysis, we can solve the integral Equations (5) and (6) in operator forms. If we define vectors $\mathbf{E}_0$ and $\mathbf{E}$ in functional space for the vectors $\mathbf{E}_{inc}$ and $\mathbf{E}$, and integral operator K_E for the integral $\int k^2\Delta\eta\underline{\mathbf{\Gamma}}\cdot dS'$, we obtain the field as follows:

$$\mathbf{E} = (\mathbf{I} - K_E)^{-1}\mathbf{E}_0 = \mathbf{E}_0 + K_E\mathbf{E}_0 + K_EK_E\mathbf{E}_0 + \cdots \tag{7}$$

Using the similar definition for the magnetic field and Eq. (6), we have

$$\mathbf{H} = \mathbf{H}_0 + K_H\mathbf{E} = \mathbf{H}_0 + K_H\mathbf{E}_0 + K_HK_E\mathbf{E}_0 + K_HK_EK_E\mathbf{E}_0 + \cdots \tag{8}$$

The optical intensity, **I**, which is the most important quantity in the optical range, i.e., the mean intensity of the energy flow in the harmonic electromagnetic field is given by the real part of the complex Poynting vector $\overline{\mathbf{S}} = \frac{1}{2}\mathbf{E}\times\mathbf{H}^*$.

$$\mathbf{I} = \mathrm{Re}\overline{\mathbf{S}} = \frac{1}{2}\mathrm{Re}\left(\mathbf{E}\times\mathbf{H}^*\right) \tag{9}$$

The conjugate of the vector **H** is indicated by the sign $\mathbf{H}^*$.

Consequently, the optical intensity as shown as

$$\begin{aligned}\mathbf{I} \;=\; &\frac{1}{2}\mathrm{Re}\left\{(\mathbf{E}_0\times\mathbf{H}_0^*) + \mathbf{E}_0\times(K_H\mathbf{E}_0)^* + (K_E\mathbf{E}_0)\times\mathbf{H}_0^* + K_E\mathbf{E}_0\times(K_H\mathbf{E}_0)^*\right.\\ &\left.+(K_EK_E\mathbf{E}_0)\times\mathbf{H}^* + \mathbf{E}_0\times(K_HK_E\mathbf{E}_0)^* + \cdots\right\}\end{aligned} \tag{10}$$

The first term of the right hand side in Eq. (10) means the optical intensity of the incident wave. The succeeding terms show the light scattering. Particularly, the terms containing multiple K_s' such as $(K_HK_E\mathbf{E}_0)^*$ and $(K_EK_E\mathbf{E}_0)$ imply the multiple scattering due to inhomogeneities.

First order scattered fields corresponding to $K_E\mathbf{E}_0$ in Eq. (7) and $K_H\mathbf{E}_0$ in Eq. (8) are

$$\begin{aligned}\mathbf{E}_{scatt} &= -\frac{j}{4}k^2\int\Delta\eta(\mathbf{r}_t')H_0^{(2)}\left(k|\mathbf{r}_t-\mathbf{r}_t'|\right)E_{inc}(\mathbf{r}_t')\mathbf{i}_y dS',\\ \mathbf{H}_{scatt} &= -\frac{j}{4}\frac{k^3}{\omega\mu}\int\Delta\eta(\mathbf{r}_t')H_0^{(2)}\left(k|\mathbf{r}_t-\mathbf{r}_t'|\right)\left(\mathbf{n}(\mathbf{r}')\times E_{inc}(\mathbf{r}_t')\right)\mathbf{i}_y dS'\end{aligned} \tag{11}$$

The second order scattered fields corresponding to $K_EK_E\mathbf{E}_0$, and $K_HK_E\mathbf{E}_0$ are

$$\begin{aligned}\mathbf{E}_{scatt} &= \left(-\frac{j}{4}k^2\right)^2\mathbf{i}_y\iint\Delta\eta(\mathbf{r}_t'')\Delta\eta(\mathbf{r}_t')H_0^{(2)}\left(k|\mathbf{r}_t'-\mathbf{r}_t''|\right)H_0^{(2)}\left(k|\mathbf{r}_t-\mathbf{r}_t'|\right)E_{inc}(\mathbf{r}_t'')dS'dS''\\ \mathbf{H}_{scatt} &= \left(-\frac{j}{4}\frac{k^3}{\omega\mu}\right)\left(-\frac{j}{4}k^2\right)\iint\Delta\eta(\mathbf{r}_t')H_0^{(2)}\left(k|\mathbf{r}_t-\mathbf{r}_t'|\right)\left(\mathbf{n}(\mathbf{r}')\times\mathbf{i}_y\right)\Delta\eta(\mathbf{r}_t'')H_0^{(2)}\left(k|\mathbf{r}_t'-\mathbf{r}_t''|\right)E_{inc}(\mathbf{r}_t'')dS'dS''\end{aligned} \tag{12}$$

In case the distance $|\mathbf{r}_t-\mathbf{r}_t'|$ of the observation point from the biological random media is large compared with the wavelength, i.e., $k|\mathbf{r}_t-\mathbf{r}_t'|\gg 1$, the Greens dyadic $\underline{\mathbf{\Gamma}}$ can be asymptotically evaluated,

$$\underline{\mathbf{\Gamma}} = G\mathbf{I} = \sqrt{\frac{2}{\pi k|\mathbf{r}_t-\mathbf{r}_t'|}}e^{-j\left(k|\mathbf{r}_t-\mathbf{r}_t'|-\frac{\pi}{4}\right)} \tag{13}$$

The unit vector $\mathbf{n}'$ are directed from the point in the random media toward the point of observation and given by $\mathbf{n}'=\mathbf{n}(\mathbf{r}_t')=(\mathbf{r}_t-\mathbf{r}_t')/|\mathbf{r}_t-\mathbf{r}_t'|$. With the help of Eq. (13), the integral operation K_E of $K_E\mathbf{E}_0$ in Eq. (11) is described as asymptotic fields,

$$K_E\mathbf{E}_0 \approx -\left(\frac{j}{4}\right)\int k^2\Delta\eta\sqrt{\frac{2}{\pi k|\mathbf{r}_t-\mathbf{r}_t'|}}e^{-j\left(k|\mathbf{r}_t-\mathbf{r}_t'|-\frac{\pi}{4}\right)}\left(E_{inc}\mathbf{i}_y\right)dS' \tag{14}$$

Similarly, the integral operation K_H of $K_H\mathbf{E}_0$ is given by

$$K_H\mathbf{E}_0 \approx \frac{-k}{\omega\mu}k^2\left(\frac{j}{4}\right)\int\Delta\eta\sqrt{\frac{2}{\pi k|\mathbf{r}_t-\mathbf{r}_t'|}}e^{-j\left(k|\mathbf{r}_t-\mathbf{r}_t'|-\frac{\pi}{4}\right)}\left(\mathbf{n}'\times E_{inc}\mathbf{i}_y\right)dS' \tag{15}$$

Hence, the dominant term of the scattered field intensity in the inhomogeneous medium, $\frac{1}{2}\mathrm{Re}(K_E\mathbf{E}_0\times(K_H\mathbf{E}_0)^*)$, can be written as

$$\begin{aligned}\frac{1}{2}\mathrm{Re}\left(K_E\mathbf{E}_0\times(K_H\mathbf{E}_0)^*\right) \;=\; &\frac{k}{\omega\mu}\left(\frac{k^2}{4}\right)^2\frac{2}{\pi k}\mathrm{Re}\iint\frac{e^{-jk_r|\mathbf{r}_t-\mathbf{r}_t'|+jk_r|\mathbf{r}_t-\mathbf{r}_t''|}}{\sqrt{|\mathbf{r}_t-\mathbf{r}_t'||\mathbf{r}_t-\mathbf{r}_t''|}}\\ &e^{-k_i|\mathbf{r}_t-\mathbf{r}_t'|}e^{-k_i|\mathbf{r}_t-\mathbf{r}_t''|}\Delta\eta(\mathbf{r}_t')\Delta\eta^*(\mathbf{r}_t'')\mathbf{n}''E_{inc}'E_{inc}''dS'dS''\end{aligned} \tag{16}$$

where we define $\mathbf{E}_0'=\mathbf{E}_0(\mathbf{r}_t')$, $\mathbf{E}_0''=\mathbf{E}_0(\mathbf{r}_t'')$ and $\mathbf{n}''=\mathbf{n}(\mathbf{r}_t,\mathbf{r}_t'')=(\mathbf{r}_t,\mathbf{r}_t'')/|\mathbf{r}_t,\mathbf{r}''|$. These equations are the fundamental formulae for the scattered field. In the next section the statistical properties of these equations are discussed.

3. STATISTICAL PROPERTIES OF SCATTERED FIELD IN BIOLOGICAL RANDOM MEDIA

The scattered field intensity of the incident wave $\mathbf{E}_0$ due to inhomogeneities $\Delta\eta$ in the medium is obtained from Eq. (10). If the statistical average $\langle\Delta\eta\rangle$ of fluctuations in random media is equal to zero, the first order terms $\langle \mathbf{E}_0 \times (K_H \mathbf{E}_0)^* \rangle$ and $\langle (K_E \mathbf{E}_0) \times \mathbf{H}_0^* \rangle$ in Eq. (10) vanish, when the symbol $\langle\rangle$ indicates an ensemble average.

In the case of stationary fluctuations $\Delta\eta(\mathbf{r}')\Delta\eta(\mathbf{r}'')$ and the isotropic case, we can introduce the correlation function by the definition

$$\left\langle \Delta\eta\left(\mathbf{r}'_t\right)\Delta\eta\left(\mathbf{r}''_t\right)\right\rangle = B_\eta\left(\mathbf{r}'_t - \mathbf{r}''_t\right)$$

Consequently, from Eq. (10) the statistical average of the total field is expressed as

$$\langle \mathbf{I} \rangle = \mathbf{I}_0 + \frac{1}{2}\mathrm{Re}\left\langle \{K_E \mathbf{E}_0 \times (K_H \mathbf{E}_0)^* + \mathbf{E}_0 \times (K_H K_E \mathbf{E}_0)^* + (K_E K_E \mathbf{E}_0) \times \mathbf{H}_0^* + \cdots\}\right\rangle + \cdots \tag{17}$$

where $\mathbf{I}_0 = \frac{1}{2}\mathrm{Re}(\mathbf{E}_0 \times \mathbf{H}_0^*)$ is the intensity of the incident wave and the other term indicates the scattering intensity. For weakly fluctuated medium the second order term $\langle K_E \mathbf{E}_0 \times (K_H \mathbf{E}_0)^* \rangle$ provides the dominant contribution to the total scattering and higher order terms of the multiple scattering concerned with higher order correlations of the inhomogeneities can be neglected. The multiple scattering, that must be considered in strongly fluctuated medium contributes to the total scattering. In this paper we study the dominant term of the X-ray scattering, which leads to the following total intensity $\langle \mathbf{I} \rangle$

$$\langle \mathbf{I} \rangle \approx \mathbf{I}_0 + \frac{1}{2}\mathrm{Re}\left\langle K_E \mathbf{E}_0 \times (K_H \mathbf{E}_0)^* \right\rangle \tag{18}$$

If the observation point is far from the sample, the condition $|\mathbf{r}_t - \mathbf{r}'_t| \gg \rho_t = |\mathbf{r}'_t - \mathbf{r}''_t|$ is satisfied. It is then convenient to use a Taylor expansion of $\mathbf{n}'' - \mathbf{n}'$ in the terms of the distance $\boldsymbol{\rho}_t = \mathbf{r}'_t - \mathbf{r}''_t$ vector between two different points $\mathbf{r}'_t$ and $\mathbf{r}''_t$ in the sample. In this case we can write approximately the scattered field in Eq. (16) as follows:

$$\langle \mathbf{I}_s \rangle \approx \frac{k}{\omega\mu}\left(\frac{k^2}{4}\right)^2 \frac{2}{\pi k}\mathrm{Re}\iint \frac{e^{-jk_r|\mathbf{r}_t - \mathbf{r}'_t| + jk_r|\mathbf{r}_t - \mathbf{r}''_t|}}{\sqrt{|\mathbf{r}_{t'} - \mathbf{r}'_t||\mathbf{r}_t - \mathbf{r}''_t|}} e^{-k_i|\mathbf{r}_t - \mathbf{r}'_t|} e^{-k_i|\mathbf{r}_t - \mathbf{r}''_t|} B_\eta\left(\mathbf{r}'_t, \mathbf{r}''_{t'}\right)\mathbf{n}'' E'_{inc} E''^*_{inc} dS' dS'' \tag{19}$$

This gives the scattered field derived, with the help of the Green's dyadic, for the autocorrelation function $B_\eta\left(\mathbf{r}'_t, \mathbf{r}''_{t'}\right)$ of random inhomogeneities and an arbitrary incident wave $\mathbf{E}_{inc}$.

The incident X-ray beam of Eq. (3) is the Gaussian TEM_{00} mode. TEM_{00} mode decays rapidly with increasing the transverse distance from the beam center. If beam has a beam width x_0 and a plane phase front at the beam waist $z = -z_0$, and is linearly polarized in the transverse direction of $\mathbf{i}_y$, the field at any given point can be written in coordinates (x, z) as

$$\mathbf{E}_{inc}(\mathbf{r}) = \mathbf{i}_y \frac{A}{\sqrt{1 - j\frac{2(z+z_0)}{kr_0}}} e^{-jk(z+z_0)} \exp\left(-\frac{j2x^2(z+z_0)}{kx_0^4\left\{1 + \frac{4(z+z_0)^2}{(kx_0^2)^2}\right\}}\right) \exp\left(-\frac{r^2}{x_0^2\left\{1 + \frac{4(z+z_0)^2}{(kx_0^2)^2}\right\}}\right) \tag{20}$$

The first exponential factor describes the phase of a plane wave and the second exponential factor is responsible for phase front curvature. The last exponential factor determines the intensity in the transverse direction.

Taylor expansions as $|\mathbf{r}_t - \mathbf{r}''_t| = |\mathbf{r}_t - \mathbf{r}'_t| + \mathbf{n}' \cdot \boldsymbol{\rho}_t + \dfrac{1}{(2|\mathbf{r}_t - \mathbf{r}'_t|)}\left\{\boldsymbol{\rho}_t^2 - (\boldsymbol{\rho}_t \cdot \mathbf{n}')^2\right\} + \cdots$, $\mathbf{r}_t''^2 = \mathbf{r}_t'^2 + \boldsymbol{\rho}_t \cdot \mathbf{i}_z(2\mathbf{r}'_t - \boldsymbol{\rho}) \cdot \mathbf{i}'_t - 2\mathbf{r}'_t \cdot \boldsymbol{\rho}_t + \boldsymbol{\rho}_t^2$, $\mathbf{n}' = \mathbf{n}_0 + \{\mathbf{n}_0(\mathbf{n}_0 \cdot \mathbf{r}'_t) - \mathbf{r}'_t\}/|\mathbf{r}|$ and $\boldsymbol{\rho}_t = \mathbf{r}'_t - \mathbf{r}''_t$, $\mathbf{n}(o) = \mathbf{n}_0$. We have scattered field intensities in random media, from Eq. (19) as

$$\begin{aligned}\langle \mathbf{I}_s \rangle \cong \; & \frac{k}{\omega\mu}\left(\frac{k^2}{4}\right)^2 \frac{2}{\pi k}\mathrm{Re}\iint \frac{e^{jk_r(\mathbf{n}' - \mathbf{i}_z)\cdot\boldsymbol{\rho}_t}}{|\mathbf{r}_t - \mathbf{r}'_t|} e^{-k_i(\mathbf{n}' - \mathbf{i}_z)\cdot\boldsymbol{\rho}_t} \\ & B_\eta(\boldsymbol{\rho})\mathbf{n}' A^2 g\left(k_i|\mathbf{r}_t - \mathbf{r}'_t|\right) e^{-2k_i(z'+z_0)} \frac{e^{-\frac{2x'^2}{x_0^2(1+\varsigma_0^2)}}}{\left\{1 + \varsigma_0^2\right\}^{1/2}} f\left(\boldsymbol{\rho}_t, \mathbf{r}'_t\right) d^2\mathbf{r}' d^2\boldsymbol{\rho}_t\end{aligned} \tag{21}$$

where wave attenuation factor along propagation axis is $g(k_i|\mathbf{r}_t-\mathbf{r}'_t|)=e^{-2k_i|\mathbf{r}_t-\mathbf{r}'_t|}$, inside random media of $\mathbf{r}_t$, $g=1$, outside, and

$$\varsigma_0=2z_0/kx_0^2\cdot f(\boldsymbol{\rho}_t)=\exp\left[-\left\{\boldsymbol{\rho}_t^2-2\mathbf{r}'_t\cdot\boldsymbol{\rho}_t+\boldsymbol{\rho}_t\cdot\mathbf{i}_z\left(2\mathbf{r}'_t-\boldsymbol{\rho}_t\right)\cdot\mathbf{i}_z\right\}/x_0^2\left\{1+j\left(2z_0/\left(kx_0^2\right)\right)\right\}\right],$$

and

$$\left(1-\mathbf{n}'\cdot\mathbf{i}_z\right)=\left(1-\left(z-z'\right)/|\mathbf{r}_t|\right)-\left(z-z'\right)\left(rr'\cos(\theta-\theta')+zz'\right)/|\mathbf{r}_t|^3+\left(z-z'\right)\left(r'^2+z'^2\right)/2|\mathbf{r}_t|^3+\cdots$$

When the correlation function is $B_\eta(\boldsymbol{\rho}_t)=\overline{\Delta\eta^2}e^{-\rho^2/\rho_0^2}$, where parameters $\overline{\Delta\eta^2}$ and ρ_0 give the variances and correlation lengths of fluctuations,

$$\int B_\eta(\boldsymbol{\rho}_t)e^{jk_r(\mathbf{n}'-\mathbf{i}_z)\cdot\boldsymbol{\rho}_t}d^2\boldsymbol{\rho}_t=\overline{\Delta\eta^2}\rho_0^2\pi e^{-k_r^2\rho_0^2\left(\frac{1-\mathbf{n}'\cdot\mathbf{i}_z}{2}\right)} \tag{22}$$

From Equations (21), (22), when $\rho_0<x_0$, length of random media Z_ℓ, we have X-ray scattering intensity in lossy random media for the cylindrical coordinate (r,θ), assuming $g(k_i|\mathbf{r}_t-\mathbf{r}'_t|)=e^{-2k_i|Z_\ell-Z'|}$ in random media of $0\le z\le z_\ell$,

$$\begin{aligned}\langle\mathbf{I}_s\rangle&=\mathbf{i}_r\frac{k}{\omega\mu}\left(\frac{k^2}{4}\right)^2\frac{\sqrt{2\pi}}{4}A^2\overline{\Delta\eta^2}x_0\rho_0^2\frac{\ell}{|\mathbf{r}|}e^{-k_r^2\rho_0^2\sin^2\frac{\theta}{2}}\\&=\mathbf{i}_r\frac{k}{\omega\mu}\frac{\sqrt{2\pi}}{16}\overline{\Delta\eta^2}A^2k^3x_0\rho_0^2\frac{1}{|\mathbf{r}|}e^{-k_r^2\rho_0^2\sin^2\frac{\theta}{2}}e^{-2k_i(z_\ell+z_0)}z_\ell\end{aligned} \tag{23}$$

4. ELECTROMAGNETIC CHARACTERISTICS IN LOSSY GRID WAVEGUIDES

Electromagnetic characteristics of X-ray and nano-meter electromagnetic waves in lossy waveguide arrays consisting of grid structures are expressed by mode expansions of lossy modes. In the region (II) of grid arrays in Fig. 3, propagation spaces are $z_{g1}\le z\le z_{g2}$, $-a/2+s(a+d)\le x\le$

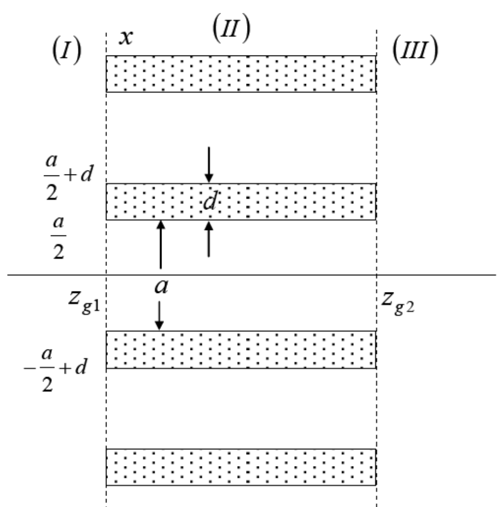


Figure 3: Waveguide type grid structure for scattering filtering.

$a/2+s(a+d)$, $s=-m,\ -(m-1),\ \cdots,\ -1,\ 0,\ 1,\ \cdots,\ m-1,\ m$ and lossy metal materials for X-rays are $z_{g1}\le z\le z_{g2}$, $a/2+s(a+d)\le x\le a/2+d+s(a+d)$.

When for the TE mode systems,

$$\psi_n=A_0\cos\frac{n\pi}{a}\left(x+\frac{a}{2}\right) \tag{24}$$

for, $-a/2+s(a+d)\le x\le a/2+s(a+d)$, 0 for $a/2+s(a+d) \le x \le a/2+d+s(a+d)$, $A_0 = \sqrt{2/a}(n\pi/a)^{-1}$, and $h_n = n\pi/a$, fields in waveguide are

$$\begin{aligned}
H_z &= \sum B_n h_n^2 \psi_n(x) e^{j\omega t-\gamma_n z} \\
E_y &= j\omega\mu \mathbf{i}_z \times \sum_{n=1}^{\infty} B_n \nabla_t \psi_n e^{j\omega t-\gamma_n z} = j\omega\mu \sum_{n=1}^{\infty} B_n \left(-\frac{n\pi}{a}\right) A_0 \sin\frac{n\pi}{a}\left(x+\frac{a}{2}\right) e^{j\omega t-\gamma_n z} \\
H_x &= -\sum_{n=1}^{\infty} B_n \gamma_n \nabla_t \psi_n e^{j\omega t-\gamma_n z} = -\sum_{n=1}^{\infty} B_n \gamma_n \left(-\frac{n\pi}{a}\right) A_0 \sin\frac{n\pi}{a}\left(x+\frac{a}{2}\right) e^{j\omega t-\gamma_n z}
\end{aligned} \tag{25}$$

X-ray intensity can be expressed as mode expansions

$$P = \frac{1}{2}\mathrm{Re}\int \mathbf{i}_z \cdot \mathbf{E}_t \times \mathbf{H}_t^* dS \tag{26}$$

In case of lossy waveguide walls of conductivity σ, impedance boundary condition is $\mathbf{n}\times\mathbf{E} = \frac{1}{2}\delta\mu\omega(1+j)\mathbf{H}$, $\delta = \sqrt{2/\omega\mu\sigma}$, propagation constant $\beta_n = \beta_n^{(r)} - j\beta_n^{(i)}$ are derived as perturbations from waveguide modes for perfect conductivity, when $R_s = \sqrt{\frac{\omega\mu}{2\sigma}}$, $\varsigma = \sqrt{\frac{\mu}{\varepsilon}}$,

$$\beta_n^{(r)} = \sqrt{k^2 - \left(\frac{n\pi}{a}\right)^2}, \qquad \beta_n^{(i)} = \frac{R_s}{\varsigma a}\frac{2k\left(h_n/k\right)^2}{\sqrt{k^2-h_n^2}} \tag{27}$$

Hence, we have

$$\gamma_n = j\beta_n^{(r)} + \beta_n^{(i)}, \qquad P = \frac{1}{2}\omega\mu\sum_{n=1}^{\infty}|B_n|^2\beta_n^{(r)} e^{-\beta_n^{(i)} z} \tag{28}$$

Using large attenuation constants of higher modes, spatial filtering of incident off-axis scattered fields can be estimated.

5. SCATTERED FIELD FILTERING BY GRID WAVEGUIDES

When scattered fields in random media expressed by Eqs. (11) and (12) are incident to grid waveguides, off-axis scattered fields with large scattering angles excite higher modes in grid waveguides and attenuate rapidly in grid waveguides with lossy walls. Spatial filtering corresponding to higher mode attenuation in lossy grid waveguides can be evaluated by mode coefficients in the sth grid waveguides.

Incident waves at the input of the grid waveguide are described by mode coefficients as, from Eq. (5), at surface $S_{g1}^{(s)}$ of the sth grid waveguides.

$$\begin{aligned}
B_n^{(s)} &= B_n^{(s),(inc)} + B_n^{(s),(scatt)}, \qquad B_n^{(s),(inc)} = \frac{1}{C_n}\int_{S_{g1}^{(s)}} \left[E_{inc}(\mathbf{r}_t''')\right] \sin\frac{n\pi}{a}\left(x_s'''+\frac{a}{2}\right) dS''' \\
B_n^{(s),(scatt)} &= \frac{1}{C_n}\int_{S_{g1}^{(s)}} \left[-\frac{j}{4}k^2\int \Delta\eta(\mathbf{r}_t')H_0^{(2)}\left(k|\mathbf{r}'''-\mathbf{r}_t'|\right)E_{inc}(\mathbf{r}_t')dS'\right] \sin\frac{n\pi}{a}\left(x_s'''\frac{a}{2}\right) dS'''
\end{aligned} \tag{29}$$

where $C_n = j\omega\mu\left(-\frac{n\pi}{a}\right)A_0\frac{a}{2}$, surface $S_{g1}^{(s)}$ is $z = z_{g1}$, $-a/2+s(a+d) \le x \le a/2+s(a+d)$, $x_s''' = x'''+s(a+d)$.

Mode power characteristics are shown as statistical coefficients as, from Eq. (28),

$$\langle|B_n^{(s)}|^2\rangle = |B_n^{(s),(inc)}|^2 + |B_n^{(s),(scatt)}|^2 \tag{30}$$

where

$$|B_n^{(s),(scatt)}|^2 = \frac{1}{|C_n|^2}\frac{A^2}{\sqrt{1+\varsigma^2}}e^{-2k_i(z_\ell+z_0)}e^{-\frac{2x_s^2}{x_0^2(1+\varsigma^2)}}\left[\frac{a}{n\pi}\left(1-(-1)^n\right)\right]^2 \tag{31}$$

$$\begin{aligned}\left\langle|B_n^{(s),(scatt)}|^2\right\rangle &= \frac{1}{|C_n|^2}\iint_{S_{g1}^{(s)}}\iint\left\langle E_{scatt}(\mathbf{r}''',\mathbf{r}'_t)E^*_{scatt}\left(\mathbf{r}^{IV},\mathbf{r}''_t\right)\right\rangle dS'dS'' \\ &\quad \sin\frac{n\pi}{a}\left(x'''_s+\frac{a}{2}\right)\sin\frac{n\pi}{a}\left(x_s^{IV}+\frac{a}{2}\right)dS'''dS^{IV} \\ &= \frac{1}{|C_n|^2}2\sqrt{\frac{\mu}{\varepsilon}}\frac{k}{\omega\mu}\frac{\sqrt{2\pi}}{16}\overline{\Delta\eta^2}A^2k^3x_0\rho_0^2\frac{1}{|\mathbf{r}|}e^{-k^2\rho_0^2\sin^2\frac{\theta_s}{2}}e^{-2k_i(z_\ell+z_0)}z_\ell|g(n,\theta)|^2\end{aligned} \tag{32}$$

Here

$$\tan\theta_s=\frac{x_s}{|\mathbf{r}|},\qquad g\left(n,\,\theta_s\right)=\begin{cases}\dfrac{n\pi/a}{(n\pi/a)^2-k^2\sin^2\theta_s} & \text{when}\quad \dfrac{n\pi}{a}\neq k\sin\theta_s \\ a/2 & \qquad\qquad \dfrac{n\pi}{a}=k\sin\theta_s\end{cases}$$

From Eqs. (27), (28) and (32), it is found that off-axis scattered fields of scattering angle θ_s excite higher n modes as $k\sin\theta_s\cong n\pi/a$ with large attenuations as $\beta_n^{(i)}\propto h_n^2=(n\pi/a)^2$. Lossy grid waveguide array may keep only transmitted waves with absorption effects due to atomic level energy structures of biological tissues, by filtering of scattered X-ray waves. Hence, lossy grid waveguides are very useful X-ray elements to improve X-ray image resolution for medical image processings.

6. CONCLUSION

The electromagnetic scattering and transmitted characteristics through X-ray grids are shown for medical image processing and computer aided diagnosis. X-ray grids have the spatial filter characteristics necessary to improve signal processing of the receiving characteristics in X-ray diagnosis. Based on this theory, the identification method of bio-medical image using transmitting and absorption characteristics of X-ray in medical media may be improved, and also it may be possible to utilize the scattering and spectroscopic characteristics effective for developing more accurate X-ray image diagnosis.

REFERENCES

1. Aichinger, H., et al., *Radiation Exposure and Image Quality in X-Ray Diagnostic Radiology*, 2003.
2. Miyazaki, Y., *Trans. IEE Japan*, Vol. 83, No. 2, 132–136, 2000.
3. Miyazaki, Y., *The Papers of Technical Meeting on Electromagnetic Theory, IEE Japan*, EMT05-77, 2005.
4. Miyazaki, Y., *Jpn. Jour. Appl. Phys.*, Vol. 13, No. 8, 1238–1248, 1974.
5. Miyazaki, Y., *Proc. of PIERS 2006*, Tokyo, 567, 2006.